
A ABORDAGEM DE CADASTRO DUPLO (DUAL FRAME): ESTIMAÇÃO
ASSISTIDA POR MODELOS LINEARES COM APLICAÇÃO EM PESQUISAS
AGROPECUÁRIAS

HEMÍLIO FERNANDES CAMPOS COELHO

Orientador: Prof. Cristiano Ferraz, PhD

Área de Concentração: Estatística Aplicada

Dissertação submetida como requerimento parcial para obtenção do
grau de Mestre em Estatística pela Universidade Federal de Pernambuco

Recife, fevereiro de 2007

Coêlho, Hemílio Fernandes Campos.

**A abordagem de cadastro duplo (dual frame):
estimação assistida por modelos lineares com aplicação
em pesquisas agropecuárias / Hemílio Fernandes Campos
Coêlho – Recife : O autor, 2007.**

xiv, 237p. : il., fig., tab., quadros.

**Dissertação (mestrado) – Universidade Federal de
Pernambuco. CCEN. Estatística, 2007.**

Inclui bibliografia e apêndices.

**1. Amostragem. 2. Cadastro duplo 3. Estimação
assistida por modelos 4. Pesquisas agropecuárias. I.
Título.**

519.93

CDD (22.ed.)

MEI2007-010

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

09 de fevereiro de 2007
(data)

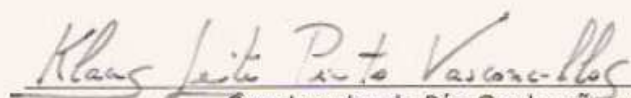
Nós recomendamos que a dissertação de mestrado de autoria de

Hemílio Fernandes Campos Coelho

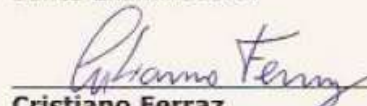
intitulada

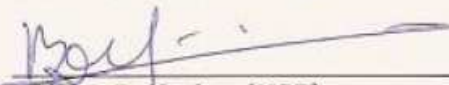
"A abordagem de cadastro duplo (dual frame): estimação assistida por modelos lineares com aplicação em pesquisas agropecuárias"

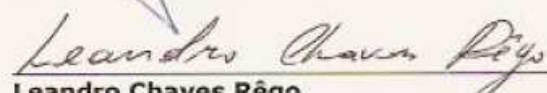
seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.


Coordenador da Pós-Graduação em Estatística

Banca Examinadora:


Cristiano Ferraz orientador


Heleno Bolfarine (USP)


Leandro Chaves Rêgo

Este documento será anexado à versão final da dissertação.

Dedico este trabalho à minha mãe.

Agradecimentos

À Deus, por ter me dado a vida, saúde e o caminho para atingir meus objetivos.

À minha mãe Wilma Arruda Campos Coelho, pelo exemplo de vida, pela luta em criar cinco filhos e dar uma boa educação, por ser uma fortaleza nas horas difíceis, e por ser a principal incentivadora dos meus estudos desde à minha opção pelo curso de Estatística.

Ao meu pai Hamilton José Costa Coelho, pelo incondicional apoio aos meus estudos desde meu primeiro dia na escola.

Aos meus irmãos Pedro, Antonio, Daniel e Danilo pelo apoio.

Ao meu tio Jadiel Sobreira (in memoriam), pelo exemplo de pai, amigo, irmão, e que de algum modo está incentivando meus trabalhos, e para sua esposa, minha tia Nilma, que o mantém vivo em nossos corações, através do amor.

À minha avó Maria José, que carinhosamente chamo de Mãe, desde a infância.

Ao meu orientador Cristiano Ferraz, pela paciência, pelos valiosos conselhos e por ter acreditado na minha capacidade de desenvolver esse trabalho.

À Merciany Rodrigues Ferreira, minha vida, minha alma e meu amor, que me proporciona momentos de alegria e amor desde o início dos meus estudos no curso de Estatística.

A todos os meus tios paternos, em especial Franklin, Gracinha, Rosário e Roberto, pela constante alegria e apoio aos meus trabalhos.

À minha prima Sílvia, pelo apoio e amizade em todos os momentos de conversa sobre os desafios da vida acadêmica.

Ao Instituto Brasileiro de Geografia e Estatística (IBGE), pela oportunidade que tive de participar do curso sobre cadastro duplo oferecido no Rio de Janeiro.

À Coordenação de Agropecuária (COAGRO), na pessoa de Flávio Bolliger.

Ao engenheiro agrônomo Marcelo Duriez, gerente de análise e planejamento da COAGRO - IBGE, pelos importantes esclarecimentos fornecidos sobre pesquisas agropecuárias.

Aos estatísticos Jeffrey Bailey e Chadd Crouse, do National Agricultural Statistics Service of the United States Department of Agriculture (NASS/USDA) pelo material fornecido, contribuição na qualidade desse trabalho e pela amizade durante o período que passei no Rio de Janeiro.

A todos os amigos de pós-graduação, em especial Themis Abensur, Luz Marina, Marcelo Rodrigo, Lílían Santos, Rejane Brito, Katya Rodrigues, Leonardo Bomfim, Ênio Lopes, Geraldo Melo, Andrea Prudente, Silvia Lemos, Abraão David, Valmir Rogério e Murilo, os quais demonstraram incondicional alegria e fraternidade nos momentos bons e difíceis.

Aos meus grandes amigos da graduação, em especial Vanessa Santos, Carlos Renato, Claudyvan Paiva, Robson Florêncio (Robinho), Romero Filho, Juliana Kátia, Raul Siqueira, Tadeu Rodrigues, Leila Rameh, Danielly Andrade, Oscar Raposo, Luiz Medeiros, Carlos Francisco, Cristiano Pastichi, Lucas Gallindo, Luciano Souza, Syntia Regina, Emmanuelle Araújo e Pedro Santana, por todos os momentos bons que vivenciamos desde a graduação e pelo contínuo apoio aos meus estudos.

À Valéria Bittencourt, pelo exemplo de profissionalismo e também pelo enorme carinho, amizade e paciência que tem pelos alunos da pós-graduação.

Aos professores do Departamento de Estatística, em especial aos professores Maria Cristina Falcão Raposo, Klaus Leite Pinto Vasconcellos, Francisco Cribari Neto e Sylvio José Pereira dos Santos, pelo apoio, orientação e amizade.

À banca examinadora, pelas valiosas sugestões que enriqueceram a qualidade desse trabalho.

Ao CNPq, pelo apoio financeiro.

*“O tempo é o tempo que não é tempo, pois quando
falamos a palavra tempo ela já virou passado. Caminhe
em busca dos seus sonhos e nunca desista enquanto
percorrer o caminho, pois o tempo mostrará o resultado
dessa busca.”*

Meu tio, Jadiel Sobreira da Costa (in Memorian)

Resumo

O presente trabalho apresenta técnicas de estimação de totais e médias populacionais sob a abordagem de cadastro duplo (dual frame), incorporando informações de variáveis auxiliares ao processo de estimação, sob a assistência de modelos lineares. A abordagem de cadastro duplo é definida como um levantamento amostral onde dois cadastros são utilizados para identificar elementos de uma única população-alvo. Considerando este contexto, diversos estimadores assistidos por modelos lineares são propostos. Seus desempenhos são comparados com estimadores encontrados na literatura e avaliados com base no método de simulação de Monte Carlo. Resultados são apresentados para o plano de amostragem aleatória simples e para o plano de amostragem estratificada. As técnicas propostas têm potencialidade de aplicação direta em pesquisas agropecuárias.

Palavras-chave: Cadastro duplo, estimador razão, pesquisas agropecuárias.

Abstract

In this thesis, the problem of estimating population totals and means, incorporating auxiliary information at the estimation stage, under a dual frame sampling approach, is considered. A dual frame sampling approach is defined as a survey sample design with two sample frames providing coverage for the same target population. Under this context, several linear model-assisted estimators are proposed for simple random sampling and stratified random sampling designs. Their performance is compared with estimators found in literature and investigated for several dual frame scenarios by a Monte Carlo simulation study. The proposed estimators are motivated by their potential application to agricultural research.

Keywords: Dual frame, model-assisted estimator, agricultural research.

Lista de Figuras	ix
Lista de Tabelas	xii
1 Introdução	1
2 Estimação sob a abordagem de cadastro duplo	5
2.1 Estimador de Hartley	8
2.2 Estimação sob o cenário 2	9
2.3 Estimação sob o cenário 3	10
2.4 Estimador de Lund	11
2.5 Estimador de Fuller & Burmeister	12
3 Aplicações da abordagem de cadastro duplo em pesquisas agropecuárias	14
4 Estimação assistida por modelos em uma abordagem de cadastro duplo	20
4.1 Estimação de N_{ab}	22
4.2 Estimadores do tipo razão sob a abordagem de cadastro duplo	24
4.2.1 Estimador razão 1	24
4.2.2 Estimador razão 2	25

4.2.3	Estimador razão 3	28
4.2.4	Estimador razão 4	29
5	Estimação assistida por modelos em uma abordagem de cadastro duplo sob estratificação	31
5.1	Estimadores do tipo razão separado sob a abordagem de cadastro duplo	33
5.1.1	Estimador razão separado 1	34
5.1.2	Estimador razão separado 2	36
5.1.3	Estimador razão separado 3	40
5.1.4	Estimador razão separado 4	42
5.2	Estimadores do tipo razão combinado sob a abordagem de cadastro duplo	44
5.2.1	Estimador razão combinado 1	44
5.2.2	Estimador razão combinado 2	46
5.2.3	Estimador razão combinado 3	48
5.2.4	Estimador razão combinado 4	49
6	Avaliação Numérica	51
6.1	Resultados	55
6.1.1	Estudo sobre a distribuição de $\hat{N}_{ab,s}$	55
6.1.2	Desempenho dos estimadores para o plano AAS	62
6.1.3	Desempenho dos estimadores separados para o plano AAE	65
6.1.4	Desempenho dos estimadores combinados para o plano AAE	68
7	Considerações Finais	126
	Apêndice A - Variâncias aproximadas dos estimadores sob o plano AAS	131
	Apêndice B - Obtenção de p para os estimadores propostos	148
	Apêndice C - Variâncias aproximadas dos estimadores sob o plano AAE	160
	Apêndice D - Programas utilizados na avaliação numérica	178
	Referências	235

Lista de Figuras

1.1	Uso de cadastro duplo para cobertura da população alvo	2
1.2	Uso de cadastro duplo aliado à relação custo-benefício	3
3.1	Exemplo de seleção de segmento de área	17
3.2	Segmento escolhido e subdivisão do segmento em estratos	18
6.1	Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 500$	56
6.2	Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 1000$	57
6.3	Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 1500$	58
6.4	Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 2000$	59
6.5	Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 2500$	60
6.6	Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 3000$	61
6.7	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a =$ $N_b = 2000$ e $N_{ab} = 500$, sob o cenário 2	72
6.8	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a =$ $N_b = 2000$ e $N_{ab} = 500$, sob o cenário 3	73
6.9	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a =$ $N_b = 1750$ e $N_{ab} = 1000$, sob o cenário 2	75
6.10	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a =$ $N_b = 1750$ e $N_{ab} = 1000$, sob o cenário 3	76

6.11	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a =$ $N_b = 1500$ e $N_{ab} = 1500$, sob o cenário 2	78
6.12	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a =$ $N_b = 1500$ e $N_{ab} = 1500$, sob o cenário 3	79
6.13	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a =$ $N_b = 2000$ e $N_{ab} = 500$ e $n_{\mathcal{A}} = 1000$, sob o cenário 2	81
6.14	Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a =$ $N_b = 2000$ e $N_{ab} = 500$ e $n_{\mathcal{A}} = 1000$, sob o cenário 3	82
6.15	Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 2000$ e $N_{ab} = N_{\mathcal{B}} = 500$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$) .	84
6.16	Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 2000$ e $N_{ab} = N_{\mathcal{B}} = 500$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	85
6.17	Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1750$ e $N_{ab} = N_{\mathcal{B}} = 1000$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	87
6.18	Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1750$ e $N_{ab} = N_{\mathcal{B}} = 1000$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	88
6.19	Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1500$ e $N_{ab} = N_{\mathcal{B}} = 1500$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	90
6.20	Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1500$ e $N_{ab} = N_{\mathcal{B}} = 1500$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	91
6.21	Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1250$ e $N_{ab} = N_{\mathcal{B}} = 2000$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	93
6.22	Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1250$ e $N_{ab} = N_{\mathcal{B}} = 2000$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	94
6.23	Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1000$ e $N_{ab} = N_{\mathcal{B}} = 2500$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	96
6.24	Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1000$ e $N_{ab} = N_{\mathcal{B}} = 2500$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	97
6.25	Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 750$ e $N_{ab} = N_{\mathcal{B}} = 3000$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$) .	99
6.26	Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 750$ e $N_{ab} = N_{\mathcal{B}} = 3000$ ($U_{\mathcal{A}} \subset U_{\mathcal{B}}$)	100
6.27	Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yH}$, $\hat{t}_{yr1s}$ e $\hat{t}_{yr2s}$, com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 2.	102
6.28	Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yFB}$, $\hat{t}_{yr3s}$ e $\hat{t}_{yr4s}$ com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 3.	103
6.29	Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yH}$, $\hat{t}_{yr1s}$ e $\hat{t}_{yr2s}$ com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 2.	105

6.30	Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yFB}$, $\hat{t}_{yrs3}$ e $\hat{t}_{yrs4}$ com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 3.	106
6.31	Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yH}$, $\hat{t}_{yr1s}$ e $\hat{t}_{yr2s}$ com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 2.	108
6.32	Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yFB}$, $\hat{t}_{yr3s}$ e $\hat{t}_{yr4s}$ com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , , sob o cenário 3.	109
6.33	Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc1}$ e $\hat{t}_{yrc2}$, com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 2.	115
6.34	Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc3}$ e $\hat{t}_{yrc4}$, com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 3.	116
6.35	Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc1}$ e $\hat{t}_{yrc2}$, com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 2.	118
6.36	Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc3}$ e $\hat{t}_{yrc4}$, com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 3.	119
6.37	Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc1}$ e $\hat{t}_{yrc2}$, com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 2.	121
6.38	Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc3}$ e $\hat{t}_{yrc4}$, com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 3.	122

Lista de Tabelas

2.1	Cenários possíveis em uma abordagem de cadastro duplo	6
2.2	Notação utilizada em uma abordagem de cadastro duplo	7
5.1	Notação utilizada em uma abordagem de cadastro duplo, sob o plano AAE	33
6.1	Populações geradas para U_a , U_b e U_{ab}	52
6.2	Populações geradas para $U_B \subset U_A$	52
6.3	Avaliação dos estimadores propostos sob os cenários 2 e 3 para o plano AAS ($N_a = N_b = 2000$ e $N_{ab} = 500$)	71
6.4	Avaliação dos estimadores propostos sob os cenários 2 e 3 para o plano AAS ($N_a = N_b = 1750$ e $N_{ab} = 1000$)	74
6.5	Avaliação dos estimadores propostos sob os cenários 2 e 3 para o plano AAS ($N_a = N_b = 1500$ e $N_{ab} = 1500$)	77
6.6	Avaliação dos estimadores sob os cenários 2 e 3 para o plano AAS, em que $n_A = 1000$ (dados populacionais: $N_a = N_b = 2000$ e $N_{ab} = 500$)	80
6.7	Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 2000$ e $N_{ab} = 500$)	83
6.8	Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1750$ e $N_{ab} = 1000$)	86
6.9	Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1500$ e $N_{ab} = 1500$)	89

6.10	Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1250$ e $N_{ab} = 2000$)	92
6.11	Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1000$ e $N_{ab} = 2500$)	95
6.12	Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 750$ e $N_{ab} = 3000$)	98
6.13	Avaliação dos estimadores razão separados sob os cenários 2 e 3, para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para j e w)	101
6.14	Avaliação dos estimadores razão separados sob os cenários 2 e 3, para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para j e w)	104
6.15	Avaliação dos estimadores razão separados sob os cenários 2 e 3, para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para j e w)	107
6.16	Avaliação dos estimadores razão separados sob os cenários 2 e 3 para o plano AAE, em que $n_{A(j)} = 1000$ (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para j e w)	110
6.17	Avaliação dos estimadores razão separados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 2000$ e $N_{ab(jw)} = 500$) .	111
6.18	Avaliação dos estimadores razão separados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 1750$ e $N_{ab(jw)} = 1000$)	112
6.19	Avaliação dos estimadores razão separados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 1500$ e $N_{ab(jw)} = 500$) .	113
6.20	Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$) . .	114
6.21	Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 500$) . .	117
6.22	Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$) .	120

6.23	Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE, em que $n_{\mathcal{A}(j)} = 1000$ (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para j e w)	123
6.24	Avaliação dos estimadores razão combinados sob o cenário 2 para o plano AAE, onde $U_{\mathcal{B}} \subset U_{\mathcal{A}}$ (dados populacionais: $N_{a(j)} = 2000$ e $N_{ab(jw)} = 500$) .	124
6.25	Avaliação dos estimadores razão combinados sob o cenário 2 para o plano AAE, onde $U_{\mathcal{B}} \subset U_{\mathcal{A}}$ (dados populacionais: $N_{a(j)} = 1750$ e $N_{ab(jw)} = 1000$)	125

CAPÍTULO 1

Introdução

A experiência com a teoria da amostragem é algo comum no cotidiano. Basta levar em consideração experimentos simples como o preparo de um pão, a temperatura do prato de sopa, o modo de preparo de um chá, a forma como um médico avalia seus pacientes, para evidenciar a necessidade de se obter informações sobre o todo, baseando-se no resultado de uma amostra. A consideração de planos amostrais para se realizar inferência a respeito de uma determinada população de interesse é de extrema importância, pois garantem a seleção de elementos para compor a amostra através de um processo de aleatorização que permite utilizar a teoria da inferência estatística. Para decidir sobre que tipo de plano amostral deve ser utilizado, é necessário ter acesso a uma lista de elementos que compõem a população de interesse, a qual é chamada de *cadastro*.

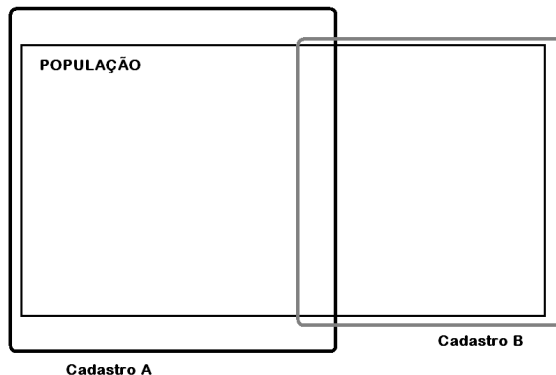
Em estudos mais sofisticados, é possível que todas as etapas referentes à forma da seleção da amostra a partir de um cadastro estejam corretos, porém todos os resultados obtidos podem estar errados não por conta de um erro amostral gerado pelo plano amostral, e sim devido a um erro não amostral, que é o erro de cobertura, ou seja, elementos que fazem parte da população alvo podem não estar incluídos no cadastro que se tem disponível, e nessa direção, a abordagem de cadastro duplo é uma alternativa para minimizar o erro de cobertura.

A abordagem de cadastro duplo é uma definida como um levantamento amostral onde dois cadastros são utilizados para identificar elementos de uma única população-alvo.

Nesta abordagem, amostras aleatórias independentes são selecionadas de cada cadastro, sem necessariamente ter que usar o mesmo esquema amostral em ambos.

Existem inúmeras situações onde é possível utilizar a abordagem de cadastro duplo. Por exemplo, quando um dos cadastros não tem um grau de cobertura desejável da população-alvo, porém existe um segundo cadastro, o qual, em conjunto com o primeiro, fornece cobertura completa. Essa situação é representada pela figura (1.1).

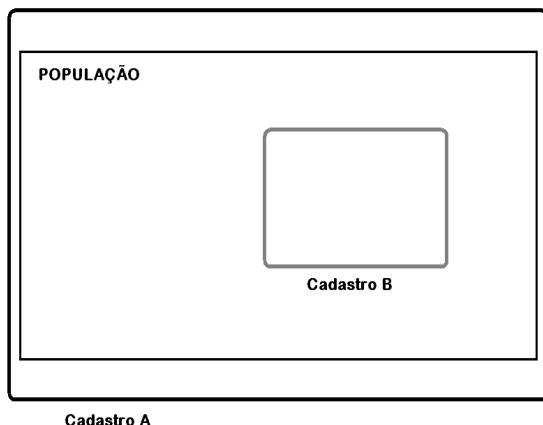
Figura 1.1: Uso de cadastro duplo para cobertura da população alvo



Uma outra situação, com grande potencial de aplicação em pesquisas agropecuárias, ocorre quando um cadastro cobre por completo a população-alvo, mas tem alto custo de seleção da amostra; um segundo cadastro disponível, com cobertura incompleta, mas de baixo custo operacional pode ser utilizado, em conjunto com o primeiro, resultando em uma melhor relação custo-benefício. Essa situação é representada pela figura (1.2). Hartley (1962) mostrou que em situações desse tipo, é possível obter a redução da variância do estimador considerado, a depender da variância dos estimadores em U_a e U_B .

Quando os cadastros utilizados não apresentam elementos em comum, a teoria de amostragem estratificada pode ser utilizada para planejar e estimar parâmetros de interesse na população em estudo. Quando há elementos presentes nos dois cadastros simultaneamente, os problemas de planejamento e estimação sob uma abordagem de cadastro duplo tornam-se mais desafiadores, devido à forma como tais informações deverão ser tratadas de modo a fornecer resultados mais precisos em relação à informação de interesse na população a ser pesquisada.

Figura 1.2: Uso de cadastro duplo aliado à relação custo-benefício



Os primeiros estudos referentes a técnicas de estimação sob a abordagem de cadastro duplo foram desenvolvidos por Hartley (1962). Desde então, vários autores passaram a dar importantes contribuições: Lund (1968) e Fuller & Burmeister (1972) realizaram melhorias nos trabalhos desenvolvidos por Hartley. Planos amostrais complexos envolvendo tal abordagem foram desenvolvidos por Bankier (1986). Problemas referentes à alocação ótima foram discutidos por Skinner & Rao (1998). Lohr e Rao (2000) realizaram comparações dos estimadores já propostos na literatura e consideraram o uso da técnica de jackknife para estimação de variâncias.

A disponibilidade de literatura a respeito da abordagem de cadastro duplo ainda é bastante limitada, quando comparada com a literatura existente sobre a utilização de um único cadastro. Quando o interesse está em estimar parâmetros através de estimadores assistidos por modelos, numa abordagem de cadastro duplo, ao que tudo indica não há literatura disponível¹.

O objetivo desta dissertação é apresentar estimadores que incorporam informação de uma variável auxiliar numa abordagem de cadastro duplo. Em particular, são propostos estimadores do tipo razão considerando os planos amostrais de amostragem aleatória simples (AAS) e amostragem aleatória estratificada (AAE). O desempenho de todos os estimadores propostos é investigado e comparado com o dos estimadores encontrados na

¹Durante os trabalhos de revisão de literatura para essa dissertação, não foram encontrados artigos abordando diretamente o assunto de estimação incorporando informações auxiliares em uma abordagem de cadastro duplo.

literatura (que não usam variáveis auxiliares no processo de estimação) através do método de simulação de Monte Carlo.

A presente dissertação está dividida em 7 capítulos. Neste primeiro capítulo foi apresentada uma introdução a respeito da abordagem de cadastro duplo. O capítulo 2 introduz os aspectos de estimação sob a abordagem de cadastro duplo, concentrando-se nos estimadores propostos por Hartley (1962) e Fuller & Burmeister (1972), os quais incorporam informações obtidas sobre o tamanho de domínios de interesse. O capítulo 3 traz considerações sobre a aplicação da abordagem de cadastros duplos em pesquisas agropecuárias, descrevendo resumidamente experiências dos Estados Unidos, Honduras e Brasil. O capítulo 4 apresenta estimadores do tipo razão para a abordagem de cadastro duplo, sob o plano AAS. O capítulo 5 apresenta estimadores do tipo separado e combinado sob o plano AAE. O capítulo 6 apresenta estudos de simulação referentes a abordagem de cadastro duplo, onde é utilizado o método de Monte Carlo, tornando possível a comparação entre os estimadores propostos e os estimadores já existentes. O capítulo 7 apresenta as considerações finais e propostas para trabalhos futuros. As linguagens de programação Ox e R constituem as plataformas computacionais utilizadas no desenvolvimento desta dissertação. A edição do texto foi realizada através do sistema de tipografia L^AT_EX.

Estimação sob a abordagem de cadastro duplo

Considere que estão disponíveis dois cadastros, $\mathcal{A}$ e $\mathcal{B}$, fornecendo a cobertura para uma população-alvo. Para a implementação da abordagem de cadastro duplo, duas condições são necessárias:

- i. todos os elementos da população-alvo devem ser identificados em pelo menos um dos cadastros;
- ii. deve ser possível identificar, para cada elemento selecionado para compor a amostra, via um cadastro $\mathcal{A}$, se ele está ou não listado no cadastro $\mathcal{B}$ e vice-versa.

Denote por U o conjunto de elementos da população-alvo, e por $U_{\mathcal{A}}$ e $U_{\mathcal{B}}$ o conjunto de elementos da população-alvo identificáveis através dos cadastros $\mathcal{A}$ e $\mathcal{B}$ respectivamente. Dessa forma, $U = U_{\mathcal{A}} \cup U_{\mathcal{B}}$. O uso de dois cadastros induz a existência de três domínios disjuntos, a , b e ab , que são subconjuntos dos conjuntos constituídos pelos cadastros $\mathcal{A}$ e $\mathcal{B}$. Um domínio é definido como uma subpopulação da população de interesse para a qual se deseja obter informações de interesse. Tratando o uso de cadastros de um modo mais amplo, numa abordagem de cadastros múltiplos, ao considerar k cadastros disponíveis, tem-se $2^k - 1$ domínios possíveis de serem identificados.

Hartley (1962) faz uma descrição dos possíveis cenários gerados por um levantamento amostral realizado através da abordagem de cadastro duplo. Tais cenários dependem basicamente da disponibilidade de informações populacionais, como mostra a tabela (2.1).

Tabela 2.1: Cenários possíveis em uma abordagem de cadastro duplo

Tipo de Informação Disponível	CENÁRIOS			
	1	2	3	4
Tamanho dos domínios e dos cadastros	Tamanhos dos domínios (cadastros) conhecidos	Tamanhos dos domínios (cadastros) conhecidos	Apenas os tamanhos dos cadastros conhecidos	Apenas as magnitudes relativas dos cadastros conhecidas
Possibilidade de alocação da amostra	Alocação de amostra aos domínios	Alocação de amostra aos cadastros	Alocação de amostra aos cadastros	Alocação de amostra aos cadastros

Os cenários 1 e 2 ilustram a situação em que os tamanhos dos domínios a , b e ab , denotados por N_a , N_b e N_{ab} respectivamente, são conhecidos, o que implica que $N_{\mathcal{A}}$ e $N_{\mathcal{B}}$ também o são. O cenário 1 permite que a alocação de amostra seja feita a cada domínio, enquanto que o cenário 2 permite a alocação apenas aos cadastros. Os cenários 3 e 4 têm em comum a situação em que a única forma de alocação de amostra possível é através dos cadastros. No cenário 3, $N_{\mathcal{A}}$ e $N_{\mathcal{B}}$ são conhecidos, mas não o tamanho dos domínios. No cenário 4, a única informação disponível é o tamanho relativo dos cadastros.

O cenário 1 ilustra a situação de um plano amostral estratificado. Já o cenário 4 fornece limitações bastante restritivas, tornando possível apenas a estimação de médias populacionais. Assim, todo o desenvolvimento apresentado nesta dissertação será referente aos cenários 2 e 3.

y_k é o valor da variável de interesse y associada ao elemento k na população-alvo. A tabela (2.2) fornece a notação utilizada para cálculos feitos a partir das amostras obtidas dos cadastros $\mathcal{A}$ e $\mathcal{B}$.

Tabela 2.2: Notação utilizada em uma abordagem de cadastro duplo

Quantidades de Interesse	CADASTRO		Domínio		
	$\mathcal{A}$	$\mathcal{B}$	a	b	ab
População	$U_{\mathcal{A}}$	$U_{\mathcal{B}}$	U_a	U_b	U_{ab}
Tamanho da População	$N_{\mathcal{A}}$	$N_{\mathcal{B}}$	N_a	N_b	N_{ab}
Total populacional	$t_{y\mathcal{A}}$	$t_{y\mathcal{B}}$	t_{ya}	t_{yb}	t_{yab}
Média populacional	$\mu_{y\mathcal{A}}$	$\mu_{y\mathcal{B}}$	μ_{ya}	μ_{yb}	μ_{yab}
Amostra	$S_{\mathcal{A}}$	$S_{\mathcal{B}}$	S_a	S_b	S_{ab}
Tamanho da amostra	$n_{\mathcal{A}}$	$n_{\mathcal{B}}$	n_a	n_b	n'_{ab} n''_{ab}
Estimador do Total	$\hat{t}_{y\mathcal{A}}$	$\hat{t}_{y\mathcal{B}}$	$\hat{t}_{ya}$	$\hat{t}_{yb}$	$\hat{t}'_{yab}$ $\hat{t}''_{yab}$
Média amostral	$\bar{y}_{\mathcal{A}}$	$\bar{y}_{\mathcal{B}}$	$\tilde{y}_a$	$\tilde{y}_b$	$\tilde{y}'_{ab}$ $\tilde{y}''_{ab}$

Denota-se que $U_a = U_{\mathcal{A}} \cap U_{\mathcal{B}}^c$, $U_b = U_{\mathcal{A}}^c \cap U_{\mathcal{B}}$ e $U_{ab} = U_{\mathcal{A}} \cap U_{\mathcal{B}}$. As quantidades n'_{ab} , $\hat{t}'_{yab}$ e $\tilde{y}'_{ab}$ são referentes a elementos na amostra que pertencem a U_{ab} e que foram obtidas do cadastro $\mathcal{A}$. As quantidades n''_{ab} , $\hat{t}''_{yab}$, e $\tilde{y}''_{ab}$ são definidas analogamente e estão relacionadas ao cadastro $\mathcal{B}$.

Dessa forma, dentre outras relações identificáveis, tem-se que:

- $N_{\mathcal{A}} = N_a + N_{ab}$;
- $N_{\mathcal{B}} = N_b + N_{ab}$ e
- $N = N_a + N_b + N_{ab} = N_a + N_{\mathcal{B}} = N_{\mathcal{A}} + N_b$.

As quantidades na tabela (2.2), referentes aos totais populacionais, são definidas da seguinte forma:

$t_y = \sum_{k \in U} y_k$ é o total populacional da variável de interesse na população.;

$t_{y\mathcal{A}} = \sum_{k \in U_{\mathcal{A}}} y_k$ é o total populacional da variável de interesse y , no cadastro $\mathcal{A}$;

$t_{y\mathcal{B}} = \sum_{k \in U_{\mathcal{B}}} y_k$ é o total populacional da variável de interesse y , no cadastro $\mathcal{B}$;

$t_{ya} = \sum_{k \in U_a} y_k$ é o total populacional da variável de interesse y , em U_a ;

$t_{yb} = \sum_{k \in U_b} y_k$ é o total populacional da variável de interesse y , em U_b ;

$t_{yab} = \sum_{k \in U_{ab}} y_k$ é o total populacional da variável de interesse y , em U_{ab} .

É possível observar que $t_y = t_{ya} + t_{yab} + t_{yb}$.

2.1 Estimador de Hartley

O método de estimação desenvolvido por Hartley (1962) foi proposto com base em uma variável y_k^* para cada cadastro, e pode ser definida da seguinte forma:

$$y_{k\mathcal{A}}^* = \begin{cases} y_k, & \text{se } k \in U_a \\ py_k, & \text{se } k \in U_{ab} \end{cases} \quad y_{k\mathcal{B}}^* = \begin{cases} y_k, & \text{se } k \in U_b \\ (1-p)y_k, & \text{se } k \in U_{ab} \end{cases}$$

onde p é uma constante de ponderação para identificação dos elementos de cada cadastro na população, tal que $0 \leq p \leq 1$. Dessa forma, é possível reescrever o total populacional, t_y , da seguinte forma:

$$\begin{aligned} t_y &= t_{ya} + t_{yb} + t_{yab} \\ &= t_{ya} + t_{yb} + (p + 1 - p)t_{yab} \\ &= t_{ya} + pt_{yab} + t_{yb} + (1 - p)t_{yab} \\ &= \sum_{k \in U_{\mathcal{A}}} y_{k\mathcal{A}}^* + \sum_{k \in U_{\mathcal{B}}} y_{k\mathcal{B}}^* \\ &= t_{y\mathcal{A}}^* + t_{y\mathcal{B}}^*. \end{aligned} \tag{2.1}$$

A expressão (2.1) mostra o total populacional como uma soma de duas parcelas, as quais estão em função das variáveis $y_{k\mathcal{A}}^*$ e $y_{k\mathcal{B}}^*$. O desenvolvimento mostra ainda que é possível propor um estimador para o total populacional que seja função linear de estimadores para os totais populacionais dos domínios a , b e ab .

2.2 Estimação sob o cenário 2

Considere o problema de obter um estimador para t_y , sob a situação do cenário 2, onde há disponibilidade de informação sobre os tamanhos populacionais de cada domínio (N_a , N_b e N_{ab}). Considere ainda que em cada cadastro foi aplicado um plano AAS, obtendo amostras denotadas por $S_{\mathcal{A}}$ e $S_{\mathcal{B}}$, de tamanhos $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$, referindo-se aos cadastros $\mathcal{A}$ e $\mathcal{B}$ respectivamente. Nestas condições, o estimador proposto por Hartley (1962) assume a seguinte forma:

$$\hat{t}_{yH} = N_a \tilde{y}_a + N_{ab}(p \tilde{y}'_{ab} + (1-p) \tilde{y}''_{ab}) + N_b \tilde{y}_b, \quad (2.2)$$

onde:

$\tilde{y}_a$ é a Média amostral do domínio a , a partir de elementos obtidos do cadastro $\mathcal{A}$;

$\tilde{y}_b$ é a Média amostral do domínio b , a partir de elementos obtidos do cadastro $\mathcal{B}$;

$\tilde{y}'_{ab}$ é a Média amostral do domínio ab , a partir de elementos obtidos do cadastro $\mathcal{A}$;

$\tilde{y}''_{ab}$ é a Média amostral do domínio ab , a partir de elementos obtidos do cadastro $\mathcal{B}$.

Ignorando os fatores de correção para populações finitas, tem-se que a variância de (2.2) é aproximadamente igual à variância do estimador para o caso em que é utilizada alocação proporcional ao tamanho, no caso de um plano de amostragem estratificada:

$$\text{Var}(\hat{t}_{yH}) = \frac{N_{\mathcal{A}}^2}{n_{\mathcal{A}}} \left[(1 - \alpha_{ab}) \sigma_a^2 + \alpha_{ab} p^2 \sigma_{ab}'^2 \right] + \frac{N_{\mathcal{B}}^2}{n_{\mathcal{B}}} \left[(1 - \beta_{ab}) \sigma_b^2 + \beta_{ab} (1-p)^2 \sigma_{ab}''^2 \right], \quad (2.3)$$

onde $\alpha_{ab} = \frac{N_{ab}}{N_{\mathcal{A}}}$ e $\beta_{ab} = \frac{N_{ab}}{N_{\mathcal{B}}}$. Tem-se ainda que σ_a^2 , σ_b^2 e σ_{ab}^2 representam as variâncias populacionais dos domínios a , b e ab , respectivamente. Além disso, valores ótimos para p , $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$, os quais minimizam (2.3) são apresentados por Hartley (1962). Um estimador para a variância do estimador pode ser obtido ao substituir as quantidades populacionais σ_a^2 , σ_b^2 , $\sigma_{ab}'^2$ e $\sigma_{ab}''^2$ pelos seus respectivos estimadores, $\hat{\sigma}_a^2$, $\hat{\sigma}_b^2$, $\hat{\sigma}_{ab}'^2$ e $\hat{\sigma}_{ab}''^2$.

2.3 Estimação sob o cenário 3

O cenário 3 corresponde à situação em que não há informação sobre N_a , N_b e N_{ab} . Desse modo, é necessário obter estimadores para essas quantidades. Através do princípio de Horvitz-Thompson, no caso de um plano AAS, a probabilidade de um elemento pertencente a um dado cadastro $\mathcal{C}$ ser selecionado para compor a amostra é $\pi_{\mathcal{C}k} = \frac{n_{\mathcal{C}}}{N_{\mathcal{C}}} = f_{\mathcal{C}}$, onde $n_{\mathcal{C}}$ é o tamanho da amostra obtida do cadastro $\mathcal{C}$, $N_{\mathcal{C}}$ é o tamanho da população identificada pelo cadastro $\mathcal{C}$ e $f_{\mathcal{C}}$ é a fração amostral, aplicada ao cadastro $\mathcal{C}$. Dado um domínio d de interesse, é possível definir um estimador para o tamanho populacional deste domínio, N_d , da seguinte forma:

$$\hat{N}_d = \sum_{k \in S_d} \pi_{\mathcal{C}k}^{-1} = \sum_{k \in S_d} \frac{N_{\mathcal{C}}}{n_{\mathcal{C}}} = \delta_d N_{\mathcal{C}},$$

onde $\delta_d = \frac{n_d}{n_{\mathcal{C}}}$, e $n_d \leq n_{\mathcal{C}}$ é o número de elementos da amostra que pertencem ao domínio d . Além disso,

$$\hat{t}_{yd} = \sum_{k \in S_d} \frac{y_k}{\pi_{\mathcal{C}k}} = \sum_{k \in S_d} \frac{N_{\mathcal{C}}}{n_{\mathcal{C}}} y_k = \delta_d N_{\mathcal{C}} \tilde{y}_d \quad \text{e} \quad \hat{\mu}_d = \frac{\hat{t}_{yd}}{\hat{N}_d} = \frac{\delta_d N_{\mathcal{C}} \tilde{y}_d}{\delta_d N_{\mathcal{C}}} = \tilde{y}_d.$$

$\hat{t}_{yd}$ é um estimador centrado para t_{yd} , pois denotando por $E_{\mathcal{P}}(\cdot)$ a esperança com respeito a um plano amostral probabilístico $\mathcal{P}$, tem-se

$$E_{\mathcal{P}}(\hat{t}_{yd}) = E_{\mathcal{P}}(\delta_d N_{\mathcal{C}} \tilde{y}_d) = \delta_d N_{\mathcal{C}} E_{\mathcal{P}}(\tilde{y}_d) = \delta_d N_{\mathcal{C}} E_{\mathcal{P}} \left(\frac{1}{n_d} \sum_{k \in U_d} y_k I_k \right).$$

onde $I_k = \begin{cases} 1, & k \in S_d \\ 0, & c.c. \end{cases}$.

Assim,

$$E_{\mathcal{P}}(\hat{t}_{yd}) = \delta_d N_{\mathcal{C}} \frac{1}{n_d} \sum_{k \in U_d} y_k E_{\mathcal{P}}(I_k) = \delta_d N_{\mathcal{C}} \frac{1}{n_d} \sum_{k \in U_d} y_k \pi_{\mathcal{C}k} = \delta_d N_{\mathcal{C}} \frac{1}{n_d} \sum_{k \in U_d} y_k \frac{n_{\mathcal{C}}}{N_{\mathcal{C}}} = \sum_{k \in U_d} y_k = t_{yd}.$$

A partir desse desenvolvimento, colocando os domínios a , b e ab no lugar de d , quando necessário, Hartley propôs o seguinte estimador para o total, sob o cenário 3:

$$\begin{aligned}
\hat{t}_{yH} &= \hat{t}_{ya} + p\hat{t}'_{yab} + \hat{t}_{yb} + (1-p)\hat{t}''_{yab} \\
&= \delta_a N_{\mathcal{A}} \tilde{y}_a + p\delta'_{ab} N_{\mathcal{A}} \tilde{y}'_{ab} + \delta_b N_{\mathcal{B}} \tilde{y}_b + (1-p)\delta''_{ab} N_{\mathcal{B}} \tilde{y}''_{ab} \\
&= \frac{N_{\mathcal{A}}}{n_{\mathcal{A}}} n_a \tilde{y}_a + p \frac{N_{\mathcal{A}}}{n_{\mathcal{A}}} n'_{ab} \tilde{y}'_{ab} + \frac{N_{\mathcal{B}}}{n_{\mathcal{B}}} n_b \tilde{y}_b + (1-p) \frac{N_{\mathcal{B}}}{n_{\mathcal{B}}} n''_{ab} \tilde{y}''_{ab} \\
&= \frac{N_{\mathcal{A}}}{n_{\mathcal{A}}} \{n_a \tilde{y}_a + p n'_{ab} \tilde{y}'_{ab}\} + \frac{N_{\mathcal{B}}}{n_{\mathcal{B}}} \{n_b \tilde{y}_b + (1-p) n''_{ab} \tilde{y}''_{ab}\} \\
&= f_{\mathcal{A}}^{-1} \{n_a \tilde{y}_a + p n'_{ab} \tilde{y}'_{ab}\} + f_{\mathcal{B}}^{-1} \{n_b \tilde{y}_b + (1-p) n''_{ab} \tilde{y}''_{ab}\}.
\end{aligned} \tag{2.4}$$

A variância de $\hat{t}_{yH}$ é dada por

$$\begin{aligned}
\text{Var}(\hat{t}_{y,H}) &= \frac{N_{\mathcal{A}}^2}{n_{\mathcal{A}}} \left\{ (1 - \alpha_{ab}) \sigma_a^2 + \alpha_{ab} p^2 \sigma_{ab}'^2 + \alpha_{ab} (1 - \alpha_{ab}) (\mu_a - p\mu_{ab})^2 \right\} \\
&+ \frac{N_{\mathcal{B}}^2}{n_{\mathcal{B}}} \left\{ (1 - \beta_{ab}) \sigma_b^2 + \beta_{ab} (1-p)^2 \sigma_{ab}''^2 + \beta_{ab} (1 - \beta_{ab}) (\mu_b - (1-p)\mu_{ab})^2 \right\}.
\end{aligned} \tag{2.5}$$

Um estimador para a variância de $\hat{t}_{y,H}$ é obtido ao substituir as quantidades populacionais pelos respectivos estimadores.

2.4 Estimador de Lund

Lund (1968) sugeriu a seguinte modificação para a forma do estimador de Hartley, sob o cenário 3:

$$\hat{t}_{y,L} = f_{\mathcal{A}}^{-1} n_a \tilde{y}_a + (p f_{\mathcal{A}}^{-1} n'_{ab} + (1-p) f_{\mathcal{B}}^{-1} n''_{ab}) \tilde{y}_{ab,L} + f_{\mathcal{B}}^{-1} n_b \tilde{y}_b \tag{2.6}$$

onde $\tilde{y}_{ab,L} = \frac{n'_{ab} \tilde{y}'_{ab} + n''_{ab} \tilde{y}''_{ab}}{n'_{ab} + n''_{ab}}$. A variância do estimador de Lund é expressa da seguinte forma:

$$\begin{aligned}
\text{Var}(\hat{t}_{yL}) &= N_a (f_{\mathcal{A}}^{-1} - 1) \sigma_a^2 + \frac{f_{\mathcal{A}}(1 - f_{\mathcal{A}}) + f_{\mathcal{B}}(1 - f_{\mathcal{B}}) N_{ab} \sigma_{ab}^2}{(f_{\mathcal{A}} + f_{\mathcal{B}})^2} + N_b (f_{\mathcal{B}}^{-1} - 1) \sigma_b^2 \\
&+ g_{\mathcal{A}} n_{\mathcal{A}}^{-1} N_{ab} N_a (\mu_a - p\mu_{ab})^2 + g_{\mathcal{B}} n_{\mathcal{B}}^{-1} N_{ab} N_b (\mu_b - (1-p)\mu_{ab})^2,
\end{aligned} \tag{2.7}$$

onde

$$g_{\mathcal{A}} = \frac{N_{\mathcal{A}} - n_{\mathcal{A}}}{N_{\mathcal{A}} - 1} \quad g_{\mathcal{B}} = \frac{N_{\mathcal{B}} - n_{\mathcal{B}}}{N_{\mathcal{B}} - 1}$$

Um estimador da variância de $\hat{t}_{yL}$ pode ser obtido ao substituir as quantidades populacionais em (2.7) pelos respectivos estimadores.

2.5 Estimador de Fuller & Burmeister

O estimador desenvolvido por Fuller & Burmeister (1972) incorpora informação sobre os tamanhos das amostras observadas no domínio ab , n'_{ab} e n''_{ab} . Definido um estimador adequado para N_{ab} , denotado por $\hat{N}_{ab}$, como será visto no capítulo 4, o estimador proposto por Fuller & Burmeister assume a forma:

$$\hat{t}_{yFB} = (N_{\mathcal{A}} - \hat{N}_{ab})\tilde{y}_a + (N_{\mathcal{B}} - \hat{N}_{ab})\tilde{y}_b + \hat{N}_{ab}\tilde{y}_{ab}, \quad (2.8)$$

onde

$$\tilde{y}_{ab} = w\tilde{y}'_{ab} + (1 - w)\tilde{y}''_{ab} \quad \text{e} \quad w = \frac{n'_{ab}(1 - f_{\mathcal{B}})}{n'_{ab}(1 - f_{\mathcal{B}}) + n''_{ab}(1 - f_{\mathcal{A}})}.$$

Quando os fatores de correção para populações finitas são negligíveis, $\tilde{y}_{ab}$ reduz-se à média amostral de $n'_{ab} + n''_{ab}$ valores observados em U_{ab} . Temos que a variância do estimador é dada por:

$$\begin{aligned} \text{Var}(\hat{t}_{yFB}) &= N_a(f_{\mathcal{A}}^{-1} - 1)\sigma_a^2 + \frac{(1 - f_{\mathcal{A}})(1 - f_{\mathcal{B}})N_{ab}\sigma_{ab}^2}{(1 - f_{\mathcal{B}})f_{\mathcal{A}} + (1 - f_{\mathcal{A}})f_{\mathcal{B}}} + N_b(f_{\mathcal{B}}^{-1} - 1)\sigma_b^2 \\ &+ (\mu_{yab} - \mu_{ya} - \mu_{yb})\delta, \end{aligned} \quad (2.9)$$

onde

$$\delta = \frac{N_{ab}N_aN_bg_{\mathcal{A}}g_{\mathcal{B}}}{n_{\mathcal{A}}N_bg_{\mathcal{B}} + n_{\mathcal{B}}N_ag_{\mathcal{A}}}$$

e μ_{ya} , μ_{yb} e μ_{yab} são as médias populacionais dos domínios U_a , U_b e U_{ab} respectivamente.

Da mesma forma que para as variâncias dos estimadores $\hat{t}_{yH}$, e $\hat{t}_{yL}$, é possível obter um estimador para a variância de $\hat{t}_{yFB}$, quando as quantidades populacionais em (2.9) são substituídas pelos respectivos estimadores.

Fuller & Burmeister mostraram que seu estimador é melhor que o estimador de Lund por apresentar menor variância, motivo pelo qual nesta dissertação apenas os estimadores de Hartley e Fuller & Burmeister serão usados nos estudos de simulação.

CAPÍTULO 3

Aplicações da abordagem de cadastro duplo em pesquisas agropecuárias

As pesquisas agropecuárias são responsáveis pelo levantamento de grande parte de informações de importância para o suporte a políticas agrícolas, como estimativas de safras, estimativas de produção animal e dados sócio-econômicos relacionados à área rural. Elas são executadas periodicamente e, via de regra, têm múltiplos propósitos.

Uma das formas mais utilizadas de aplicação da abordagem de cadastro duplo em pesquisas agropecuárias consiste na combinação de um cadastro de área (area-frame) - listagem de pequenos segmentos de área cobrindo um determinado território - com um cadastro específico (list-frame) - listagem de elementos pertencentes a uma população específica, como por exemplo, listagem de produtores de grãos.

Uma pesquisa agropecuária que faça uso de apenas um cadastro de área tem a vantagem de garantir a total cobertura da população-alvo, além de garantir adequidade à sua natureza de múltiplos propósitos. No entanto, dentre as desvantagens, é possível citar o alto custo para composição da amostra e o fato de que estimativas obtidas a partir de levantamentos com cadastros do tipo área são sensíveis a outliers (Carfagna, 2004). Além disso, quando a forma de obtenção dos dados é através de entrevistas, a pesquisa torna-se muito vulnerável a altas taxas de não-respostas, devido ao difícil acesso aos informantes.

Pesquisas agropecuárias que usam apenas cadastros específicos, quando disponíveis, apresentam como vantagem o baixo custo na implementação de um plano amostral.

Porém, cadastros específicos apresentam em geral um baixo nível de cobertura da população-alvo. Além disso, é necessário verificar até que ponto as informações desse tipo de cadastro estão atualizadas.

A abordagem de cadastro duplo, combinando cadastros de área com cadastros específicos, apresenta-se como uma alternativa de grande potencial para unir as vantagens de cada tipo de cadastro. Neste caso, definindo $\mathcal{A}$ como sendo o cadastro do tipo área, e $\mathcal{B}$ como o cadastro específico, tem-se então que $U_{\mathcal{B}} \subset U_{\mathcal{A}}$. Sob tais condições, a forma dos estimadores apresentados na capítulo 2 sofre considerável simplificação, pois $N_{ab} = N_{\mathcal{B}}$ e $N_b = 0$. Por exemplo, o estimador (2.2) reduz-se a

$$\hat{t}_{y,H} = N_a \tilde{y}_a + N_{ab}(p \tilde{y}'_{ab} + (1-p) \tilde{y}''_{ab}). \quad (3.1)$$

Em muitas situações práticas, o valor de p em (3.1) é considerado nulo, o que resulta em um tipo de estimador chamado de “screening”, que tem essa denominação devido à necessidade de identificar quais elementos da amostra obtida do cadastro de área que estão no cadastro de lista.

Vários países têm tido experiências de aplicação da abordagem de cadastro duplo, como descrito em livro editado e publicado pela FAO (1998), com o objetivo de compartilhar e difundir as técnicas de amostragem em uso, aplicadas à agricultura.

Nos Estados Unidos, as pesquisas no setor agropecuário são desenvolvidas pelo *United States Department of Agriculture* (USDA) através do *National Agricultural Statistics Service* (NASS), o qual fornece oficialmente toda a informação estatística referente à agricultura.

O NASS utiliza métodos de pesquisa tomando como base cadastros múltiplos e, desde 1950, abordagens como a de cadastro duplo fornecem grande contribuição nas pesquisas agropecuárias do país.

O programa desenvolvido pelo NASS é intitulado *Multiple Frame Agricultural Survey*, iniciado em 1986 em 27 estados daquele país e é conduzido nos meses de junho, julho, agosto, setembro, janeiro e março, com o intuito de obter estimativas referentes a áreas produtivas, número de unidades agrícolas em atividade, área total ocupada por unidades agrícolas, produção total das unidades agrícolas, etc...

Na construção dos cadastros, o cadastro do tipo área é resultante da combinação dos

cadastros de área construídos para cada estado, independentemente. A cada 10 ou 20 anos o cadastro é atualizado. Já os cadastros específicos são regularmente atualizados, e utilizados em combinação com o cadastro de área. Existem aproximadamente 2 milhões de unidades agrícolas nos Estados Unidos, e os cadastros específicos, quando construídos em grande escala, conseguem uma cobertura de aproximadamente 1,1 milhões de unidades agrícolas. Os cadastros do tipo listas considerados pelo NASS são construídos a partir das informações referentes à fazendeiros, agronegócios e áreas agrícolas (em hectares).

As informações coletadas para composição do cadastro específico são divididas em duas partes. A primeira é referente aos dados para identificar, localizar e contactar fazendeiros e empresários do setor: nome, endereço, telefone, estado, distrito, município, razão social, número de identificação de empregador, etc. A segunda parte contém dados sobre a fazenda e/ou negócio: total de hectares, total de hectares destinado a determinado produto, capacidade de armazenamento de grãos, quantidade de cabeças de gado, número de empregados, entre outros.

A construção do cadastro do tipo lista é feita a partir das informações de identificação oriundas de sindicatos, fazendas ligadas à programas do setor, estado de origem e impostos cobrados pelo estado, licença de funcionamento no Estado, cadastramento através de agências federais, estaduais e identificação em revistas e notícias ligadas ao setor. Uma vez construído, o cadastro é constantemente atualizado através de pesquisas feitas pelo NASS ao longo do ano, como também através do censo agropecuário. O cadastro específico é construído com múltiplos propósitos, dentre os quais é possível citar os seguintes:

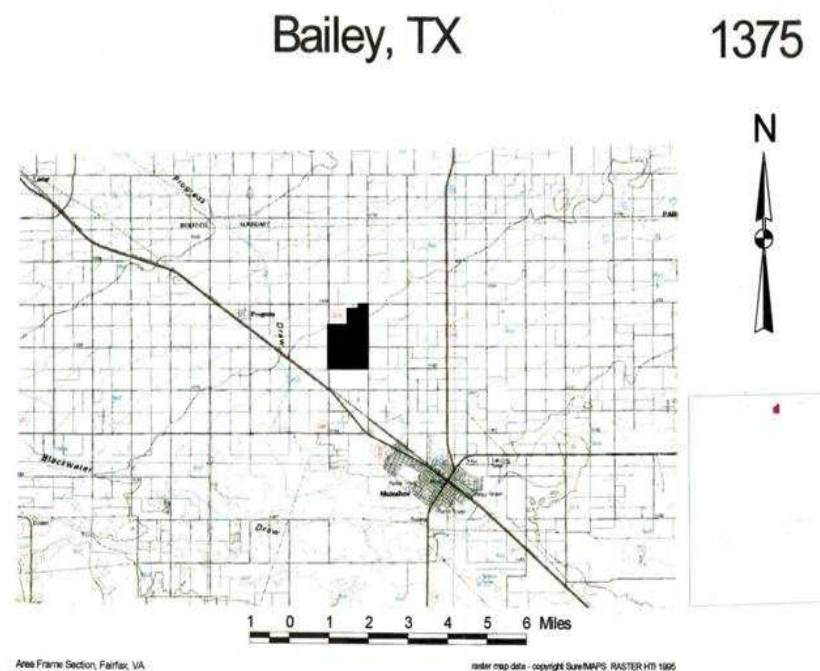
- i. identificação e classificação de fazendeiros e/ou empresários do setor que estão diretamente ligados a determinados produtos de interesse;
- ii. estratificação para identificar grupos de fazendeiros e produtores que possuem determinado produto em comum, para que se possa coletar amostras em cada grupo;
- iii. pesquisas específicas, ou seja a coleta de dados é feita de forma a atender uma população-alvo em particular.
- iv. possibilidade de se atribuir probabilidades de seleção para cada elemento do cadastro

Dessa forma, os cadastros específicos desenvolvidos pelo NASS tornam possíveis a

redução de custos através de métodos específicos de coleta de dados (correio e telefone, por exemplo) pesquisas de produtos específicos e redução da variância de estimadores construídos a partir de dados amostrais.

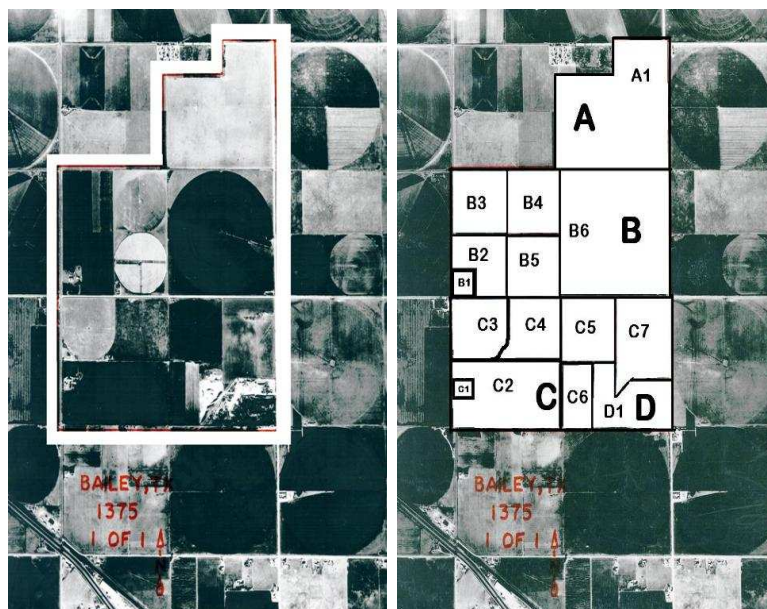
O cadastro do tipo área do NASS é construído a partir de toda a área dos Estados Unidos, que é dividida em segmentos de terra através de delimitações, associando fazendas, produção, animais, etc. com a área pertencente a cada segmento. Imagens de satélite, mapas topográficos, sistemas de informação geográfico e fotos aéreas também são utilizados para construção dos cadastros de área. A vantagem desse cadastro está na cobertura completa de todas as unidades agrícolas em atividade. As figuras (3.1) e (3.2) ilustram um exemplo de segmento de área selecionado pelo NASS.

Figura 3.1: Exemplo de seleção de segmento de área



Fonte: National Agricultural Statistics Service (NASS)

Figura 3.2: Segmento escolhido e subdivisão do segmento em estratos



Fonte: National Agricultural Statistics Service (NASS)

Os Estados Unidos utilizam técnicas envolvendo cadastros múltiplos para muitos produtos agrícolas, pois um grande percentual da produção total é concentrado em um número pequeno de grandes unidades agrícolas. Com o passar dos anos, a tecnologia se tornou uma grande aliada na composição dos cadastros de área, que passaram a ser formulados através de imagens de satélite.

Em Honduras, no ano de 1998, de um total de 112.490 km² no país, aproximadamente 35.000 km² estiveram destinados à agricultura. A maioria da região produtiva do país era destinada à produção de grãos (feijão e arroz por exemplo) para consumo interno, armazenamento e alimentação de animais. Cerca de 90% das unidades agrícolas produziam ou utilizavam alguns tipos de grãos. As pesquisas em Honduras naquela época também estavam focadas em áreas de colheita, vegetais, colheita de um tipo de produto permanentemente cultivado, como por exemplo, café, cacau e banana.

Os cadastros do tipo área em Honduras eram obtidos através de foto aérea e mapas topográficos, enquanto que os cadastros específicos consistiam de uma listagem dos grandes produtores de grãos no país. No processo de estimação de informações de interesse, foram utilizadas informações obtidas a partir do cadastro de área, obtendo bons

resultados para alguns tipos de grãos.

No Brasil, a abordagem de cadastro duplo foi utilizada no período 1986-1997, em pesquisas agropecuárias focando os estados de Paraná, Santa Catarina, São Paulo e Distrito Federal. Ao todo, os estados considerados somam uma área total de 550.000 Km². O programa foi desenvolvido pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em cooperação com o Instituto Brasileiro de Pesquisa Espacial (INPE). No método, um cadastro do tipo área foi utilizado. Esse cadastro era composto apenas uma vez e utilizado repetidas vezes ao longo dos anos. Um cadastro específico, que é utilizado como complemento a este cadastro e anualmente é atualizado a fim de garantir a melhor coleta possível de informações.

A maior dificuldade em realizar uma pesquisa por amostragem no setor agropecuário, no Brasil, é obter informações referentes aos municípios, pois o custo operacional é bastante elevado.

Apesar da existência de cadastros específicos e de área, as informações ligadas à agropecuária que são coletadas pelo IBGE estão baseadas apenas em função de opiniões de especialistas dos produtos (ou commodities) de interesse. Atualmente, o IBGE busca seguir o modelo de pesquisa que é praticado pelos Estados Unidos, baseado na opinião de especialistas aliada aos resultados obtidos através de pesquisas por amostragem, com o objetivo de propor um sistema de pesquisas agropecuárias que seja eficiente e aplicável à realidade brasileira. Para tanto, a abordagem de cadastro duplo surge como uma alternativa com potencial para ser explorada. O desenvolvimento desta dissertação é motivado pela sua potencial contribuição ao desenvolvimento de pesquisas agropecuárias no Brasil.

CAPÍTULO 4

Estimação assistida por modelos em uma abordagem de cadastro duplo

O uso de informações auxiliares no processo de estimação de um determinado parâmetro de interesse está consolidado na literatura considerando a disponibilidade e utilização de um único cadastro, sob a suposição de que este fornece cobertura completa da população-alvo. Quando é considerada a abordagem de cadastro duplo, no entanto, a literatura existente, ao que tudo indica, é praticamente nenhuma. Neste capítulo, o desenvolvimento de estimadores assistidos por modelos, numa abordagem de cadastro duplo, é apresentado, como parte da contribuição original dessa dissertação.

Considere a situação em que há informação de uma variável auxiliar que está disponível no cadastro $\mathcal{A}$, sob os cenários 2 e 3. Com isso, é possível ter acesso à informação sobre o total da variável auxiliar, ou seja:

$$t_{X_{\mathcal{A}}} = \sum_{k \in U_{\mathcal{A}}} x_k,$$

onde x_k é o valor observado da variável auxiliar referente ao k -ésimo elemento da população pertencente no cadastro $\mathcal{A}$.

Uma vez que há interseção (ou cruzamento de informações) entre os cadastros $\mathcal{A}$ e $\mathcal{B}$, é natural considerar que o total da variável auxiliar pode ser escrito da seguinte forma:

$$t_{X_{\mathcal{A}}} = t_{x_a} + t'_{x_{ab}}, \quad \text{onde}$$

t_{x_a} é o total da variável auxiliar em U_a ;

$t'_{x_{ab}}$ é o total da variável auxiliar em U_{ab} , a partir do cadastro $\mathcal{A}$.

Situações onde variáveis auxiliares são utilizadas podem ser motivadas pela realizações de censos no passado, os quais disponibilizam informações da variável de interesse para toda a população. Se há evidência de que a população da variável de interesse tenha uma relação linear com a variável auxiliar e ainda, se é verificado que a variância da variável de interesse é proporcional à variável auxiliar, estimadores do tipo razão podem ser utilizados no processo de estimação.

Nesta dissertação, serão considerados estimadores desse tipo para o cadastro $\mathcal{A}$, para cada domínio de interesse. Nesse contexto, a relação entre a variável de interesse e a variável auxiliar, na população do cadastro $\mathcal{A}$, será assistida por um modelo da forma:

$$\xi : y_{k\mathcal{A}} = \begin{cases} \beta_a x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}} & \text{se } k \in U_a \\ \beta_{ab} x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}} & \text{se } k \in U_{ab} \end{cases} \quad (4.1)$$

onde $E_{\xi}(\varepsilon_{k\mathcal{A}}) = 0$, $\forall k \in U_{\mathcal{A}}$. Além disso, tem-se que

$$E_{\xi}(y_{k\mathcal{A}}) = \begin{cases} \beta_a x_{k\mathcal{A}}, & k \in U_a \\ \beta_{ab} x_{k\mathcal{A}}, & \text{se } k \in U_{ab} \end{cases} ; \quad \text{Var}_{\xi}(y_{k\mathcal{A}}) = \lambda x_{k\mathcal{A}}, \quad \forall k \in U_{\mathcal{A}}$$

A forma geral do estimador razão para os domínios é dada por

$$\hat{t}_{y_a, R} = \frac{\hat{t}_{y_a}}{\hat{t}_{x_a}} t_{x_a} \quad \text{e} \quad \hat{t}_{y_{ab}, R} = \frac{\hat{t}_{y_{ab}}}{\hat{t}_{x_{ab}}} t_{x_{ab}} \quad (4.2)$$

Sob um plano de AAS, tem-se que:

$$\hat{t}_{y_a,R} = \frac{\tilde{y}_a}{\tilde{x}_a} t_{x_a} \quad \text{e} \quad \hat{t}_{y_{ab},R} = \frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}} t_{x_{ab}}. \quad (4.3)$$

4.1 Estimação de N_{ab}

Considere o cenário 3, onde apenas $N_{\mathcal{A}}$ e $N_{\mathcal{B}}$ são conhecidos. Quando não se tem conhecimento do tamanho populacional em cada domínio, o problema de estimação concentra-se em determinar o tamanho populacional do domínio em que há cruzamento de informações, o qual é resultante da combinação dos cadastros $\mathcal{A}$ e $\mathcal{B}$. A justificativa para isso é simples, pois uma vez que há interseção entre tais cadastros e obtém-se uma estimativa de N_{ab} , denotada por $\hat{N}_{ab}$, automaticamente são obtidas estimativas para os demais tamanhos populacionais, fazendo-se

$$\hat{N}_a = N_{\mathcal{A}} - \hat{N}_{ab} \quad \text{e} \quad \hat{N}_b = N_{\mathcal{B}} - \hat{N}_{ab}.$$

Hartley (1962) propôs o estimador

$$\hat{N}_{ab,H} = p \frac{n'_{ab} N_{\mathcal{A}}}{n_{\mathcal{A}}} + (1 - p) \frac{n''_{ab} N_{\mathcal{B}}}{n_{\mathcal{B}}}. \quad (4.4)$$

O estimador de Hartley não leva em consideração o conhecimento de duplicatas, ou seja, informações que estão presentes tanto em $n_{\mathcal{A}}$ quanto em $n_{\mathcal{B}}$, pois considera n'_{ab} e n''_{ab} sem realizar uma “varredura” das informações, a fim de verificar possíveis repetições. A variância de $\hat{N}_{ab,H}$ é dada por

$$\text{Var}(\hat{N}_{ab,H}) = p^2 f_{\mathcal{A}}^{-2} \text{Var}(n'_{ab}) + (1 - p)^2 f_{\mathcal{B}}^{-2} \text{Var}(n''_{ab}). \quad (4.5)$$

Verifica-se que n'_{ab} e n''_{ab} são variáveis aleatórias com distribuição hipergeométrica. Com isso, o valor de p que minimiza a variância (Hartley 1962), é dado por

$$p_H = \frac{n_{\mathcal{A}} N_b g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}}, \text{ onde} \quad (4.6)$$

$$g_{\mathcal{A}} = \frac{N_{\mathcal{A}} - n_{\mathcal{A}}}{N_{\mathcal{A}} - 1} \quad \text{e} \quad g_{\mathcal{B}} = \frac{N_{\mathcal{B}} - n_{\mathcal{B}}}{N_{\mathcal{B}} - 1}.$$

Para este valor de p , a variância do estimador para o domínio N_{ab} é dada por

$$\text{Var}(\hat{N}_{ab,H}) = \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \quad (4.7)$$

Para o caso em que N_{ab} é desconhecido, Fuller & Burmeister (1972) propuseram um estimador que é função dos tamanhos de amostra nos cadastros e domínios e das quantidades $N_{\mathcal{A}}$ e $N_{\mathcal{B}}$. Primeiramente, foi considerado que:

$$\hat{N}_a = N_{\mathcal{A}} - \hat{N}_{ab} \quad \text{e} \quad \hat{N}_b = N_{\mathcal{B}} - \hat{N}_{ab}.$$

Substituindo esses valores em (4.6), tem-se que

$$\hat{p}_{H,FB} = \frac{n_{\mathcal{A}} \hat{N}_b g_{\mathcal{B}}}{n_{\mathcal{A}} \hat{N}_b g_{\mathcal{B}} + n_{\mathcal{B}} \hat{N}_a g_{\mathcal{A}}}, \quad (4.8)$$

onde $\hat{p}_{H,FB}$ é a estimativa da constante p de Hartley proposta por Fuller & Burmeister, sob o cenário 3. Substituindo a estimativa de p_H em (4.4), tem-se que o valor de $\hat{N}_{ab}$ pode ser obtido como sendo a menor raiz da seguinte equação quadrática

$$\begin{aligned} (n_{\mathcal{A}} g_{\mathcal{B}} + n_{\mathcal{B}} g_{\mathcal{A}}) \hat{N}_{ab,s}^2 - (n_{\mathcal{A}} N_{\mathcal{B}} g_{\mathcal{B}} + n_{\mathcal{B}} N_{\mathcal{A}} g_{\mathcal{A}} + n'_{ab} N_{\mathcal{A}} g_{\mathcal{B}} + n''_{ab} N_{\mathcal{B}} g_{\mathcal{A}}) \hat{N}_{ab,s} \\ + (n'_{ab} g_{\mathcal{B}} + n''_{ab} g_{\mathcal{A}}) N_{\mathcal{A}} N_{\mathcal{B}} = 0 \end{aligned} \quad (4.9)$$

Fuller & Burmeister mostraram ainda que a variância de $\hat{N}_{ab,s}$ é tal que

$$\text{Var} \left(\hat{N}_{ab,s} \right) = \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} + O(1) \quad \text{e} \quad E(\hat{N}_{ab,s}) = N_{ab} + O \left(\frac{1}{n} \right).$$

O uso de $\hat{N}_{ab,s}$ é recomendado quando não é praticável identificar elementos da amostra que pertencem aos dois cadastros.

Com isso, são propostos 4 estimadores assistidos por modelos. Os estimadores 1 e 2 são propostos para o cenário 2, enquanto os estimadores 3 e 4 são propostos para o cenário 3. Os estimadores serão denotados por $\hat{t}_{yr1}$, $\hat{t}_{yr2}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ e são escritos da seguinte forma:

$$\begin{aligned}
\text{i.} \quad \hat{t}_{yr1} &= \frac{\tilde{y}_a}{\tilde{x}_a} t_{xa} + p \frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}} t'_{xab} + (1-p) N_{ab} \tilde{y}''_{ab} + N_b \tilde{y}_b; \\
\text{ii.} \quad \hat{t}_{yr2} &= \left(\frac{N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}}{N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}} \right) t_{X_A} + N_b \tilde{y}_b + (1-p) N_{ab} \tilde{y}''_{ab}; \\
\text{iii.} \quad \hat{t}_{yr3} &= \frac{\tilde{y}_a}{\tilde{x}_a} t_{xa} + p \frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}} t'_{xab} + (1-p) \hat{N}_{ab,s} \tilde{y}''_{ab} + \hat{N}_b \tilde{y}_b; \\
\text{iv.} \quad \hat{t}_{yr4} &= \left(\frac{\hat{N}_a \tilde{y}_a + p \hat{N}_{ab,s} \tilde{y}'_{ab}}{\hat{N}_a \tilde{x}_a + p \hat{N}_{ab,s} \tilde{x}'_{ab}} \right) t_{X_A} + \hat{N}_{ab,s} (1-p) \tilde{y}''_{ab} + \hat{N}_b \tilde{y}_b.
\end{aligned}$$

4.2 Estimadores do tipo razão sob a abordagem de cadastro duplo

A seguir serão apresentados os estimadores propostos originalmente no âmbito dessa dissertação. As provas de todas as variâncias aproximadas apresentadas estão no apêndice A, pág. 130.

4.2.1 Estimador razão 1

O estimador razão 1 foi proposto sob a condição em que existe informação da variável auxiliar disponível para os domínios pertencentes ao cadastro $\mathcal{A}$, ou seja, que t_{xa} e t_{xab} são conhecidos e a população é assistida de acordo com o modelo (4.1).

$$\begin{aligned}
\hat{t}_{yr1} &= \hat{t}_{y_{a,R}} + \hat{t}_{y_b} + p \hat{t}'_{y_{ab,R}} + (1-p) \hat{t}''_{y_{ab,R}} \\
&= N_a \frac{\tilde{y}_a}{\tilde{x}_a} \mu_{x_a} + N_{ab} \left(p \frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}} \mu'_{x_{ab}} + (1-p) \tilde{y}''_{ab} \right) + N_b \tilde{y}_b,
\end{aligned} \tag{4.10}$$

onde $\mu_{x_i} = \frac{t_{x_i}}{N_i}$.

O uso do método de linearização de Taylor possibilita que a variância aproximada de (4.10) seja obtida. Tem-se que

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr1}) = & N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2 + p^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}^{'2} \\
& + (R'_a)^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2 + p^2 (R'_{ab})^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}^{'2} \\
& - 2(R'_a)^2 \left(\frac{N_a}{n_a} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{ya} \sigma_{xa} \\
& - 2p^2 R'_{ab} \left(\frac{N_{ab}}{n'_{ab}} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab} \\
& + N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 + (1-p)^2 N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}^{'2}, \tag{4.11}
\end{aligned}$$

onde ρ_{xya} e ρ_{xyab} representam a correlação entre as observações da amostra pertencentes à variável resposta e à variável auxiliar, que estão em U_a e U_{ab} , ou seja, $(Y_a; X_a)$ e $(Y_{ab}; X_{ab})$. Tem-se ainda que $P_a = \frac{N_a}{N_{\mathcal{A}}}$, $P'_{ab} = \frac{N_{ab}}{N_{\mathcal{A}}}$, $P_b = \frac{N_b}{N_{\mathcal{B}}}$, $P''_{ab} = \frac{N_{ab}}{N_{\mathcal{B}}}$, $R'_a = \frac{t_{ya}}{t_{xa}}$, $R'_{ab} = \frac{t_{yab}}{t_{xab}}$ e σ_{td}^2 representa a variância populacional da variável T pertencente ao domínio d . O valor de p que minimiza a variância de $\hat{t}_{yr1}$ é obtido no apêndice B, pág. 147.

Um estimador para p , denotado por $\hat{p}$ é obtido ao substituir as quantidades populacionais desconhecidas pelos seus respectivos estimadores. A partir de $\hat{p}$, um estimador da variância é obtido de maneira análoga.

4.2.2 Estimador razão 2

Ainda sob o cenário 2, o estimador a seguir é utilizado sob o ponto de vista em que apenas $t_{X_{\mathcal{A}}}$ é conhecido, ou seja, a informação auxiliar é conhecida apenas para o cadastro completo. Tem-se que

$$\hat{t}_{yr2} = \left[\frac{N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}}{N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}} \right] t_{X_{\mathcal{A}}} + N_b \tilde{y}_b + (1-p) N_{ab} \tilde{y}''_{ab} \tag{4.12}$$

A variância aproximada pode ser obtida a partir da aproximação linear do estimador, pelo método de linearização de Taylor. A variância de $\hat{t}_{yr2}$ é dada por

$$A\text{Var}(\hat{t}_{yr2}) = t_{X_{\mathcal{A}}}^2 A\text{Var}(\hat{G}_{\mathcal{A}}) + \text{Var}(\hat{G}_{\mathcal{B}}) \quad (4.13)$$

onde

$$\begin{aligned} A\text{Var}(\hat{G}_{\mathcal{A}}) &= \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 \left\{ N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2 + p^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}^{'2} \right. \\ &+ G_{\mathcal{A}}^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2 + p^2 G_{\mathcal{A}}^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}^{'2} \\ &- 2G_{\mathcal{A}} N_a^2 \frac{n_{\mathcal{A}}}{n_a^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{ya} \sigma_{xa} \\ &\left. - 2p^2 G_{\mathcal{A}} N_{ab}^2 \frac{n_{\mathcal{A}}}{n_{ab}^{'2}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab} \right\} \end{aligned}$$

$$\text{Var}(\hat{G}_{\mathcal{B}}) = N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 + (1-p)^2 N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}^{'2}$$

Tem-se ainda que $G_{\mathcal{A}} = \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*}$, com $t_{y\mathcal{A}}^* = N_a \mu_{ya} + p N_{ab} \mu'_{yab}$, $t_{x\mathcal{A}}^* = N_a \mu_{xa} + p N_{ab} \mu'_{xab}$ e σ_{td}^2 é a variância da variável T no domínio d .

O valor de p que minimiza a variância dada em 4.13 é fornecido no apêndice B, pág. 147. Para obter um estimador para p , basta substituir as quantidades populacionais desconhecidas pelos respectivos estimadores em cada termo k_i . Uma vez obtido um estimador para p , um estimador para a variância aproximada de $\hat{t}_{yr2}$ é obtido de maneira análoga.

Admitindo a situação que $U_B \subset U_A$, o estimador razão 2 reduz-se à:

$$\hat{t}_{yr2} = \left(\frac{N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}}{N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}} \right) t_{X_A} + (1-p) N_{ab} \tilde{y}''_{ab}, \quad (4.14)$$

Com variância aproximada igual a

$$\begin{aligned} A\text{Var}(\hat{G}_A) &= \left(\frac{1}{t_{x_A}^*} \right)^2 \left\{ N_A^2 \left(\frac{1-f_A}{n_A} \right) P_a \sigma_{ya}^2 + p^2 N_A^2 \left(\frac{1-f_A}{n_A} \right) P'_{ab} \sigma_{yab}^{'2} \right. \\ &+ G_A^2 N_A^2 \left(\frac{1-f_A}{n_A} \right) P_a \sigma_{xa}^2 + p^2 G_A^2 N_A^2 \left(\frac{1-f_A}{n_A} \right) P'_{ab} \sigma_{xab}^{'2} \\ &- 2 G_A t_{x_A}^{*2} N_a^2 \frac{n_A}{n_a^2} \left(1 - \frac{n_A}{N_A} \right) \rho_{xya} \sigma_{ya} \sigma_{xa} \\ &- \left. 2 p^2 G_A N_{ab}^2 \frac{n_A}{n_{ab}^{'2}} \left(1 - \frac{n_A}{N_A} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab} \right\} \\ &+ (1-p)^2 N_B^2 \left(\frac{1-f_B}{n_B} \right) P''_{ab} \sigma_{yab}^{'2}. \end{aligned} \quad (4.15)$$

O valor de p que minimiza (4.15) também é obtido a partir da análise das soluções de um polinômio de grau 6, onde os termos k_i , $i = 1, \dots, 6$ sob a são fornecidos no apêndice B, pág. 147. Um estimador para p pode ser obtido ao substituir as quantidades populacionais desconhecidas pelos respectivos estimadores em cada termo k_i . Uma vez obtido um estimador para p , um estimador para a variância aproximada do estimador é obtido de maneira análoga.

4.2.3 Estimador razão 3

Da mesma forma que $\hat{t}_{yr1}$, é admitido que a informação referente à variável auxiliar é conhecida para os domínios. Sob o cenário 3, o estimador para um total de um certo domínio d pode ser dado pelo seguinte estimador:

$$\tilde{t}_d = \hat{N}_d \tilde{y}_d$$

Com isso, é possível propor estimadores que estão em função de quantidades desse tipo. O estimador razão 3 é definido da seguinte forma:

$$\hat{t}_{yr3} = \frac{\tilde{y}_a}{\tilde{x}_a} t_{xa} + p \frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}} t'_{xab} + (1-p) \hat{N}_{ab,s} \tilde{y}''_{ab} + \left(N_B - \hat{N}_{ab,s} \right) \tilde{y}_b, \quad (4.16)$$

onde $\hat{N}_{ab,s}$ é o estimador para o tamanho do domínio U_{ab} , proposto por Fuller & Burmeister (1972). Note que $N_B - \hat{N}_{ab,s} = \hat{N}_b$. Com isso, o estimador razão 3 está em função de quantidades que não levam em consideração a presença de duplicatas no processo de estimação. Note ainda que estimadores do tipo $\tilde{t}_d$ não são lineares, e mais uma vez é necessário utilizar a técnica de linearização de Taylor para obter uma expressão aproximada da variância de 4.16. Logo, a variância aproximada do estimador $\hat{t}_{yr3}$ é dada por

$$\begin{aligned} A\text{Var}(\hat{t}_{yr3}) &= t_{ya}^2 \left(\frac{1-f_A}{n_A P_a} \right) \sigma_{xa}^2 + p^2 t_{xab}^2 \left(\frac{1-f_A}{n_A P'_{ab}} \right) \sigma_{xab}^2 + N_a^2 \left(\frac{1-f_A}{n_A P_a} \right) \sigma_{ya}^2 \\ &+ p^2 N_{ab}^2 \left(\frac{1-f_A}{n_A P'_{ab}} \right) \sigma_{yab}^2 + N_b^2 \left(\frac{1-f_B}{n_B P_b} \right) \sigma_{yb}^2 + N_{ab}^2 (1-p)^2 \left(\frac{1-f_B}{n_B P''_{ab}} \right) \sigma_{yab}^2 \\ &+ \left[(1-p)^2 (\mu_{yab}^2) + \mu_{yb}^2 \right] \frac{N_a N_b N_{ab} g_A g_B}{n_A N_b g_B + n_B N_a g_A} - 2 N_a t_{ya} \frac{n_A}{n_a^2} \left(1 - \frac{n_A}{N_A} \right) \rho_{xya} \sigma_{xa} \sigma_{ya} \\ &- 2 p^2 N_{ab} t'_{xab} \frac{n_A}{n_{ab}^2} \left(1 - \frac{n_A}{N_A} \right) \rho'_{xyab} \sigma'_{xab} \sigma'_{yab}. \end{aligned} \quad (4.17)$$

O valor de p que minimiza 4.17 é obtido no apêndice B, pág. 147.

Como $\hat{t}_{yr3}$ é proposto sob o cenário 3, as quantidades P_a , P'_{ab} , P_b e P''_{ab} são desconhecidas. Logo, para obter um estimador para p , também é necessário obter estimativas para estas quantidades. Dessa forma, tem-se que

$$\hat{P}_a = \frac{n_a}{n_{\mathcal{A}}} ; \quad \hat{P}'_{ab} = \frac{n'_{ab}}{n_{\mathcal{A}}} ; \quad \hat{P}_b = \frac{n_b}{n_{\mathcal{B}}} ; \quad \hat{P}''_{ab} = \frac{n''_{ab}}{n_{\mathcal{B}}}.$$

4.2.4 Estimador razão 4

O estimador razão 4 é proposto da mesma forma que $\hat{t}_{yr2}$. Porém, sua adoção deve ser levada em consideração para o caso em que a estrutura da população está sob o cenário 3. Com isso, tem-se que

$$\hat{t}_{yr4} = \left[\frac{\hat{N}_a \tilde{y}_a + p \hat{N}_{ab,s} \tilde{y}'_{ab}}{\hat{N}_a \tilde{x}_a + p \hat{N}_{ab,s} \tilde{x}'_{ab}} \right] t_{X_{\mathcal{A}}} + \hat{N}_{ab,s} (1-p) \tilde{y}''_{ab} + \hat{N}_b \tilde{y}_b, \quad (4.18)$$

com variância aproximada igual a

$$A\text{Var}(\hat{t}_{yr4}) = t_{X_{\mathcal{A}}}^2 A\text{Var}(\hat{G}_{\mathcal{A}}) + A\text{Var}(\hat{G}_{\mathcal{B}}), \quad (4.19)$$

$$\begin{aligned} A\text{Var}(\hat{G}_{\mathcal{A}}) &= \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 \left\{ \left(\mu_{ya}^2 + p^2 \mu'_{yab} + G_{\mathcal{A}}^2 \mu_{xa}^2 + p^2 G_{\mathcal{A}}^2 \mu_{xab}^{'2} \right) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \right. \\ &+ N_a^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P_a} \right) \sigma_{ya}^2 + p^2 N_{ab}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P'_{ab}} \right) \sigma_{yab}^{'2} + G_{\mathcal{A}}^2 N_a^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P_a} \right) \sigma_{xa}^2 \\ &+ p^2 G_{\mathcal{A}}^2 N_{ab}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P'_{ab}} \right) \sigma_{xab}^{'2} - 2 G_{\mathcal{A}} N_a^2 \frac{n_{\mathcal{A}}}{n_a^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{xa} \sigma_{ya} \\ &\left. - 2 p^2 G_{\mathcal{A}} N_{ab}^2 \frac{n_{\mathcal{A}}}{n_{ab}^{'2}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{xab} \sigma'_{yab} \right\}, \end{aligned}$$

onde $G_{\mathcal{A}} = \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*}$, com $t_{y\mathcal{A}}^* = N_a\mu_{ya} + pN_{ab}\mu'_{yab}$, $t_{x\mathcal{A}}^* = N_a\mu_{xa} + pN_{ab}\mu'_{xab}$ e σ_{td}^2 é a variância da variável T no domínio d .

$$\begin{aligned} A\text{Var}(\hat{G}_{\mathcal{B}}) &= G_{\mathcal{B}} + \left\{ (1-p)^2(\mu_{yab}''^2) + (\mu_{yb}^2) \right\} \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \\ &+ (1-p)^2 N_{ab}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}} P_{ab}''} \right) \sigma_{yab}''^2 + N_b^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}} P_b} \right) \sigma_{yb}^2 \end{aligned}$$

onde $G_{\mathcal{B}} = N_b\mu_{yb} + (1-p)N_{ab}\mu_{yab}''$. Da mesma forma que a vista para o estimador razão 2, o valor de p que minimiza 4.19 é obtido no apêndice B, pág 147. Aqui, as quantidades P_a , P'_{ab} , P_b e P''_{ab} também são desconhecidas. Logo, para obter um estimador para p , também é necessário obter estimativas para estas quantidades.

Estimação assistida por modelos em uma abordagem de cadastro duplo sob estratificação

Neste capítulo será apresentado o desenvolvimento de estimadores aplicáveis a um contexto de estratificação com base nos estimadores apresentados no capítulo 4. Seja U o conjunto de índices que representam os elementos pertencentes à população-alvo. Seja N o número de elementos de U . Tem-se que U é dividido em H estratos, pertencentes ao estrato h , onde $h = 1, 2, \dots, H$. Uma vez que o plano de amostragem aleatória simples é utilizado em cada estrato, n_h observações são aleatoriamente selecionadas em cada estrato.

Sob a situação em que são considerados domínios de interesse, é possível ampliar a perspectiva referente à estratificação. Seja U_d o conjunto de elementos da população que representa um domínio d e seja $\mathcal{E}$ o número de estratos que contêm elementos de U_d . Dessa forma,

$$U_d = U_{d,1} \cup U_{d,2} \cup \dots \cup U_{d,\mathcal{E}} \dots \cup U_{d,H} \quad \text{e} \quad N_d = N_{d(1)} + N_{d(2)} + \dots + N_{d(\mathcal{E})} + \dots + N_{d(H)}$$

onde $U_{d,e}$ representa o conjunto de elementos pertencentes ao domínio d , dentro do estrato e , onde $e = 1, \dots, \mathcal{E}$. $N_{d(e)}$ representa o tamanho populacional do domínio d dentro do estrato e .

Sob a abordagem de cadastro duplo, sejam $\mathcal{J}$ e $\mathcal{W}$ o número de estratos pertencentes aos cadastros $\mathcal{A}$ e $\mathcal{B}$, respectivamente. Note que

$$\begin{aligned}
U_{\mathcal{A}} &= U_a \bigcup U_{ab} \\
&= \left(U_{a,1} \bigcup \dots \bigcup U_{a,\mathcal{J}} \right) \bigcup \left(U_{ab,1} \bigcup \dots \bigcup U_{ab,\mathcal{J}} \right) \\
&= \left(U_{a,1} \bigcup U_{ab,1} \right) \bigcup \left(U_{a,2} \bigcup U_{ab,2} \right) \bigcup \dots \bigcup \left(U_{a,\mathcal{J}} \bigcup U_{ab,\mathcal{J}} \right) \\
&= U_{\mathcal{A},1} \bigcup U_{\mathcal{A},2} \bigcup \dots \bigcup U_{\mathcal{A},\mathcal{J}}
\end{aligned} \tag{5.1}$$

$$\begin{aligned}
U_{\mathcal{B}} &= U_b \bigcup U_{ab} \\
&= \left(U_{b,1} \bigcup \dots \bigcup U_{b,\mathcal{W}} \right) \bigcup \left(U_{ab,1} \bigcup \dots \bigcup U_{ab,\mathcal{W}} \right) \\
&= \left(U_{b,1} \bigcup U_{ab,1} \right) \bigcup \left(U_{b,2} \bigcup U_{ab,2} \right) \bigcup \dots \bigcup \left(U_{b,\mathcal{W}} \bigcup U_{ab,\mathcal{W}} \right) \\
&= U_{\mathcal{B},1} \bigcup U_{\mathcal{B},2} \bigcup \dots \bigcup U_{\mathcal{B},\mathcal{W}}
\end{aligned} \tag{5.2}$$

Temos ainda que $N_{ab(jw)}$ representa o tamanho da população no domínio ab que pertence à interseção do estrato j de $\mathcal{A}$ com o estrato w de $\mathcal{B}$, onde $j = 1, \dots, \mathcal{J}$ e $w = 1, \dots, \mathcal{W}$. Sob a perspectiva da estratificação, se um cadastro é utilizado para aplicação do plano de amostragem estratificada, e este encontra-se incompleto no sentido de que elementos não listados no cadastro constituírem um estrato, tem-se a situação conceitual em que é aplicado um plano amostral na população sem obter nenhuma amostra desse estrato de elementos não listados, mas que fazem parte da população-alvo. Dessa forma, ao se aplicar planos de amostragem estratificada em dois cadastros de forma independente, tem-se a situação em que duas amostras independentes são obtidas da população-alvo.

Para estabelecer uma notação apropriada para a abordagem de cadastro duplo aliada ao plano AAE, a tabela (5.1) ilustra a notação utilizada. Dessa forma, serão apresentados estimadores dos tipos separados e combinados, que são adaptações dos estimadores apresentados no capítulo 4, sob a perspectiva de estratificação. Sob o cenário 2,

- (i) $N_{\mathcal{A}(1)}, N_{\mathcal{A}(2)}, \dots, N_{\mathcal{A}(\mathcal{J})}$ são conhecidos;
- (ii) $N_{\mathcal{B}(1)}, N_{\mathcal{B}(2)}, \dots, N_{\mathcal{B}(\mathcal{W})}$ são conhecidos;
- (iii) $N_{a(1)}, N_{a(2)}, \dots, N_{a(\mathcal{J})}$ são conhecidos;
- (iv) $N_{b(1)}, N_{b(2)}, \dots, N_{b(\mathcal{W})}$ são conhecidos.

Tabela 5.1: Notação utilizada em uma abordagem de cadastro duplo, sob o plano AAE

Quantidades de Interesse	CADASTRO		Domínio		
	$\mathcal{A}$	$\mathcal{B}$	a	b	ab
População	$U_{\mathcal{A}(j)}$	$U_{\mathcal{B}(w)}$	$U_{a(j)}$	$U_{b(w)}$	$U_{ab(jw)}$
Tamanho da População	$N_{\mathcal{A}(j)}$	$N_{\mathcal{B}(w)}$	$N_{a(j)}$	$N_{b(w)}$	$N_{ab(jw)}$
Total populacional	$t_{y\mathcal{A}(j)}$	$t_{y\mathcal{B}(w)}$	$t_{ya(j)}$	$t_{yb(w)}$	$t_{yab(jw)}$
Média populacional	$\mu_{y\mathcal{A}(j)}$	$\mu_{y\mathcal{B}(w)}$	$\mu_{ya(j)}$	$\mu_{yb(w)}$	$\mu_{yab(jw)}$
Amostra	$S_{\mathcal{A}(j)}$	$S_{\mathcal{B}(w)}$	$S_{a(j)}$	$S_{b(w)}$	$S_{ab(jw)}$
Tamanho da amostra	$n_{\mathcal{A}(j)}$	$n_{\mathcal{B}(w)}$	$n_{a(j)}$	$n_{b(w)}$	$n'_{ab(jw)} \quad n''_{ab(jw)}$
Estimador do Total	$\hat{t}_{y\mathcal{A}(j)}$	$\hat{t}_{y\mathcal{B}(w)}$	$\hat{t}_{ya(j)}$	$\hat{t}_{yb(w)}$	$\hat{t}'_{yab(jw)} \quad \hat{t}''_{yab(jw)}$
Média amostral	$\bar{y}_{\mathcal{A}(j)}$	$\bar{y}_{\mathcal{B}(w)}$	$\bar{y}_{a(j)}$	$\bar{y}_{b(w)}$	$\bar{y}'_{ab(jw)} \quad \bar{y}''_{ab(jw)}$

Sob o cenário 3, apenas $N_{\mathcal{A}}$, $N_{\mathcal{B}}$ e as quantidades (i) e (ii) são conhecidos. As quantidades populacionais da variável auxiliar no cadastro $\mathcal{A}$, $t_{X_{\mathcal{A}}}$, $t_{xa} = t_{xa(1)} + \dots + t_{xa(\mathcal{J})}$ e $t_{xab} = t_{xab(1)} + \dots + t_{xab(\mathcal{W})}$ são conhecidas.

5.1 Estimadores do tipo razão separado sob a abordagem de cadastro duplo

Os estimadores do tipo separado serão utilizados quando houver evidência de que a relação entre a variável de interesse e a variável auxiliar na população, para cada estrato, é descrita por um modelo da forma

$$\xi : y_{k\mathcal{A}} = \begin{cases} \beta_{a(j)}x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}} & \text{se } k \in U_{a(j)} \\ \beta_{ab(jw)}x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}} & \text{se } k \in U_{ab(jw)} \end{cases} \quad (5.3)$$

onde

$$E_{\xi}(\varepsilon_{k\mathcal{A}}) = 0, \forall k \in U_{\mathcal{A}}; \quad E_{\xi}(y_{k\mathcal{A}}) = \begin{cases} \beta_{a(j)}x_{k\mathcal{A}}, & k \in U_{a(j)} \\ \beta_{ab(jw)}x_{k\mathcal{A}}, & \text{se } k \in U_{ab(jw)} \end{cases}$$

$$\text{Var}_{\xi}(y_{k\mathcal{A}}) = \lambda x_{k\mathcal{A}}, \forall k \in U_{\mathcal{A}(j)}$$

Tem-se que os coeficientes $\beta_{a(j)}$ e $\beta_{ab(jw)}$, são diferentes entre os estratos. Para cálculo das quantidades referentes aos estimadores, seja $k_{d(e)}$ o conjunto não-vazio de elementos pertencentes ao domínio d dentro do estrato e . A seguir são apresentados os estimadores separados propostos originalmente no âmbito dessa dissertação. As provas das variâncias aproximadas encontram-se no apêndice C, pág. 159.

5.1.1 Estimador razão separado 1

Sob o modelo (5.3), o estimador separado 1 foi obtido a partir do estimador razão 1 proposto no capítulo 4. A forma do estimador separado 1 leva em consideração a soma das estimativas de totais para cada domínio, e é dada pela seguinte expressão

$$\hat{t}_{yrs1} = \sum_{j=1}^{\mathcal{J}} \tilde{\delta}_{a1(j)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} p_{(jw)} \tilde{\delta}'_{ab1(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b1(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab1(jw)}, \quad (5.4)$$

onde

$$\tilde{\delta}_{a1(j)} = \begin{cases} \frac{\tilde{y}_{a(j)}}{\tilde{x}_{a(j)}} t_{xa(j)}, & \text{se } k_{a(j)} \text{ existe} \\ 0, & \text{c.c.} \end{cases} \quad \tilde{\delta}_{b1(w)} = \begin{cases} N_{b(w)} \tilde{y}_{b(w)}, & \text{se } k_{b(w)} \text{ existe} \\ 0, & \text{c.c.} \end{cases}$$

$$\tilde{\delta}'_{ab1(jw)} = \begin{cases} \frac{\tilde{y}'_{ab(jw)}}{\tilde{x}'_{ab(jw)}} t_{xab(jw)}, & \text{se } k_{ab(jw)} \text{ existe, cadastro } \mathcal{A} \\ 0, & \text{c.c.} \end{cases}$$

$$\tilde{\delta}''_{ab1(jw)} = \begin{cases} \frac{\tilde{y}''_{ab(jw)}}{\tilde{x}''_{ab(jw)}} t_{xab(jw)}, & \text{se } k_{ab(jw)} \text{ existe, cadastro } \mathcal{B} \\ 0, & \text{c.c.} \end{cases}$$

tem-se que $p_{(jw)}$ é uma constante de ponderação tal que $0 \leq p_{(jw)} \leq 1$ para identificação dos elementos de cada cadastro e estrato na população. Se o domínio a ou b for considerado

um estrato, tanto a soma em j referente ao domínio a quanto a soma em w referente ao domínio b terá apenas um termo, e os demais componentes da soma são nulos. Se a e b forem considerados estratos, a situação a ser abordada recai no cenário 1, onde cada domínio é considerado um estrato e demais estratos são identificados dentro do domínio ab . Note que essa mesma situação ocorre se o domínio ab for indicado como um estrato na fase de planejamento. A variância do estimador separado 1 é dada por:

$$\begin{aligned}
A\text{Var}(\hat{t}_{yrs1}) &= \sum_{j=1}^{\mathcal{J}} [\text{Var}(\tilde{t}_{ya(j)}) + R_{a(j)}^2 \text{Var}(\tilde{t}_{xa(j)}) - 2R_{a(j)} \text{ACov}(\tilde{t}_{ya(j)}; \tilde{t}_{xa(j)})] \\
&+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} [p_{(jw)}^2 \text{Var}(\tilde{t}'_{yab(jw)}) + p_{(jw)}^2 R_{ab(jw)}^2 \text{Var}(\tilde{t}'_{xab(jw)}) \\
&- 2R_{ab(jw)} p_{(jw)}^2 \text{ACov}(\tilde{t}_{yab(jw)}; \tilde{t}_{xab(jw)})] \\
&+ \sum_{w=1}^{\mathcal{W}} \text{Var}(\tilde{t}_{yb(w)}) + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 \text{Var}(\tilde{t}''_{yab(jw)}) \\
&= \sum_{j=1}^{\mathcal{J}} \left[N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 + R_{a(j)}^2 N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \right. \\
&- \left. 2R_{a(j)}^2 \left(\frac{N_{a(j)}}{n_{a(j)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \right] \\
&+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left[p_{(jw)}^2 N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \right. \\
&+ p_{(jw)}^2 R_{ab(jw)}^2 N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\
&- \left. 2R_{ab(jw)} p_{(jw)}^2 \left(\frac{N_{ab(jw)}}{n'_{ab(jw)}} \right) n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right] \\
&+ \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\
&+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2 \tag{5.5}
\end{aligned}$$

Os termos que compõem as somas em j e em w são similares aos termos do estimador razão 1, sob o plano AAS. Neste caso, as constantes de ponderação $p_{(jw)}$ são os pontos que minimizam a variância aproximada do estimador para dois estratos j e w , e são obtidas através do mesmo procedimento descrito para o estimador razão 1.

5.1.2 Estimador razão separado 2

O estimador separado 2 foi obtido a partir do estimador razão 2. A forma do estimador leva em consideração a soma de todos os componentes $G_{\mathcal{A}(jw)}$ dos estratos em $\mathcal{A}$, onde

$$G_{\mathcal{A}(jw)} = \frac{N_{a(j)}\tilde{y}_a + p_{(jw)}N_{ab(jw)}\tilde{y}'_{ab(jw)}}{N_{a(j)}\tilde{x}_a + p_{(jw)}N_{ab(jw)}\tilde{x}'_{ab(jw)}}, \quad \text{para } w \text{ fixo}$$

Verifica-se que $p_{(jw)}$ é uma constante de ponderação para identificação dos elementos na população, onde $0 \leq p_{(jw)} \leq 1$. A forma do estimador é dada por

$$\begin{aligned} \hat{t}_{yrs2} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab2(jw)} t_{\mathcal{I}(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b2(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab2(jw)} \\ \hat{t}_{yrs2} &= \hat{\theta}_{\mathcal{A}(jw)} + \hat{\theta}_{\mathcal{B}(jw)}, \end{aligned} \quad (5.6)$$

onde

$$\hat{\theta}_{\mathcal{A}(jw)} = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab2(jw)} t_{\mathcal{I}(jw)} \quad \hat{\theta}_{\mathcal{B}(jw)} = \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b2(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab2(jw)}$$

$$\tilde{\delta}'_{ab2(jw)} = \begin{cases} \frac{N_{a(j)}\tilde{y}_a + p_{(jw)}N_{ab(jw)}\tilde{y}'_{a(jw)}}{N_{a(j)}\tilde{x}_a + p_{(jw)}N_{ab(jw)}\tilde{x}'_{a(jw)}}, & \text{se } k_{a(j)} \text{ existe,} & N_{ab(jw)} \neq 0 \\ \frac{\tilde{y}_a}{\tilde{x}_a}, & \text{se } k_{a(j)} \text{ existe,} & N_{ab(jw)} = 0, N_{a(j)} \neq 0 \\ p_{(jw)} \frac{\tilde{y}'_{ab(jw)}}{\tilde{x}'_{ab(jw)}}, & \text{se } k_{ab(jw)} \text{ existe, cad. } \mathcal{A} & N_{ab(jw)} \neq 0, N_{a(j)} = 0 \\ 0, & \text{se } N_{a(j)} = 0, & N_{ab(jw)} = 0 \end{cases}$$

$$\tilde{\delta}_{b2(w)} = \begin{cases} N_{b(w)}\tilde{y}_{b(w)}, & \text{se } k_{b(w)} \text{ existe} \\ 0, & \text{c.c.} \end{cases} \quad \tilde{\delta}_{ab2(jw)}'' = \begin{cases} N_{ab(jw)}\tilde{y}_{ab(jw)}'', & \text{se } k_{ab(jw)} \text{ existe, cad. } \mathcal{B} \\ 0, & \text{c.c.} \end{cases}$$

$$t_{\mathcal{I}(jw)} = \begin{cases} t_{X_{\mathcal{A}(j)}}, & \text{se } N_{ab(jw)} \neq 0 \\ t_{xa(j)}, & \text{se } N_{ab(jw)} = 0 \\ t_{xab(jw)}, & \text{se } N_{a(j)} = 0 \end{cases} \quad (5.7)$$

Simplificações em $\tilde{\delta}_{ab(jw)}'$ podem ser verificadas no caso em que todos os estratos em $\mathcal{A}$ possuem elementos pertencentes ao domínio a . Neste caso,

$$\tilde{\delta}_{ab2(jw)}' = \begin{cases} \frac{N_{a(j)}\tilde{y}_{a(j)} + p_{(jw)}N_{ab(jw)}\tilde{y}_{a(jw)}'}{N_{a(j)}\tilde{x}_{a(j)} + p_{(jw)}N_{ab(jw)}\tilde{x}_{a(jw)}'}, & \text{se } k_{a(j)} \text{ existe, } N_{ab(jw)} \neq 0 \\ \frac{\tilde{y}_{a(j)}}{\tilde{x}_{a(j)}}, & \text{se } k_{a(j)} \text{ existe, } N_{ab(jw)} = 0, N_{a(j)} \neq 0 \\ 0, & \text{se } N_{a(j)} = 0, N_{ab(jw)} = 0 \end{cases}$$

verifica-se que quantidades populacionais conhecidas e referentes à variável auxiliar dentro de cada estrato j são utilizadas. Porém, dependendo da forma como os estratos são constituídos em $\mathcal{A}$ e $\mathcal{B}$, a razão $\frac{\tilde{y}_a}{\tilde{x}_a}t_{xa(j)}$ pode ser repetida z_j vezes em $\hat{\theta}_{\mathcal{A}(jw)}$, onde $z_j \geq 1$, quando $N_{ab(jw)} = 0$, para j e w fixos. Assim, de modo a evitar que tais repetições apareçam, a forma do estimador corrigido ($\hat{t}_{yrs2,corr}$) é dada por

$$\begin{aligned} \hat{t}_{yrs2,corr} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{ab2(jw)}' t_{\mathcal{I}(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b2(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}_{ab2(jw)}'' \\ &- \sum_{j=1}^{\mathcal{J}} z_j \frac{\tilde{y}_{a(j)}}{\tilde{x}_{a(j)}} t_{xa(j)} I_{(j)} \end{aligned} \quad (5.8)$$

onde

$$I_{(j)} = \begin{cases} 1, & \text{se } N_{ab(jw)} \neq 0, \forall w \\ 0, & \text{se } N_{ab(jw)} = 0, \forall w \end{cases} \quad (5.9)$$

A variância aproximada de $\hat{t}_{yrs2}$ sem correção é dada por

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs2}) &= \sum_{j=1}^J \sum_{w=1}^W t_{\mathcal{I}(jw)}^2 \left(\frac{1}{t_{x\mathcal{A}(jw)}^*} \right)^2 \left\{ N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ &+ p_{(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2G_{\mathcal{A}(jw)} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \\ &\left. - 2p^2 G_{\mathcal{A}(jw)} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ &+ \sum_{w=1}^W N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\ &+ \sum_{j=1}^J \sum_{w=1}^W (1 - p_{(jw)})^2 N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2, \end{aligned} \quad (5.10)$$

onde

$$t_{x\mathcal{A}(jw)}^* = t_{xa(j)} + p t'_{xab(jw)}$$

Os valores $p_{(jw)}$ são obtidos individualmente da mesma forma que a descrita para o estimador razão 2. Cada $p_{(jw)}$ é o valor que minimiza a variância do estimador combinado para dois cadastros j e w . Como a forma da variância foi obtida para o estimador separado 2 sem correção, é possível incorporar uma correção na forma da variância do estimador.

Sob a situação em que $U_{\mathcal{B}} \subset U_{\mathcal{A}}$, o estimador $\hat{t}_{yrs2}$ reduz-se à:

$$\hat{t}_{yrs2} = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab2(jw)} t_{\mathcal{I}(jw)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab2(jw)}, \quad (5.11)$$

com variância aproximada da seguinte forma:

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs2}) &= \sum_{j=1}^J \sum_{w=1}^W t_{\mathcal{I}(jw)}^2 \left(\frac{1}{t_{x\mathcal{A}(jw)}^*} \right)^2 \left\{ N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ &+ p_{(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2G_{\mathcal{A}(jw)} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \\ &- \left. 2p^2 G_{\mathcal{A}(j)} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ &+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2 \end{aligned} \quad (5.12)$$

5.1.3 Estimador razão separado 3

Sob o cenário 3, o estimador razão separado 3 é a versão estratificada do estimador razão 3 proposto no capítulo 4. Neste caso, o estimador é calculado sob a situação em que os tamanhos dos conjuntos de elementos para cada domínio são desconhecidos em cada estrato. A forma do estimador separado 3 é dada por

$$\begin{aligned}\hat{t}_{yrs3} &= \sum_{j=1}^{\mathcal{J}} \tilde{\delta}_{a3(j)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} p_{(jw)} \tilde{\delta}'_{ab3(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b3(w)} \\ &+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab3(jw)},\end{aligned}\quad (5.13)$$

onde

$$\tilde{\delta}_{a3(j)} = \tilde{\delta}_{a1(j)} \quad \tilde{\delta}'_{ab3(jw)} = \tilde{\delta}'_{ab1(jw)}$$

$$\tilde{\delta}_{b3(w)} = \begin{cases} \hat{N}_{b(w)} \tilde{y}_{b(w)}, & \text{se } k_{b(w)} \text{ existe} \\ 0, & \text{c.c.} \end{cases}$$

$$\tilde{\delta}''_{ab3(jw)} = \begin{cases} \hat{N}_{ab,s(jw)} \tilde{y}''_{ab(jw)}, & \text{se } k_{ab(jw)} \text{ existe} \rightarrow \text{cadastro } \mathcal{B} \\ 0, & \text{c.c.} \end{cases}$$

Sendo $N_{ab(jw)}$ desconhecido, um estimador pode ser apresentado da mesma forma que a apresentada em (4.9), ou seja:

$$q_1 \hat{N}_{ab,s(jw)}^2 - q_2 \hat{N}_{ab,s(jw)} + q_3 = 0 \quad (5.14)$$

onde:

$$q_1 = n_{\mathcal{A}(j)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} g_{\mathcal{A}(j)}$$

$$q_2 = n_{\mathcal{A}(j)} N_{\mathcal{B}(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{\mathcal{A}(j)} g_{\mathcal{A}(j)} + n'_{ab(jw)} N_{\mathcal{A}(j)} g_{\mathcal{B}(w)} + n''_{ab(jw)} N_{\mathcal{B}(w)} g_{\mathcal{A}(j)}$$

$$q_3 = (n'_{ab(jw)}g_{\mathcal{B}(w)} + n''_{ab(jw)}g_{\mathcal{A}(j)}) N_{\mathcal{A}(j)}N_{\mathcal{B}(w)}$$

$$g_{\mathcal{A}(j)} = \frac{N_{\mathcal{A}(j)} - n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)} - 1} \quad g_{\mathcal{B}(w)} = \frac{N_{\mathcal{B}(w)} - n_{\mathcal{B}(w)}}{N_{\mathcal{B}(w)} - 1}$$

Com isso, tem-se que

$$\hat{N}_{a(j)} = N_{\mathcal{A}(j)} - \sum_{w=1}^W \hat{N}_{ab,s(jw)}$$

$$\hat{N}_{b(w)} = N_{\mathcal{B}(w)} - \sum_{j=1}^J \hat{N}_{ab,s(jw)}$$

A variância aproximada de $\hat{t}_{yrs3}$ é dada por

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs3}) &= \sum_{j=1}^J \left[t_{ya(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{xa(j)}^2 + N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{ya(j)}^2 \right. \\ &\quad \left. - 2N_{a(j)} t_{ya(j)} \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{xa(j)} \sigma_{ya(j)} \right] \\ &\quad + \sum_{j=1}^J \sum_{w=1}^W \left[p_{(jw)}^2 t_{xab(jw)}'^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{xab(jw)}'^2 \right. \\ &\quad + p_{(jw)}^2 N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{yab(jw)}'^2 \\ &\quad \left. - 2p_{(jw)}^2 N_{ab(jw)} t_{xab(jw)}' \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{xab(jw)} \sigma'_{yab(jw)} \right] \\ &\quad + \sum_{w=1}^W N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P_b} \right) \sigma_{yb}^2 + \sum_{j=1}^J \sum_{w=1}^W \mu_{yb(w)}^2 \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ &\quad + \sum_{j=1}^J \sum_{w=1}^W \left[N_{ab(jw)}^2 (1 - p_{(jw)})^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P''_{ab(jw)}} \right) \sigma_{yab(jw)}''^2 \right. \\ &\quad \left. + (1 - p_{(jw)})^2 (\mu_{yab(jw)}''^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \right] \end{aligned} \quad (5.15)$$

5.1.4 Estimador razão separado 4

O estimador razão separado 4 está em função de estimadores para quantidades desconhecidas na população ($N_{a(j)}$, $N_{b(w)}$ e $N_{ab(jw)}$). A forma do estimador razão separado 4 é análoga a expressão (7.2), pois

$$\begin{aligned}\hat{t}_{yrs4} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab4(jw)} t_{\mathcal{I}(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b4(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab4(jw)} \\ &= \hat{\theta}_{\mathcal{A}(jw)} + \hat{\theta}_{\mathcal{B}(jw)},\end{aligned}\tag{5.16}$$

onde

$$\begin{aligned}\hat{\theta}_{\mathcal{A}(jw)} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab4(jw)} t_{\mathcal{I}(jw)} & \hat{\theta}_{\mathcal{B}(jw)} &= \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b4(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab4(jw)} \\ \tilde{\delta}_{b4(w)} &= \tilde{\delta}_{b3(w)} & \delta''_{ab4(jw)} &= \delta''_{ab3(jw)} \\ \tilde{\delta}'_{ab4(jw)} &= \begin{cases} \frac{\hat{N}_{a(j)} \tilde{y}_{a(j)} + p_{(jw)} \hat{N}_{ab,s(jw)} \tilde{y}'_{a(jw)}}{\hat{N}_{a(j)} \tilde{x}_{a(j)} + p_{(jw)} \hat{N}_{ab,s(jw)} \tilde{x}'_{a(jw)}}, & \text{se } k_{a(j)} \text{ existe, } \mathcal{S}_{\mathcal{A}(j)} \cap \mathcal{S}_{\mathcal{B}(w)} = \mathcal{S}_{ab(jw)} \\ \frac{\tilde{y}_{a(j)}}{\tilde{x}_{a(j)}}, & \text{se } k_{a(j)} \text{ existe, } \mathcal{S}_{ab(jw)} = \emptyset \\ p_{(jw)} \frac{\tilde{y}'_{ab(jw)}}{\tilde{x}'_{ab(jw)}}, & \text{se } k_{ab(jw)} \text{ existe, } \mathcal{S}_{a(j)} = \emptyset \\ 0, & \text{se } \mathcal{S}_{a(j)} = \emptyset, \mathcal{S}_{ab(jw)} = \emptyset \end{cases}\end{aligned}$$

$t_{\mathcal{I}(jw)}$ é definido em (5.7) e $\hat{N}_{a(j)}$, $\hat{N}_{b(w)}$ e $\hat{N}_{ab(jw)}$ são obtidos de maneira análoga a do estimador 3. Neste caso, $\hat{N}_{ab(jw)} = 0$ no caso em que não houver interseção entre estratos de $\mathcal{A}$ e $\mathcal{B}$. Neste caso, também será necessário uma correção do estimador 4, com o objetivo de evitar que termos $\frac{\tilde{y}_{a(j)}}{\tilde{x}_{a(j)}} t_{xa(j)}$ sejam repetidos z_j vezes. Dessa forma,

$$\begin{aligned}
\hat{t}_{yrs4,corr} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab4(jw)} t_{\mathcal{I}(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b4(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab4(jw)} \\
&- \sum_{j=1}^{\mathcal{J}} z_j \tilde{\delta}_{a4(j)}
\end{aligned} \tag{5.17}$$

onde $\tilde{\delta}_{a4(j)} = \tilde{\delta}_{a1(j)}$

A variância aproximada para $\hat{t}_{yrs4}$ sem correção é dada por

$$\begin{aligned}
A\text{Var}(\hat{t}_{yrs4}) &= \sum_{j=1}^J \sum_{w=1}^W t_{\mathcal{I}(jw)}^2 \left(\frac{1}{t_{x\mathcal{A}(jw)}^*} \right)^2 \left\{ \left(\mu_{ya(j)}^2 + p^2 \mu_{yab(jw)}'^2 + G_{\mathcal{A}(j)}^2 \mu_{xa(j)}^2 \right. \right. \\
&+ \left. p^2 G_{\mathcal{A}(jw)}^2 \mu_{xab(jw)}'^2 \right) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\
&+ N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{ya(j)}^2 + p^2 N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{yab(jw)}'^2 \\
&+ G_{\mathcal{A}(jw)}^2 N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{xa(j)}^2 + p^2 G_{\mathcal{A}(j)}^2 N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{xab(jw)}'^2 \\
&- 2 G_{\mathcal{A}(jw)} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{xa(j)} \sigma_{ya(j)} \\
&- \left. 2 p^2 G_{\mathcal{A}(jw)} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{xab(jw)} \sigma'_{yab(jw)} \right\} \\
&+ \sum_{w=1}^W N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P_b} \right) \sigma_{yb}^2 + \sum_{j=1}^J \sum_{w=1}^W \mu_{yb(w)}^2 \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\
&+ \sum_{j=1}^J \sum_{w=1}^W \left[N_{ab(jw)}^2 (1 - p_{(jw)})^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P''_{ab(jw)}} \right) \sigma_{yab(jw)}''^2 \right. \\
&+ \left. (1 - p_{(jw)})^2 (\mu_{yab(jw)}''^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \right]
\end{aligned} \tag{5.18}$$

Os valores $p_{(jw)}$ são obtidos separadamente, e cada $p_{(jw)}$ é o valor que minimiza a variância do estimador para dois estratos j e w .

5.2 Estimadores do tipo razão combinado sob a abordagem de cadastro duplo

Analogamente aos estimadores do tipo razão separado, os estimadores do tipo razão combinado 1 e 2 estão sob o cenário 2. Os demais encontram-se na situação do cenário 3, com quantidades desconhecidas (tamanho da população nos domínios). Na composição dos estimadores combinados, para um dado estrato e considere que a informação da variável auxiliar referente ao total populacional está disponível para os domínios a e ab , e que haja evidência de que a relação entre a variável de interesse e a variável auxiliar na população, para cada estrato, é descrita por um modelo da forma

$$\xi : y_{k\mathcal{A}} = \begin{cases} \beta_{a(j)}x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}}, & \beta_{a(j)} = \beta_a \quad k \in U_{a(j)} \\ \beta_{ab(jw)}x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}}, & \beta_{ab(jw)} = \beta_{ab} \quad k \in U_{ab(jw)} \end{cases} \quad (5.19)$$

onde

$$E_{\xi}(\varepsilon_{k\mathcal{A}}) = 0, \quad \forall k; \quad E_{\xi}(y_{k\mathcal{A}}) = \begin{cases} \beta_{a(j)}x_{k\mathcal{A}}, & \beta_{a(j)} = \beta_a, \quad k \in U_{a(j)} \\ \beta_{ab(jw)}x_{k\mathcal{A}}, & \beta_{ab(jw)} = \beta_{ab}, \quad k \in U_{ab(jw)} \end{cases}$$

$$\text{Var}_{\xi}(y_{k\mathcal{A}}) = \lambda x_{k\mathcal{A}}, \quad \forall k \in U_{\mathcal{A}(j)}$$

Os estimadores combinados propostos originalmente no âmbito dessa dissertação são apresentados a seguir. As provas das variâncias aproximadas estão no apêndice C, pág. 159.

5.2.1 Estimador razão combinado 1

O estimador combinado 1 é outra modificação do estimador razão 1, dessa vez, sob a situação em que a população dentro dos estratos é assistida pelo modelo (5.19). Para a construção do estimador, é necessário definir o estimador do total de uma variável de interesse V , no domínio d , dentro do estrato e , sob o cenário 2, da seguinte forma:

$$\tilde{\alpha}_{Vd(e)} = \begin{cases} N_{d(e)} \tilde{y}_{d(e)} , & \text{se } k_{d(e)} \text{ existe} \\ 0 , & \text{c.c.} \end{cases}$$

Logo, o estimador combinado 1 é definido como

$$\begin{aligned} \hat{t}_{yrc1} &= \left(\frac{\sum_{j=1}^{\mathcal{J}} \tilde{\alpha}_{ya(j)}}{\sum_{j=1}^{\mathcal{J}} \tilde{\alpha}_{xa(j)}} \right) t_{xa} + p \left(\frac{\sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}'_{yab(jw)}}{\sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}'_{xab(jw)}} \right) t_{xab} \\ &+ \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}_{yb(w)} + (1-p) \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}''_{yab(jw)} \end{aligned} \quad (5.20)$$

A variância aproximada do estimador razão combinado 1 é dada pela seguinte expressão:

$$\begin{aligned} A\text{Var}(\hat{t}_{yrc1}) &= \sum_{j=1}^J N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \\ &+ p^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ R_{a(j)}^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 R_{ab(jw)}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2R_{a(j)}^2 \sum_{j=1}^J \left(\frac{N_{a(j)}}{n_{a(j)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \\ &- 2p^2 R_{ab(jw)}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left(\frac{N_{ab(jw)}}{n'_{ab(jw)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \\ &+ \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)}^2 \left(\frac{1-f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\ &+ (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)}^2 \left(\frac{1-f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2 \end{aligned} \quad (5.21)$$

5.2.2 Estimador razão combinado 2

O estimador 2 é derivado a partir do estimador razão 2. Nesse caso, o termo $G_{\mathcal{A}}$ envolverá todos os estimadores de totais para cada domínio dentro de cada estrato. A forma do estimador é dada por

$$\begin{aligned} \hat{t}_{yrc2} = & \left(\frac{\sum_{j=1}^{\mathcal{J}} \tilde{\alpha}_{ya(j)} + p \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}'_{yab(jw)}}{\sum_{j=1}^{\mathcal{J}} \tilde{\alpha}_{xa(j)} + p \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}'_{xab(jw)}} \right) t_{X_{\mathcal{A}}} + \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}_{yb(w)} \\ & + (1-p) \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}''_{yab(jw)} \end{aligned} \quad (5.22)$$

A forma da variância aproximada do estimador $\hat{t}_{yrc2}$ é dada por

$$\begin{aligned} AVar(\hat{t}_{yrc2}) = & t_{X_{\mathcal{A}}}^2 \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 \left\{ \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ & + p^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ & + G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ & + p^2 G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ & - 2G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya} \sigma_{ya(j)} \sigma_{xa(j)} \\ & \left. - 2p^2 G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ & + \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\ & + (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2 \end{aligned} \quad (5.23)$$

Sob a situação em que $U_{\mathcal{B}} \subset U_{\mathcal{A}}$ a forma de $\hat{t}_{yrc2}$ reduz-se à:

$$\hat{t}_{yrc2} = \left(\frac{\sum_{j=1}^{\mathcal{J}} \tilde{\alpha}_{ya(j)} + p \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}'_{yab(jw)}}{\sum_{j=1}^{\mathcal{J}} \tilde{\alpha}_{xa(j)} + p \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}'_{xab(jw)}} \right) t_{X_{\mathcal{A}}} + (1-p) \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\alpha}''_{yab(jw)}, \quad (5.24)$$

com variância aproximada igual a

$$\begin{aligned} A\text{Var}(\hat{t}_{yrc2}) = & t_{X_{\mathcal{A}}}^2 \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 \left\{ \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ & + p^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ & + G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ & + p^2 G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ & - 2G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya} \sigma_{ya(j)} \sigma_{xa(j)} \\ & \left. - 2p^2 G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ & + \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1-f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\ & + (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1-f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2 \end{aligned}$$

5.2.3 Estimador razão combinado 3

Sob cenário 3, a quantidade considere a quantidade $\tilde{\gamma}_{Vd(e)}$, definida da seguinte forma:

$$\tilde{\gamma}_{Vd(e)} = \begin{cases} \hat{N}_{d(e)} \tilde{y}_{d(e)} , & \text{se } k_{d(e)} \text{ existe} \\ 0 , & \text{c.c.} \end{cases}$$

A forma do estimador razão combinado 3 é dada por

$$\begin{aligned} \hat{t}_{yrc3} = & \left(\frac{\sum_{j=1}^{\mathcal{J}} \tilde{\gamma}_{ya(j)}}{\sum_{j=1}^{\mathcal{J}} \tilde{\gamma}_{xa(j)}} \right) t_{xa} + p \left(\frac{\sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}'_{yab(jw)}}{\sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}'_{xab(jw)}} \right) t_{xab} \\ & + \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}_{yb(w)} + (1-p) \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}''_{yab(jw)} \end{aligned} \quad (5.25)$$

A variância aproximada do estimador razão combinado 3 é dada pela expressão:

$$\begin{aligned} A\text{Var}(\hat{t}_{yrc3}) = & \sum_{j=1}^{\mathcal{J}} \left\{ \mu_{ya(j)}^2 \text{Var}(\hat{N}_{a(j)}) + N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{ya(j)}^2 \right\} \\ & + p^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}^2 \text{Var}(\hat{N}_{ab,s(jw)}) + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{yab(jw)}'^2 \right\} \\ & + R_a^2 \sum_{j=1}^{\mathcal{J}} \left\{ \mu_{xa(j)}^2 \text{Var}(\hat{N}_{a(j)}) + N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{xa(j)}^2 \right\} \\ & + p^2 R_{ab}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{xab(jw)}^2 \text{Var}(\hat{N}_{ab,s(jw)}) + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{xab(jw)}'^2 \right\} \\ & - 2R_a \sum_{j=1}^{\mathcal{J}} \left\{ \mu_{ya(j)} \mu_{xa(j)} \text{Var}(\hat{N}_{a(j)}) + N_{a(j)}^2 \text{ACov}(\tilde{y}_{a(j)} ; \tilde{x}_{a(j)}) \right\} \\ & - 2p^2 R_{ab} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)} \mu_{xab(jw)} \text{Var}(\hat{N}_{ab,s(jw)}) \right. \\ & \left. + N_{ab(jw)}^2 \text{ACov}(\tilde{y}'_{ab(jw)} ; \tilde{x}'_{ab(jw)}) \right\} \end{aligned}$$

$$\begin{aligned}
& + \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yb(w)}^2 \text{Var}(\hat{N}_{b(w)}) + N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) \sigma_{yb(w)}^2 \right\} \\
& + (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}^2 \text{Var}(\hat{N}_{ab,s(jw)}) \right. \\
& \left. + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) \sigma_{yab(jw)}''^2 \right\}
\end{aligned} \tag{5.26}$$

5.2.4 Estimador razão combinado 4

O estimador razão combinado 4 é similar ao estimador razão combinado 2, o que o difere é o fato de estar em função de tamanhos populacionais que são desconhecidos. A forma do estimador razão 4 é dada por

$$\begin{aligned}
\hat{t}_{yrc4} = & \left(\frac{\sum_{j=1}^{\mathcal{J}} \tilde{\gamma}_{ya(j)} + p \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}'_{yab(jw)}}{\sum_{j=1}^{\mathcal{J}} \tilde{\gamma}_{xa(j)} + p \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}'_{yab(jw)}} \right) t_{X_{\mathcal{A}}} + \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}_{yb(w)} \\
& + (1-p) \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\gamma}''_{yab(jw)}
\end{aligned} \tag{5.27}$$

A variância aproximada do estimador 4 é definida pela expressão

$$\begin{aligned}
A\text{Var}(\hat{t}_{yrc4}) = & \left(\frac{t_{X_{\mathcal{A}_C}}}{t_{x_{\mathcal{A}_C}}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}_C}^2 A_3 + p^2 G_A^2 A_4 - G_{\mathcal{A}_C} A_5 - p^2 G_{\mathcal{A}_C} A_6 \\
& + M_1 + p^2 M_2 + G_{\mathcal{A}_C}^2 M_3 + p^2 G_{\mathcal{A}_C}^2 M_4) + B_1 + (1-p)^2 B_2 \\
& + M_5 + (1-p)^2 M_6
\end{aligned} \tag{5.28}$$

onde

$$A_1 = \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{ya(j)}^2; \quad A_2 = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma'_{yab(jw)}{}^2;$$

$$A_3 = \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{xa(j)}^2; \quad A_4 = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma'_{xab(jw)}{}^2$$

$$A_5 = \sum_{j=1}^{\mathcal{J}} 2 N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya} \sigma_{xa(j)} \sigma_{ya(j)};$$

$$A_6 = 2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{xab(jw)} \sigma'_{yab(jw)};$$

$$B_1 = \sum_{w=1}^{\mathcal{W}} N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}} P_{b(w)}} \right) \sigma_{yb(w)}^2 \quad B_2 = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}} P''_{ab(jw)}} \right) \sigma''_{yab(jw)}{}^2;$$

$$\begin{aligned} M_1 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{ya(j)}^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ M_2 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yab(jw)}'^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ M_3 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{xa(j)}^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ M_4 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{xab(jw)}'^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ M_5 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yb(w)}^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ M_6 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yab(jw)}''^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \end{aligned}$$

Maiores detalhes a respeito das quantidades presentes nos termos das variâncias dos estimadores podem ser vistos no apêndice C, pág. 159.

CAPÍTULO 6

Avaliação Numérica

Os desempenhos dos estimadores propostos nos capítulos 4 e 5 foram avaliados através do método de simulação de Monte Carlo. O estudo foi realizado utilizando a linguagem matricial de programação `0x`, versão 4.04 e os gráficos apresentados foram construídos através do pacote estatístico `R`, versão 2.4.0.

Admitindo-se que apenas o cadastro $\mathcal{A}$ possui informação auxiliar disponível, o cenário considerado na avaliação numérica foi que a relação entre a variável de interesse e a variável auxiliar no cadastro $\mathcal{A}$ é descrita pelo modelo (4.1). O cadastro $\mathcal{B}$ foi gerado de modo a ter informação do cadastro $\mathcal{A}$ apenas no domínio ab , e a informação referente ao domínio b foi gerada de modo a não ter relação com a variável auxiliar. Para a avaliação dos estimadores sob o plano AAS, foram considerados os cenários referentes à população-alvo descritos na tabela (6.1), sob a situação em que $U_a = U_{\mathcal{A}} \cap U_{\mathcal{B}}^c$, $U_b = U_{\mathcal{A}}^c \cap U_{\mathcal{B}}$ e $U_{ab} = U_{\mathcal{A}} \cap U_{\mathcal{B}}$. Sob a situação em que $U_{\mathcal{B}} \subset U_{\mathcal{A}}$, os estimadores $\hat{t}_{yH}$ e $\hat{t}_{yr2}$ são avaliados de acordo com a estrutura da tabela (6.2), pois como $N_{ab} = N_{\mathcal{B}}$ é conhecido, estes são os únicos estimadores que podem ser comparados para essa situação.

Tabela 6.1: Populações geradas para U_a , U_b e U_{ab}

POPULAÇÃO	N_a	N_b	N_{ab}	N
1	2000	2000	500	4500
2	1750	1750	1000	4500
3	1500	1500	1500	4500

Tabela 6.2: Populações geradas para $U_B \subset U_A$

POPULAÇÃO	N_a	N_{ab}	N
1	2000	500	2500
2	1750	1000	2750
3	1500	1500	3000
4	1250	2000	3250
5	1000	2500	3500
6	750	3000	3750

Sob o plano AAS, para a situação em que $U_a = U_A \cap U_B^c$, $U_b = U_A^c \cap U_B$ e $U_{ab} = U_A \cap U_B$, foram considerados os casos em que $n_A = n_B = 125, 250, 500, 750, 1000, 1250$. Além disso, para população 1 (tabela 6.1) foi fixado $n_A = 1000$ enquanto n_B variava, tomando como valores 125, 250, 500, 750, 1000 e 1250.

Sob a situação em que $U_B \subset U_A$, foram considerados os casos em que $n_A = n_B = 125, 250, 500, 750, 1000, 1250, 1750, 2000$. Também foi fixado $n_A = 500$ enquanto n_B variava, tomando como valores 25, 50, 100, 150, 200, 250, 300 e 350.

Para o cálculo e avaliação dos estimadores separados, sob a situação em que $U_a = U_A \cap U_B^c$, $U_b = U_A^c \cap U_B$ e $U_{ab} = U_A \cap U_B$, foram considerados estratos de mesmo tamanho para os cadastros $\mathcal{A}$ e $\mathcal{B}$, ou seja, $N_{\mathcal{A}(j)} = 2500$, $j = 1, 2, 3$ e $N_{\mathcal{B}(w)} = 2500$, $w = 1, 2, 3$. Foi considerado que cada estrato em $\mathcal{A}$ possui interseção com apenas um estrato em $\mathcal{B}$. A relação entre a variável de interesse e a variável auxiliar em cada estrato, no cadastro $\mathcal{A}$ é descrita pelo modelo descrito em (5.3). Para cada estrato, foram considerados os casos em que $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 125, 250, 500, 750, 1000, 1250, 1750, 2000$ para todo j e w . Aqui, a situação em que $n_{\mathcal{A}(j)} = 1000$ para todo j e $n_{\mathcal{B}(w)} = 125, 250, 500, 750, 1000, 1250, 1750, 2000$, para todo w , também foi considerada

Para o caso em que $U_{\mathcal{B}} \subset U_{\mathcal{A}}$, foram considerados os seguintes tamanhos de população para cada estrato: $N_{\mathcal{A}(j)} = 2500, 2750, 3000, 3250, 3500, 3750$ e $N_{ab(jw)} = N_{\mathcal{B}(w)} = 500, 1000, 1500, 2000, 2500, 3000$. Para os tamanhos de amostra, foram considerados os casos em que $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 125, 250, 500, 750, 1000, 1250, 1750, 2000$. Foi considerado ainda que $n_{\mathcal{A}(j)} = 500$ e $n_{\mathcal{B}(w)} = 25, 50, 100, 150, 200, 250, 300, 350$.

Para o cálculo e avaliação dos estimadores combinados, foram considerados os mesmos cenários populacionais descritos para os estimadores do tipo separado, como também os tamanhos de amostra em cada estrato. Porém, a população do cadastro $\mathcal{A}$ foi gerada de modo a ser assistida pelo modelo descrito em (5.19)

Para a avaliação dos estimadores foram considerados o viés relativo, viés, desvio padrão e erro quadrático médio dos estimadores da média populacional em cada cenário descrito na tabela (6.1). O viés de um estimador $\hat{\theta}$ é dado por

$$V(\hat{\theta}) = E(\hat{\theta} - \theta) = E(\hat{\theta}) - \theta$$

e o erro quadrático médio por

$$EQM(\hat{\theta}) = \text{Var}(\hat{\theta}) + V(\hat{\theta})^2.$$

O viés relativo é dado por

$$VR(\hat{\theta}) = \frac{|E(\hat{\theta}) - \theta|}{\theta},$$

onde θ é o parâmetro de interesse e $\hat{\theta}$ é o estimador de θ . O viés estimado e o viés relativo estimado são obtidos a partir da estimativa de $E(\hat{\theta})$, via Monte Carlo.

O esquema de geração de dados foi conduzido da seguinte forma:

Passo 1. Gerar i observações da variável auxiliar para o domínio a , onde $i = 1, \dots, N_a$, seguindo distribuição $\mathcal{N}(12, 4)$. Para a estrutura referente aos estimadores separados e combinados, foram geradas três populações seguindo distribuição $\mathcal{N}(8, 1)$, $\mathcal{N}(12, 1)$ e $\mathcal{N}(16, 1)$, respectivamente;

Passo 2. Gerar i observações da variável auxiliar para o domínio ab , onde $i = 1, \dots, N_{ab}$, seguindo distribuição $\mathcal{N}(6, 1)$. Para a estrutura referente aos estimadores separados e combinados, foram geradas três populações para o domínio ab seguindo distribuição $\mathcal{N}(5, 1)$, $\mathcal{N}(6, 1)$ e $\mathcal{N}(7, 1)$, respectivamente;

Passo 3. Gerar i observações da variável de interesse para o domínio a , onde $i = 1, \dots, N_a$, seguindo distribuição $\mathcal{N}(\beta_a x_{kA}, \lambda x_{kA})$, onde $\beta_a = 15$ e $\lambda = 1$. Para a estrutura referente aos estimadores separados, foram geradas três populações para o domínio a seguindo distribuição $\mathcal{N}(16x_{kA}, x_{kA})$, $\mathcal{N}(24x_{kA}, x_{kA})$ e $\mathcal{N}(32x_{kA}, x_{kA})$, respectivamente. Para os estimadores combinados, todos os estratos foram gerados de acordo com uma distribuição $\mathcal{N}(10x_{kA}, x_{kA})$.

Passo 4. Gerar i observações da variável de interesse para o domínio ab , onde $i = 1, \dots, N_{ab}$, seguindo distribuição $\mathcal{N}(\beta_{ab} x_{kA}, \lambda x_{kA})$, onde $\beta_{ab} = 6$ e $\lambda = 1$. Para a estrutura referente aos estimadores separados, foram geradas três populações para o domínio a seguindo distribuição $\mathcal{N}(10x_{kA}, x_{kA})$, $\mathcal{N}(12x_{kA}, x_{kA})$ e $\mathcal{N}(14x_{kA}, x_{kA})$, respectivamente. Para os estimadores combinados, todos os estratos foram gerados de acordo com uma distribuição $\mathcal{N}(5x_{kA}, x_{kA})$.

Passo 5. Gerar i observações da variável de interesse para o domínio b , onde $i = 1, \dots, N_b$, seguindo distribuição $\mathcal{N}(8, 9)$. Para a estrutura referente aos estratos, foram geradas três populações seguindo distribuição $\mathcal{N}(7, 1)$, $\mathcal{N}(9, 1)$ e $\mathcal{N}(2, 1)$, respectivamente;

Passo 6. Na estrutura de população gerada nos passos de 1 a 5, coletar uma amostra de tamanho especificado através dos planos amostrais considerados e obter estimativas para o parâmetro de interesse através dos estimadores propostos;

As populações geradas nos passos de 1 a 5 permaneceram fixas em todas as etapas da simulação. No processo de simulação de Monte Carlo, o passo 6 foi repetido $r = 1000$ vezes, para cada tamanho de amostra especificado.

Os resultados das quantidades obtidas para cada estimador são descritos pelas tabelas e figuras a seguir. Além das tabelas, as figuras ilustram gráficos de viés relativo e desvio padrão para cada estimador, com o objetivo de verificar o comportamento de cada estimador para os tamanhos de N_{ab} e para os tamanhos de amostra especificados.

6.1 Resultados

Nesta seção são apresentados um estudo sobre a distribuição empírica do estimador $\hat{N}_{ab,s}$ e os resultados das simulações referentes ao desempenho dos estimadores propostos.

6.1.1 Estudo sobre a distribuição de $\hat{N}_{ab,s}$

O estimador de N_{ab} , denotado por $\hat{N}_{ab,s}$, é definido como a solução de uma equação do segundo grau e é utilizado nos casos em que não é possível identificar duplicatas nas amostras obtidas de $\mathcal{A}$ e $\mathcal{B}$. Além disso, não está em função dos valores das variáveis utilizadas no processo de estimação. As figuras (6.1) a (6.6) ilustram a distribuição empírica do estimador de N_{ab} . Para cada uma das populações geradas na tabela (6.1), foram obtidas 1000 estimativas de N_{ab} e com estas foram construídos gráficos de densidade estimada.

Pela figura (6.1), para $N_{ab} = 500$, verifica-se que à medida em que os tamanhos de amostra aumentam, a distribuição empírica das estimativas fornece evidência de que $\hat{N}_{ab,s}$ converge para uma distribuição centrada no real valor de N_{ab} . Além disso, verifica-se também que há evidência de redução da variância do estimador para tamanhos de amostra elevados, como o observado para $n_{\mathcal{A}} = n_{\mathcal{B}} = 1250$. Para $N_{ab} = 1000$, como mostra a figura (6.2), observa-se que o comportamento da distribuição empírica das estimativas de N_{ab} fornece evidência de que novamente o estimador $\hat{N}_{ab,s}$ possui distribuição centrada no real valor de N_{ab} . Para os demais valores de referência para N_{ab} , o comportamento observado foi o mesmo, o que evidencia que $\hat{N}_{ab,s}$ é um estimador que garante uma boa aproximação para N_{ab} , e essa aproximação se torna ainda melhor à medida em que os tamanhos de amostra nos cadastros aumentam. Os histogramas fornecem evidência de que para grandes amostras, $\hat{N}_{ab,s}$ converge assintoticamente para a distribuição normal. Este mesmo estudo também foi avaliado para $N_{ab} = 2000, 2500, 3000$, e o resultado foi o mesmo quando comparado aos demais valores de referência.

Figura 6.1: Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 500$

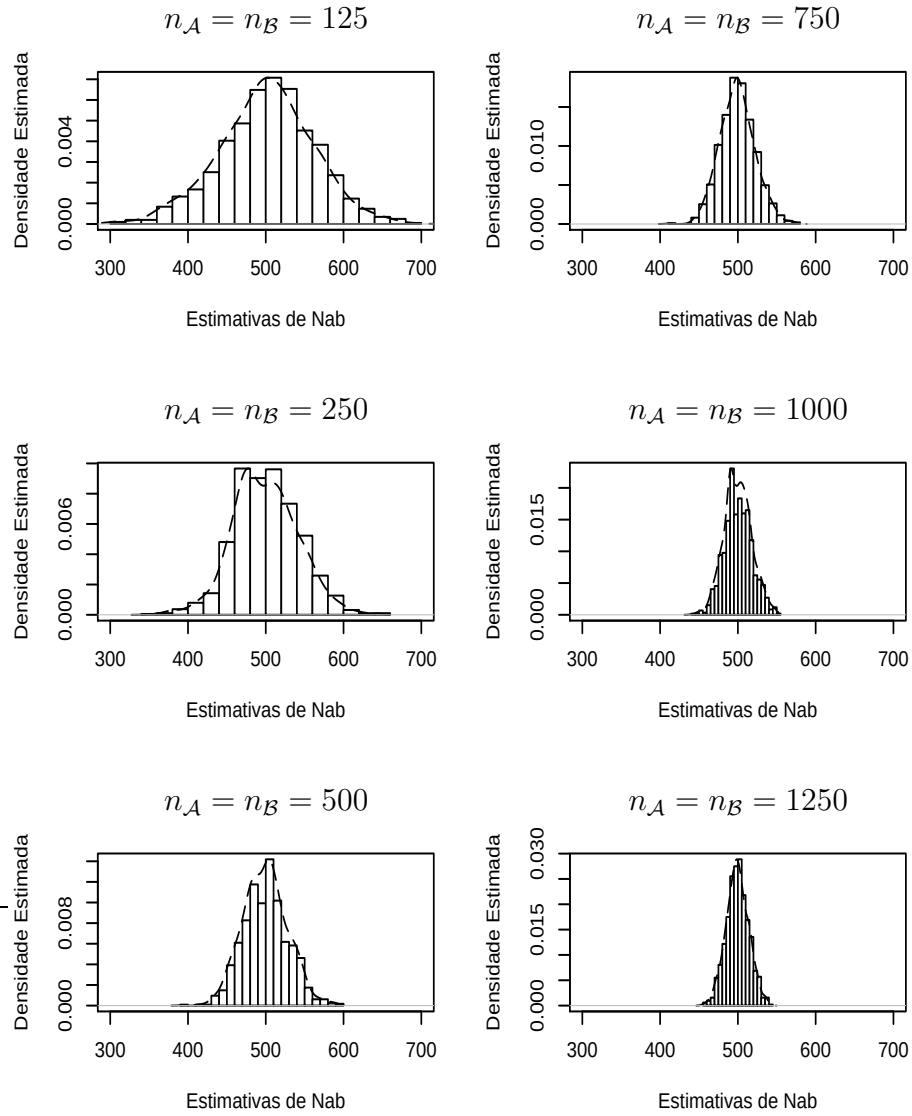


Figura 6.2: Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 1000$

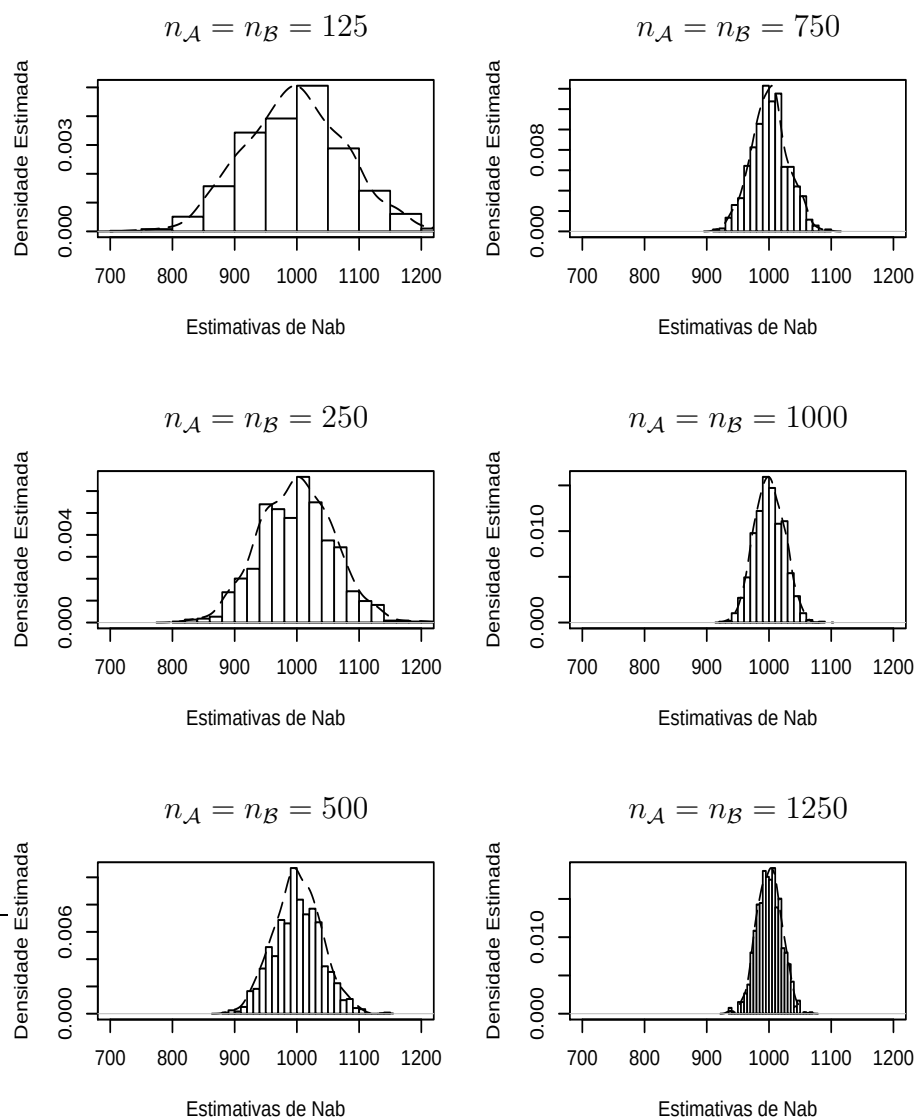


Figura 6.3: Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 1500$

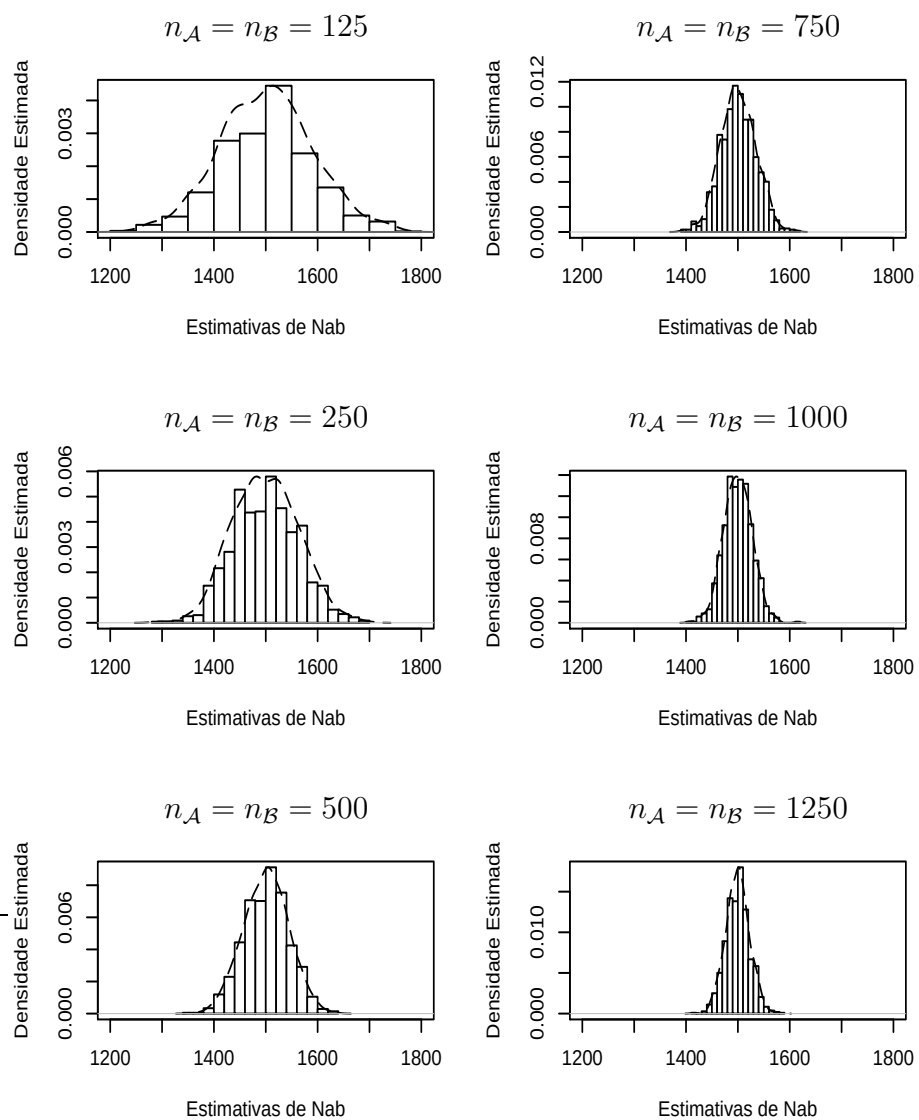


Figura 6.4: Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 2000$

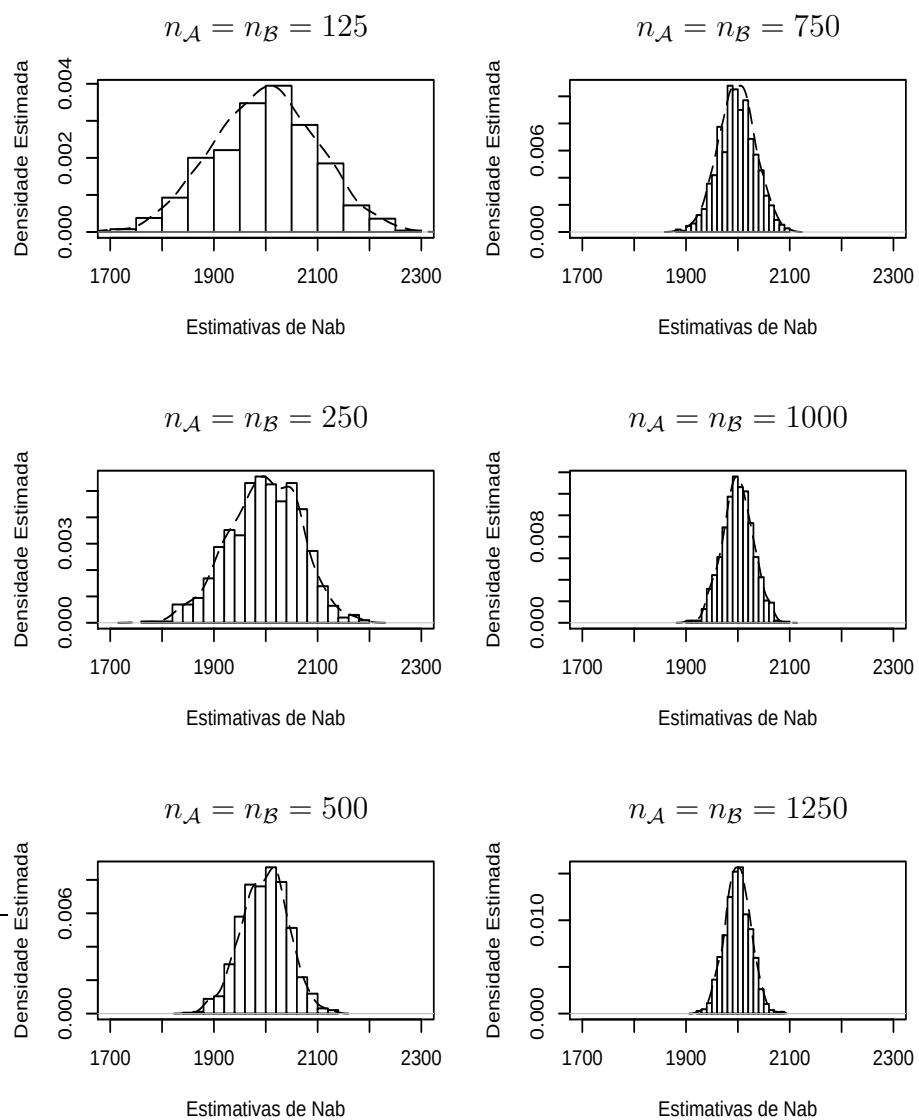


Figura 6.5: Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 2500$

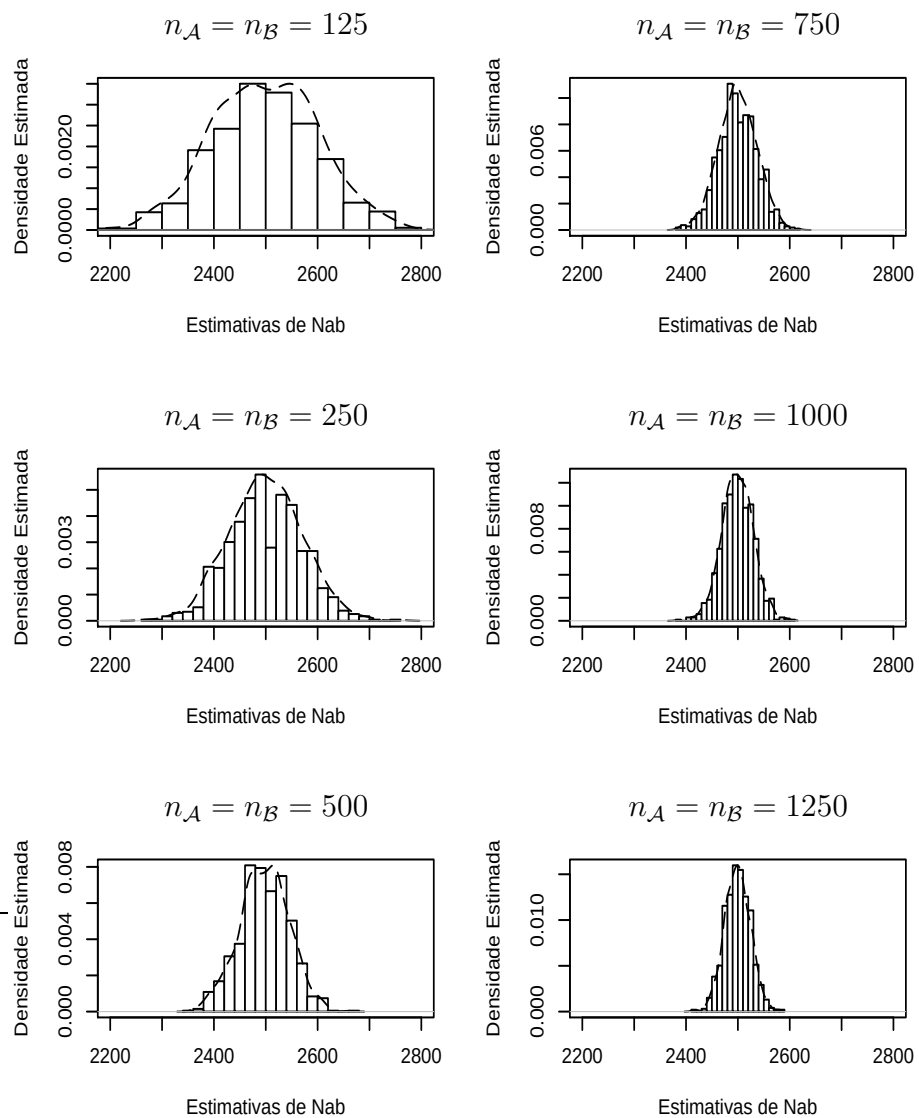
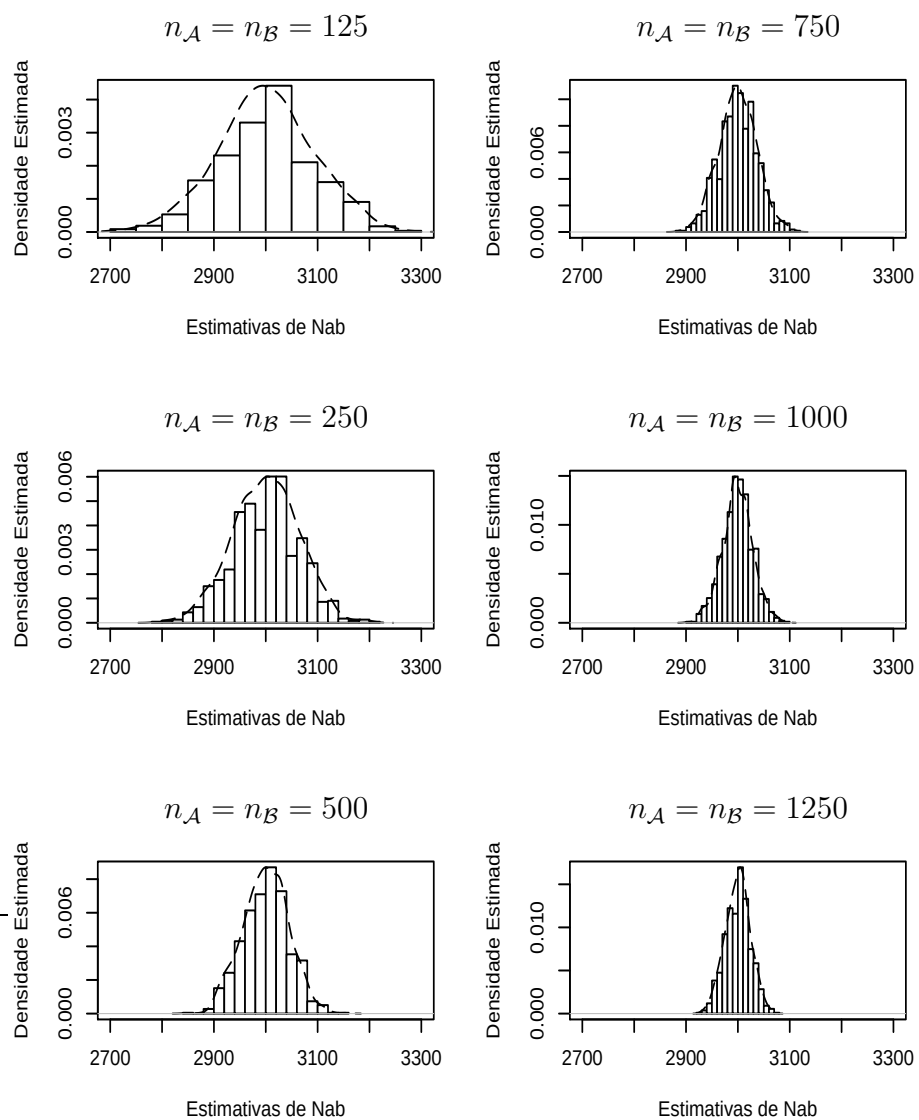


Figura 6.6: Densidade Estimada de $\hat{N}_{ab,s}$, com valor de referência $N_{ab} = 3000$



6.1.2 Desempenho dos estimadores para o plano AAS

O desempenho dos estimadores propostos sob o plano AAS é analisado a seguir, considerando os cenários 2 e 3.

As tabelas (6.3) a (6.5) e as figuras (6.7) a (6.14) descrevem os resultados referentes aos estimadores propostos sob o plano de AAS para a situação em que são geradas populações para U_a , U_b e U_{ab} .

Cenário 2

Para $N_{ab} = 500$ (tabela (6.3) e figura (6.7)), verifica-se que o estimador $\hat{t}_{yr1}$ é aproximadamente não-viesado, e que seu viés relativo é menor que o viés relativo de $\hat{t}_{yH}$. Além disso, o valor do viés relativo não sofreu alterações significativas com o aumento do tamanho de amostra, pois para todos os tamanhos de amostra considerados, os valores sempre estiveram próximos de zero. Verifica-se também que $\hat{t}_{yr1}$ apresenta valores do desvio padrão e erro quadrático médio menores que $\hat{t}_{yH}$, e que o aumento dos tamanhos de amostra de $\mathcal{A}$ e $\mathcal{B}$ contribuiu para a diminuição desses valores (tabela 6.3 e figura 6.7). O estimador $\hat{t}_{yr2}$ possui maiores valores de viés relativo, desvio padrão e erro quadrático médio em comparação aos outros estimadores que são utilizados sob o cenário 2 (tabela (6.3)). Porém, à medida em que $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$ aumentam, há evidência de que, para $N_{ab} = 500$, há uma diminuição do viés relativo de $\hat{t}_{yr2}$ e que apesar de ocorrer uma diminuição dos valores do desvio padrão, os valores correspondentes ao desvio padrão de $\hat{t}_{yr2}$ para cada tamanho de amostra foram maiores em relação aos valores de $\hat{t}_{yH}$ e $\hat{t}_{yr1}$. Para o caso onde $n_{\mathcal{A}} = 1000$ (tabela 6.6 e figura 6.13), verifica-se o mesmo comportamento, com exceção de que a redução do desvio padrão de $\hat{t}_{yr2}$ não foi tão evidente em relação aos casos onde $n_{\mathcal{A}}$ varia conjuntamente com $n_{\mathcal{B}}$.

Para $N_{ab} = 1000$ (tabela (6.4) e figura (6.9)), verificou-se que $\hat{t}_{yr1}$ apresentou aumento do viés relativo para os tamanhos de amostra iguais a 125 e 250. Em relação ao desvio padrão e erro quadrático médio, o estimador apresentou aumento quando comparado a $\hat{t}_{yH}$, para $n_{\mathcal{A}} = n_{\mathcal{B}} = 125$. Porém, à medida em que os tamanhos de amostra aumentam (tabela 6.4 e figura 6.8), verifica-se uma redução do erro quadrático médio. Para $n_{\mathcal{A}} = n_{\mathcal{B}} = 1250$ os valores são aproximadamente iguais aos valores apresentados no caso em

que $N_{ab} = 500$. Para o estimador $\hat{t}_{yr2}$ verifica-se o aumento no valor do viés relativo e erro quadrático médio. Porém, houve maior redução no valor do desvio padrão em comparação ao caso onde $N_{ab} = 500$.

Para $N_{ab} = 1500$ (tabelas 6.5 e figura 6.11), $\hat{t}_{yr1}$ apresentou aumento nos valores de viés relativo, desvio padrão e erro quadrático médio para os tamanhos de amostra $n_{\mathcal{A}} = n_{\mathcal{B}} = 125, 250$. Porém, quando os tamanhos de amostra aumentam, verifica-se o mesmo comportamento quando comparado aos casos em que $N_{ab} = 500, 1000$. Isso fornece evidência de que independentemente do tamanho de N_{ab} , $\hat{t}_{yr1}$ tem um desempenho melhor que o estimador de Hartley, para tamanhos de amostra superiores a 250 em cada cadastro. Da mesma forma que os casos anteriores, o estimador $\hat{t}_{yr2}$ continua apresentado valores elevados de viés relativo e erro quadrático médio. Porém, a partir de $n_{\mathcal{A}} = n_{\mathcal{B}} = 250$, verifica-se que o estimador possui menor desvio padrão em relação ao estimador $\hat{t}_{yH}$, o que evidencia que o aumento de N_{ab} contribui para a redução do desvio padrão e conseqüentemente, da variância do estimador.

Para o caso em que $U_{\mathcal{B}} \subset U_{\mathcal{A}}$, as tabelas (6.7) a (6.12) e as figuras (6.15) a (6.26) ilustram o comportamento dos estimadores $\hat{t}_{yH}$ e $\hat{t}_{yr2}$. Para $N_{ab} = 500$ (tabela (6.7) e figuras (6.15) e (6.16)), verifica-se que o estimador $\hat{t}_{yr2}$ apresenta viés relativo maior que o estimador $\hat{t}_{yH}$, que é aproximadamente não viesado. Quando $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$ aumentam, verifica-se ainda que não há grande redução no valor do viés relativo de $\hat{t}_{yr2}$ a partir de $n_{\mathcal{A}} = 750$, os valores são aproximadamente iguais. Verifica-se também que os valores do desvio padrão e erro quadrático médio de $\hat{t}_{yr2}$ são maiores em relação a $\hat{t}_{yH}$, e isso é mais evidente quando $n_{\mathcal{A}} = 125$ e $n_{\mathcal{B}} = 25$. Para o caso em que $n_{\mathcal{A}} = 500$ o comportamento em termos de viés relativo foi o mesmo. Porém, a redução nos valores do desvio padrão e erro quadrático médio permaneceu aproximadamente constante para todos os tamanhos de amostra, não sendo tão evidente quanto o caso em que $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$ variam conjuntamente.

Para $N_{ab} = 1000$ (tabela (6.8) e figuras (6.17) e (6.18)) verifica-se que houve um aumento do viés relativo e erro quadrático médio de $\hat{t}_{yr2}$, quando comparado ao caso em que $N_{ab} = 500$. Em relação ao desvio padrão, os valores continuaram aproximadamente iguais para os tamanhos de amostra especificados. Isso também foi verificado para o caso em que $n_{\mathcal{A}} = 500$.

Para $N_{ab} = 1500$ (tabela (6.9) e figuras (6.19) e (6.20)), o que representa neste

caso 50% do valor da população, verifica-se que apesar do aumento do erro quadrático médio e viés relativo, o estimador $\hat{t}_{yr2}$ já apresenta menores valores de desvio padrão em comparação a $\hat{t}_{yH}$. Para n_A o comportamento foi o mesmo, porém o ganho em relação ao estimador $\hat{t}_{yH}$ no que diz respeito à redução no valor do desvio padrão só foi observado a partir de $n_B = 100$, permanecendo aproximadamente constante para os demais tamanhos de amostra.

Para os demais tamanhos de N_{ab} , foi observado o mesmo comportamento, quando comparados a $N_{ab} = 1500$. Apesar de apresentar viés relativo e erro quadrático médio maiores que $\hat{t}_{yH}$, o estimador $\hat{t}_{yr2}$ apresentou maior redução do desvio padrão, e consequentemente, da variância. Para $n_A = 500$, o que se observa é que os valores do desvio padrão são menores que os valores de desvio padrão de $\hat{t}_{yH}$, para todos os tamanhos de amostra n_B considerados. Os resultados evidenciam que, se $N_{ab} = N_B$ é maior que N_a , ou seja, quanto maior for o tamanho do cadastro B , ou ainda, quanto maior for o nível de cobertura do cadastro B , melhor serão as chances do estimador $\hat{t}_{yr2}$ de garantir estimativas que, apesar de apresentar um certo viés, são originadas a partir de um estimador com menor variância.

Cenário 3

Para $N_{ab} = 500$ (tabela (6.3) e figura (6.8)), verificou-se que o estimador $\hat{t}_{yr3}$ é aproximadamente não viesado, e à medida em que o tamanho de amostra aumenta, verifica-se que o viés relativo tende a ficar menor quando comparado ao estimador $\hat{t}_{yFB}$. Além disso, $\hat{t}_{yr3}$ apresentou menores valores do desvio padrão e erro quadrático médio que os valores em $\hat{t}_{FB}$ quando os tamanhos de amostra foram maiores que 500, e que o aumento dos tamanhos de amostra contribui para a redução desses valores. O estimador $\hat{t}_{yr4}$, apresentou valores de viés relativo, desvio padrão e erro quadrático médio maiores quando comparado aos estimadores $\hat{t}_{yr3}$ e $\hat{t}_{yFB}$, exceto no caso em que $n_A = n_B = 125$, onde o valor do desvio padrão foi menor que o desvio padrão de $\hat{t}_{yr3}$. Verifica-se ainda que à medida em que o tamanho de amostra aumenta, há evidência de que os valores de desvio padrão e erro quadrático médio tendem a ficar próximos aos valores dos outros estimadores, o que pode ser observado para $n_A = n_B = 1250$. Para $n_A = 1000$ (tabela (6.6) e figura (6.14)), verifica-se que o comportamento dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ é o mesmo quando n_B é

grande.

Para $N_{ab} = 1000$ (tabela (6.4) e figura (6.10)), verifica-se o mesmo comportamento para $\hat{t}_{yr3}$ em termos de viés relativo. Em relação aos valores de desvio padrão e erro quadrático médio, verificou-se que houve uma diminuição se comparado ao caso em que $N_{ab} = 500$, quando $n_A = n_B = 125, 250$. Esses valores ainda são elevados para esses tamanhos de amostra, porém já são menores que os valores do estimador $\hat{t}_{FB}$, e que o aumento no tamanho da amostra contribui para uma redução ainda maior dos valores do desvio padrão e erro quadrático médio. O estimador $\hat{t}_{yr4}$ apresentou um aumento no valor do viés relativo e erro quadrático médio em relação ao caso em que $N_{ab} = 500$, porém os valores do desvio padrão estão bem próximos aos valores do estimador $\hat{t}_{yFB}$, alguns destes sendo menores para determinados tamanhos de amostra.

Para $N_{ab} = 1500$ (tabela (6.5) e figura (6.12)), o estimador $\hat{t}_{yr3}$ apresentou o mesmo comportamento quando comparado ao caso em que $N_{ab} = 1000$, em termos de viés relativo, desvio padrão e erro quadrático médio. Para o estimador $\hat{t}_{yr4}$, quando comparado aos demais valores de N_{ab} , verifica-se que apesar de apresentar viés relativo e erro quadrático médio elevados, os valores do desvio padrão ficaram menores que os valores do estimador $\hat{t}_{yFB}$, o que evidencia que o aumento de N_{ab} contribui para redução do desvio padrão e conseqüentemente da variância do estimador $\hat{t}_{yr4}$.

6.1.3 Desempenho dos estimadores separados para o plano AAE

O desempenho dos estimadores propostos sob o plano AAE é analisado a seguir, considerando os cenários 2 e 3.

As tabelas (6.13) a (6.15) e as figuras (6.27) a (6.32) descrevem os resultados referentes aos estimadores separados, propostos para a situação em que são geradas populações para U_a , U_b e U_{ab} .

Cenário 2

Para $N_{ab(jw)} = 500$ (tabela (6.13) e figura (6.27)), verifica-se que o estimador $\hat{t}_{yrs1}$ apresentou viés relativo menor que $\hat{t}_{yH}$, e o que se observa à medida em que o tamanho

de amostra aumenta, é que os estimadores $\hat{t}_{yrs1}$ e $\hat{t}_{yH}$ são aproximadamente não viesados. Em relação ao desvio padrão e erro quadrático médio, $\hat{t}_{yrs1}$ apresentou menores valores em relação a $\hat{t}_{yH}$ e aumento dos tamanhos de amostra $n_{\mathcal{A}(j)}$ e $n_{\mathcal{B}(w)}$ contribuiu para a redução no valor do desvio padrão e erro quadrático médio de $\hat{t}_{yrs1}$. O estimador $\hat{t}_{yrs2}$ possui valores de viés relativo, desvio padrão e erro quadrático médio maiores em relação a $\hat{t}_{yrs2}$ e $\hat{t}_{yH}$. Verifica-se que o viés relativo de $\hat{t}_{yrs2}$ não é afetado pelo aumento dos tamanhos de amostra em cada estrato, os valores permanecem aproximadamente iguais. Em relação ao desvio padrão, o aumento do tamanho de amostra em cada estrato contribui para que haja uma diminuição do desvio, e conseqüentemente, da variância do estimador. Para o caso em que $n_{\mathcal{A}(j)} = 1000$, verifica-se o mesmo comportamento, porém a redução no desvio padrão dos estimadores é mais lenta em comparação ao caso onde $n_{\mathcal{A}(j)}$ também varia.

Para $N_{ab(jw)} = 1000$ (tabela (6.14) e figura (6.29)) verifica-se que $\hat{t}_{yrs1}$ apresentou o mesmo comportamento referente ao viés relativo, desvio padrão e erro quadrático médio, sendo melhor que o estimador $\hat{t}_{yH}$, exceto para o caso em que $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 125$. À medida em que os tamanhos de amostra aumentam em cada estrato, verifica-se novamente que o estimador é aproximadamente não viesado, e que os valores do desvio padrão e erro quadrático médio são aproximadamente iguais aos do caso onde $N_{ab(jw)} = 500$. O estimador $\hat{t}_{yrs2}$ apresentou comportamento similar ao caso em que $N_{ab} = 500$ para os valores do erro quadrático médio e viés relativo. Em relação ao desvio padrão, houve um aumento, e apesar de haver uma diminuição do desvio padrão à medida em que $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$ aumentam, não houve evidência de que o $\hat{t}_{yrs2}$ tenha menor desvio padrão que $\hat{t}_{yH}$ e $\hat{t}_{yrs1}$.

Para $N_{ab(jw)} = 1500$ (tabela (6.15) e figura (6.31)), o estimador $\hat{t}_{yrs1}$ apresentou aumento do viés relativo, desvio padrão e erro quadrático médio quando $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 125, 250$, e à medida em que o tamanho de amostra aumenta, o comportamento do estimador foi o mesmo quando comparado aos casos em que $N_{ab(jw)} = 500, 1000$, o que evidencia que $\hat{t}_{yrs1}$ é aproximadamente não viesado e possui menor variância quando comparado ao estimador $\hat{t}_{yH}$. O estimador $\hat{t}_{yrs2}$ não apresentou bom desempenho, pois teve aumento dos valores de viés relativo, desvio padrão e erro quadrático médio, não apresentando nenhuma evidência favorável ao seu uso.

Sob a situação em que $U_B \subset U_A$ (tabelas (6.17) a (6.19)), não houve evidência de que

o estimador $\hat{t}_{yrs2}$ apresenta melhor desempenho quando comparado à $\hat{t}_{yH}$ para nenhum valor de $N_{ab(j)}$. Esse fato provavelmente está relacionado ao modelo descrito em (5.3), o qual considera valores β_i diferentes para cada domínio. Apesar de $\hat{t}_{yrs2}$ ter apresentando um desempenho ruim em relação ao estimador $\hat{t}_{yH}$, há evidência de que esse desempenho seja devido ao modelo que descreve a população de interesse. Na forma dos estimadores, seja no plano AAS ou AE, os estimadores que não apresentaram o desempenho desejável fazem uso da informação auxiliar que é conhecida para o cadastro inteiro, e no processo de estimação incorpora a razão entre o total da variável de interesse para o cadastro $\mathcal{A}$ e o total para a variável auxiliar para o cadastro $\mathcal{A}$. Já os estimadores que apresentaram bom desempenho são aqueles em que as informações da variável auxiliar são conhecidas para os domínios, e o estimador para total é calculado com base na soma de estimadores do tipo razão para cada domínio. Como a população foi gerada de modo que cada domínio é assistido por um modelo distinto, há evidência para afirmar que se a população fosse assistida por um único modelo, o desempenho dos estimadores que não apresentaram bom desempenho poderia ser melhor.

Cenário 3

Para $N_{ab(jw)} = 500$ (tabela (6.13) e figura (6.28)), verifica-se que o estimador $\hat{t}_{yrs3}$ apresentou valores de viés relativo aproximadamente igual ao estimador $\hat{t}_{yFB}$, para cada tamanho de amostra especificado e os resultados mostram ainda que ambos são aproximadamente não viesados. Em relação aos valores de desvio padrão e erro quadrático médio, observou-se que o estimador $\hat{t}_{yrs3}$ obteve melhor desempenho que o estimador $\hat{t}_{yFB}$ quando $n_{\mathcal{A}} = n_{\mathcal{B}} = 1250$, o que não fornece evidência do ganho em utilizar $\hat{t}_{yrs3}$. O estimador $\hat{t}_{yrs4}$ apresentou maior valor de viés relativo, desvio padrão e erro quadrático médio em relação a $\hat{t}_{yrs3}$ e $\hat{t}_{yFB}$. Quando o tamanho de amostra aumenta, o desvio padrão de $\hat{t}_{yrs4}$ tende a se aproximar dos valores de desvio padrão de $\hat{t}_{yrs3}$ e $\hat{t}_{yFB}$, e isso é mais evidente a partir de $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 750$. Sob a situação em que $n_{\mathcal{A}(j)} = 1000$ (tabela 6.16), para todo j , verifica-se que o comportamento dos estimadores é o mesmo, exceto no que diz respeito ao valor do desvio padrão dos estimadores, que diminui mais lentamente em comparação aos valores da tabela (6.13).

Para $N_{ab(jw)} = 1000$ (tabela (6.14) e figura (6.30)), verifica-se que não há grande diferença entre de $\hat{t}_{yrs3}$ e $\hat{t}_{yFB}$, pois o valor do viés relativo, desvio padrão e erro quadrático médio para os tamanhos de amostra considerados são aproximadamente iguais, o que não fornece evidência sobre qual dos dois estimadores apresenta melhor desempenho. O estimador $\hat{t}_{yrs4}$ apresentou um aumento no viés relativo e erro quadrático médio quando comparado ao caso em que $N_{ab(jw)} = 500$, porém ao comparar os valores de desvio padrão, verifica-se que houve uma redução, e os valores estão próximos aos valores de $\hat{t}_{yFB}$ e $\hat{t}_{yrs3}$, porém sem evidenciar o ganho em utilizar o estimador $\hat{t}_{yrs4}$, no que diz respeito à redução do valor da variância do estimador.

Para $N_{ab(jw)} = 1500$ (tabela (6.14) e figura (6.32)), verifica-se que o estimador $\hat{t}_{yrs3}$ apresentou melhor desempenho que o estimador $\hat{t}_{yFB}$, pois apresentou menores valores de desvio padrão e erro quadrático médio. Em relação ao viés relativo, o estimador continua sendo aproximadamente não viesado. Isso fornece condições suficientes para afirmar que o estimador $\hat{t}_{yrs3}$ será melhor que $\hat{t}_{yFB}$ quando houver evidência de que o domínio ab contenha um grande percentual de elementos da população dentro de cada estrato. O estimador $\hat{t}_{yrs4}$ apresentou maiores valores de erro quadrático médio e desvio padrão em relação aos estimadores $\hat{t}_{yrs3}$ e $\hat{t}_{yFB}$.

6.1.4 Desempenho dos estimadores combinados para o plano AAE

As tabelas (6.20) a (6.23) e as figuras (6.33) a (6.38) descrevem os resultados referentes aos estimadores combinados propostos para a situação em que são geradas populações para U_a , U_b e U_{ab} .

Cenário 2

Para $N_{ab(jw)} = 500$ (tabela (6.20) e figura (6.33)), verifica-se que o estimador $\hat{t}_{yrc1}$ apresentou viés relativo menor que o estimador de $\hat{t}_{yH}$ e próximo de zero, o que evidencia o fato de que $\hat{t}_{yrc1}$ é aproximadamente não-viesado. Além disso, $\hat{t}_{yrc1}$ apresentou menor valor de desvio padrão e erro quadrático médio em comparação ao estimador $\hat{t}_{yH}$, e que esses valores diminuem à medida em que $n_{\mathcal{A}(j)}$ e $n_{\mathcal{B}(w)}$ aumentam dentro de cada estrato.

O estimador $\hat{t}_{yrc2}$ apresentou viés relativo próximo de 0.1 para todos os tamanhos de amostra considerados, e apresentou maior valor do erro quadrático médio em relação a $\hat{t}_{yrc1}$ e $\hat{t}_{yH}$. Porém, há evidência de que o estimador $\hat{t}_{yrc2}$ apresenta menor desvio padrão que o estimador $\hat{t}_{yH}$ (figura (6.33)). Para os casos em que $n_{\mathcal{A}(j)} = 1000$ (tabela (5.23)), o mesmo comportamento é observado para todos os estimadores, porém a redução no desvio padrão é mais lenta quando comparada ao caso em que $n_{\mathcal{A}(j)}$ e $n_{\mathcal{B}(w)}$ variam conjuntamente.

Para $N_{ab(jw)} = 1000$ (tabela (6.21) e figura (6.35)), verifica-se que o estimador $\hat{t}_{yrc1}$ apresentou o mesmo desempenho quando comparado ao caso em que $N_{ab(jw)} = 500$. Para $\hat{t}_{yrc2}$, verifica-se que o aumento de $N_{ab(jw)}$ em cada estrato contribuiu para o aumento do viés relativo, desvio padrão e erro quadrático médio.

Para $N_{ab(jw)} = 1500$ (tabela (6.22) e figura (6.37)), verifica-se que em comparação aos casos em que $N_{ab(jw)} = 500, 1000$, o estimador $\hat{t}_{yrc1}$ apresentou um aumento nos valores do desvio padrão e erro quadrático médio, os quais são maiores que os valores referentes a $\hat{t}_{yH}$, apenas para $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 125$. Porém, à medida em que os tamanhos de amostra aumentam em cada estrato, verifica-se uma diminuição do desvio padrão e erro quadrático médio, e nos casos onde $n_{\mathcal{A}(j)} = n_{\mathcal{B}(w)} = 1000, 1250$, observa-se que os valores estão próximos aos valores referentes a $N_{ab(jw)} = 500, 1000$.

Sob a situação em que $U_B \subset U_A$ (tabelas (6.24) e (6.25)), não houve evidência de que o estimador $\hat{t}_{yrc2}$ apresenta melhor desempenho quando comparado a $\hat{t}_{yH}$ para nenhum valor de $N_{ab(jw)}$. Esse fato provavelmente está relacionado novamente à forma como a população está sendo assistida.

Cenário 3

Para $N_{ab(jw)} = 500$ (tabela (6.20) e figura (6.34)), verifica-se que o estimador $\hat{t}_{yrc3}$ é aproximadamente não viesado, como mostra os valores do viés relativo. Além disso, $\hat{t}_{yrc3}$ apresentou menor desvio padrão e erro quadrático médio em relação a $\hat{t}_{yFB}$, para todos os tamanhos de amostra considerados. O estimador $\hat{t}_{yrc4}$ apresentou viés relativo próximo de 0.1 e possui valor do desvio padrão e erro quadrático médio maiores que os valores de $\hat{t}_{yFB}$ e $\hat{t}_{yrc3}$. À medida em que os tamanhos de amostra em cada estrato aumentam,

há uma diminuição do desvio padrão, mas não a ponto de fornecer evidência que $\hat{t}_{yrc4}$ possui menor desvio padrão que $\hat{t}_{yFB}$ e $\hat{t}_{yrc3}$. Para $n_{\mathcal{A}(j)} = 1000$, para todo j , verifica-se que o comportamento dos estimadores é o mesmo, exceto no que diz respeito ao valor do desvio padrão dos estimadores, que diminui mais lentamente em comparação aos valores da tabela (6.20).

Para $N_{ab(jw)} = 1000$ (tabela (6.21) e figura (6.36)), o estimador $\hat{t}_{yrc3}$ apresentou o mesmo desempenho quando comparado ao caso em que $N_{ab(jw)} = 500$. Para $\hat{t}_{yrc4}$, verifica-se que o aumento de $N_{ab(jw)}$ em cada estrato contribuiu para o aumento do viés relativo e erro quadrático médio para todos os tamanhos de amostra considerados, porém houve diminuição no valor do desvio padrão em relação a $\hat{t}_{yFB}$. Neste caso, também é possível constatar que à medida em que os tamanhos de amostra em cada estrato aumentam há uma redução nos valores do desvio padrão do estimador $\hat{t}_{yrc4}$, porém, $\hat{t}_{yrc3}$ ainda apresentou menores valores do desvio padrão para cada tamanho de amostra.

Para $N_{ab(jw)} = 1500$ (tabela (6.22) e figura (6.38)), o estimador $\hat{t}_{yrc3}$ apresentou o mesmo desempenho quando comparado com os casos em que $N_{ab(jw)} = 500, 1000$. Em relação ao estimador $\hat{t}_{yrc4}$, houve uma diminuição dos valores do erro quadrático médio e viés relativo, porém houve um aumento do desvio padrão para cada tamanho de amostra. Apesar do estimador $\hat{t}_{yrc4}$ ter apresentado menores valores do desvio padrão em relação a $\hat{t}_{yH}$ quando $N_{ab(jw)} = 1000$, os resultados não terem fornecido uma forte evidência de que esse estimador tenha um melhor desempenho que $\hat{t}_{FB}$.

Tabela 6.3: Avaliação dos estimadores propostos sob os cenários 2 e 3 para o plano AAS ($N_a = N_b = 2000$ e $N_{ab} = 500$)

Cenário	Estimador	$n_{\mathcal{A}}=n_{\mathcal{B}}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	0.04423	0.00050	1.33145	1.77470
		250	0.02703	0.00031	0.95107	0.90527
		500	0.00539	0.00006	0.60667	0.36808
		750	0.01896	0.00022	0.47029	0.22153
		1000	0.01166	0.00013	0.37641	0.14182
		1250	0.00493	0.00006	0.30383	0.09234
	$\hat{t}_{yr1}$	125	0.00089	0.00001	0.22202	0.04930
		250	-0.00164	0.00002	0.15823	0.02504
		500	-0.00180	0.00002	0.10610	0.01126
		750	-0.00068	0.00001	0.08122	0.00660
		1000	0.00035	0.00000	0.06703	0.00449
		1250	-0.00183	0.00002	0.05290	0.00280
	$\hat{t}_{yr2}$	125	-2.93249	0.03338	8.63730	83.20236
		250	-1.29061	0.01469	5.32753	30.04822
		500	-0.62153	0.00707	3.32198	11.42186
		750	-0.36952	0.00421	2.36155	5.71345
		1000	-0.25147	0.00286	1.83004	3.41227
		1250	-0.17060	0.00194	1.48975	2.24846
3	$\hat{t}_{yFB}$	125	0.03466	0.00039	1.62514	2.64229
		250	0.03308	0.00038	1.13674	1.29327
		500	0.01669	0.00019	0.73850	0.54565
		750	0.03038	0.00035	0.57627	0.33301
		1000	0.01380	0.00016	0.45671	0.20877
		1250	0.01393	0.00016	0.36943	0.13667
	$\hat{t}_{yr3}$	125	-0.21349	0.00243	6.76718	45.84027
		250	0.39528	0.00450	3.99021	16.07800
		500	0.08015	0.00091	0.75080	0.57013
		750	0.02998	0.00034	0.23446	0.05587
		1000	0.01835	0.00021	0.16784	0.02851
		1250	0.00358	0.00004	0.13325	0.01777
	$\hat{t}_{yr4}$	125	4.28057	0.04872	3.06852	27.73910
		250	3.37851	0.03846	2.89219	19.77907
		500	2.39848	0.02730	2.35462	11.29695
		750	1.99248	0.02268	1.96945	7.84869
		1000	1.69196	0.01926	1.45578	4.98205
		1250	1.49409	0.01701	0.99945	3.23119

Figura 6.7: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a = N_b = 2000$ e $N_{ab} = 500$, sob o cenário 2

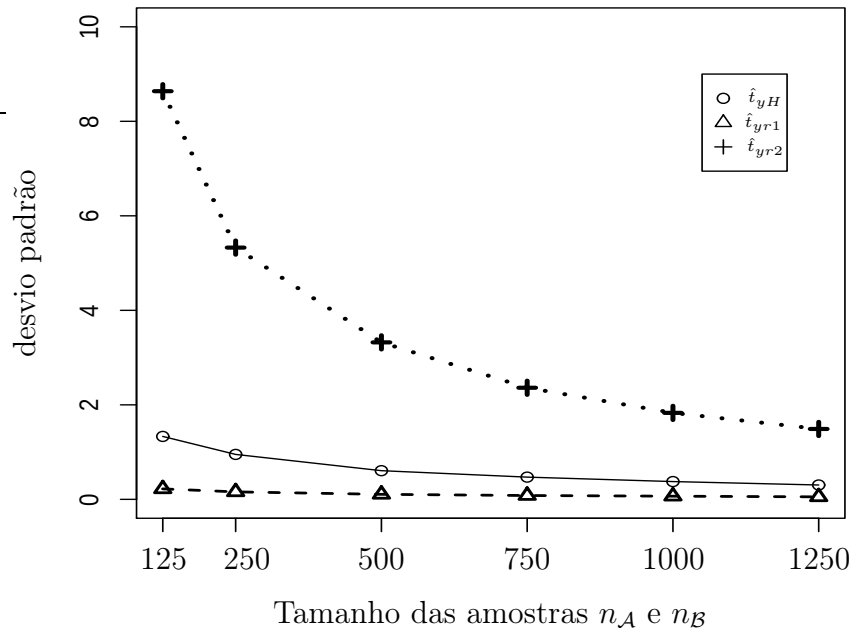
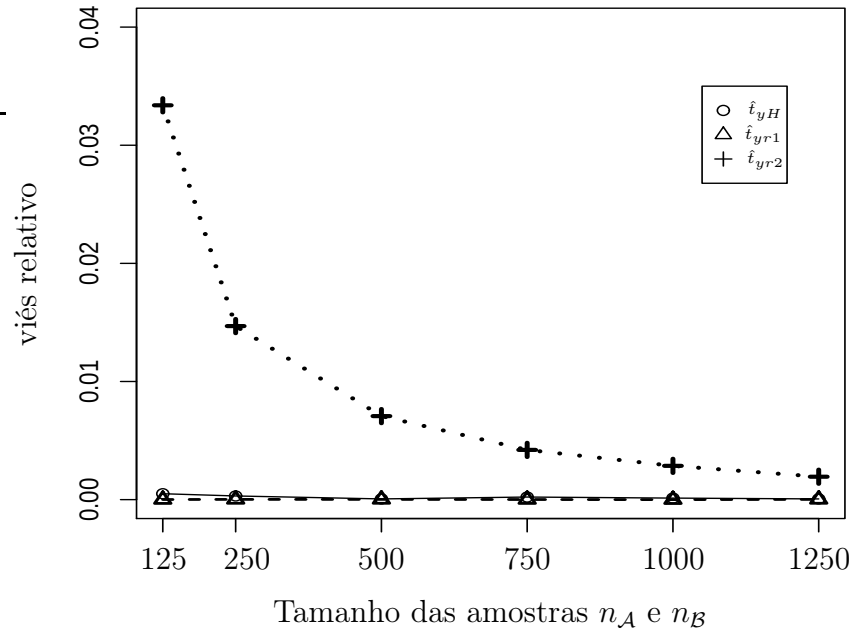


Figura 6.8: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a = N_b = 2000$ e $N_{ab} = 500$, sob o cenário 3

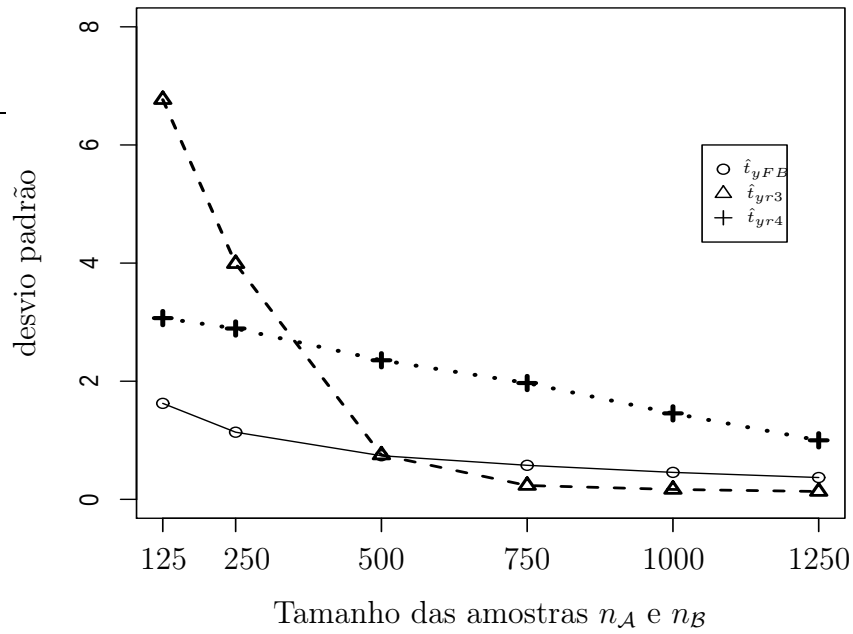
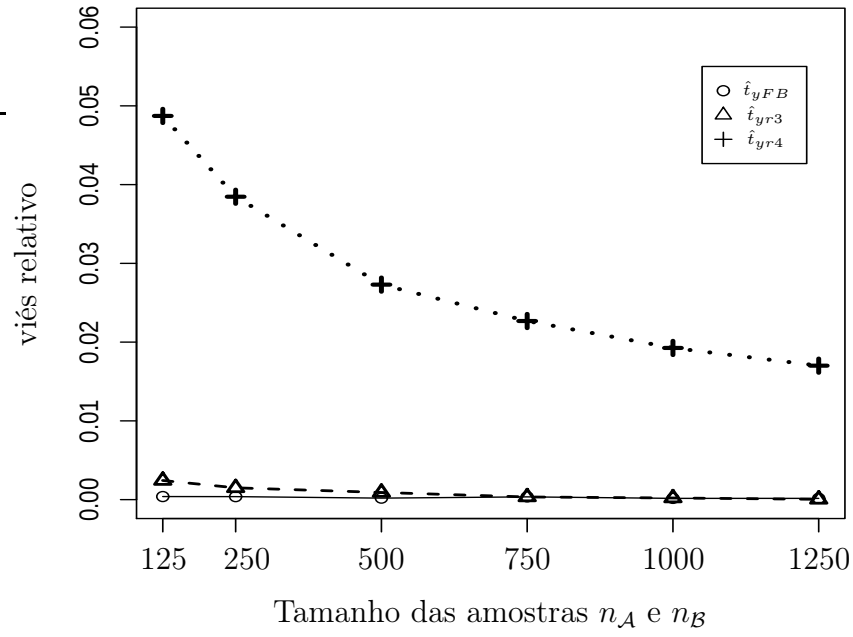


Tabela 6.4: Avaliação dos estimadores propostos sob os cenários 2 e 3 para o plano AAS ($N_a = N_b = 1750$ e $N_{ab} = 1000$)

Cenário	Estimador	$n_{\mathcal{A}}=n_{\mathcal{B}}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	-0.04756	0.00058	1.33616	1.78758
		250	-0.03031	0.00037	0.93865	0.88198
		500	-0.00387	0.00005	0.61165	0.37413
		750	0.00225	0.00003	0.47986	0.23027
		1000	0.00336	0.00004	0.37775	0.14271
		1250	0.00299	0.00004	0.31690	0.10044
	$\hat{t}_{yr1}$	125	0.10308	0.00127	3.23842	10.49799
		250	-0.01710	0.00021	0.17608	0.03130
		500	-0.00259	0.00003	0.10954	0.01201
		750	-0.00198	0.00002	0.08367	0.00700
		1000	-0.00247	0.00003	0.06871	0.00473
		1250	-0.00032	0.00000	0.05669	0.00321
	$\hat{t}_{yr2}$	125	-21.00958	0.25803	8.95244	521.54842
		250	-23.36930	0.28702	5.96937	581.75764
		500	-25.06966	0.30790	3.41189	640.12896
		750	-25.67420	0.31532	2.14637	663.77127
		1000	-25.99603	0.31928	1.37380	677.68076
		1250	-26.19655	0.32174	0.80790	686.91220
3	$\hat{t}_{yFB}$	125	-0.02380	0.00029	1.91549	3.66967
		250	-0.02482	0.00030	1.34954	1.82187
		500	-0.02275	0.00028	0.86022	0.74049
		750	-0.01374	0.00017	0.66878	0.44745
		1000	-0.00561	0.00007	0.52511	0.27577
		1250	-0.00204	0.00003	0.44518	0.19819
	$\hat{t}_{yr3}$	125	0.04256	0.00052	1.04812	1.10037
		250	0.01107	0.00014	0.75688	0.57299
		500	0.02487	0.00031	0.49942	0.25004
		750	0.02059	0.00025	0.38635	0.14969
		1000	0.01251	0.00015	0.31278	0.09799
		1250	0.00844	0.00010	0.26043	0.06790
	$\hat{t}_{yr4}$	125	15.20313	0.18672	1.88136	234.67478
		250	15.54234	0.19089	1.31159	243.28453
		500	15.67046	0.19246	0.90440	246.38113
		750	15.67370	0.19250	0.68823	246.13844
		1000	15.69064	0.19271	0.54462	246.49285
		1250	15.69349	0.19274	0.45151	246.48944

Figura 6.9: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a = N_b = 1750$ e $N_{ab} = 1000$, sob o cenário 2

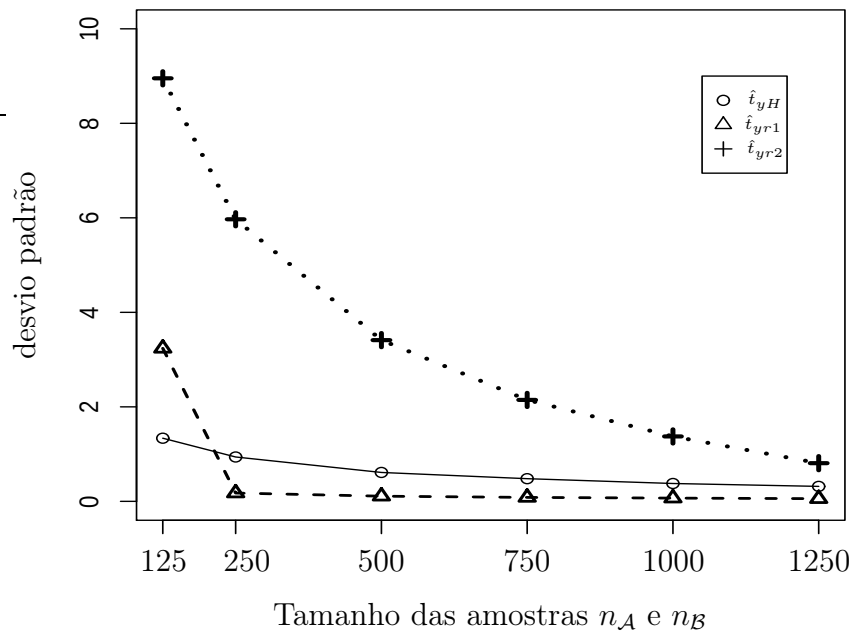
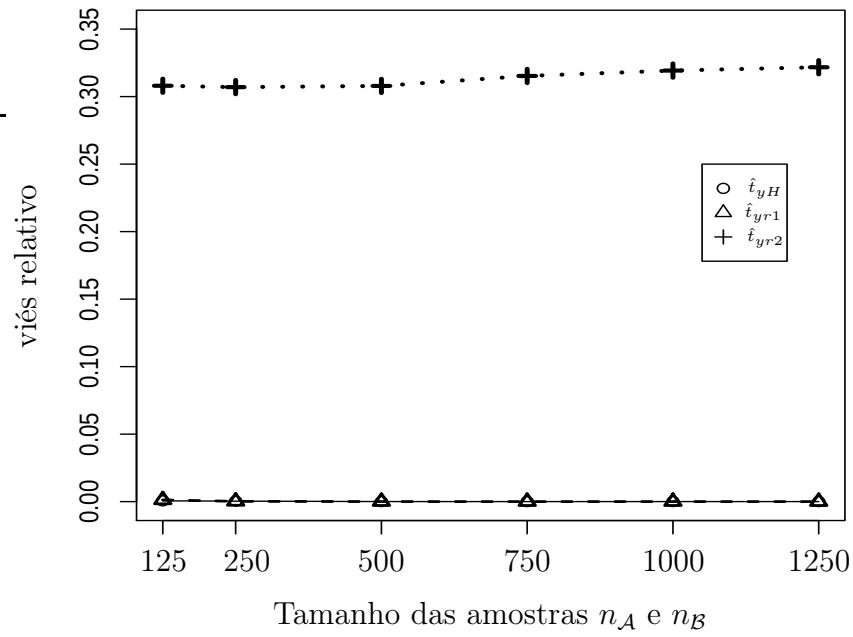


Figura 6.10: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a = N_b = 1750$ e $N_{ab} = 1000$, sob o cenário 3

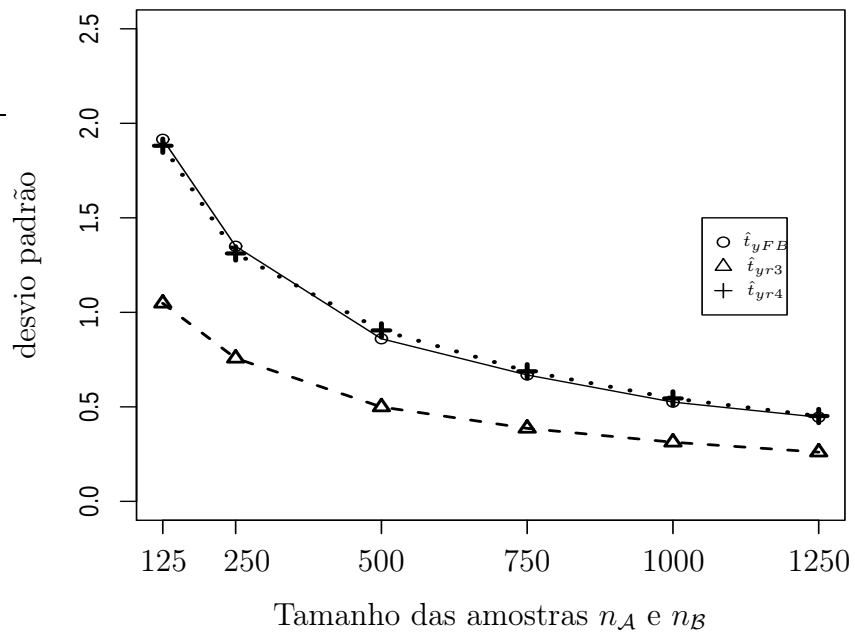
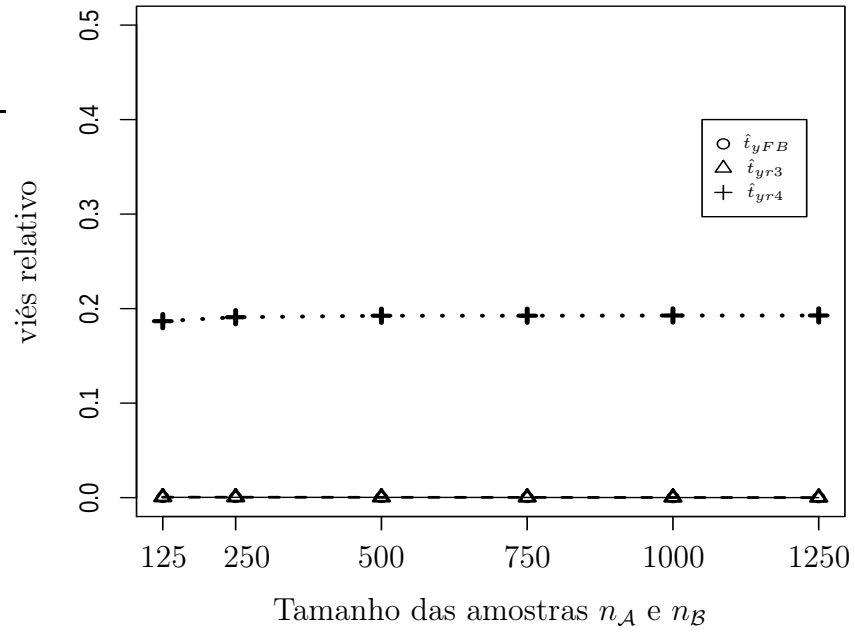


Tabela 6.5: Avaliação dos estimadores propostos sob os cenários 2 e 3 para o plano AAS ($N_a = N_b = 1500$ e $N_{ab} = 1500$)

Cenário	Estimador	$n_{\mathcal{A}}=n_{\mathcal{B}}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	0.07481	0.00100	1.30545	1.70978
		250	0.02487	0.00033	0.87922	0.77364
		500	0.01287	0.00017	0.58546	0.34293
		750	0.02505	0.00033	0.48391	0.23480
		1000	0.01316	0.00018	0.38603	0.14919
		1250	0.02091	0.00028	0.31782	0.10145
	$\hat{t}_{yr1}$	125	3.49366	0.04666	124.67583	15556.26847
		250	0.60316	0.00806	18.51545	343.18563
		500	-0.00673	0.00009	0.18180	0.03310
		750	-0.00261	0.00003	0.10068	0.01014
		1000	-0.00411	0.00005	0.07578	0.00576
		1250	-0.00434	0.00006	0.06134	0.00378
	$\hat{t}_{yr2}$	125	-11.82976	0.15801	1.54365	142.3261
		250	-11.82788	0.15798	0.85444	140.62872
		500	-11.80324	0.15765	0.56738	139.63843
		750	-11.80039	0.15762	0.45381	139.45517
		1000	-11.78228	0.15737	0.36860	138.95802
		1250	-11.78384	0.15740	0.31287	138.95678
3	$\hat{t}_{yFB}$	125	0.02090	0.00028	1.95265	3.81327
		250	0.01682	0.00022	1.38398	1.91568
		500	-0.00170	0.00002	0.94938	0.90133
		750	0.01579	0.00021	0.76289	0.58225
		1000	0.01724	0.00023	0.61043	0.37293
		1250	0.02817	0.00038	0.50550	0.25633
	$\hat{t}_{yr3}$	125	0.05306	0.00071	1.16911	1.36964
		250	0.01305	0.00017	0.83432	0.69626
		500	0.01174	0.00016	0.56808	0.32285
		750	0.00661	0.00009	0.44022	0.19383
		1000	-0.00504	0.00007	0.35644	0.12707
		1250	-0.00827	0.00011	0.29989	0.09000
	$\hat{t}_{yr4}$	125	9.09491	0.12148	1.59403	85.25838
		250	8.90073	0.11889	1.17148	80.59536
		500	8.75939	0.11700	0.84948	77.44856
		750	8.73209	0.11663	0.68070	76.71266
		1000	8.70312	0.11625	0.54491	76.04129
		1250	8.69413	0.11613	0.45864	75.79827

Figura 6.11: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a = N_b = 1500$ e $N_{ab} = 1500$, sob o cenário 2

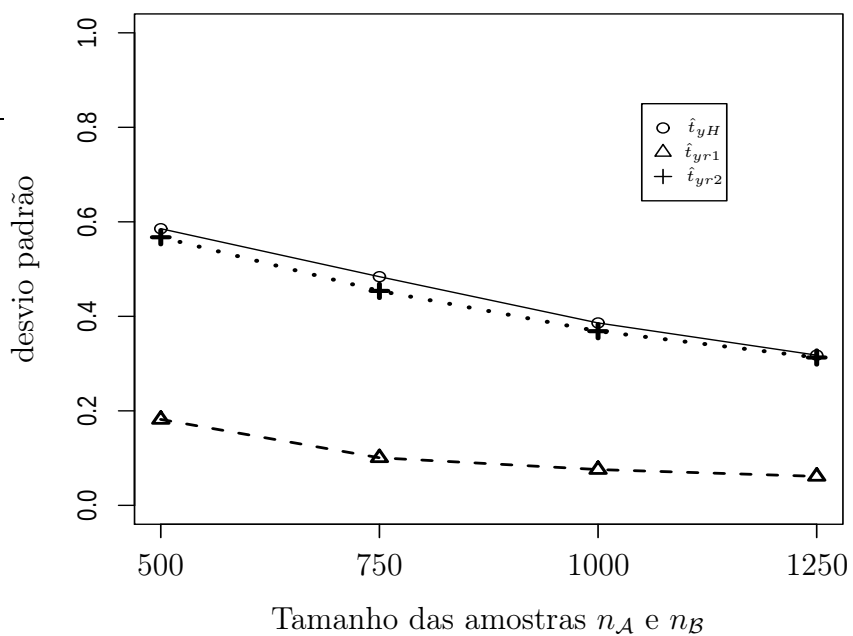
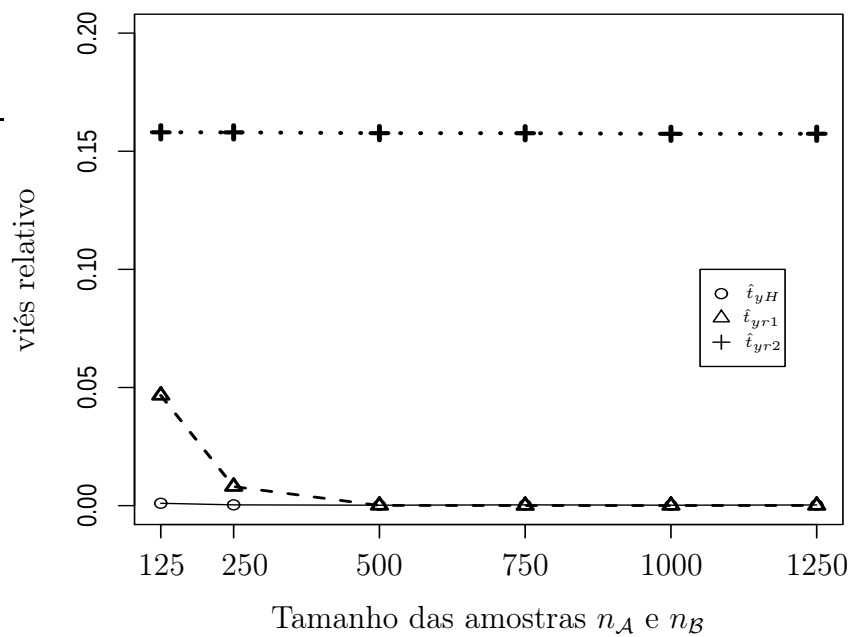


Figura 6.12: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yr3}$ e $\hat{t}_{yr4}$ com $N_a = N_b = 1500$ e $N_{ab} = 1500$, sob o cenário 3

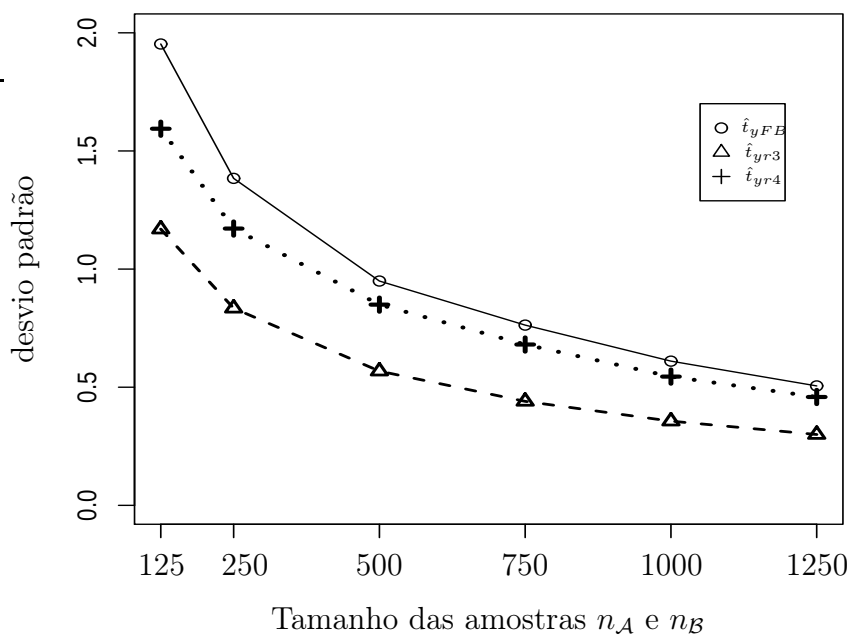
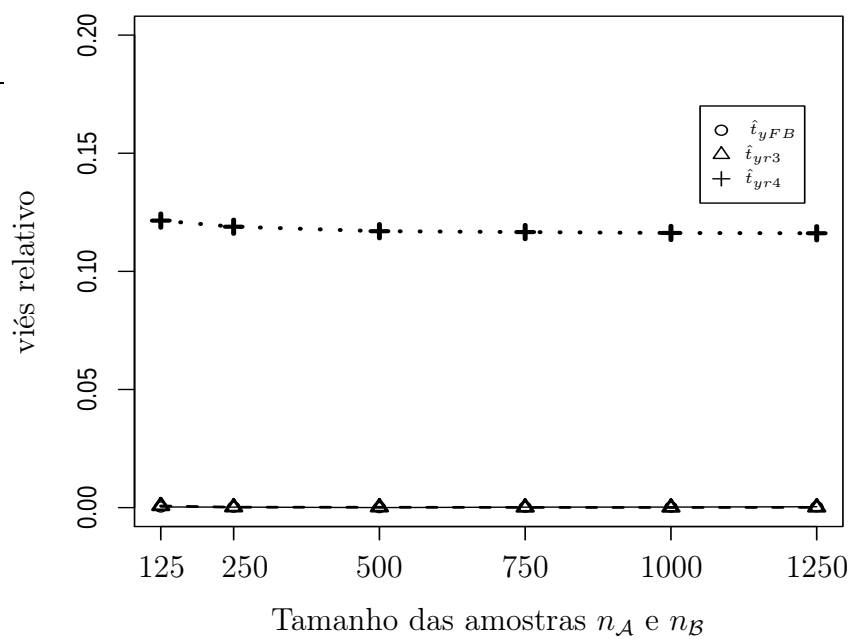


Tabela 6.6: Avaliação dos estimadores sob os cenários 2 e 3 para o plano AAS, em que $n_{\mathcal{A}} = 1000$ (dados populacionais: $N_a = N_b = 2000$ e $N_{ab} = 500$)

Cenário	Estimador	$n_{\mathcal{A}}$	$n_{\mathcal{B}}$	Estatística			
				Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	1000	200	0.01208	0.00014	0.39134	0.15329
		1000	300	0.01341	0.00015	0.38829	0.15095
		1000	400	0.01134	0.00013	0.38497	0.14833
		1000	500	0.01241	0.00014	0.38351	0.14723
		1000	700	0.01287	0.00015	0.37994	0.14452
		1000	900	0.01210	0.00014	0.37776	0.14285
	$\hat{t}_{yr1}$	1000	200	0.27416	0.00312	14.13538	199.88411
		1000	300	-0.06618	0.00075	1.97652	3.91101
		1000	400	-0.00085	0.00001	0.08805	0.00775
		1000	500	0.00032	0.00000	0.08096	0.00655
		1000	700	0.00091	0.00001	0.07483	0.00560
		1000	900	0.00053	0.00001	0.06891	0.00475
	$\hat{t}_{yr2}$	1000	200	-0.52078	0.00593	2.60279	7.04575
		1000	300	-0.36269	0.00413	2.16650	4.82526
		1000	400	-0.31675	0.00361	2.02762	4.21159
		1000	500	-0.29058	0.00331	1.95318	3.89936
		1000	700	-0.26598	0.00303	1.87906	3.60160
		1000	900	-0.25465	0.00290	1.84316	3.46208
3	$\hat{t}_{yFB}$	1000	200	0.01522	0.00017	0.51693	0.26745
		1000	300	0.01778	0.00020	0.51195	0.26241
		1000	400	0.01890	0.00022	0.50259	0.25296
		1000	500	0.02084	0.00024	0.49309	0.24357
		1000	700	0.02281	0.00026	0.47752	0.22854
		1000	900	0.01658	0.00019	0.46252	0.21420
	$\hat{t}_{yr3}$	1000	200	0.00325	0.00004	0.37614	0.14149
		1000	300	0.00519	0.00006	0.34953	0.12220
		1000	400	0.00120	0.00001	0.32374	0.10481
		1000	500	0.00304	0.00003	0.29963	0.08979
		1000	700	0.00611	0.00007	0.24940	0.06224
		1000	900	0.00368	0.00016	0.19610	0.03864
	$\hat{t}_{yr4}$	1000	200	3.82313	0.04352	1.56250	17.05771
		1000	300	2.60385	0.02964	1.34766	8.59624
		1000	400	2.08364	0.02372	1.47803	6.52613
		1000	500	1.93095	0.02198	1.78471	6.91375
		1000	700	1.70143	0.01937	1.83540	6.26355
		1000	900	1.63381	0.01860	1.58416	5.17889

Figura 6.13: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yH}$, $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ com $N_a = N_b = 2000$ e $N_{ab} = 500$ e $n_{\mathcal{A}} = 1000$, sob o cenário 2

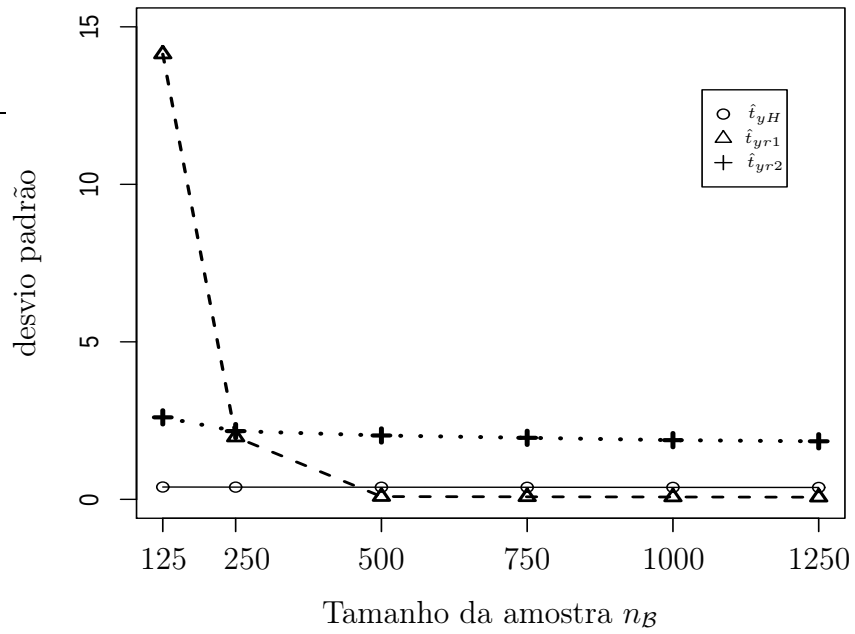
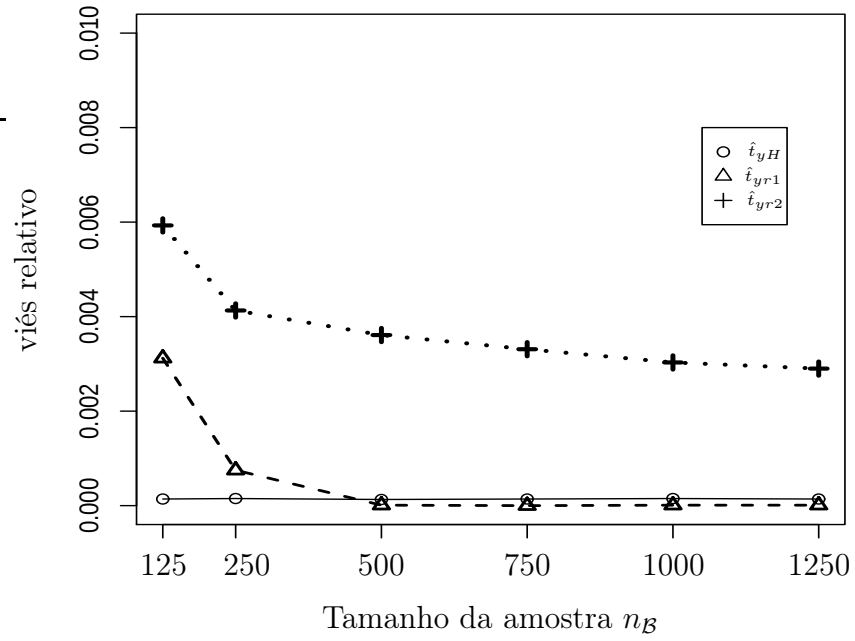


Figura 6.14: Viés Relativo e Desvio Padrão dos estimadores $\hat{t}_{yFB}$, $\hat{t}_{yR3}$ e $\hat{t}_{yR4}$ com $N_a = N_b = 2000$ e $N_{ab} = 500$ e $n_{\mathcal{A}} = 1000$, sob o cenário 3

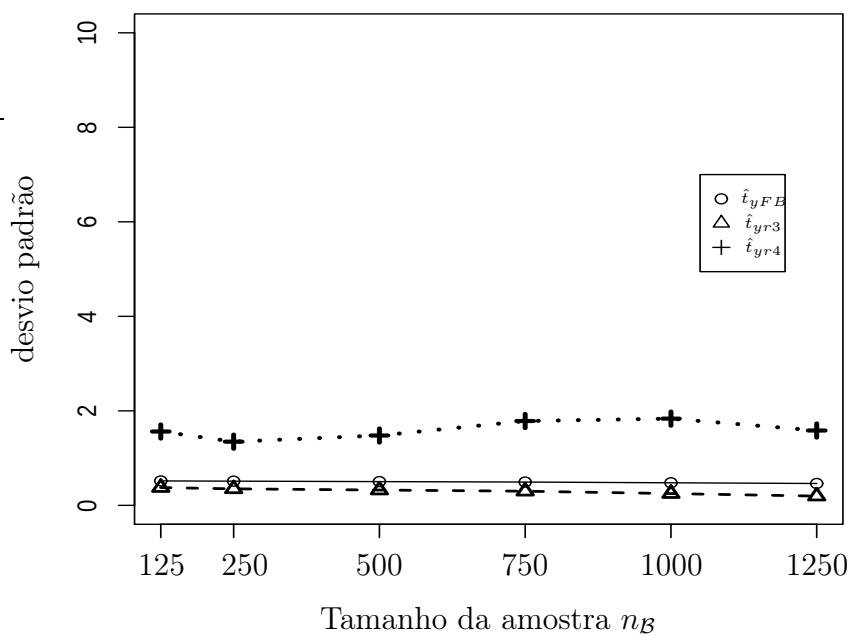
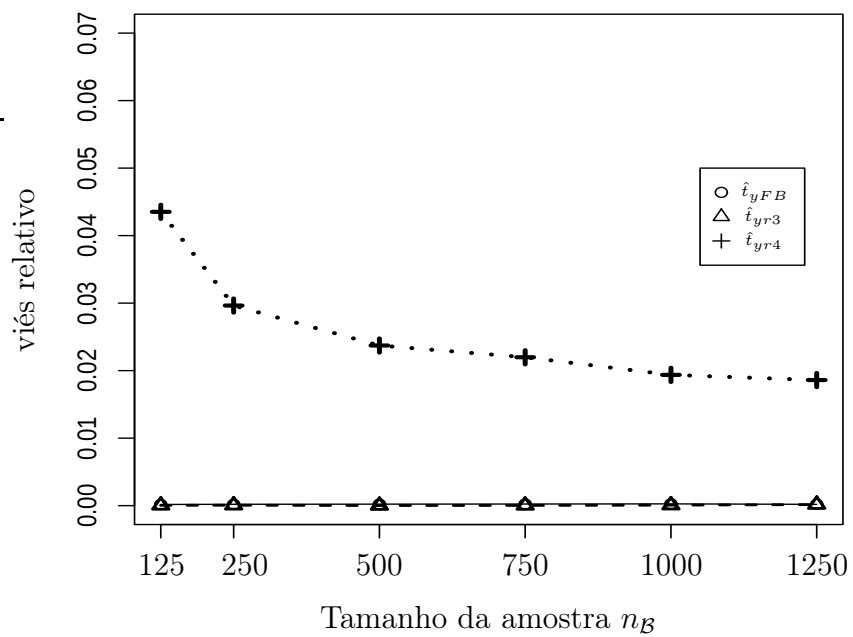


Tabela 6.7: Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 2000$ e $N_{ab} = 500$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.03932	0.00026	2.31891	5.37887
	250	50	0.00508	0.00003	1.65100	2.72582
	500	100	-0.01688	0.00011	1.10177	1.21418
	750	150	0.03726	0.00025	0.84572	0.71663
	1000	200	0.03723	0.00025	0.68042	0.46436
	1250	250	0.02385	0.00016	0.54196	0.29429
	1750	300	-0.00020	0.00000	0.36332	0.13200
	2000	350	-0.01173	0.00008	0.27237	0.07432
$\hat{t}_{yr2}$	125	25	-6.67628	0.04399	15.37367	280.92244
	250	50	-4.41814	0.02911	9.73697	114.32860
	500	100	-3.09016	0.02036	5.84491	43.71207
	750	150	-2.51822	0.01659	4.41678	25.84943
	1000	200	-2.17278	0.01432	3.51097	17.04787
	1250	250	-2.13704	0.01408	2.80075	12.41115
	1750	300	-2.05583	0.01355	1.87278	7.73373
	2000	350	-2.03084	0.01338	1.44623	6.21589
$\hat{t}_{yH}$	500	25	-0.01703	0.00011	1.10092	1.21232
	500	50	-0.01766	0.00012	1.10008	1.21050
	500	100	-0.01688	0.00011	1.10177	1.21418
	500	150	-0.01719	0.00011	1.10038	1.21113
	500	200	-0.01564	0.00010	1.10167	1.21391
	500	250	-0.01598	0.00011	1.09904	1.20815
	500	300	-0.01684	0.00011	1.09810	1.20610
	500	350	-0.01645	0.00011	1.09814	1.20619
$\hat{t}_{yr2}$	500	25	-4.22370	0.02783	7.96482	81.27805
	500	50	-3.36962	0.02220	6.33107	51.43687
	500	100	-3.09016	0.02036	5.84491	43.71207
	500	150	-3.01057	0.01984	5.71284	41.70007
	500	200	-2.97283	0.01959	5.64975	40.75739
	500	250	-2.95073	0.01944	5.61284	40.21072
	500	300	-2.93576	0.01934	5.58840	39.84892
	500	350	-2.92541	0.01928	5.57168	39.60161

Figura 6.15: Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 2000$ e $N_{ab} = N_B = 500$ ($U_A \subset U_B$)

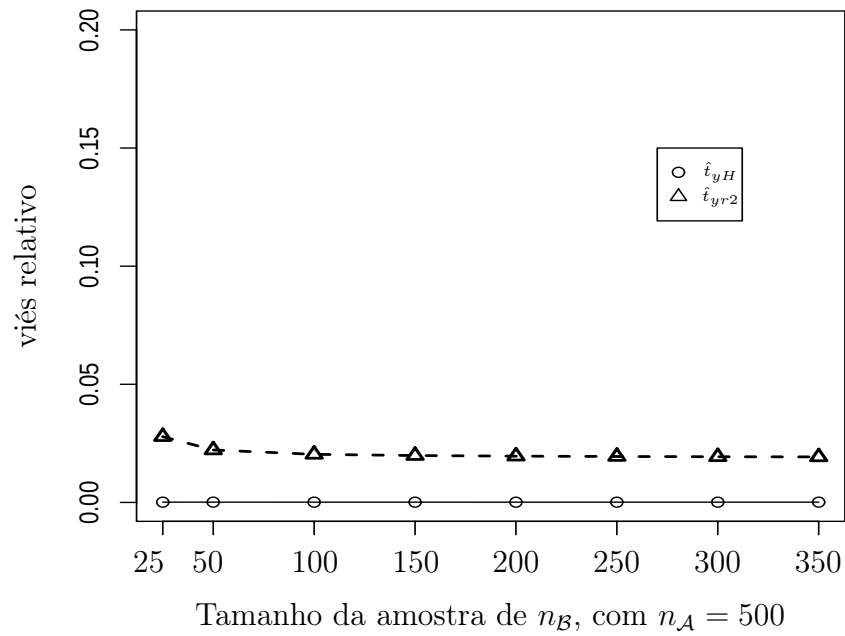
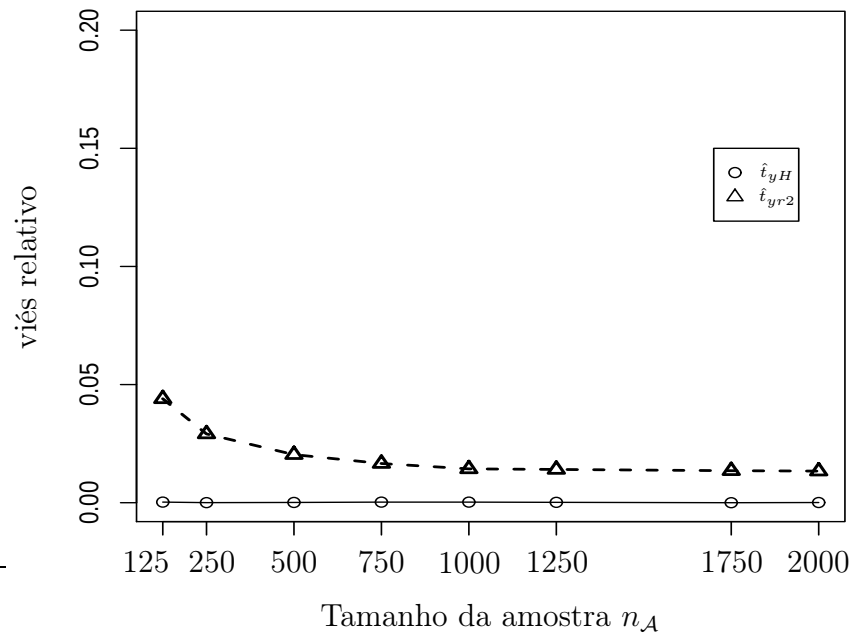


Figura 6.16: Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 2000$ e $N_{ab} = N_B = 500$ ($U_A \subset U_B$)

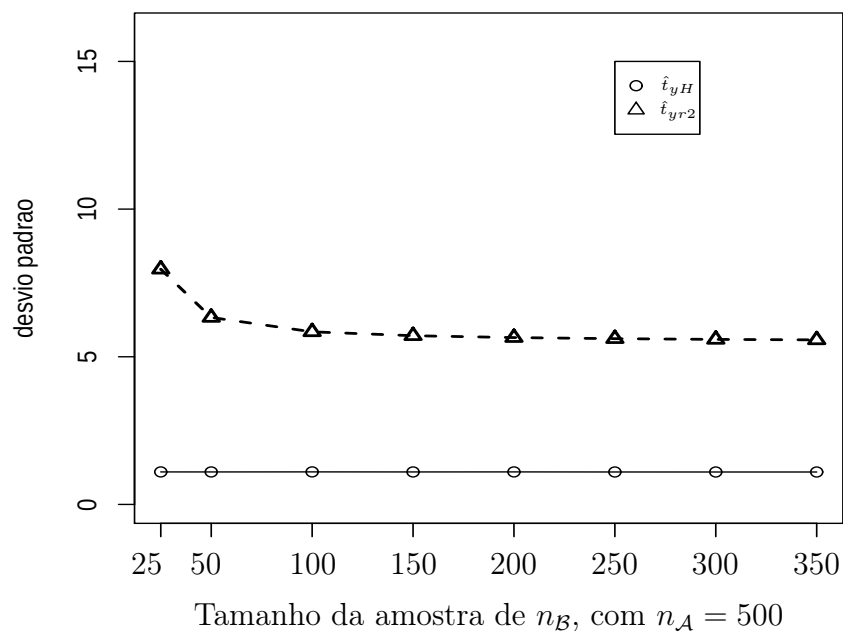
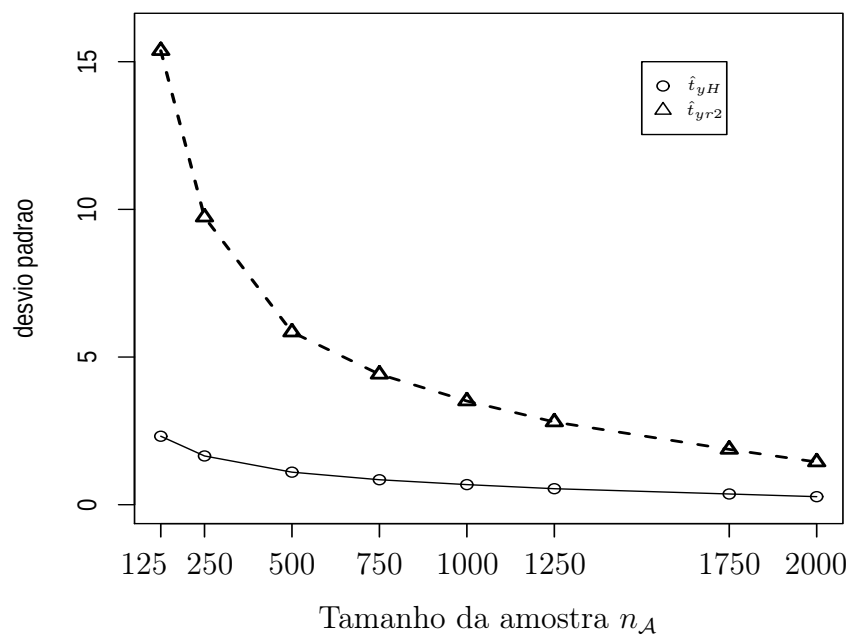


Tabela 6.8: Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1750$ e $N_{ab} = 1000$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	-0.00613	0.00005	2.25932	5.10458
	250	50	0.02977	0.00023	1.54120	2.37619
	500	100	0.00531	0.00004	0.98979	0.97972
	750	150	0.01429	0.00011	0.74691	0.55808
	1000	200	0.01630	0.00013	0.58819	0.34624
	1250	250	0.00577	0.00004	0.49001	0.24014
	1750	300	-0.00030	0.00000	0.33944	0.11522
	2000	350	-0.00545	0.00004	0.28603	0.08184
$\hat{t}_{yr2}$	125	25	-39.37630	0.30732	12.31409	1702.12978
	250	50	-42.82892	0.33427	6.74675	1879.83492
	500	100	-44.16756	0.34472	2.86155	1958.96229
	750	150	-43.66750	0.34082	2.43292	1912.76940
	1000	200	-42.22246	0.32954	2.02693	1786.84414
	1250	250	-40.26322	0.31425	1.78521	1624.31382
	1750	300	-30.11450	0.23504	1.43201	908.93359
	2000	350	-26.29052	0.20519	1.24655	692.74553
$\hat{t}_{yH}$	500	25	0.00581	0.00005	0.99608	0.99221
	500	50	0.00601	0.00005	0.99524	0.99054
	500	100	0.00531	0.00004	0.98979	0.97972
	500	150	0.00393	0.00003	0.98878	0.97771
	500	200	0.00522	0.00004	0.98862	0.97739
	500	250	0.00518	0.00004	0.98878	0.97772
	500	300	0.00578	0.00005	0.98767	0.97553
	500	350	0.00460	0.00004	0.98601	0.97223
$\hat{t}_{yr2}$	500	25	-17.30961	0.13510	5.68676	331.96167
	500	50	-30.35394	0.23691	4.24471	939.37915
	500	100	-44.16756	0.34472	3.86155	1958.96229
	500	150	-45.51805	0.35526	3.39632	2083.42793
	500	200	-45.12430	0.35219	4.00479	2052.24051
	500	250	-44.82869	0.34988	4.36977	2028.70650
	500	300	-44.62848	0.34832	4.62098	2013.05441
	500	350	-44.47173	0.34709	4.78987	2000.67770

Figura 6.17: Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1750$ e $N_{ab} = N_B = 1000$ ($U_A \subset U_B$)

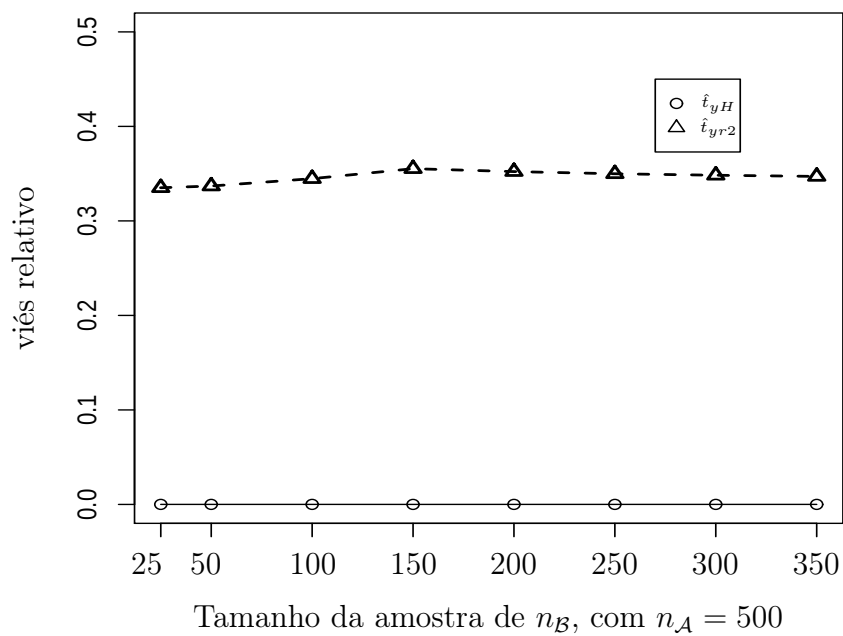
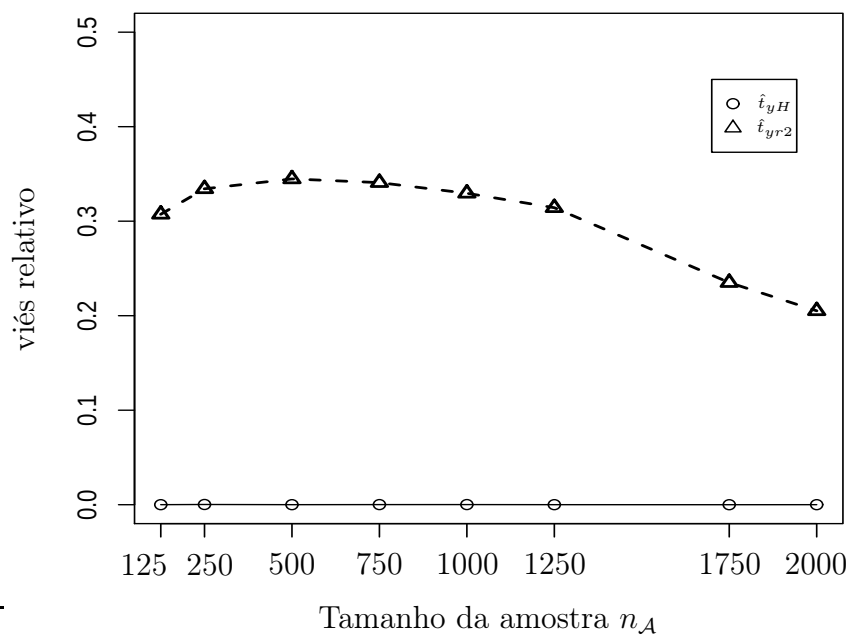


Figura 6.18: Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1750$ e $N_{ab} = N_B = 1000$ ($U_A \subset U_B$)

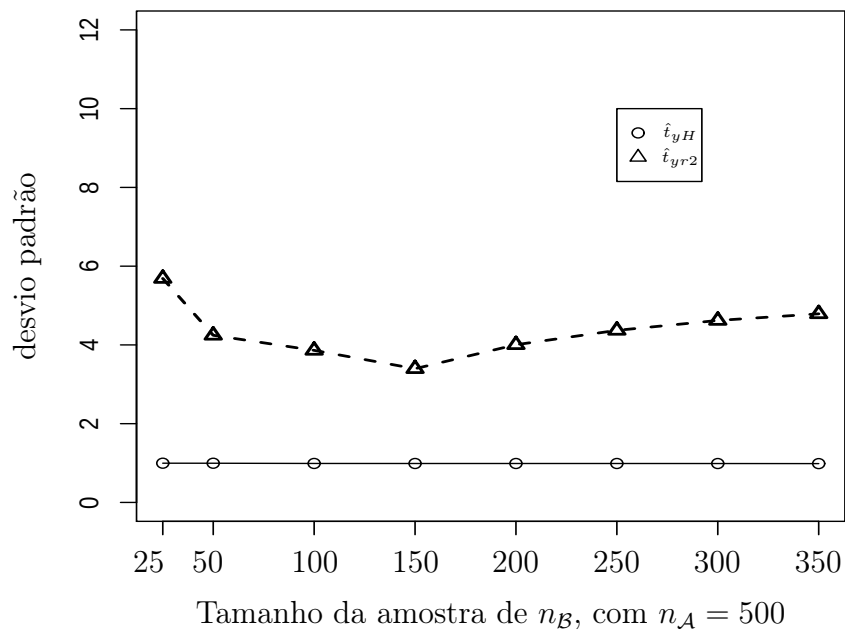
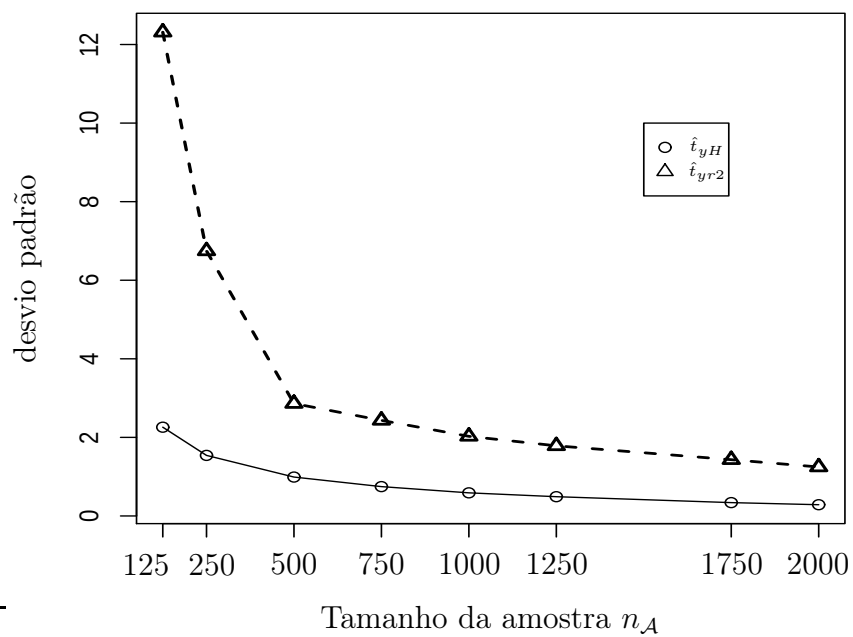


Tabela 6.9: Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1500$ e $N_{ab} = 1500$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	-0.03661	0.00034	1.86024	3.46183
	250	50	0.01866	0.00017	1.35236	1.82922
	500	100	0.01950	0.00018	0.90371	0.81707
	750	150	0.04017	0.00037	0.70975	0.50536
	1000	200	0.01241	0.00011	0.57323	0.32874
	1250	250	0.01097	0.00010	0.46961	0.22066
	1750	300	0.00290	0.00003	0.33597	0.11289
	2000	350	0.00089	0.00001	0.27899	0.07784
$\hat{t}_{yr2}$	125	25	-23.28459	0.21495	1.66235	544.93571
	250	50	-23.28172	0.21492	0.86184	542.78122
	500	100	-23.28269	0.21493	0.62612	542.47590
	750	150	-23.29608	0.21505	0.48714	542.94474
	1000	200	-23.29053	0.21500	0.39629	542.60594
	1250	250	-23.29457	0.21504	0.33939	542.75219
	1750	300	-22.71484	0.20969	0.27049	516.70428
	2000	350	-20.62807	0.19042	0.21581	427.11960
$\hat{t}_{yH}$	500	25	0.02198	0.00020	0.91100	0.83040
	500	50	0.02184	0.00020	0.90621	0.82170
	500	100	0.01950	0.00018	0.90371	0.81707
	500	150	0.02165	0.00020	0.90260	0.81515
	500	200	0.01941	0.00018	0.89791	0.80662
	500	250	0.01908	0.00018	0.89496	0.80132
	500	300	0.02135	0.00020	0.89255	0.79709
	500	350	0.02039	0.00019	0.89228	0.79659
$\hat{t}_{yr2}$	500	25	-12.37558	0.11424	4.76575	175.86739
	500	50	-21.63410	0.19971	2.37758	473.68711
	500	100	-23.28269	0.21493	0.62612	542.47590
	500	150	-23.28950	0.21499	0.60970	542.77261
	500	200	-23.28632	0.21496	0.60560	542.61966
	500	250	-23.28631	0.21496	0.59883	542.61101
	500	300	-23.28971	0.21499	0.59641	542.76610
	500	350	-23.28885	0.21499	0.59445	542.72389

Figura 6.19: Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1500$ e $N_{ab} = N_B = 1500$ ($U_A \subset U_B$)

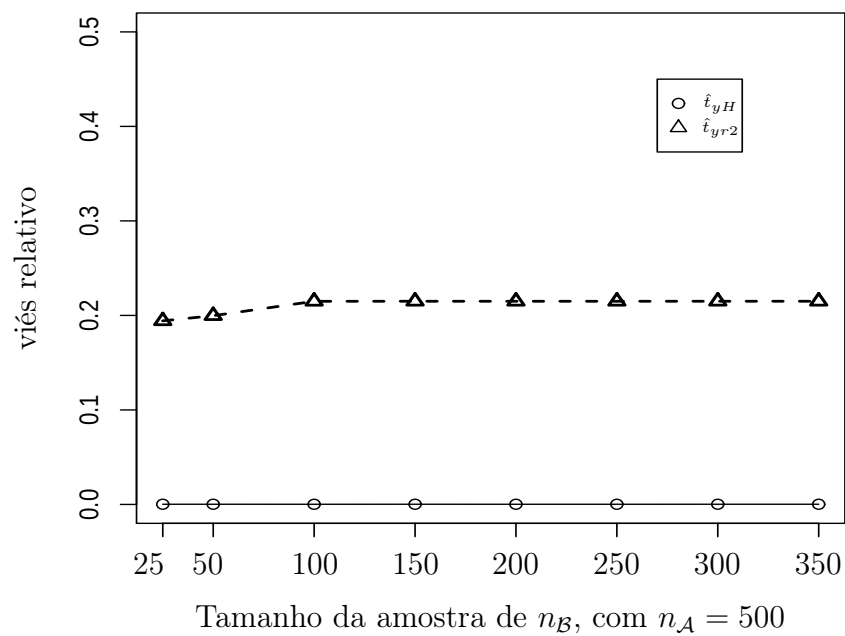
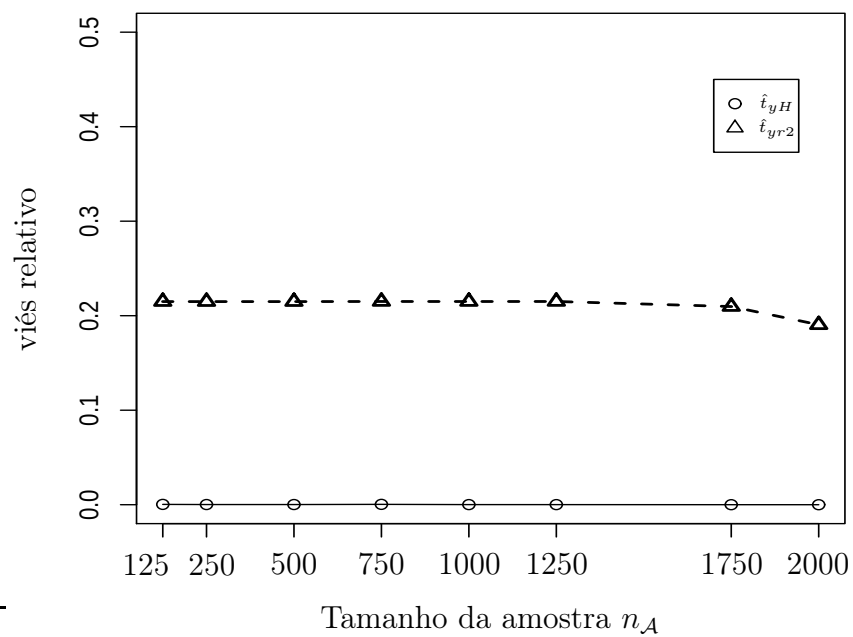


Figura 6.20: Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1500$ e $N_{ab} = N_B = 1500$ ($U_A \subset U_B$)

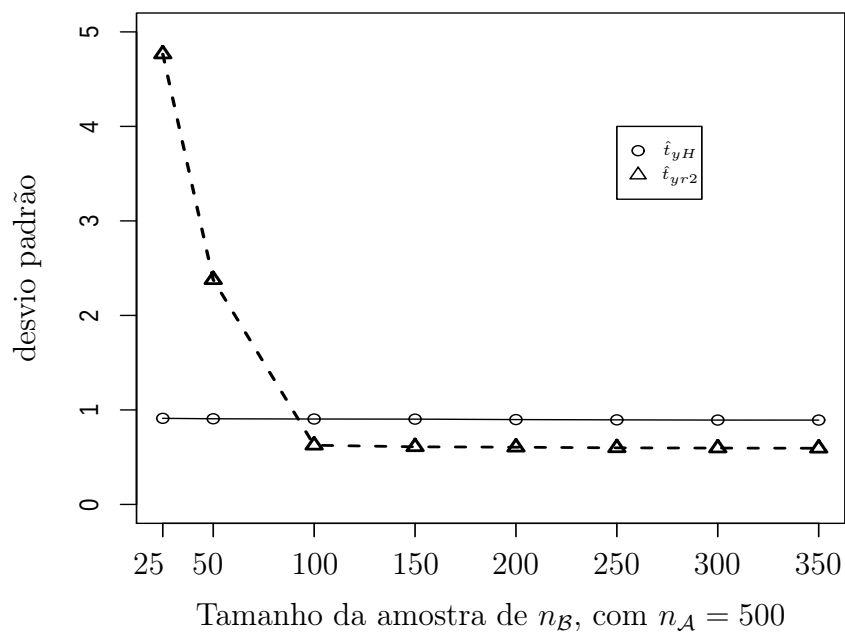
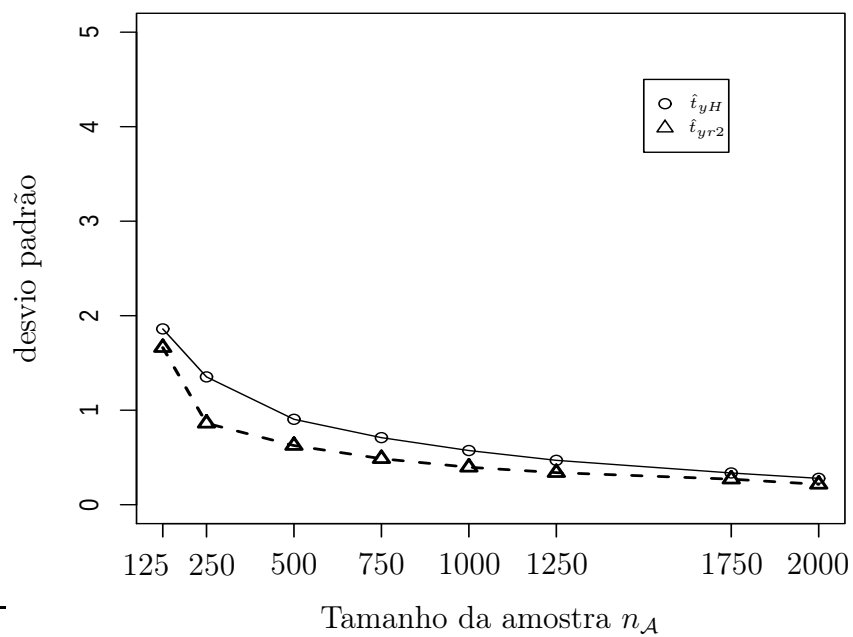


Tabela 6.10: Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1250$ e $N_{ab} = 2000$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.01770	0.00019	1.71389	2.93772
	250	50	0.02076	0.00023	1.21541	1.47764
	500	100	-0.00891	0.00010	0.81313	0.66126
	750	150	0.00859	0.00009	0.65204	0.42523
	1000	200	0.01705	0.00019	0.53186	0.28316
	1250	250	0.01034	0.00011	0.44483	0.19798
	1750	300	-0.00225	0.00002	0.31284	0.09787
	2000	350	-0.00847	0.00009	0.27393	0.07511
$\hat{t}_{yr2}$	125	25	-4.51680	0.04931	1.24356	21.94795
	250	50	-4.56472	0.04984	0.86245	21.58053
	500	100	-4.59553	0.05017	0.58166	21.45724
	750	150	-4.60559	0.05028	0.46851	21.43098
	1000	200	-4.61180	0.05035	0.37679	21.41065
	1250	250	-4.61224	0.05036	0.32103	21.37578
	1750	300	-4.60848	0.05031	0.23044	21.29118
	2000	350	-4.61363	0.05037	0.19564	21.32388
$\hat{t}_{yH}$	500	25	-0.00698	0.00008	0.82188	0.67553
	500	50	-0.00789	0.00009	0.81846	0.66994
	500	100	-0.00891	0.00010	0.81313	0.66126
	500	150	-0.01149	0.00013	0.81094	0.65775
	500	200	-0.01077	0.00012	0.80632	0.65027
	500	250	-0.01183	0.00013	0.80406	0.64665
	500	300	-0.01374	0.00015	0.80299	0.64498
	500	350	-0.01570	0.00017	0.80068	0.64134
$\hat{t}_{yr2}$	500	25	-4.50741	0.04921	0.78297	20.92974
	500	50	-4.59550	0.05017	0.58307	21.45863
	500	100	-4.59553	0.05017	0.58166	21.45724
	500	150	-4.59479	0.05016	0.58034	21.44884
	500	200	-4.59510	0.05017	0.57915	21.45036
	500	250	-4.59515	0.05017	0.57913	21.45078
	500	300	-4.59492	0.05017	0.57921	21.44877
	500	350	-4.59453	0.05016	0.57848	21.44432

Figura 6.21: Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1250$ e $N_{ab} = N_B = 2000$ ($U_A \subset U_B$)

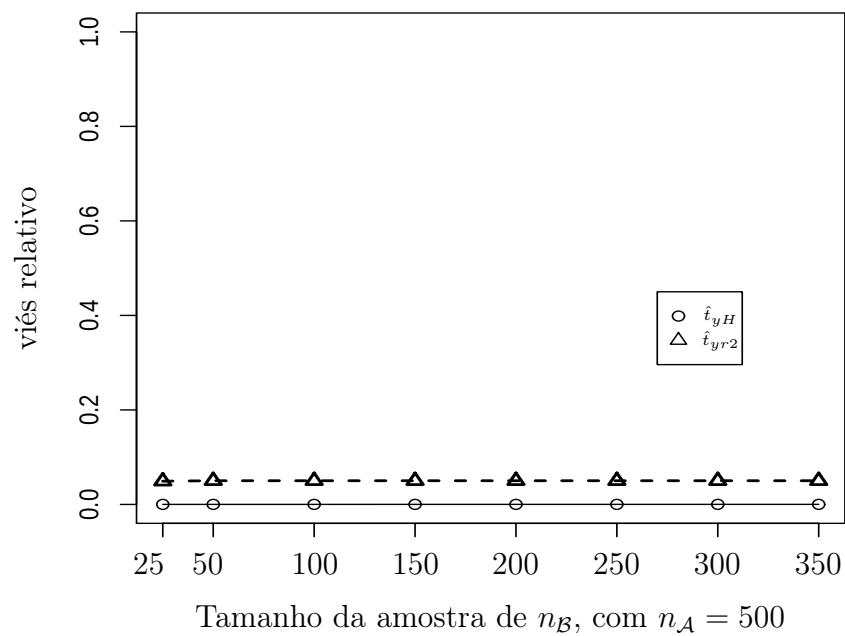
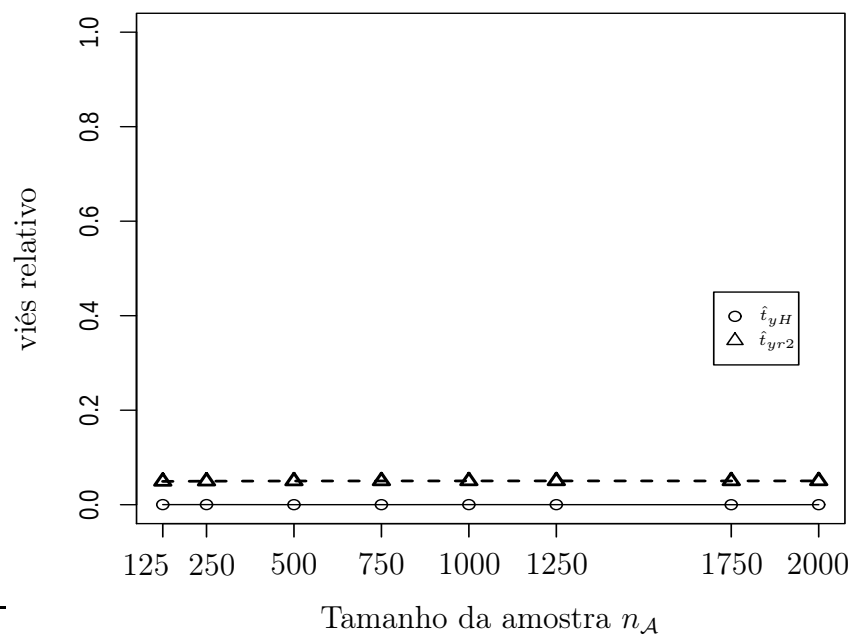


Figura 6.22: Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1250$ e $N_{ab} = N_B = 2000$ ($U_A \subset U_B$)

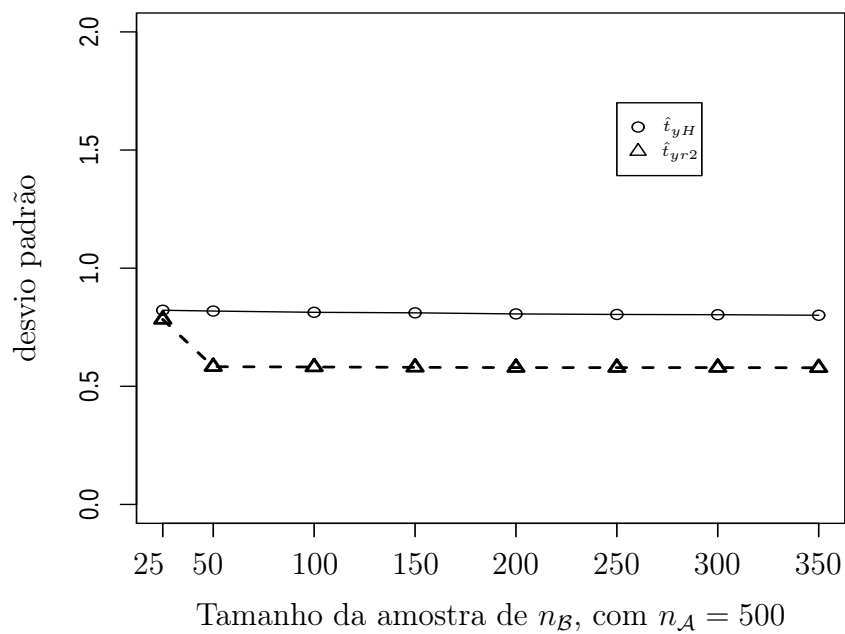
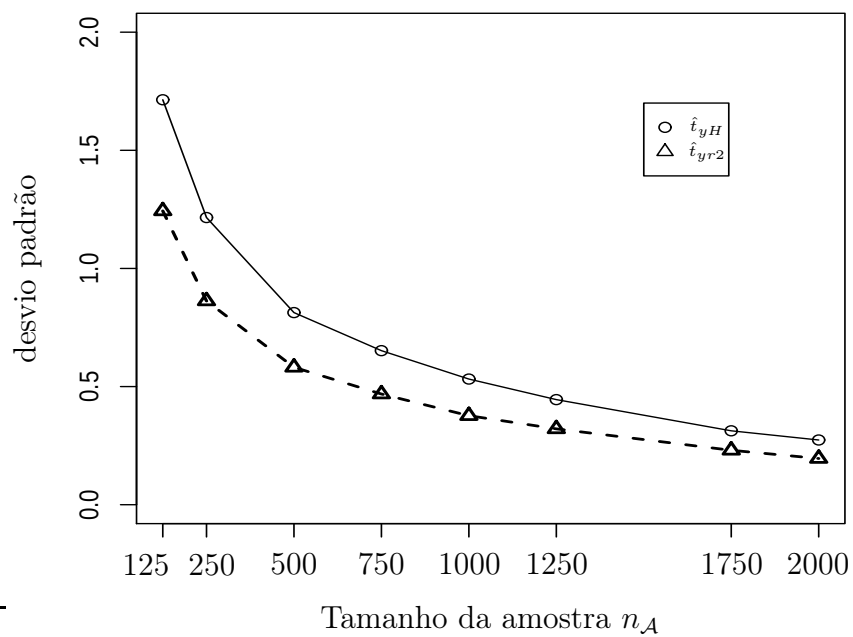


Tabela 6.11: Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 1000$ e $N_{ab} = 2500$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.06765	0.00088	1.54894	2.40379
	250	50	0.04782	0.00062	1.06220	1.13056
	500	100	0.01175	0.00015	0.69933	0.48920
	750	150	0.00852	0.00011	0.57191	0.32716
	1000	200	0.02129	0.00028	0.47630	0.22732
	1250	250	0.01421	0.00018	0.39782	0.15846
	1750	300	0.01541	0.00020	0.29790	0.08898
	2000	350	0.01835	0.00024	0.26198	0.06897
$\hat{t}_{yr2}$	125	25	11.95150	0.15461	1.06589	143.97442
	250	50	11.87946	0.15367	0.72791	141.65132
	500	100	11.85569	0.15337	0.49113	140.79864
	750	150	11.85363	0.15334	0.38531	140.65709
	1000	200	11.85271	0.15333	0.32167	140.59020
	1250	250	11.86410	0.15348	0.27818	140.83436
	1750	300	11.86319	0.15346	0.21055	140.77969
	2000	350	11.86455	0.15348	0.18264	140.80088
$\hat{t}_{yH}$	500	25	0.00544	0.00007	0.70341	0.49482
	500	50	0.00780	0.00010	0.70084	0.49123
	500	100	0.01175	0.00015	0.69933	0.48920
	500	150	0.01507	0.00019	0.69762	0.48691
	500	200	0.01886	0.00024	0.69416	0.48222
	500	250	0.01473	0.00019	0.69135	0.47819
	500	300	0.01490	0.00019	0.69021	0.47661
	500	350	0.01513	0.00020	0.68872	0.47456
$\hat{t}_{yr2}$	500	25	11.84579	0.15324	0.54556	140.62046
	500	50	11.85133	0.15331	0.50726	140.71128
	500	100	11.85569	0.15337	0.49113	140.79864
	500	150	11.85693	0.15338	0.48946	140.82627
	500	200	11.85806	0.15340	0.48667	140.85053
	500	250	11.85437	0.15335	0.48405	140.76044
	500	300	11.85400	0.15335	0.48271	140.75035
	500	350	11.85378	0.15334	0.48273	140.74510

Figura 6.23: Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1000$ e $N_{ab} = N_B = 2500$ ($U_A \subset U_B$)

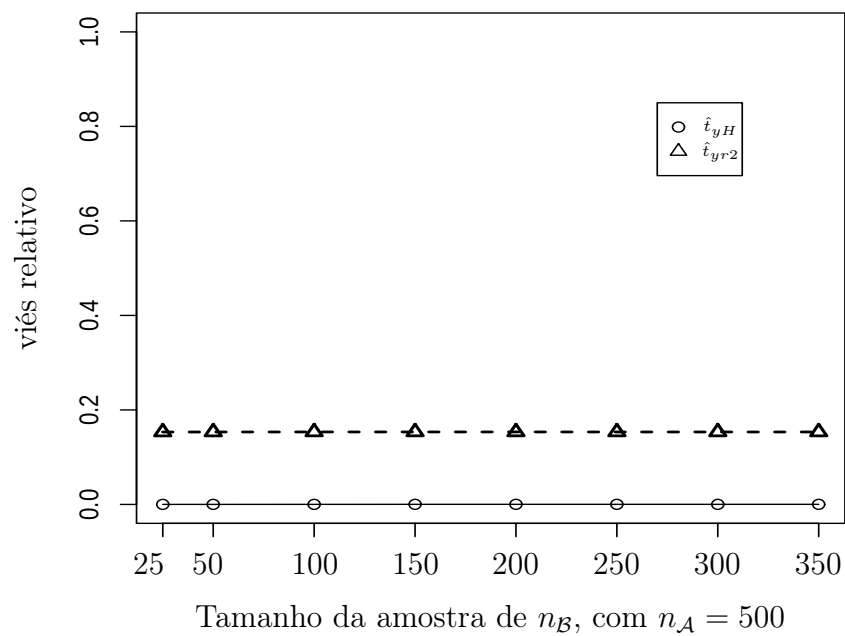
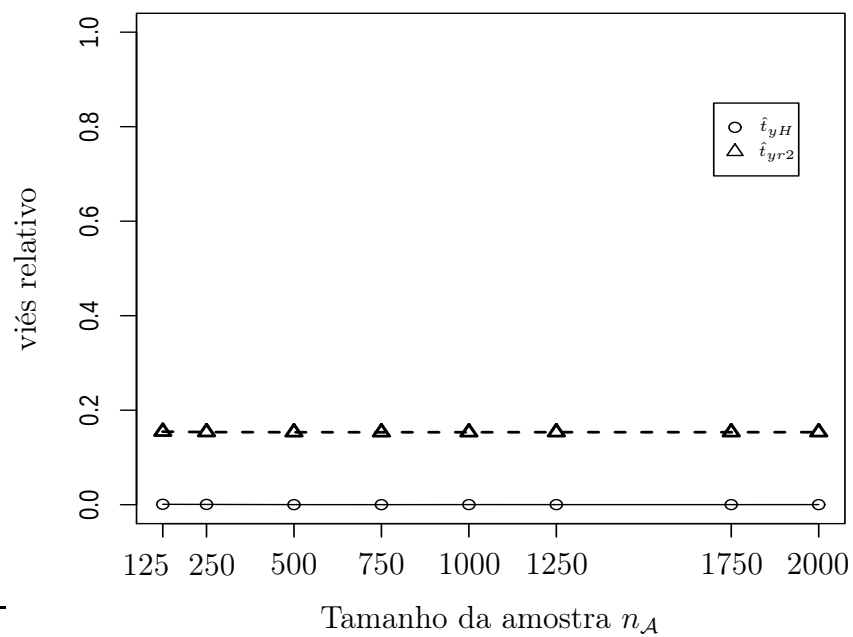


Figura 6.24: Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 1000$ e $N_{ab} = N_B = 2500$ ($U_A \subset U_B$)

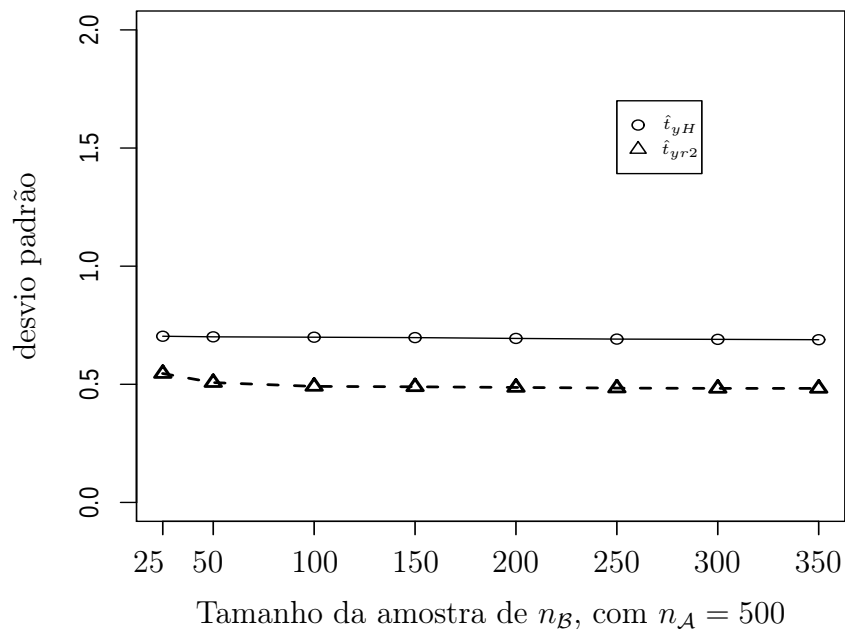
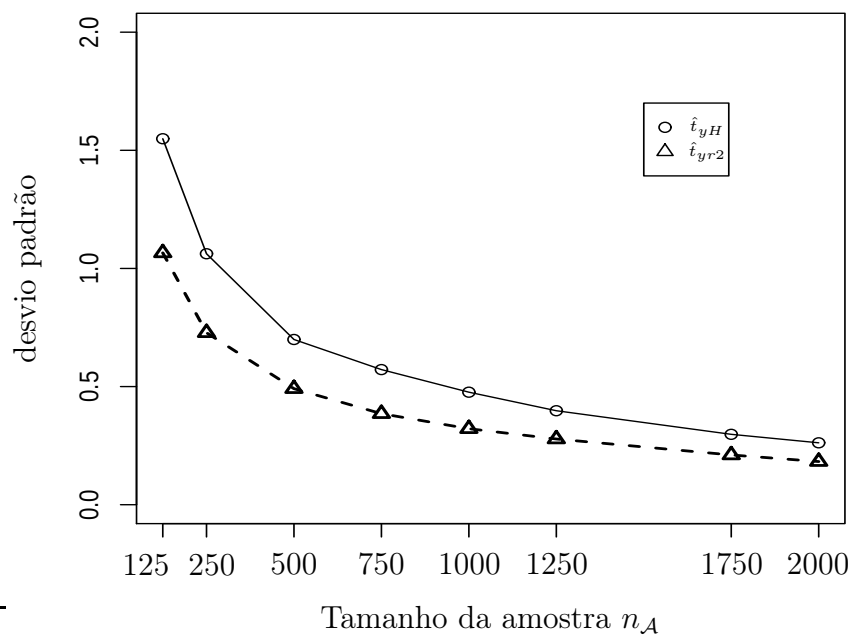


Tabela 6.12: Avaliação dos estimadores sob o cenário 2 para o plano AAS, onde $U_B \subset U_A$ (dados populacionais: $N_a = 750$ e $N_{ab} = 3000$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.03309	0.00051	1.28471	1.65158
	250	50	-0.00685	0.00011	0.89631	0.80342
	500	100	-0.00224	0.00003	0.60567	0.36684
	750	150	0.00184	0.00003	0.46401	0.21531
	1000	200	0.00127	0.00002	0.39738	0.15791
	1250	250	-0.00086	0.00001	0.33351	0.11123
	1750	300	-0.00742	0.00011	0.25770	0.06646
	2000	350	-0.00735	0.00011	0.22703	0.05160
$\hat{t}_{yr2}$	125	25	26.93758	0.41570	0.89491	726.43421
	250	50	26.89471	0.41504	0.65015	723.74788
	500	100	26.91489	0.41535	0.42950	724.59583
	750	150	26.89906	0.41510	0.34660	723.67950
	1000	200	26.88448	0.41488	0.28518	722.85679
	1250	250	26.88314	0.41486	0.25162	722.76675
	1750	300	26.88720	0.41492	0.20887	722.96491
	2000	350	26.88581	0.41490	0.19356	722.88415
$\hat{t}_{yH}$	500	25	-0.00781	0.00012	0.60465	0.36566
	500	50	-0.00860	0.00013	0.60200	0.36248
	500	100	-0.00224	0.00003	0.60567	0.36684
	500	150	-0.00130	0.00002	0.60303	0.36365
	500	200	-0.00609	0.00009	0.60304	0.36369
	500	250	-0.00631	0.00010	0.59995	0.35998
	500	300	-0.00595	0.00009	0.59480	0.35383
	500	350	-0.00582	0.00009	0.59232	0.35088
$\hat{t}_{yr2}$	500	25	26.88630	0.41491	0.64954	723.29483
	500	50	26.89323	0.41501	0.51894	723.51534
	500	100	26.91489	0.41535	0.42950	724.59583
	500	150	26.91393	0.41533	0.40042	724.51992
	500	200	26.90473	0.41519	0.38492	724.01277
	500	250	26.90461	0.41519	0.37571	723.99917
	500	300	26.90504	0.41519	0.37736	724.02351
	500	350	26.90512	0.41520	0.37423	724.02544

Figura 6.25: Viés Relativo de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 750$ e $N_{ab} = N_B = 3000$ ($U_A \subset U_B$)

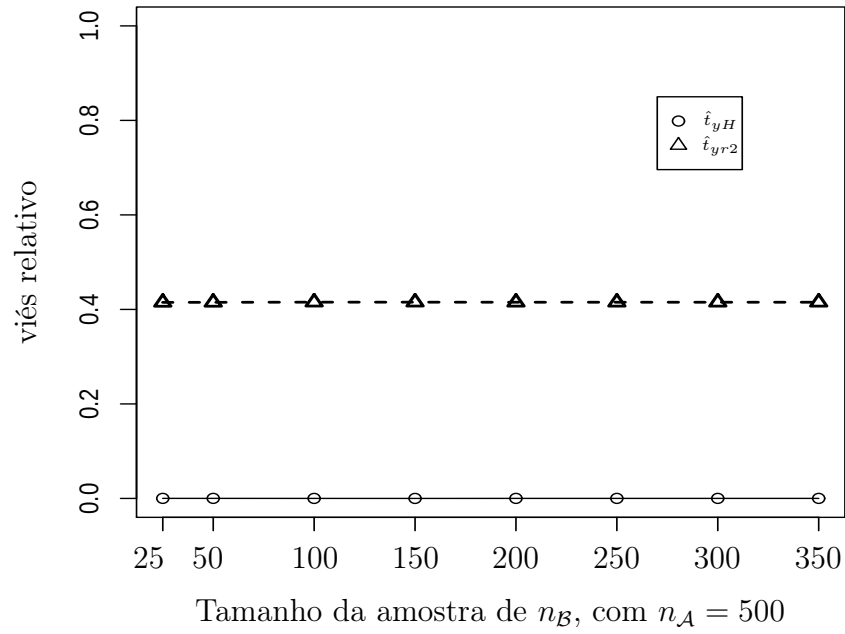
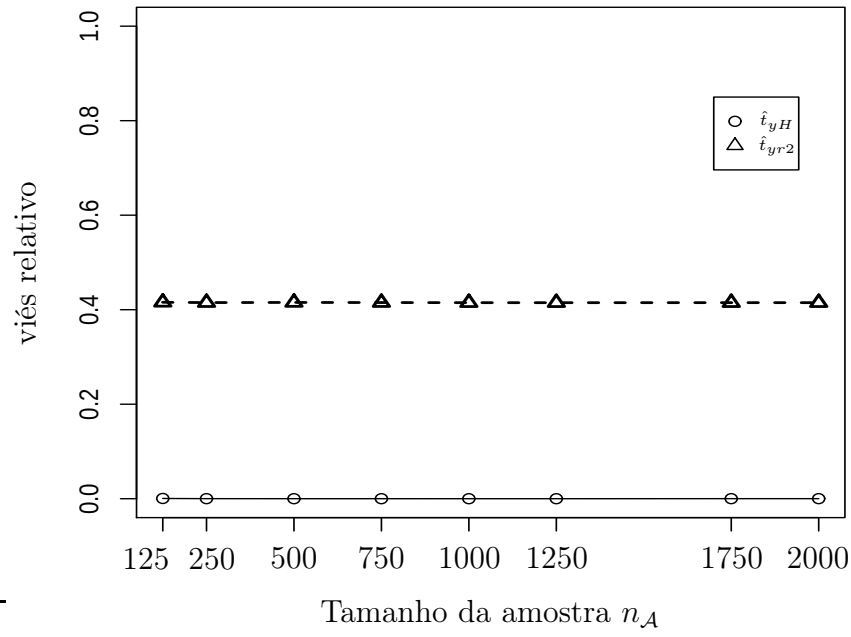


Figura 6.26: Desvio Padrão de $\hat{t}_{yH}$ e $\hat{t}_{yr2}$, para $N_a = 750$ e $N_{ab} = N_B = 3000$ ($U_A \subset U_B$)

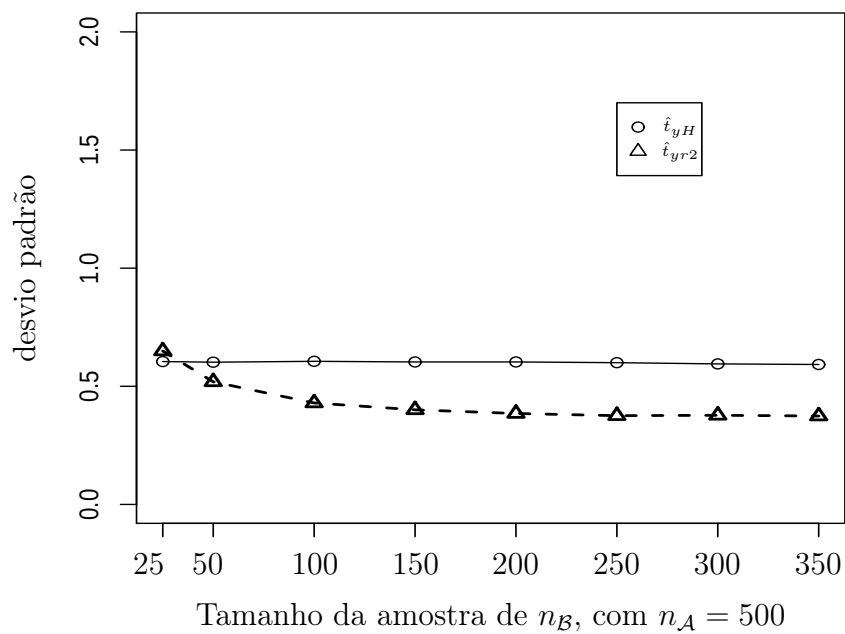
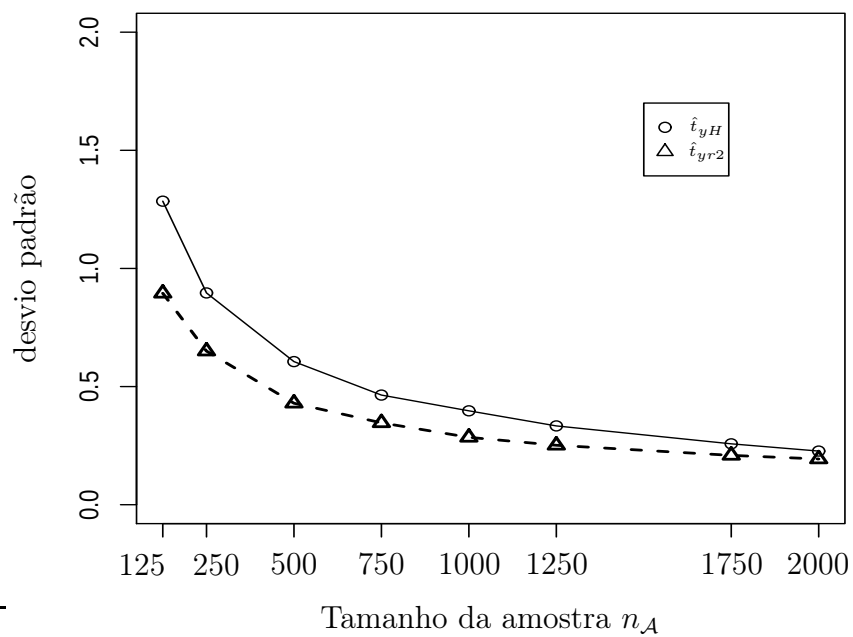


Tabela 6.13: Avaliação dos estimadores razão separados sob os cenários 2 e 3, para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para j e w)

Cenário	Estimador	$n_{\mathcal{A}(j)}=n_{\mathcal{B}(w)}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	0.00897	0.00006	0.66954	0.44836
		250	-0.00540	0.00004	0.45273	0.20499
		500	-0.01870	0.00013	0.29785	0.08907
		750	0.00522	0.00004	0.22702	0.05157
		1000	0.00183	0.00001	0.18187	0.03308
		1250	-0.00572	0.00004	0.14387	0.02073
	$\hat{t}_{yr1}$	125	-0.00598	0.00004	0.15984	0.02558
		250	-0.00341	0.00002	0.10942	0.01198
		500	-0.00689	0.00005	0.07212	0.00525
		750	-0.00058	0.00000	0.05424	0.00294
		1000	-0.00105	0.00001	0.04491	0.00202
		1250	0.00112	0.00001	0.03594	0.00129
	$\hat{t}_{yr2}$	125	11.89780	0.08029	2.11428	146.02776
		250	12.43554	0.08392	1.20691	156.09926
		500	12.67974	0.08557	0.72912	161.30749
		750	12.79726	0.08636	0.53247	164.05334
		1000	12.78559	0.08628	0.42774	163.65434
		1250	12.83129	0.08659	0.35175	164.76584
3	$\hat{t}_{yFB}$	125	0.00991	0.00007	1.13739	1.29375
		250	-0.04880	0.00033	0.78718	0.62204
		500	-0.03225	0.00022	0.50690	0.25799
		750	0.01737	0.00012	0.38038	0.14499
		1000	-0.01052	0.00007	0.30395	0.09249
		1250	0.00218	0.00001	0.25018	0.06259
	$\hat{t}_{yr3}$	125	0.12702	0.00086	1.32415	1.76952
		250	0.09556	0.00064	0.78741	0.62914
		500	0.03164	0.00021	0.49547	0.24649
		750	-0.00559	0.00004	0.38096	0.14517
		1000	0.02096	0.00004	0.31524	0.09981
		1250	-0.00436	0.00003	0.24535	0.06022
	$\hat{t}_{yr4}$	125	7.50563	0.05065	3.02664	65.49498
		250	7.68334	0.05185	2.20354	63.88929
		500	7.45899	0.05034	1.37569	57.52911
		750	7.46427	0.05037	1.02843	56.77298
		1000	7.45121	0.05028	0.85919	56.25868
		1250	7.40002	0.04994	0.67170	55.21146

Figura 6.27: Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yH}$, $\hat{t}_{yrs1}$ e $\hat{t}_{yrs2}$, com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 2.

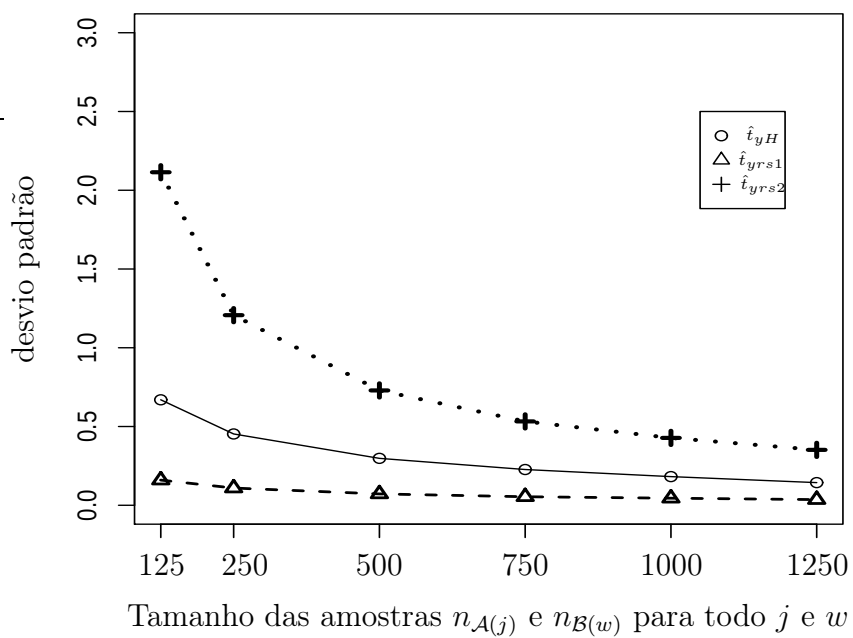
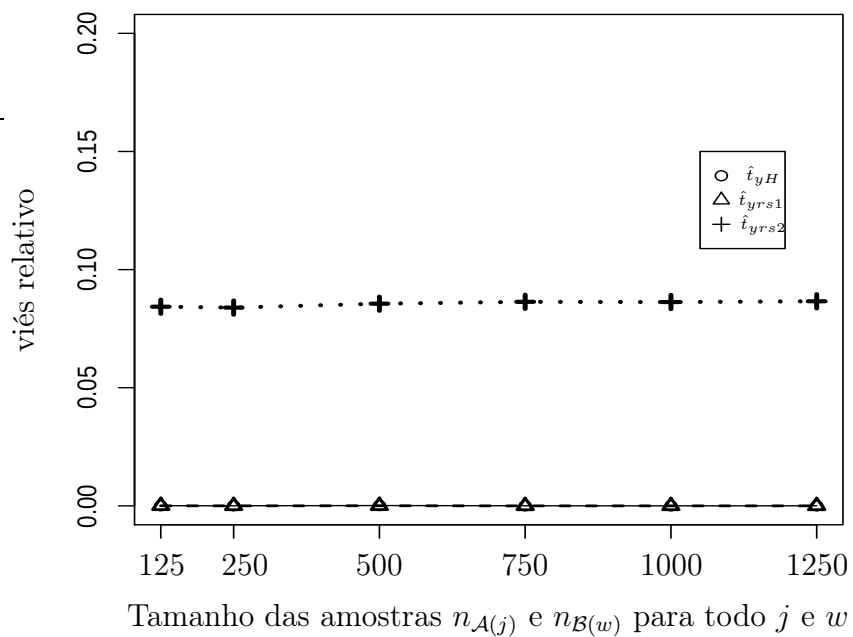


Figura 6.28: Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yFB}$, $\hat{t}_{yr3s}$ e $\hat{t}_{yr4s}$ com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 3.

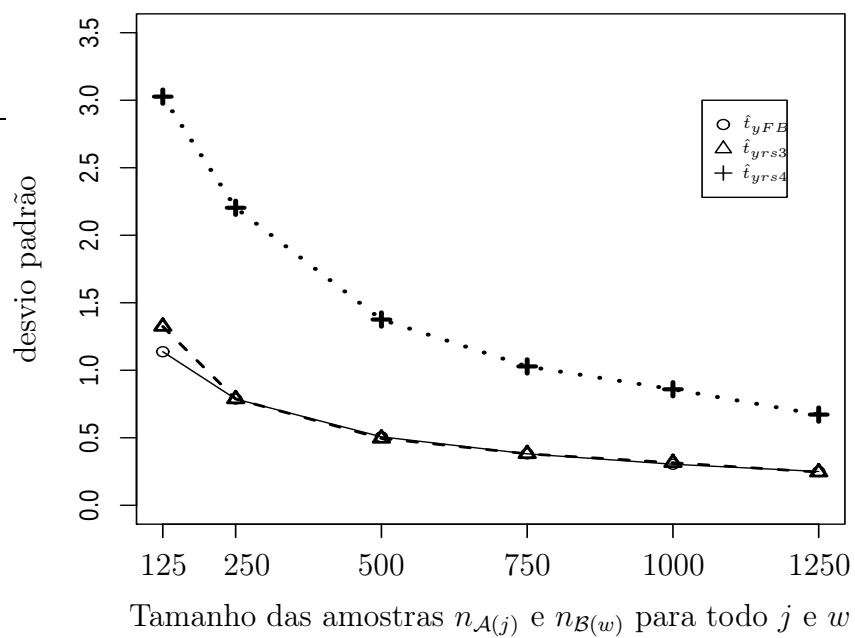
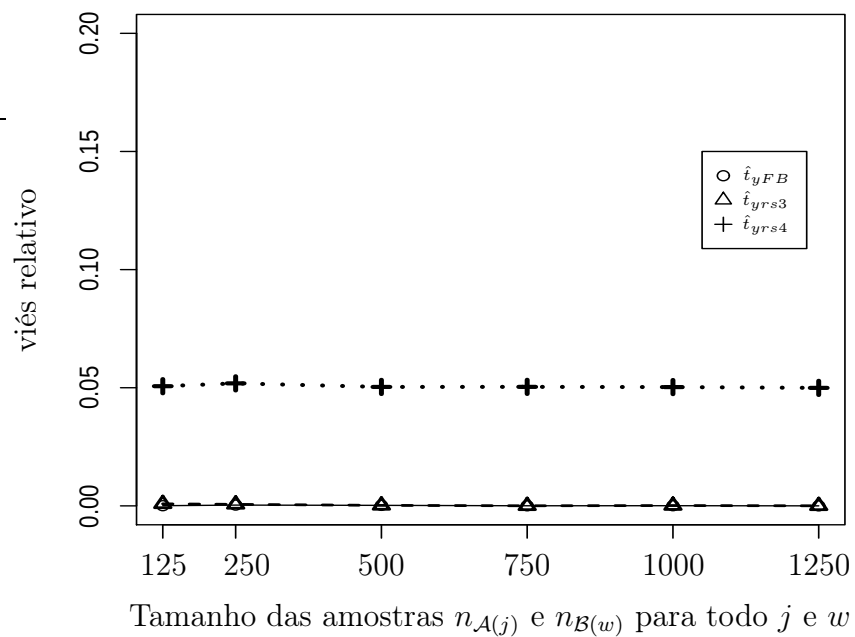


Tabela 6.14: Avaliação dos estimadores razão separados sob os cenários 2 e 3, para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para j e w)

Cenário	Estimador	$n_{\mathcal{A}(j)}=n_{\mathcal{B}(w)}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	0.01492	0.00011	0.62682	0.39313
		250	0.00150	0.00001	0.42976	0.18470
		500	0.00398	0.00003	0.28506	0.08128
		750	-0.00817	0.00006	0.21181	0.04493
		1000	0.00458	0.00003	0.18399	0.03387
		1250	-0.00208	0.00001	0.15252	0.02327
	$\hat{t}_{yr1}$	125	-0.03664	0.00026	0.84001	0.70695
		250	-0.00394	0.00003	0.11335	0.01286
		500	-0.00323	0.00002	0.07473	0.00560
		750	-0.00306	0.00002	0.05608	0.00315
		1000	-0.00230	0.00002	0.04643	0.00216
		1250	-0.00108	0.00001	0.03893	0.00152
	$\hat{t}_{yr2}$	125	8.91797	0.06416	9.33430	166.65943
		250	10.75693	0.07739	5.69087	148.09754
		500	11.64911	0.08381	3.42298	147.41847
		750	12.08742	0.08697	2.64183	153.08490
		1000	12.31317	0.08859	2.12122	156.11382
		1250	12.17692	0.08761	1.83447	151.64270
3	$\hat{t}_{yFB}$	125	-0.05375	0.00039	1.44772	2.09880
		250	-0.02075	0.00015	1.01686	1.03443
		500	0.01290	0.00009	0.63073	0.39799
		750	-0.01654	0.00012	0.48520	0.23569
		1000	0.02248	0.00016	0.42322	0.17962
		1250	-0.00711	0.00005	0.34826	0.12134
	$\hat{t}_{yr3}$	125	0.11003	0.00079	1.45920	2.14138
		250	0.04301	0.00031	1.03647	1.07613
		500	-0.00036	0.00000	0.66352	0.44026
		750	0.01310	0.00009	0.50476	0.25495
		1000	-0.01853	0.00003	0.43024	0.18545
		1250	0.00745	0.00005	0.34933	0.12209
	$\hat{t}_{yr4}$	125	17.42497	0.12537	1.57561	306.11223
		250	17.58905	0.12655	1.10970	310.60597
		500	17.67457	0.12716	0.71003	312.89463
		750	17.65536	0.12703	0.52799	311.99049
		1000	17.64617	0.12696	0.43543	311.57691
		1250	17.69206	0.12729	0.39330	313.16372

Figura 6.29: Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yH}$, $\hat{t}_{yrs1s}$ e $\hat{t}_{yrs2s}$ com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 2.

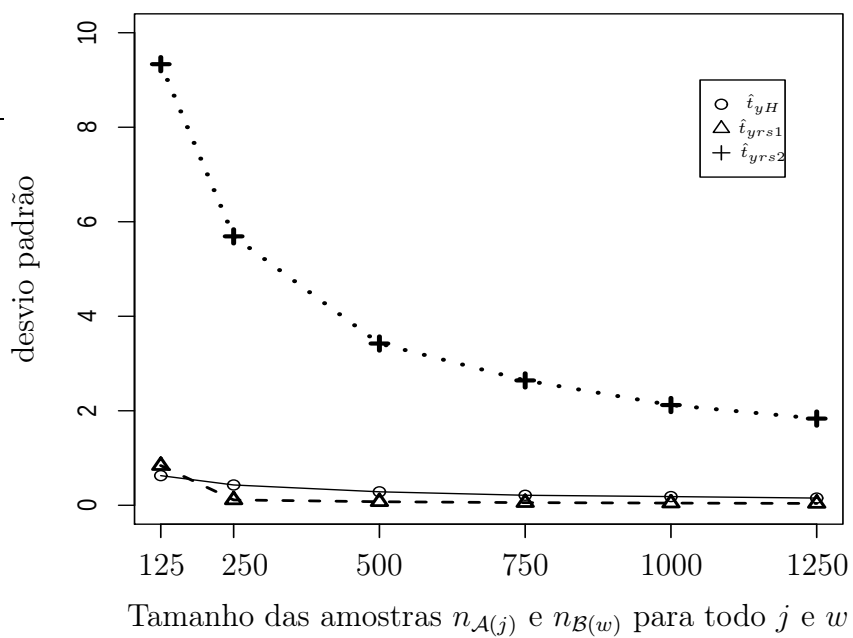
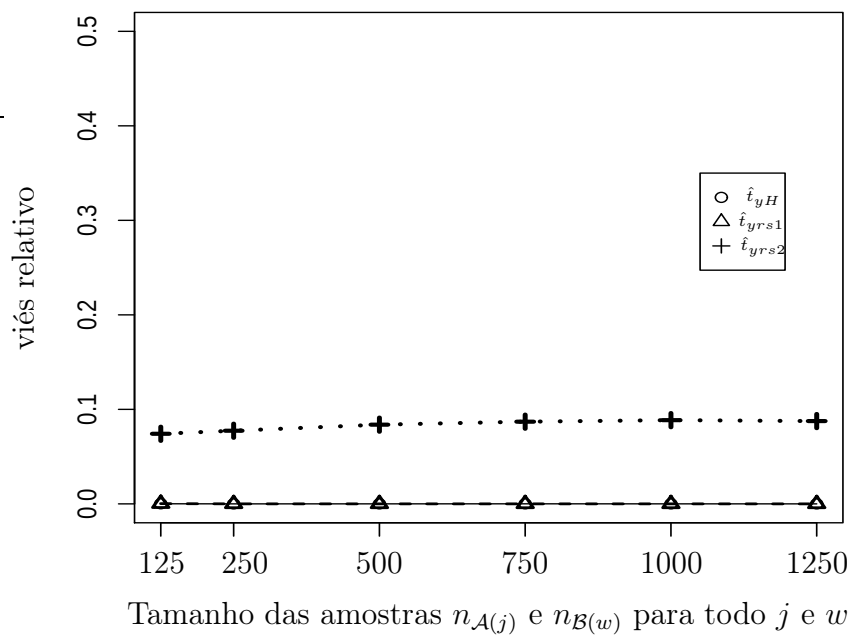


Figura 6.30: Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yFB}$, $\hat{t}_{yrs3}$ e $\hat{t}_{yrs4}$ com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 3.

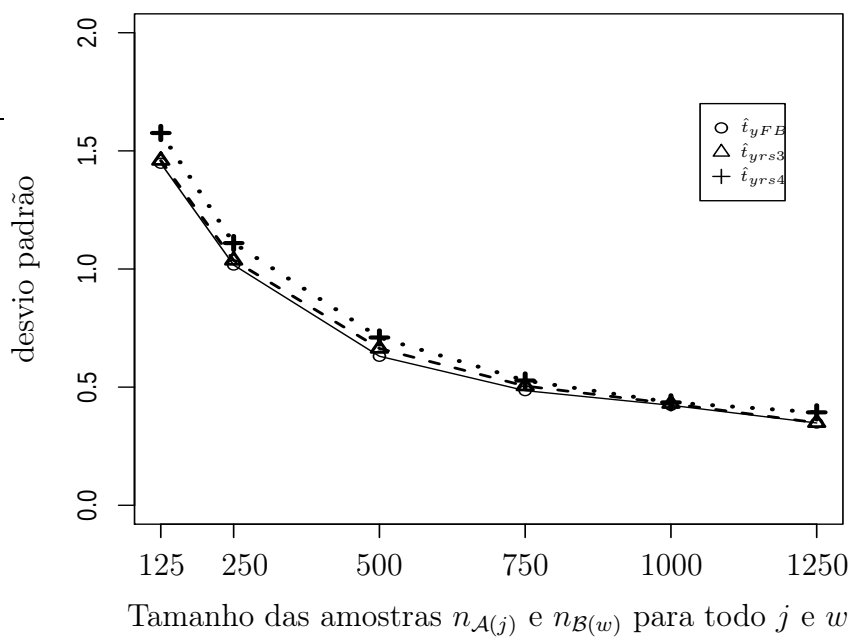
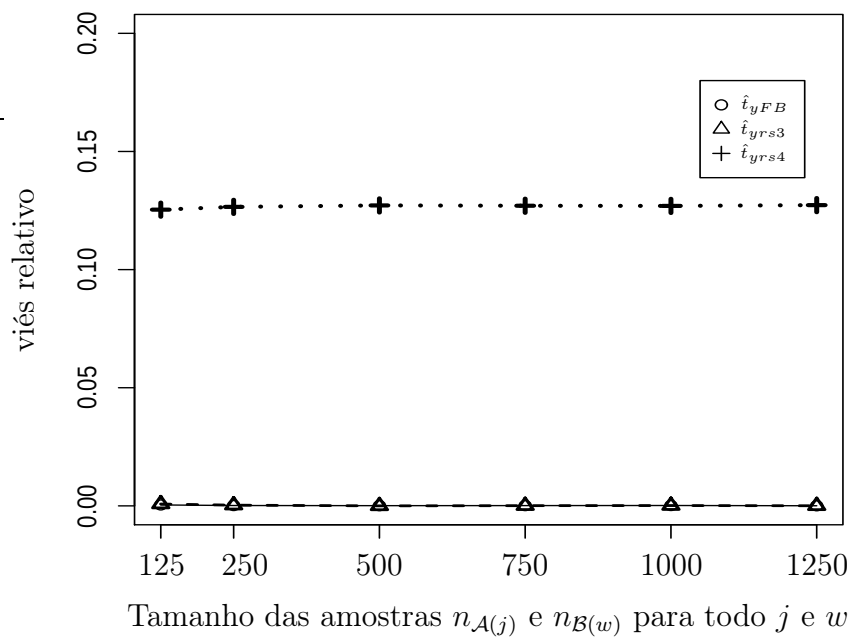


Tabela 6.15: Avaliação dos estimadores razão separados sob os cenários 2 e 3, para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para j e w)

Cenário	Estimador	$n_{\mathcal{A}(j)}=n_{\mathcal{B}(w)}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	-0.00218	0.00002	0.64829	0.42029
		250	-0.00168	0.00001	0.44892	0.20153
		500	0.01282	0.00010	0.29832	0.08916
		750	-0.00887	0.00007	0.22585	0.05109
		1000	-0.00232	0.00002	0.18569	0.03448
		1250	0.00128	0.00001	0.15813	0.02501
	$\hat{t}_{yr1}$	125	-0.11243	0.00087	5.42635	29.45787
		250	-0.10892	0.00084	3.81994	14.60379
		500	0.01224	0.00009	0.88710	0.78710
		750	0.00165	0.00001	0.06368	0.00406
		1000	0.00077	0.00001	0.04339	0.00188
		1250	-0.00035	0.00000	0.03150	0.00099
	$\hat{t}_{yr2}$	125	-12.30674	0.09506	10.74869	266.99012
		250	-14.87547	0.11490	9.17279	305.41950
		500	-17.46426	0.13490	7.43531	360.28432
		750	-18.14833	0.14018	6.49022	371.48468
		1000	-19.03238	0.14701	5.57444	393.30596
		1250	-19.56520	0.15113	5.16568	409.48137
3	$\hat{t}_{yFB}$	125	0.03821	0.00030	1.73968	3.02794
		250	-0.04511	0.00035	1.22003	1.49050
		500	-0.00573	0.00004	0.77757	0.60465
		750	0.00079	0.00001	0.62307	0.38821
		1000	-0.02653	0.00020	0.50646	0.25720
		1250	0.00981	0.00008	0.43766	0.19164
	$\hat{t}_{yr3}$	125	-0.00487	0.00044	1.58606	2.51562
		250	0.05257	0.00041	1.09732	1.20688
		500	0.02659	0.00021	0.70179	0.49322
		750	-0.00791	0.00006	0.56432	0.31852
		1000	0.02807	0.00006	0.45646	0.20914
		1250	-0.00743	0.00002	0.39378	0.15512
	$\hat{t}_{yr4}$	125	19.47766	0.15045	1.97498	383.27965
		250	19.17587	0.14812	1.31685	369.44823
		500	19.04420	0.14710	0.83164	363.37304
		750	18.94065	0.14630	0.69682	359.23371
		1000	18.96862	0.14652	0.55016	360.11109
		1250	18.91370	0.14609	0.46940	357.94832

Figura 6.31: Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yH}$, $\hat{t}_{yrs1}$ e $\hat{t}_{yrs2}$ com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 2.

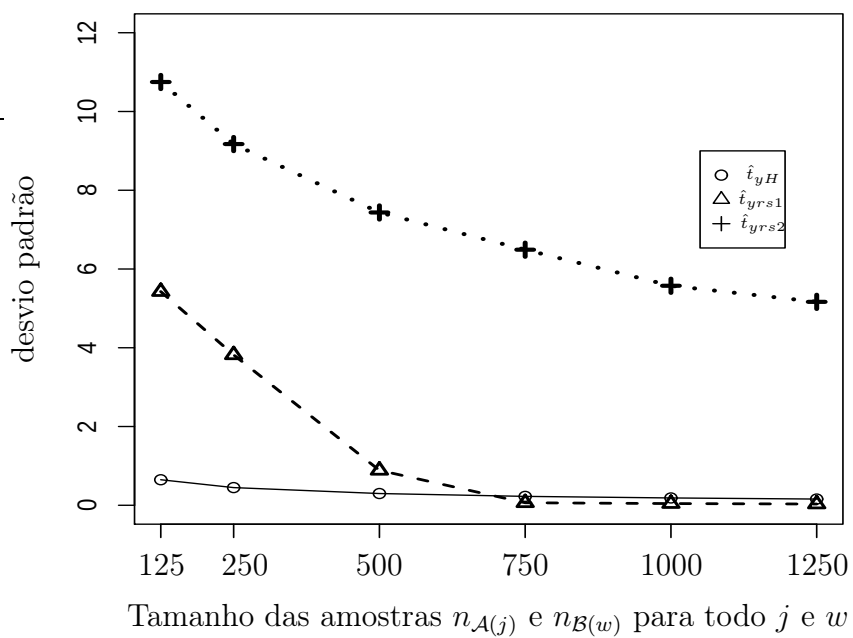
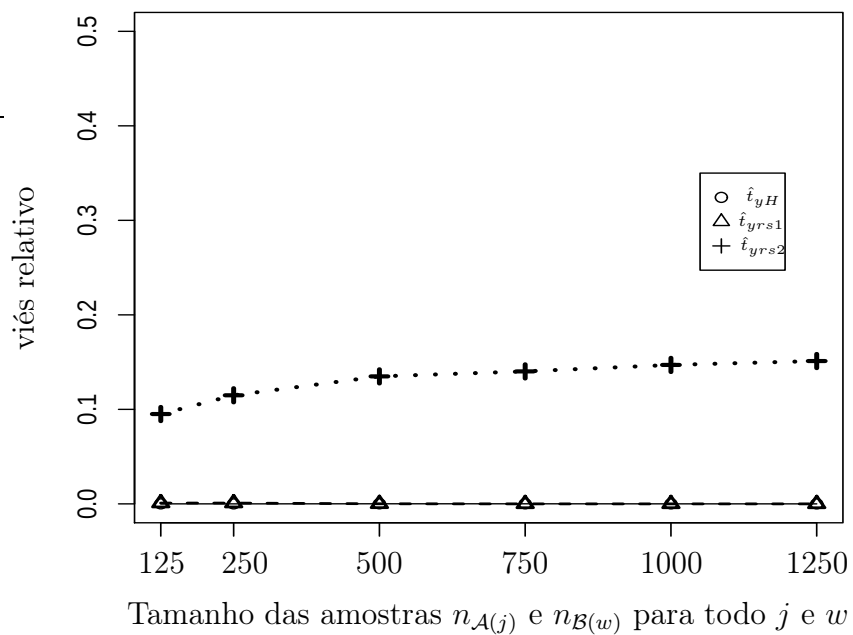


Figura 6.32: Viés Relativo e Desvio Padrão dos estimadores separados $\hat{t}_{yFB}$, $\hat{t}_{yrs3}$ e $\hat{t}_{yrs4}$ com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 3.

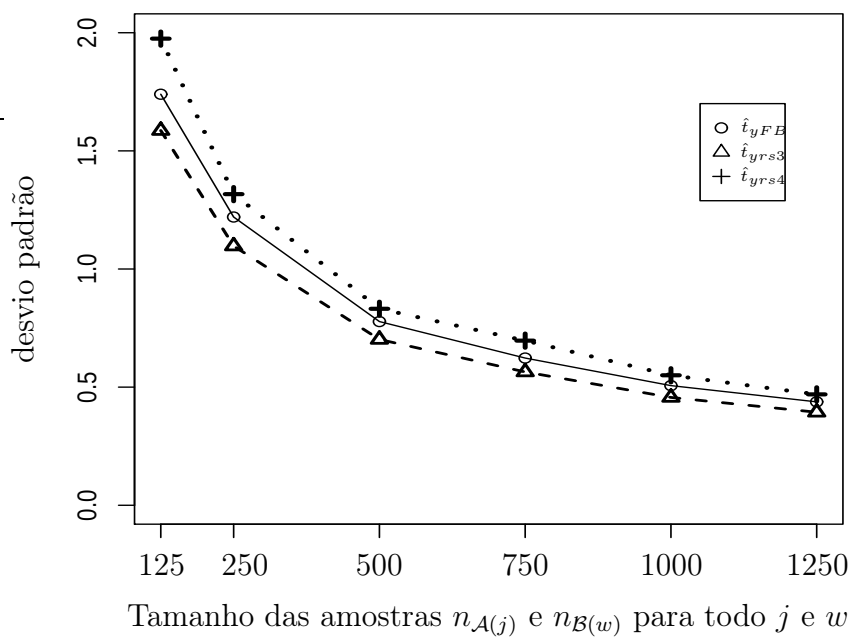
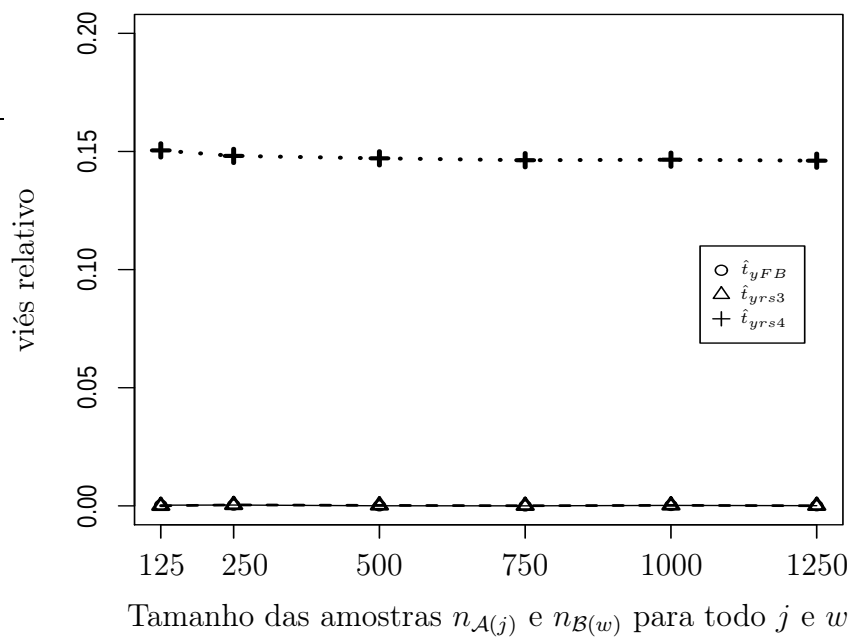


Tabela 6.16: Avaliação dos estimadores razão separados sob os cenários 2 e 3 para o plano AAE, em que $n_{\mathcal{A}(j)} = 1000$ (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para j e w)

Cenário	Estimador	$n_{\mathcal{A}}$	$n_{\mathcal{B}}$	Estatística			
				Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	1000	200	0.00685	0.00005	0.18603	0.03465
		1000	300	-0.01156	0.00008	0.17628	0.03121
		1000	400	-0.01097	0.00007	0.19184	0.03692
		1000	500	0.00123	0.00001	0.18839	0.03549
		1000	700	0.00256	0.00002	0.18291	0.03346
		1000	900	-0.00923	0.00006	0.17619	0.03113
	$\hat{t}_{yr1}$	1000	200	3.80899	0.02570	82.87477	6882.73616
		1000	300	0.23561	0.00159	11.41958	130.46234
		1000	400	-0.00558	0.00004	0.04714	0.00225
		1000	500	-0.00082	0.00001	0.04346	0.00189
		1000	700	0.00012	0.00000	0.04859	0.00236
		1000	900	0.00114	0.00001	0.04618	0.00213
	$\hat{t}_{yr2}$	1000	200	15.18026	0.10244	0.49206	230.68235
		1000	300	14.22956	0.09603	0.41761	202.65471
		1000	400	13.70046	0.09246	0.39153	187.85595
		1000	500	13.40086	0.09043	0.39188	179.73662
		1000	700	13.03416	0.08796	0.41822	170.06426
		1000	900	12.88445	0.08695	0.41726	166.18309
3	$\hat{t}_{yFB}$	1000	200	0.01428	0.00010	0.38782	0.15061
		1000	300	-0.00829	0.00006	0.36906	0.13627
		1000	400	-0.01304	0.00009	0.36228	0.13142
		1000	500	0.01055	0.00007	0.35722	0.12772
		1000	700	-0.00986	0.00007	0.32885	0.10824
		1000	900	0.00466	0.00003	0.31739	0.10076
	$\hat{t}_{yr3}$	1000	200	-0.00860	0.00006	0.46040	0.21204
		1000	300	0.00105	0.00001	0.42881	0.18388
		1000	400	0.00703	0.00005	0.40695	0.16565
		1000	500	-0.00410	0.00003	0.38485	0.14812
		1000	700	0.02139	0.00004	0.36505	0.13372
		1000	900	-0.01107	0.00007	0.31932	0.10209
	$\hat{t}_{yr4}$	1000	200	8.32854	0.05620	1.40815	71.34748
		1000	300	8.04194	0.05427	1.12868	65.94672
		1000	400	7.92571	0.05349	1.03382	63.88562
		1000	500	7.83519	0.05288	0.93504	62.26449
		1000	700	7.69335	0.05192	0.88977	59.97936
		1000	900	7.49259	0.05056	0.80049	56.77967

Tabela 6.17: Avaliação dos estimadores razão separados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 2000$ e $N_{ab(jw)} = 500$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.03612	0.00014	1.18301	1.40082
	250	50	-0.01184	0.00005	0.77944	0.60767
	500	100	-0.02093	0.00008	0.52691	0.27808
	750	150	-0.01165	0.00004	0.40880	0.16725
	1000	200	0.00790	0.00003	0.31963	0.10223
	1250	250	-0.00011	0.00000	0.25578	0.06542
	1500	300	0.00823	0.00003	0.21967	0.04832
	1750	350	-0.00539	0.00002	0.17375	0.03022
$\hat{t}_{yrs2}$	125	25	22.09601	0.08425	3.35497	499.48964
	250	50	23.03532	0.08783	2.07045	534.91259
	500	100	23.53856	0.08975	1.13012	555.34097
	750	150	23.64489	0.09015	0.87765	559.85119
	1000	200	23.73248	0.09049	0.69320	563.71138
	1250	250	23.74899	0.09055	0.56905	564.33818
	1500	300	23.76659	0.09062	0.44387	565.04776
	1750	350	23.78980	0.09070	0.36338	566.08640
$\hat{t}_{yH}$	500	25	-0.02043	0.00008	0.52913	0.28040
	500	50	0.02543	0.00010	0.53189	0.28355
	500	100	0.02612	0.00010	0.53099	0.28263
	500	150	0.01534	0.00006	0.52067	0.27133
	500	200	-0.01446	0.00006	0.54471	0.29692
	500	250	0.00414	0.00002	0.51791	0.26824
	500	300	0.00026	0.00000	0.54541	0.29747
	500	350	-0.01355	0.00005	0.54034	0.29215
$\hat{t}_{yrs2}$	500	25	25.99478	0.09911	1.15355	677.05950
	500	50	24.28701	0.09260	1.17669	591.24364
	500	100	23.45129	0.08941	1.17062	551.33330
	500	150	23.20193	0.08846	1.20356	539.77792
	500	200	23.12675	0.08818	1.19181	536.26702
	500	250	22.97734	0.08761	1.23186	529.47556
	500	300	22.97990	0.08762	1.21910	529.56213
	500	350	22.92499	0.08741	1.21375	527.02825

Tabela 6.18: Avaliação dos estimadores razão separados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 1750$ e $N_{ab(jw)} = 1000$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.02595	0.00012	1.09840	1.20716
	250	50	-0.03032	0.00014	0.73229	0.53717
	500	100	-0.00160	0.00001	0.48446	0.23470
	750	150	-0.00728	0.00003	0.37158	0.13812
	1000	200	0.00119	0.00001	0.31318	0.09809
	1250	250	-0.00069	0.00000	0.25721	0.06616
	1500	300	0.00491	0.00002	0.21593	0.04665
	1750	350	-0.00333	0.00001	0.17867	0.03193
$\hat{t}_{yrs2}$	125	25	12.84211	0.05747	15.45505	403.77833
	250	50	15.24468	0.06823	11.22928	358.49688
	500	100	17.63791	0.07894	7.12942	361.92439
	750	150	18.55582	0.08304	4.94603	368.78174
	1000	200	19.12160	0.08558	3.88105	380.69819
	1250	250	19.80962	0.08865	3.18099	402.53986
	1500	300	20.20096	0.09041	2.63336	415.01331
	1750	350	20.93346	0.09368	2.29060	443.45676
$\hat{t}_{yH}$	500	25	0.00196	0.00001	0.49606	0.24608
	500	50	-0.00046	0.00000	0.48573	0.23593
	500	100	-0.00207	0.00001	0.49929	0.24930
	500	150	-0.00141	0.00001	0.48955	0.23966
	500	200	-0.00963	0.00004	0.48526	0.23557
	500	250	-0.00322	0.00001	0.47267	0.22343
	500	300	0.00214	0.00001	0.48942	0.23953
	500	350	-0.01574	0.00007	0.46601	0.21741
$\hat{t}_{yrs2}$	500	25	28.62167	0.12809	7.83405	880.57236
	500	50	20.73767	0.09281	7.36984	484.36558
	500	100	17.59178	0.07873	7.01204	358.63940
	500	150	17.09688	0.07651	6.16340	330.29084
	500	200	16.48782	0.07379	6.26404	311.08647
	500	250	16.46106	0.07367	6.19209	309.30846
	500	300	16.03003	0.07174	6.25340	296.06690
	500	350	15.71817	0.07034	6.17815	285.23041

Tabela 6.19: Avaliação dos estimadores razão separados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 1500$ e $N_{ab(jw)} = 500$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.00335	0.00002	0.99466	0.98936
	250	50	-0.00862	0.00005	0.66918	0.44788
	500	100	-0.01055	0.00006	0.45404	0.20627
	750	150	0.02715	0.00014	0.35123	0.12410
	1000	200	0.00773	0.00004	0.29273	0.08575
	1250	250	-0.00193	0.00001	0.24972	0.06236
	1500	300	0.00718	0.00004	0.20589	0.04244
	1750	350	0.00090	0.00000	0.17311	0.02997
$\hat{t}_{yrs2}$	125	25	-13.51840	0.07064	15.07615	410.03757
	250	50	-16.31079	0.08523	11.75305	404.17628
	500	100	-19.08938	0.09975	7.88809	426.62627
	750	150	-19.74655	0.10318	5.61540	421.45889
	1000	200	-19.09165	0.09976	3.93216	379.95281
	1250	250	-17.09508	0.08933	2.45558	298.27147
	1500	300	-14.23659	0.07439	1.77015	205.81397
	1750	350	-10.47935	0.05476	1.25761	111.39843
$\hat{t}_{yH}$	500	25	-0.01063	0.00006	0.46971	0.22074
	500	50	-0.00757	0.00004	0.47296	0.22375
	500	100	-0.00072	0.00000	0.47322	0.22394
	500	150	0.04306	0.00023	0.43267	0.18906
	500	200	-0.01776	0.00009	0.46582	0.21731
	500	250	-0.00447	0.00002	0.45032	0.20281
	500	300	-0.02493	0.00013	0.44985	0.20299
	500	350	-0.02002	0.00010	0.44082	0.19472
$\hat{t}_{yrs2}$	500	25	16.50814	0.08626	7.39030	327.13514
	500	50	-2.33071	0.01218	5.57671	36.53187
	500	100	-18.97758	0.09917	8.25903	428.36002
	500	150	-26.58265	0.13891	10.63840	819.81282
	500	200	-26.84511	0.14028	11.31568	848.7044
	500	250	-25.64142	0.13399	11.09841	780.65700
	500	300	-23.81935	0.12447	11.28536	694.72107
	500	350	-23.43571	0.12246	11.10076	672.45928

Tabela 6.20: Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$)

Cenário	Estimador	$n_{\mathcal{A}(j)}=n_{\mathcal{B}(w)}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	0.00105	0.00002	0.28789	0.08288
		250	-0.00235	0.00004	0.19320	0.03733
		500	-0.00865	0.00015	0.12781	0.01641
		750	0.00182	0.00003	0.09666	0.00935
		1000	-0.00044	0.00001	0.07746	0.00600
		1250	-0.00132	0.00002	0.06276	0.00394
	$\hat{t}_{yr1}$	125	-0.00407	0.00007	0.11294	0.01277
		250	-0.00241	0.00004	0.07433	0.00553
		500	-0.00501	0.00008	0.05061	0.00259
		750	-0.00029	0.00000	0.03760	0.00141
		1000	-0.00077	0.00001	0.02988	0.00089
		1250	0.00132	0.00002	0.02394	0.00057
	$\hat{t}_{yr2}$	125	5.73737	0.09675	0.23936	32.97476
		250	5.73886	0.09677	0.16041	32.96030
		500	5.73914	0.09678	0.10434	32.94859
		750	5.74559	0.09688	0.07685	33.01774
		1000	5.74090	0.09681	0.06514	32.96212
		1250	5.74628	0.09690	0.05282	33.02253
3	$\hat{t}_{yFB}$	125	0.00721	0.00012	0.46510	0.21637
		250	-0.01607	0.00027	0.32299	0.10458
		500	-0.01311	0.00022	0.20760	0.04327
		750	0.00742	0.00013	0.15401	0.02377
		1000	-0.00497	0.00008	0.12436	0.01549
		1250	0.00212	0.00004	0.10354	0.01073
	$\hat{t}_{yr3}$	125	0.18123	0.00306	10.72646	115.08976
		250	0.06581	0.00111	0.90776	0.82835
		500	0.03049	0.00051	0.22123	0.04987
		750	0.01632	0.00028	0.08380	0.00729
		1000	0.00984	0.00017	0.06220	0.00397
		1250	0.00857	0.00014	0.04994	0.00257
	$\hat{t}_{yr4}$	125	3.78267	0.06378	2.11679	18.78941
		250	4.14594	0.06991	2.24911	22.24727
		500	4.93566	0.08323	1.97341	28.25506
		750	5.45859	0.09205	1.47304	31.96611
		1000	5.64827	0.09524	1.26856	33.51222
		1250	5.82025	0.09814	0.74282	34.42705

Figura 6.33: Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc1}$ e $\hat{t}_{yrc2}$, com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 2.

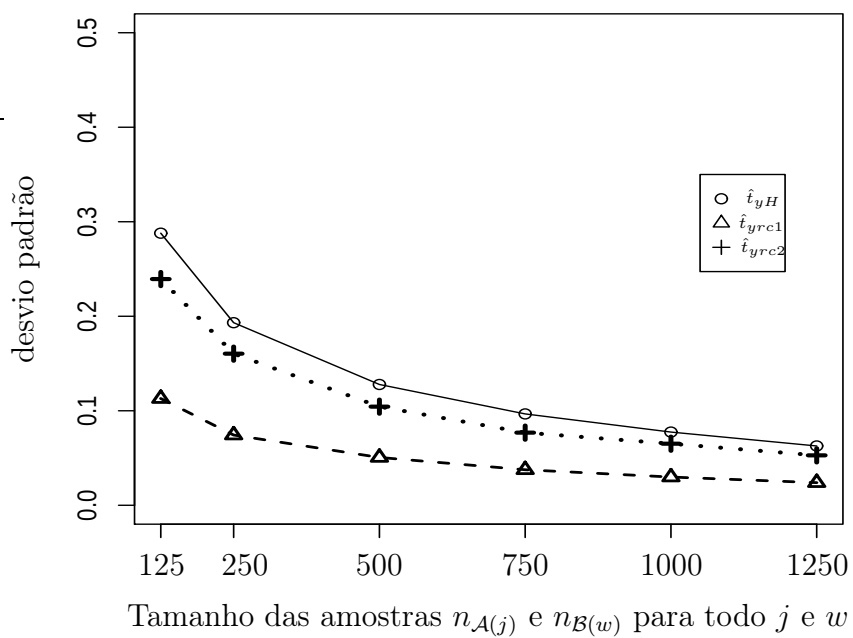
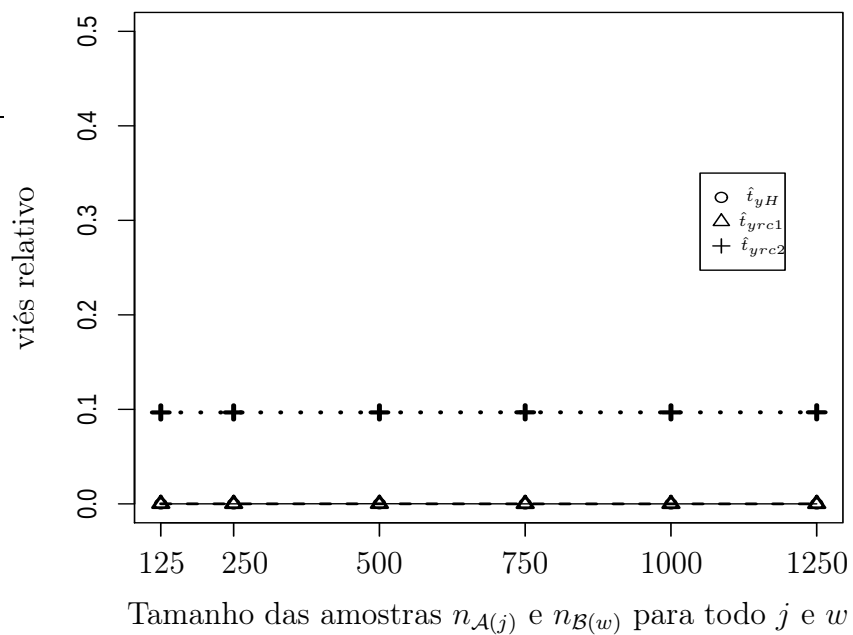


Figura 6.34: Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc3}$ e $\hat{t}_{yrc4}$, com $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para todo estrato j e w , sob o cenário 3.

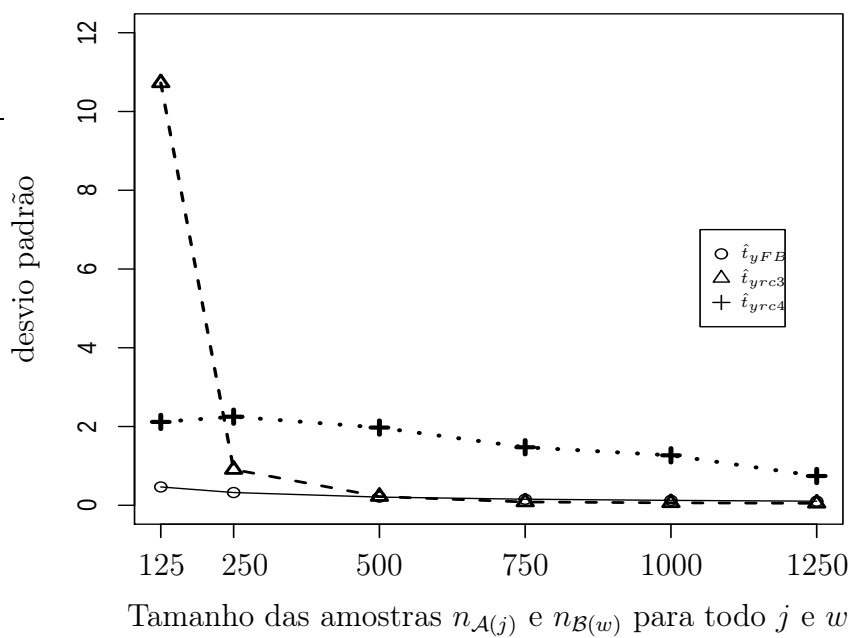
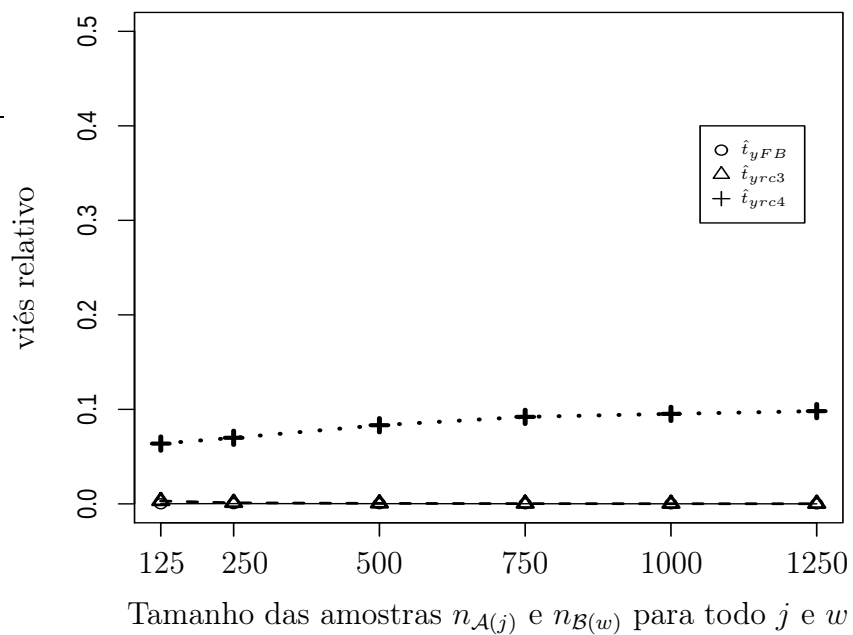


Tabela 6.21: Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 500$)

Cenário	Estimador	$n_{\mathcal{A}(j)}=n_{\mathcal{B}(w)}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	125	0.01359	0.00024	0.26405	0.06991
		250	0.00466	0.00008	0.19116	0.03657
		500	-0.00137	0.00002	0.12899	0.01664
		750	0.00002	0.00000	0.09546	0.00911
		1000	-0.00135	0.00002	0.08026	0.00644
		1250	0.00083	0.00001	0.06346	0.00403
	$\hat{t}_{yr1}$	125	0.00076	0.00001	0.10576	0.01119
		250	-0.00043	0.00001	0.07323	0.00536
		500	0.00003	0.00000	0.05061	0.00256
		750	-0.00034	0.00001	0.03907	0.00153
		1000	0.00013	0.00000	0.03125	0.00098
		1250	0.00006	0.00000	0.02510	0.00063
	$\hat{t}_{yr2}$	125	11.01722	0.19773	0.40299	121.54156
		250	11.00136	0.19745	0.24718	121.09112
		500	11.01065	0.19762	0.17474	121.26501
		750	11.01423	0.19768	0.14097	121.33323
		1000	11.01989	0.19778	0.11271	121.45067
		1250	11.01842	0.19775	0.09455	121.41443
3	$\hat{t}_{yFB}$	125	0.02267	0.00041	0.56609	0.32097
		250	0.00420	0.00008	0.39311	0.15455
		500	-0.00823	0.00015	0.27235	0.0742
		750	-0.00058	0.00001	0.21341	0.04554
		1000	0.00315	0.00006	0.16775	0.02815
		1250	0.00664	0.00012	0.13686	0.01877
	$\hat{t}_{yr3}$	125	0.02152	0.00039	0.35888	0.12926
		250	0.02145	0.00038	0.23928	0.05771
		500	0.00799	0.00014	0.16685	0.02790
		750	0.00654	0.00012	0.13045	0.01706
		1000	0.00039	0.00001	0.10438	0.01089
		1250	-0.00233	0.00004	0.08511	0.00725
	$\hat{t}_{yr4}$	125	7.96972	0.14304	0.47628	63.74335
		250	7.97383	0.14311	0.32869	63.68997
		500	7.96895	0.14302	0.21370	63.54982
		750	7.97387	0.14311	0.16801	63.61080
		1000	7.97852	0.14320	0.14161	63.67688
		1250	7.98095	0.14324	0.11082	63.70781

Figura 6.35: Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc1}$ e $\hat{t}_{yrc2}$, com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 2.

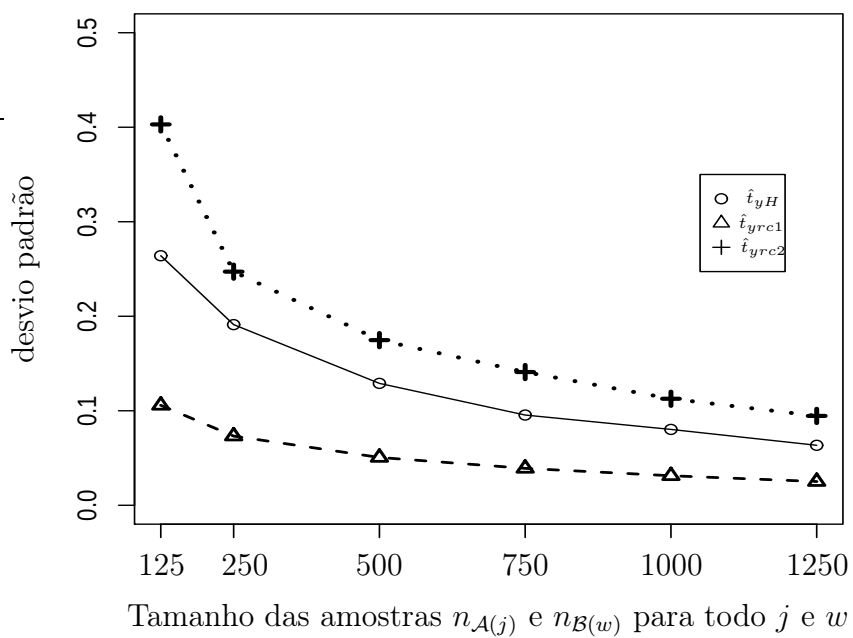
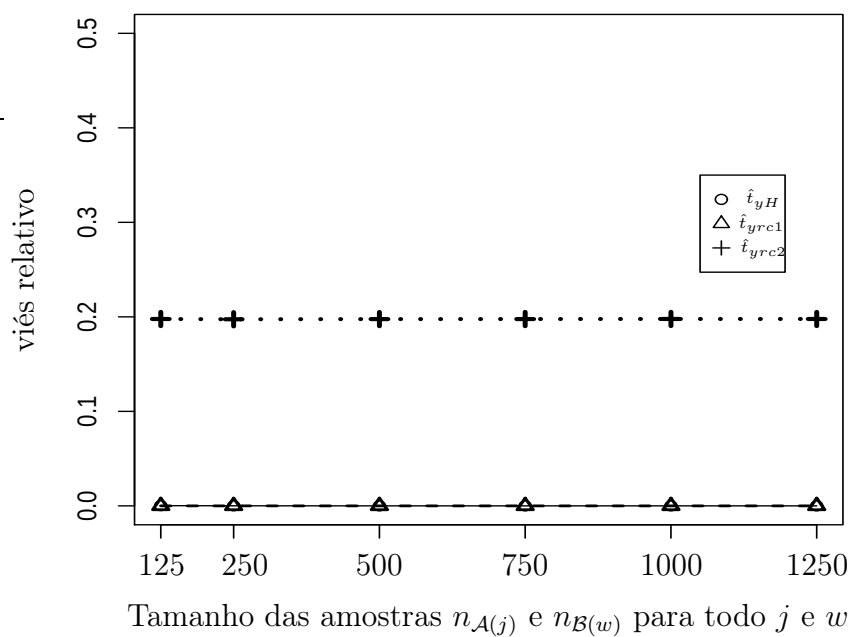


Figura 6.36: Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc3}$ e $\hat{t}_{yrc4}$, com $N_{a(j)} = N_{b(w)} = 1750$ e $N_{ab(jw)} = 1000$, para todo estrato j e w , sob o cenário 3.

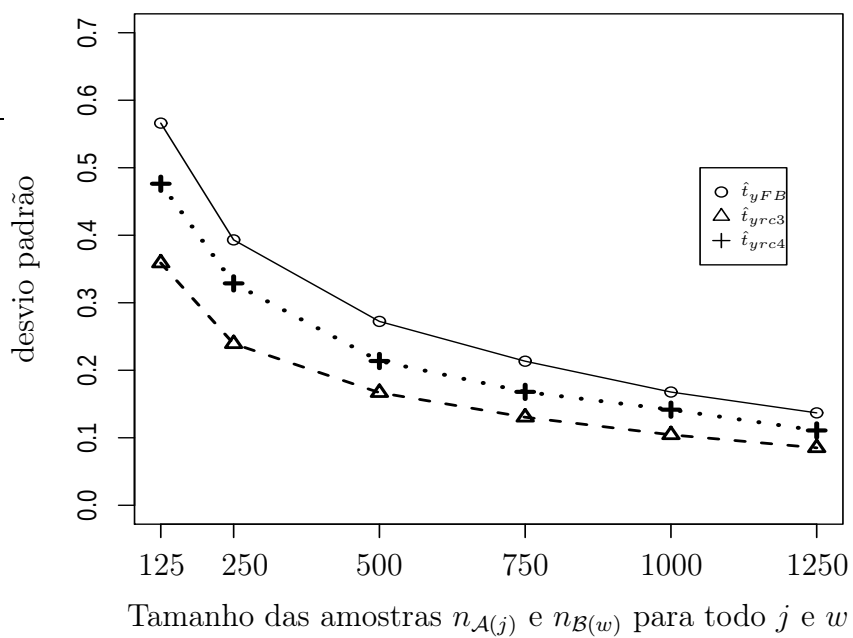
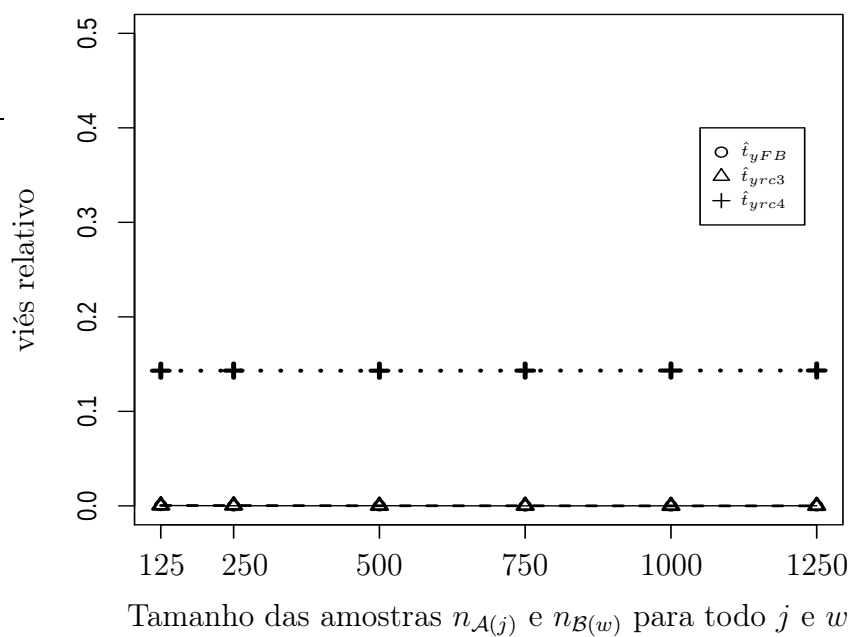


Tabela 6.22: Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE (dados populacionais: $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$)

Cenário	Estimador	$n_{\mathcal{A}(j)}=n_{\mathcal{B}(w)}$	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	0.00174	0.00003	0.26512	0.07029	
	250	-0.00196	0.00004	0.19015	0.03616	
	500	0.00017	0.00000	0.12478	0.01557	
	750	0.00080	0.00002	0.09868	0.00974	
	1000	-0.00282	0.00005	0.08143	0.00664	
	1250	-0.00081	0.00002	0.06481	0.00420	
	125	-0.00036	0.00001	0.40441	0.16355	
	250	0.00363	0.00007	0.13428	0.01805	
	500	0.00011	0.00000	0.06263	0.00392	
	750	-0.00134	0.00003	0.04565	0.00209	
	1000	-0.00172	0.00003	0.03850	0.00149	
	1250	0.00160	0.00003	0.03131	0.00098	
	125	15.71316	0.30251	0.62516	247.29426	
	250	15.76662	0.30354	0.43061	248.77173	
	500	15.79484	0.30408	0.29348	249.56297	
	750	15.78904	0.30397	0.22342	249.34371	
	1000	15.78937	0.30398	0.17763	249.33576	
	1250	15.78890	0.30397	0.15435	249.31331	
$\hat{t}_{yrc1}$	125	-0.01225	0.00024	0.66434	0.44150	
	250	0.00099	0.00002	0.46719	0.21827	
	500	0.00335	0.00006	0.31388	0.09853	
	750	-0.00146	0.00003	0.25153	0.06327	
	1000	0.00094	0.00002	0.20646	0.04263	
	1250	-0.00936	0.00018	0.16705	0.02799	
	125	0.02260	0.00044	0.43363	0.18855	
	250	-0.00555	0.00011	0.30309	0.09189	
	500	-0.00426	0.00008	0.20109	0.04045	
	750	-0.00130	0.00003	0.16519	0.02729	
	1000	-0.00034	0.00001	0.12987	0.01687	
	1250	0.00769	0.00015	0.11324	0.01288	
	125	2.82797	0.05444	1.00863	9.01474	
	250	2.83265	0.05453	0.54058	8.31616	
	500	2.83950	0.05467	0.35888	8.19156	
	750	2.83186	0.05452	0.29232	8.10490	
	1000	2.83369	0.05455	0.23838	8.08662	
	1250	2.82032	0.05430	0.19952	7.99399	
$\hat{t}_{yrc2}$	125	-0.01225	0.00024	0.66434	0.44150	
	250	0.00099	0.00002	0.46719	0.21827	
	500	0.00335	0.00006	0.31388	0.09853	
	750	-0.00146	0.00003	0.25153	0.06327	
	1000	0.00094	0.00002	0.20646	0.04263	
	1250	-0.00936	0.00018	0.16705	0.02799	
	125	0.02260	0.00044	0.43363	0.18855	
	250	-0.00555	0.00011	0.30309	0.09189	
	500	-0.00426	0.00008	0.20109	0.04045	
	750	-0.00130	0.00003	0.16519	0.02729	
	1000	-0.00034	0.00001	0.12987	0.01687	
	1250	0.00769	0.00015	0.11324	0.01288	
	125	2.82797	0.05444	1.00863	9.01474	
	250	2.83265	0.05453	0.54058	8.31616	
	500	2.83950	0.05467	0.35888	8.19156	
	750	2.83186	0.05452	0.29232	8.10490	
	1000	2.83369	0.05455	0.23838	8.08662	
	1250	2.82032	0.05430	0.19952	7.99399	
$\hat{t}_{yrc3}$	125	-0.01225	0.00024	0.66434	0.44150	
	250	0.00099	0.00002	0.46719	0.21827	
	500	0.00335	0.00006	0.31388	0.09853	
	750	-0.00146	0.00003	0.25153	0.06327	
	1000	0.00094	0.00002	0.20646	0.04263	
	1250	-0.00936	0.00018	0.16705	0.02799	
	125	0.02260	0.00044	0.43363	0.18855	
	250	-0.00555	0.00011	0.30309	0.09189	
	500	-0.00426	0.00008	0.20109	0.04045	
	750	-0.00130	0.00003	0.16519	0.02729	
	1000	-0.00034	0.00001	0.12987	0.01687	
	1250	0.00769	0.00015	0.11324	0.01288	
	125	2.82797	0.05444	1.00863	9.01474	
	250	2.83265	0.05453	0.54058	8.31616	
	500	2.83950	0.05467	0.35888	8.19156	
	750	2.83186	0.05452	0.29232	8.10490	
	1000	2.83369	0.05455	0.23838	8.08662	
	1250	2.82032	0.05430	0.19952	7.99399	
$\hat{t}_{yrc4}$	125	-0.01225	0.00024	0.66434	0.44150	
	250	0.00099	0.00002	0.46719	0.21827	
	500	0.00335	0.00006	0.31388	0.09853	
	750	-0.00146	0.00003	0.25153	0.06327	
	1000	0.00094	0.00002	0.20646	0.04263	
	1250	-0.00936	0.00018	0.16705	0.02799	
	125	0.02260	0.00044	0.43363	0.18855	
	250	-0.00555	0.00011	0.30309	0.09189	
	500	-0.00426	0.00008	0.20109	0.04045	
	750	-0.00130	0.00003	0.16519	0.02729	
	1000	-0.00034	0.00001	0.12987	0.01687	
	1250	0.00769	0.00015	0.11324	0.01288	
	125	2.82797	0.05444	1.00863	9.01474	
	250	2.83265	0.05453	0.54058	8.31616	
	500	2.83950	0.05467	0.35888	8.19156	
	750	2.83186	0.05452	0.29232	8.10490	
	1000	2.83369	0.05455	0.23838	8.08662	
	1250	2.82032	0.05430	0.19952	7.99399	

Figura 6.37: Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc1}$ e $\hat{t}_{yrc2}$, com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 2.

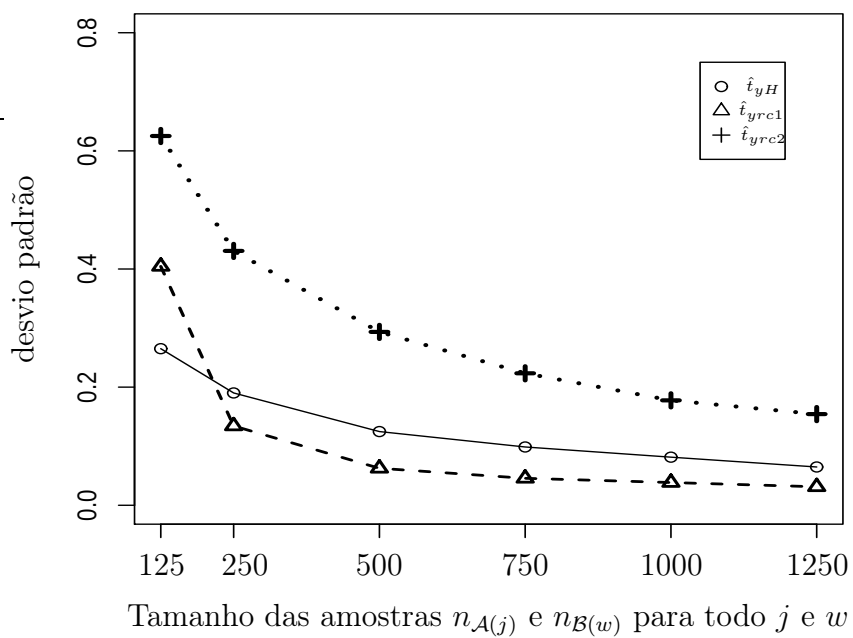
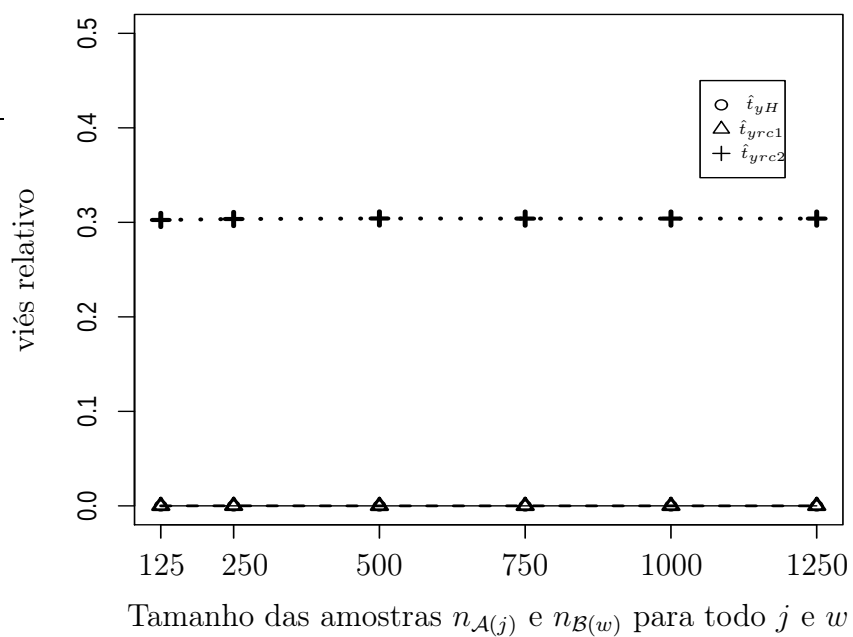


Figura 6.38: Viés Relativo e Desvio Padrão dos estimadores combinados $\hat{t}_{yH}$, $\hat{t}_{yrc3}$ e $\hat{t}_{yrc4}$, com $N_{a(j)} = N_{b(w)} = 1500$ e $N_{ab(jw)} = 1500$, para todo estrato j e w , sob o cenário 3.

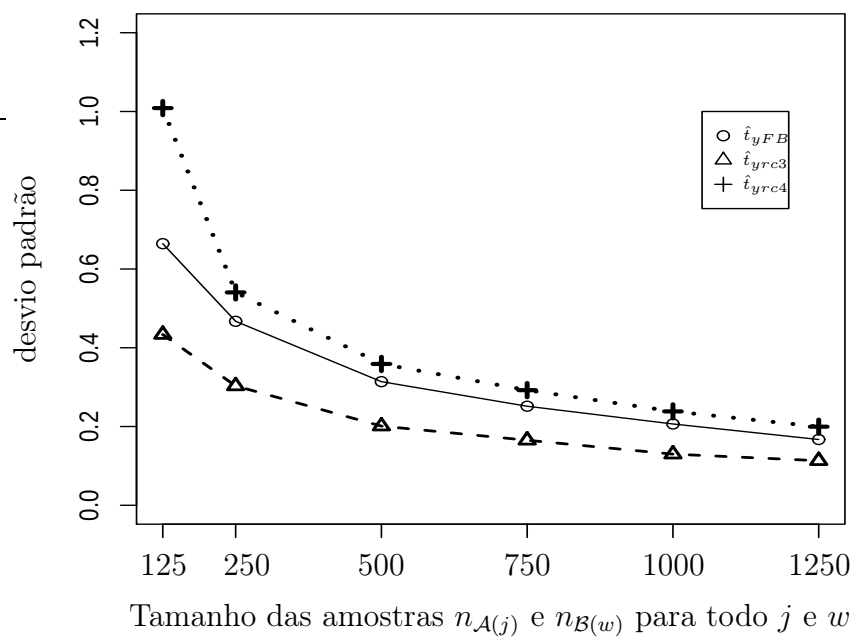
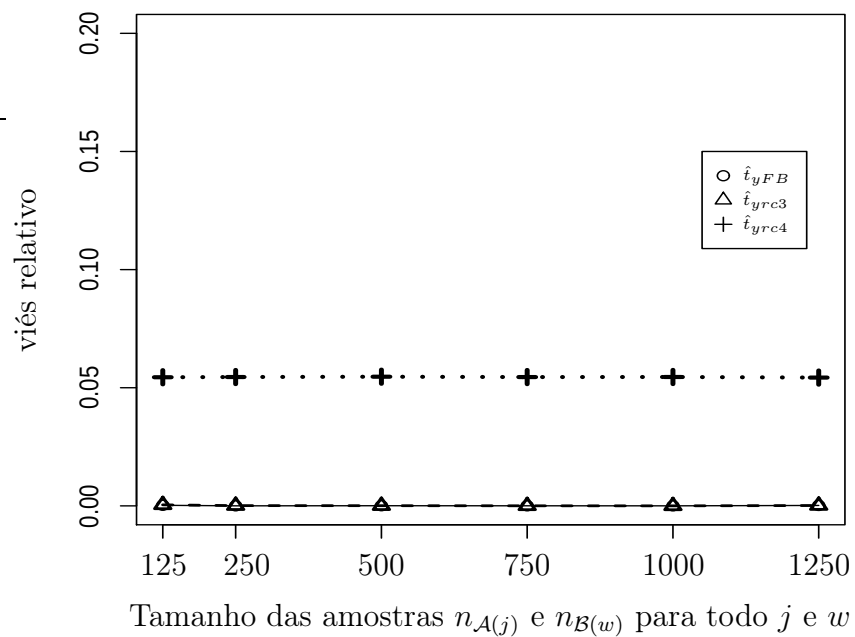


Tabela 6.23: Avaliação dos estimadores razão combinados sob os cenários 2 e 3 para o plano AAE, em que $n_{\mathcal{A}(j)} = 1000$ (dados populacionais: $N_{a(j)} = N_{b(w)} = 2000$ e $N_{ab(jw)} = 500$, para j e w)

Cenário	Estimador	$n_{\mathcal{A}}$	$n_{\mathcal{B}}$	Estatística			
				Viés	Viés Relativo	D.P.	E.Q.M.
2	$\hat{t}_{yH}$	1000	200	0.00082	0.00001	0.08467	0.00717
		1000	300	-0.00043	0.00001	0.07847	0.00616
		1000	400	0.00070	0.00001	0.07993	0.00639
		1000	500	0.00012	0.00000	0.07812	0.00610
		1000	700	-0.00026	0.00000	0.07762	0.00603
		1000	900	-0.00163	0.00003	0.07771	0.00604
	$\hat{t}_{yr1}$	1000	200	-0.15223	0.00257	6.52652	42.61863
		1000	300	-0.00675	0.00011	0.10234	0.01052
		1000	400	-0.00057	0.00001	0.02993	0.00090
		1000	500	-0.00041	0.00001	0.03097	0.00096
		1000	700	-0.00028	0.00000	0.03241	0.00105
		1000	900	-0.00054	0.00001	0.03061	0.00094
	$\hat{t}_{yr2}$	1000	200	6.06773	0.10232	0.09809	36.82699
		1000	300	5.91307	0.09971	0.07674	34.97029
		1000	400	5.85217	0.09868	0.06949	34.25270
		1000	500	5.81445	0.09805	0.06924	33.81268
		1000	700	5.77284	0.09734	0.06531	33.32995
		1000	900	5.75138	0.09698	0.06620	33.08276
3	$\hat{t}_{yFB}$	1000	200	-0.00270	0.00005	0.16509	0.02726
		1000	300	0.00135	0.00002	0.15338	0.02353
		1000	400	0.00228	0.00004	0.15264	0.02330
		1000	500	0.00029	0.00000	0.14904	0.02221
		1000	700	0.00204	0.00003	0.13960	0.01949
		1000	900	-0.00341	0.00006	0.12857	0.01654
	$\hat{t}_{yr3}$	1000	200	0.00291	0.00005	0.12746	0.01625
		1000	300	-0.00076	0.00001	0.11624	0.01351
		1000	400	0.00135	0.00002	0.10548	0.01113
		1000	500	0.00034	0.00001	0.09406	0.00885
		1000	700	0.00101	0.00002	0.06985	0.00488
		1000	900	0.00690	0.00002	0.05415	0.00298
	$\hat{t}_{yr4}$	1000	200	1.33827	0.02257	1.18756	3.20127
		1000	300	2.27525	0.03837	2.62299	12.05686
		1000	400	4.07977	0.06879	2.72750	24.08381
		1000	500	5.13558	0.08660	2.04482	30.55549
		1000	700	5.58081	0.09411	1.30187	32.84028
		1000	900	5.64128	0.09513	1.18483	33.22791

Tabela 6.24: Avaliação dos estimadores razão combinados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 2000$ e $N_{ab(jw)} = 500$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.00906	0.00009	0.51114	0.26134
	250	50	-0.00385	0.00004	0.33290	0.11084
	500	100	-0.00849	0.00008	0.22530	0.05083
	750	150	-0.00523	0.00005	0.17251	0.02979
	1000	200	0.00342	0.00003	0.13618	0.01856
	1250	250	-0.00223	0.00002	0.10832	0.01174
	1500	300	0.00173	0.00002	0.09417	0.00887
	1750	350	-0.00276	0.00003	0.07379	0.00545
$\hat{t}_{yrc2}$	125	25	1.98892	0.01950	1.73474	6.96511
	250	50	2.00014	0.01961	1.12322	5.26218
	500	100	2.09033	0.02049	0.76666	4.95725
	750	150	2.08673	0.02046	0.59464	4.70804
	1000	200	2.10936	0.02068	0.46317	4.66392
	1250	250	2.09585	0.02055	0.38952	4.54430
	1500	300	2.10329	0.02062	0.29360	4.51003
	1750	350	2.10054	0.02059	0.22892	4.46466
$\hat{t}_{yH}$	500	25	-0.00730	0.00007	0.22585	0.05106
	500	50	0.01221	0.00012	0.22972	0.05292
	500	100	0.01236	0.00012	0.22774	0.05202
	500	150	0.00442	0.00004	0.22139	0.04903
	500	200	-0.00462	0.00005	0.22853	0.05225
	500	250	0.00304	0.00003	0.22293	0.04971
	500	300	-0.00261	0.00003	0.23117	0.05345
	500	350	-0.00664	0.00007	0.22827	0.05215
$\hat{t}_{yrc2}$	500	25	0.59338	0.00582	0.40631	0.51719
	500	50	-0.36669	0.00360	0.54482	0.43129
	500	100	2.08649	0.02046	0.78682	4.97253
	500	150	4.83086	0.04736	0.83025	24.02653
	500	200	7.19908	0.07058	0.86747	52.57931
	500	250	9.41354	0.09229	0.96636	89.54855
	500	300	11.13695	0.10919	0.63980	124.44109
	500	350	10.35743	0.10155	0.97218	108.22151

Tabela 6.25: Avaliação dos estimadores razão combinados sob o cenário 2 para o plano AAE, onde $U_B \subset U_A$ (dados populacionais: $N_{a(j)} = 1750$ e $N_{ab(jw)} = 1000$)

Estimador	n_A	n_B	Estatística			
			Viés	Viés Relativo	D.P.	E.Q.M.
$\hat{t}_{yH}$	125	25	0.01123	0.00013	0.45858	0.21042
	250	50	-0.01271	0.00015	0.30864	0.09542
	500	100	-0.00398	0.00005	0.20601	0.04245
	750	150	-0.00459	0.00005	0.15671	0.02458
	1000	200	-0.00065	0.00001	0.13377	0.01789
	1250	250	-0.00034	0.00000	0.10904	0.01189
	1500	300	0.00238	0.00003	0.09208	0.00848
	1750	350	-0.00157	0.00002	0.07648	0.00585
$\hat{t}_{yrc2}$	125	25	13.16711	0.15088	1.30640	175.07946
	250	50	13.04970	0.14953	0.86561	171.04394
	500	100	12.95322	0.14842	0.55958	168.09913
	750	150	12.83841	0.14711	0.42889	165.00864
	1000	200	12.66066	0.14507	0.33858	160.40684
	1250	250	12.48211	0.14303	0.26027	155.87076
	1500	300	12.24082	0.14026	0.20413	149.87939
	1750	350	11.98083	0.13728	0.15398	143.56403
$\hat{t}_{yH}$	500	25	0.00256	0.00003	0.20934	0.04383
	500	50	0.00049	0.00001	0.20197	0.04079
	500	100	-0.00102	0.00001	0.21082	0.04445
	500	150	-0.00025	0.00000	0.21298	0.04536
	500	200	-0.00460	0.00005	0.20407	0.04166
	500	250	0.00089	0.00001	0.19760	0.03905
	500	300	-0.00210	0.00002	0.20785	0.04321
	500	350	-0.00642	0.00007	0.19594	0.03843
$\hat{t}_{yrc2}$	500	25	10.81087	0.12388	0.28749	116.95758
	500	50	11.58482	0.13275	0.39238	134.36208
	500	100	12.97492	0.14867	0.58101	168.68622
	500	150	14.15790	0.16223	0.67921	200.90737
	500	200	15.25294	0.17478	0.78262	233.26457
	500	250	16.25047	0.18621	0.83730	264.77879
	500	300	17.23633	0.19750	0.92245	297.94199
	500	350	18.15231	0.20800	0.97148	330.45028

Considerações Finais

No desenvolvimento desta dissertação, foram apresentados vários estimadores assistidos por modelos lineares sob a abordagem de cadastro duplo. Foram considerados estimadores para uso sob amostragem aleatória simples e estratificada. No contexto de estratificação, foram apresentadas versões de estimadores do tipo separado e combinado. Para todos os casos, a forma da variância aproximada dos estimadores foi derivada através do método de linearização de Taylor.

O desempenho dos estimadores propostos foi avaliado através de um estudo de simulação de Monte Carlo, comparando-os com os estimadores de Hartley(1962) e Fuller & Burmeister (1972), disponíveis na literatura. Os resultados das simulações mostraram que os estimadores propostos apresentaram melhor desempenho em comparação aos estimadores propostos na literatura, como era esperado.

No caso da amostragem aleatória simples, sob o cenário 2, foram estudados os estimadores $\hat{t}_{yH}$, $\hat{t}_{yFB}$, $\hat{t}_{yr1}$, $\hat{t}_{yr2}$, $\hat{t}_{yr3}$, $\hat{t}_{yr4}$. O estudo mostrou que os estimadores $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ têm bom desempenho quando N_{ab} é grande na população. No estudo de simulação, $N = 4500$, e os resultados referentes ao bom desempenho dos estimadores $\hat{t}_{yr1}$ e $\hat{t}_{yr2}$ ocorrem quando $N_{ab} = 1500$, o que representa aproximadamente 33% da população. O estudo revela ainda que o melhor estimador sob o cenário 2 foi $\hat{t}_{yr1}$, pois apresentou menores valores de erro quadrático médio e desvio padrão, quando comparado ao estimador de Hartley e estes valores diminuem à medida em que o tamanho de amostra em cada cadastro aumenta.

Ainda, como os estimadores apresentaram menor desvio padrão que o estimador de Hartley, a escolha por um destes vai depender do tipo de informação auxiliar disponível no cadastro $\mathcal{A}$, e da evidência a respeito de como a população se comporta, ou seja, se há evidência de que a população é assistida pelo modelo descrito em (4.1). Se houver evidência de que o modelo é dessa forma e há informação auxiliar disponível para os domínios, é razoável considerar o estimador $\hat{t}_{yr1}$, e ainda admitir tamanhos de amostra elevados em cada cadastro. No estudo, o estimador se comportou bem para valores de $n_{\mathcal{A}}$ e $n_{\mathcal{B}}$ maiores que 250 cada, as quais juntas representam um percentual de 11% de toda a população que é coberta pelos dois cadastros.

Quando $U_{\mathcal{B}} \subset U_{\mathcal{A}}$, o estimador de Hartley apresentou menor erro quadrático médio que $\hat{t}_{yr2}$. Porém, à medida em que o conjunto de elementos em $\mathcal{B}$ aumenta, o estimador $\hat{t}_{yr2}$ passa a apresentar bons resultados em termos do desvio padrão, que apresentou-se menor que o valor do desvio padrão para o estimador de Hartley. A escolha por um destes vai depender do tipo de informação auxiliar disponível e da opinião do responsável pela condução da pesquisa, pois apesar de ser um pouco viesado, o estimador $\hat{t}_{yr2}$ possui menor valor da variância na situação em que o cadastro $\mathcal{B}$ (cadastro de lista) possui um grande percentual de elementos de $\mathcal{A}$ cadastro de área. Uma vez que essa situação tem aplicação direta em pesquisas agropecuárias, a evidência a respeito da importância do nível de cobertura do cadastro de lista era de se esperar.

Em relação ao cenário 3, o estimador $\hat{t}_{yr3}$ apresentou melhor desempenho que o estimador de Fuller & Burmeister, pois apresentou menor valor de desvio padrão e erro quadrático médio para tamanhos de amostra elevados nos dois cadastros, pois o estudo mostrou que o bom desempenho de $\hat{t}_{yr3}$ ocorre para $n_{\mathcal{A}} = n_{\mathcal{B}} = 500$. O estimador $\hat{t}_{yr4}$ apresentou menor valor do desvio padrão quando $N_{ab} = 1500$, o que representa aproximadamente 33% da população. Porém, $\hat{t}_{yr4}$ apresenta ainda maior valor de viés e erro quadrático médio quando comparado aos estimadores $\hat{t}_{yr3}$ e $\hat{t}_{yFB}$. Quando houver evidência de que a população é assistida pelo modelo descrito em (4.1), e na utilização do plano amostral for evidenciado o cenário 3, o estimador $\hat{t}_{yr3}$ é o mais indicado.

No caso do plano de amostragem estratificada e sob o cenário 2, o estimador do tipo separado $\hat{t}_{yrs1}$ apresentou melhor desempenho do que o estimador de Hartley, por apresentar menor desvio padrão e menor erro quadrático médio. Verifica-se ainda que à medida

que $N_{ab(jw)}$ aumenta para cada estrato j e w , é recomendável que o estimador seja utilizado quando os tamanhos de amostra em cada estrato forem relativamente grandes. Como não houve evidência de que $\hat{t}_{yrs2}$ apresente melhor desempenho, é recomendável utilizar o estimador $\hat{t}_{yrs1}$ quando a informação da variável auxiliar em cada domínio do cadastro $\mathcal{A}$ estiver disponível e quando houver evidência de que a população em cada estrato é assistida pelo modelo descrito em (5.3). No cenário 3, o estimador $\hat{t}_{yrs3}$, apresentou menor desvio padrão e erro quadrático médio que o estimador de Fuller & Burmeister. Neste caso, à medida em que há evidência de que $N_{ab(jw)}$ representa grande parte ou até mesmo a maioria de elementos na população, e quando houver evidência de que a população é assistida pelo modelo descrito em (5.3) é recomendável utilizar $\hat{t}_{yrs3}$. O estimador $\hat{t}_{yrs4}$ não forneceu evidência de desvio padrão e erro quadrático médio menores que o estimador de Fuller & Burmeister.

Em relação aos estimadores combinados, sob o cenário 2, o estimador $\hat{t}_{yrc1}$ apresentou melhor desempenho quando comparado com o estimador de Hartley, por apresentar menores valores da variância e erro quadrático médio. O estimador $\hat{t}_{yrc2}$ pode ser considerado quando há evidência de que o tamanho populacional de $N_{ab(jw)} = 500$ ou menor, em cada estrato, ou seja, $\hat{t}_{yrc2}$ é indicado quando a quantidade de informações pertencentes ao domínio ab for pequena na população. Sob essa situação, o uso de um ou outro estimador do tipo combinado vai depender da informação auxiliar que se tem disponível para os estratos dentro do cadastro $\mathcal{A}$. Quando houver evidência de que a população em cada estrato é assistida pelo modelo descrito em (5.19), é preferível utilizar $\hat{t}_{yrc1}$. No cenário 3, o estimador $\hat{t}_{yrc3}$, apresentou menor desvio padrão e erro quadrático médio que o estimador de Fuller & Burmeister, independentemente do valor das quantidades $N_{ab(jw)}$ na população. Quando há evidência que a população é assistida pelo modelo descrito em (5.19) é recomendável utilizar $\hat{t}_{yrc3}$.

No caso em que $U_B \subset U_A$, apesar dos resultados sob o plano de amostragem aleatória simples terem sido favoráveis ao estimador $\hat{t}_{yr2}$, verificou-se que $\hat{t}_{yrs2}$ e $\hat{t}_{yrc2}$ no caso estratificado não apresentaram evidência de que sejam melhores, quando comparados com o estimador de Hartley. Como o tamanho de amostra nos estratos do cadastro $\mathcal{B}$ representam frações dos tamanhos de amostras dos estratos do cadastro $\mathcal{A}$, o desempenho tenha sido afetado por isso. Outro fator que pode ter contribuído para o fraco desempenho dos

estimadores é a forma como a população está assistida, ou seja, a população está sendo assistida por modelos distintos para os domínios.

Por fim, é importante evidenciar que o fato dos estimadores $\hat{t}_{yr2}$, $\hat{t}_{yr4}$, $\hat{t}_{yrs2}$ e $\hat{t}_{yrc2}$ não terem apresentado o desempenho esperado possivelmente está associado ao fato de que estes utilizam no processo de estimação toda a informação auxiliar do cadastro $\mathcal{A}$, que é a quantidade $t_{X_{\mathcal{A}}}$. Como a razão que multiplica $t_{X_{\mathcal{A}}}$ nos estimadores $\hat{t}_{yr2}$, $\hat{t}_{yr4}$, $\hat{t}_{yrs2}$ e $\hat{t}_{yrc2}$ é referente à todo o conjunto de elementos do cadastro $\mathcal{A}$, é provável que os estimadores atinjam o desempenho esperado se a população inteira no cadastro $\mathcal{A}$ for assistida por um modelo da forma

$$\eta : y_{k\mathcal{A}} = \beta_{\mathcal{A}}x_{k\mathcal{A}} + \varepsilon_{k\mathcal{A}}$$

onde $\beta_{\mathcal{A}}$ é o impacto na média da variável de interesse que é comum aos domínios a e ab . Além disso,

$$E_{\eta}(\varepsilon_{k\mathcal{A}}) = 0; \quad E_{\eta}(y_{k\mathcal{A}}) = \beta_{\mathcal{A}}x_k; \quad \text{Var}_{\eta}(y_{k\mathcal{A}}) = \lambda x_{k\mathcal{A}}$$

A evidência a respeito desse comportamento na população é uma das propostas para trabalhos futuros, de modo a realizar um estudo de simulação sobre quais cenários populacionais são mais indicados para uso de cada um dos estimadores propostos. Por fim, outras propostas para as pesquisas referentes ao uso de variáveis auxiliares na abordagem de cadastro duplo são as seguintes:

- Estender toda a teoria desenvolvida para o caso em que são considerados planos amostrais complexos, como por exemplo, amostragem em dois estágios;
- Propor estimadores do tipo razão para a abordagem de duplo cadastro onde há informação auxiliar disponível para os dois cadastros;
- Propor formas e estimadores das variâncias dos estimadores do tipo razão para o caso em que há informação auxiliar disponível para os dois cadastros;
- Propor estimadores do tipo regressão para a abordagem, onde mais de uma variável auxiliar é utilizada no processo de estimação, dentro do cadastro $\mathcal{A}$;

- Propor estimadores das variâncias dos estimadores do tipo regressão, para o caso onde mais de uma variável auxiliar é utilizada no processo de estimação apenas no cadastro $\mathcal{A}$;
- Propor estimadores do tipo regressão para o caso em que há informação auxiliar disponível para os dois cadastros;
- Propor formas e estimadores das variâncias dos estimadores do tipo regressão para o caso em que há informação auxiliar disponível para os dois cadastros;
- Comparar os estimadores propostos para planos amostrais complexos com os estimadores existentes, os quais não abordam o uso de variáveis auxiliares;
- Neste trabalho, o estimador $\hat{N}_{ab,s}$ é utilizado devido ao fato de que as duplicatas não são identificadas no processo de estimação. É possível realizar todo o estudo desenvolvido neste trabalho e nos trabalhos propostos utilizando o estimador para N_{ab} que leva em consideração a existência e identificação de duplicatas. Esse estimador foi proposto por Fuller & Burmeister (1972) e pode ser aplicado a todos os estimadores utilizados no cenário 3.
- Realizar estudos de alocação ótima a partir de um estimador escolhido.

Apêndice A

Variâncias aproximadas dos estimadores sob o plano AAS

Cenário 2:

Estimador razão 1

$$\begin{aligned}\hat{t}_{yr1} &= \hat{t}_{ya,R} + \hat{t}_{yb} + p\hat{t}'_{yab,R} + (1-p)\hat{t}''_{yab,R} \\ &= N_a \frac{\tilde{y}_a}{\tilde{x}_a} \mu_{xa} + N_{ab} \left(p \frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}} \mu'_{xab} + (1-p) \tilde{y}''_{ab} \right) + N_b \tilde{y}''_b, \quad \text{onde } \mu_{xi} = \frac{t_{xi}}{N_i}\end{aligned}$$

Com isso, tem-se que

$$\hat{t}_{yr1} = \hat{G}_{\mathcal{A}} + \hat{G}_{\mathcal{B}}, \quad \text{onde}$$

$$\begin{aligned}\hat{G}_{\mathcal{A}} &= N_a \frac{\tilde{y}_a}{\tilde{x}_a} \mu_{xa} + N_{ab} p \frac{\tilde{y}_a}{\tilde{x}_a} \mu'_{xab} \quad \text{e} \\ \hat{G}_{\mathcal{B}} &= N_b \tilde{y}_b + N_{ab} \tilde{y}''_{ab}.\end{aligned}$$

$\hat{G}_{\mathcal{A}}$ é um estimador não-linear, proveniente do cadastro $\mathcal{A}$, enquanto $\hat{G}_{\mathcal{B}}$, proveniente do cadastro $\mathcal{B}$ é linear. É possível reescrever $\hat{G}_{\mathcal{A}}$ da seguinte forma:

$$\hat{G}_{\mathcal{A}} = \hat{R}_a t_{xa} + p \hat{R}_{ab} t'_{xab}$$

Desse modo, tem-se $\hat{G}_{\mathcal{A}} \approx \hat{G}_{\mathcal{AL}}$, onde:

$$\hat{G}_{\mathcal{AL}} = G_{\mathcal{A}} + \left[\frac{t_{ya}}{t_{xa}} + \frac{1}{t_{xa}} \left(\tilde{t}_{ya} - \frac{t_{ya}}{t_{xa}} \tilde{t}_{xa} \right) \right] t_{xa} + p \left[\frac{t'_{yab}}{t'_{xab}} + \frac{1}{t'_{xab}} \left(\tilde{t}'_{yab} - \frac{t'_{yab}}{t'_{xab}} \tilde{t}'_{xab} \right) \right] t'_{xab}$$

$$\begin{aligned}
&= G_{\mathcal{A}} + t_{ya} + \tilde{t}_{ya} - R'_a \tilde{t}_{xa} + p t'_{yab} + p \tilde{t}'_{yab} - p R'_{ab} \tilde{t}'_{xab} \\
&= G_{\mathcal{A}} + t_{ya} + p t'_{yab} + \tilde{t}_{ya} + p \tilde{t}'_{yab} - R'_a \tilde{t}_{xa} - p R'_{ab} \tilde{t}'_{xab} ,
\end{aligned}$$

onde $\tilde{t}_d = N_d \tilde{y}_d$ e $G_{\mathcal{A}} = R'_a t_{xa} + p R'_{ab} t'_{xab} = t_{ya} + p t'_{yab}$. Logo, o estimador $\hat{t}_{yr1}$ pode ser aproximado da seguinte forma:

$$\hat{t}_{yr1} \approx \hat{t}_{yr1L}$$

$$\begin{aligned}
\hat{t}_{yr1L} &= \hat{G}_{\mathcal{A}L} + \hat{G}_{\mathcal{B}} \\
\hat{t}_{yr1L} &= G_{\mathcal{A}} + t_{ya} + p t'_{yab} + \tilde{t}_{ya} + p \tilde{t}'_{yab} - R'_a \tilde{t}_{xa} - p R'_{ab} \tilde{t}'_{xab} \\
&\quad + \tilde{t}_{yb} + (1-p) \tilde{t}''_{yab}
\end{aligned}$$

A variância aproximada de $\hat{t}_{yr1}$ é dada por:

$$A\text{Var}(\hat{t}_{yr1}) = A\text{Var}(\hat{G}_{\mathcal{A}L}) + \text{Var}(\hat{G}_{\mathcal{B}}) + 2A\text{Cov}(\hat{G}_{\mathcal{A}}, \hat{G}_{\mathcal{B}})$$

Uma vez que são aplicados planos amostrais de forma independente para cada um dos cadastros, $A\text{Cov}(\hat{G}_{\mathcal{A}}, \hat{G}_{\mathcal{B}}) = 0$.

Assim:

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr1}) &= A\text{Var}(\hat{G}_{\mathcal{A}}) + \text{Var}(\hat{G}_{\mathcal{B}}) \\
&= \text{Var}\left(G_{\mathcal{A}} + t_{ya} + p t'_{yab} + \tilde{t}_{ya} + p \tilde{t}'_{yab} - R'_a \tilde{t}_{xa} - p R'_{ab} \tilde{t}'_{xab}\right) \\
&\quad + \text{Var}\left(\tilde{t}_{yb} + (1-p) \tilde{t}''_{yab}\right).
\end{aligned}$$

Logo:

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr1}) &= \text{Var}(\tilde{t}_{ya}) + p^2 \text{Var}(\tilde{t}'_{yab}) + (R'_a)^2 \text{Var}(\tilde{t}_{xa}) + p^2 (R'_{ab})^2 \text{Var}(\tilde{t}'_{xab}) \\
&\quad + 2p A\text{Cov}(\tilde{t}_{ya}, \tilde{t}'_{yab}) - 2(R'_a) A\text{Cov}(\tilde{t}_{ya}, \tilde{t}_{xa}) \\
&\quad - 2p(R'_{ab}) A\text{Cov}(\tilde{t}_{ya}, \tilde{t}'_{xab}) - 2p(R'_a) A\text{Cov}(\tilde{t}'_{yab}, \tilde{t}_{xa}) \\
&\quad - 2p^2(R'_{ab}) A\text{Cov}(\tilde{t}'_{yab}, \tilde{t}'_{xab}) + 2p(R'_a)(R'_{ab}) A\text{Cov}(\tilde{t}_{xa}, \tilde{t}'_{xab})
\end{aligned}$$

$$+ \text{Var}(\tilde{t}_{yb}) + (1-p)^2 \text{Var}(\tilde{t}_{yab}'') + 2(1-p) \text{ACov}(\tilde{t}_{yb}, \tilde{t}_{yab}'').$$

Neste caso, N_a , N_b e N_{ab} são conhecidos. Com isso, as seguintes quantidades de interesse no processo de estimação devem ser consideradas:

- $\tilde{t}_{ya} = N_a \tilde{y}_a$;
- $\tilde{t}_{xa} = N_a \tilde{x}_a$;
- $\tilde{t}_{yab}' = N_{ab} \tilde{y}_{ab}'$;
- $\tilde{t}_{xab}' = N_{ab} \tilde{y}_{ab}'$;
- $\tilde{t}_{yb} = N_b \tilde{y}_b$;
- $\tilde{t}_{yab}'' = N_{ab} \tilde{y}_{ab}''$;

No processo de obtenção de expressão para a variância aproximada, também é necessário obter as seguintes covariâncias:

1. $\text{ACov}(\tilde{t}_{ya}, \tilde{t}_{yab}') = \text{ACov}(N_a \tilde{y}_a, N_{ab} \tilde{y}_{ab}') = N_a N_{ab} \text{ACov}(\tilde{y}_a, \tilde{y}_{ab}')$;
2. $\text{ACov}(\tilde{t}_{ya}, \tilde{t}_{xa}) = \text{ACov}(N_a \tilde{y}_a, N_a \tilde{x}_a) = N_a^2 \text{ACov}(\tilde{y}_a, \tilde{x}_a)$;
3. $\text{ACov}(\tilde{t}_{ya}, \tilde{t}_{xab}') = \text{ACov}(N_a \tilde{y}_a, N_{ab} \tilde{x}_{ab}') = N_a N_{ab} \text{ACov}(\tilde{y}_a, \tilde{x}_{ab}')$;
4. $\text{ACov}(\tilde{t}_{yab}', \tilde{t}_{xa}) = \text{ACov}(N_{ab} \tilde{y}_{ab}', N_a \tilde{x}_a) = N_{ab} N_a \text{ACov}(\tilde{y}_{ab}', \tilde{x}_a)$;
5. $\text{ACov}(\tilde{t}_{yab}', \tilde{t}_{xab}') = \text{ACov}(N_{ab} \tilde{y}_{ab}', N_{ab} \tilde{x}_{ab}') = N_{ab}^2 \text{ACov}(\tilde{y}_{ab}', \tilde{x}_{ab}')$;
6. $\text{ACov}(\tilde{t}_{xa}, \tilde{t}_{xab}') = \text{ACov}(N_a \tilde{x}_a, N_{ab} \tilde{x}_{ab}') = N_a N_{ab} \text{ACov}(\tilde{x}_a, \tilde{x}_{ab}')$;
7. $\text{ACov}(\tilde{t}_{yb}, \tilde{t}_{yab}'') = \text{ACov}(N_b \tilde{y}_b, N_{ab} \tilde{y}_{ab}'') = N_b N_{ab} \text{ACov}(\tilde{y}_b, \tilde{y}_{ab}'')$;

A expressão geral da covariância aproximada entre médias de dois domínios é a seguinte (Särndal, 1992):

$$\text{ACov}(\tilde{y}_{S_1}, \tilde{y}_{S_2}) = \frac{1}{N_1 N_2} \sum_{k \in U_1} \sum_{l \in U_2} \Delta_{kl} \left(\frac{y_k - \bar{y}_{U_1}}{\pi_k} \right) \left(\frac{y_l - \bar{y}_{U_2}}{\pi_l} \right)$$

Sob o caso AAS, tem-se

$$\begin{aligned}\pi_k &= \frac{n_c}{N_c} \text{ e} \\ \Delta_{kl} &= \pi_{kl} - \pi_k \pi_l = \frac{n_c(n_c - 1)}{N_c(N_c - 1)} - \left(\frac{n_c}{N_c}\right)^2 = -\left(\frac{1}{N_c - 1}\right) \left(\frac{n_c}{N_c}\right) \left(1 - \frac{n_c}{N_c}\right) = k_c,\end{aligned}$$

onde N_c e n_c indicam o tamanho da população e o número de elementos da amostra do cadastro $\mathcal{C}$, respectivamente.

Logo,

$$\begin{aligned}ACov(\tilde{y}_{S_1}, \tilde{y}_{S_2}) &= \frac{k_c}{N_1 N_2} \sum_{k \in U_1} \sum_{l \in U_2} \left(\frac{y_k - \bar{y}_{U_1}}{n_c/N_c}\right) \left(\frac{y_l - \bar{y}_{U_2}}{n_c/N_c}\right) \\ &= \frac{k_c}{N_1 N_2} \left(\frac{N_c}{n_c}\right)^2 \sum_{k \in U_1} \sum_{l \in U_2} (y_k - \bar{y}_{U_1})(y_l - \bar{y}_{U_2})\end{aligned}$$

U_1 e U_2 são disjuntos, ou seja, $k \neq l$. Daí tem-se que

$$\begin{aligned}ACov(\tilde{y}_{S_1}, \tilde{y}_{S_2}) &= \frac{k_c}{N_1 N_2} \left(\frac{N_c}{n_c}\right)^2 \sum_{k \in U_1} (y_k - \bar{y}_{U_1}) \sum_{l \in U_2} (y_l - \bar{y}_{U_2}) \\ &= \frac{k_c}{N_1 N_2} \left(\frac{N_c}{n_c}\right)^2 \left\{ \sum_{k \in U_1} y_k - N_1 \bar{y}_{U_1} \right\} \left\{ \sum_{l \in U_2} y_l - N_2 \bar{y}_{U_2} \right\} \\ &= \frac{k_c}{N_1 N_2} \left(\frac{N_c}{n_c}\right)^2 \{N_1 \bar{y}_{U_1} - N_1 \bar{y}_{U_1}\} \{N_2 \bar{y}_{U_1} - N_2 \bar{y}_{U_2}\} \\ &= 0\end{aligned}$$

Dessa forma, independentemente do conhecimento ou não de N_a , N_b e N_{ab} , e devido ao fato de que as probabilidades de inclusão de primeira ordem são todas iguais a uma constante, devido ao plano amostral considerado, as covariâncias 1, 3, 4, 6 e 7 serão nulas. As covariâncias que restaram não serão nulas, devido ao fato de que os termos presentes em cada uma pertencem a um mesmo domínio, além de que um dos termos em cada expressão constitui uma variável auxiliar, e que de fato se espera possuir relação com a variável resposta presente no mesmo domínio.

O objetivo agora é obter expressões para as covariâncias 2 e 5. Sejam $\tilde{y}_d$ e $\tilde{x}_d$ as médias amostrais da variável resposta e da variável auxiliar, pertencentes ao domínio d . Sabemos que a covariância entre duas médias amostrais, $\bar{x}$ e $\bar{y}$ é dada da seguinte forma:

$$\text{Cov}(\bar{x}, \bar{y}) = \frac{1}{n} \left(1 - \frac{n}{N}\right) \sum_{i=1}^N (x_i - \bar{x}_U)(y_i - \bar{y}_U).$$

Para obter as covariâncias de interesse, sejam:

$$u_k = \begin{cases} y_k, & \text{se } k \in d \\ 0, & \text{c.c.} \end{cases} \quad \text{e} \quad w_k = \begin{cases} x_k, & \text{se } k \in d \\ 0, & \text{c.c.} \end{cases}$$

Tem-se que $\bar{y}_C = \frac{\sum_{k \in S_C} y_k}{n_C}$. Assim,

$$\bar{u}_C = \frac{\sum_{k \in S_C} u_k}{n_C} = \frac{\sum_{k \in S_d} y_k}{n_C} \quad \text{e} \quad \bar{w}_C = \frac{\sum_{k \in S_C} w_k}{n_C} = \frac{\sum_{k \in S_d} x_k}{n_C}$$

Dessa forma, é possível considerar que a covariância entre $\bar{u}_C$ e $\bar{w}_C$ é dada da seguinte forma:

$$\begin{aligned} \text{Cov}(\bar{u}_C, \bar{w}_C) &= \frac{1}{n_C} \left(1 - \frac{n_C}{N_C}\right) \sum_{k \in U_C} (u_k - \bar{u}_{U_C})(w_k - \bar{w}_{U_C}) \\ &= \frac{1}{n_C} \left(1 - \frac{n_C}{N_C}\right) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})(x_k - \bar{x}_{U_d}). \end{aligned}$$

Logo, a covariância entre duas médias pertencentes ao mesmo domínio pode ser expressa como:

$$\begin{aligned} \text{ACov}(\tilde{y}_d, \tilde{x}_d) &= \frac{1}{n_d^2} \text{Cov} \left(\sum_{k \in S_d} y_k, \sum_{k \in S_d} x_k \right) = \left(\frac{n_C}{n_d} \right)^2 \text{Cov}(\bar{u}_C, \bar{w}_C) \\ &= \left(\frac{n_C}{n_d} \right)^2 \frac{1}{n_C} \left(1 - \frac{n_C}{N_C}\right) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})(x_k - \bar{x}_{U_d}) \\ &= \frac{n_C}{n_d^2} \left(1 - \frac{n_C}{N_C}\right) \rho_{xyd} \sigma_{yd} \sigma_{xd}, \end{aligned}$$

onde ρ_{xyd} é a correlação entre as variáveis pertencentes ao domínio d . Além de obter uma expressão para a covariância entre médias pertencentes ao mesmo domínio, é necessário

ainda obter a expressão para a variância para estimadores do tipo $\tilde{t}_{yd} = N_d \tilde{y}_d$. Com esse objetivo, tem-se que

$$\begin{aligned}
A\text{Var}(\tilde{t}_{yd}) &= \sum_{k \in U_d} \sum_{l \in U_d} \Delta_{kl} \left(\frac{y_k - \bar{y}_{U_d}}{\pi_k} \right) \left(\frac{y_l - \bar{y}_{U_d}}{\pi_l} \right) \\
&= \sum_{k \in U_d} \Delta_{kk} \left(\frac{y_k - \bar{y}_{U_d}}{\pi_k} \right)^2 + \sum_{k \in U_d} \Delta_{kl} \left(f \frac{y_k - \bar{y}_{U_d}}{\pi_k} \right) \left(\frac{y_k - \bar{y}_{U_d}}{\pi_l} \right) \\
A\text{Var}(\tilde{t}_{yd}) &= \sum_{k \in U_d} \left(\frac{N_c}{n_c} \right) \left(1 - \frac{n_c}{N_c} \right) (y_k - \bar{y}_{U_d})^2 \\
&\quad - \sum_{k \in U_d} \frac{1}{N_c - 1} \left(\frac{N_c}{n_c} \right) \left(1 - \frac{n_c}{N_c} \right) (y_k - \bar{y}_{U_d}) (y_k - \bar{y}_{U_d}) \\
&= \frac{1}{N_c - 1} \left(\frac{N_c}{n_c} \right) \left(1 - \frac{n_c}{N_c} \right) \left\{ (N_c - 1) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 \right. \\
&\quad \left. - \sum_{\substack{k \in U_d \\ k \neq l}} \sum_{l \in U_d} (y_k - \bar{y}_{U_d}) (y_k - \bar{y}_{U_d}) \right\} \\
&= -k_c \left\{ (N_c - 1) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 - \left(\sum_{k \in U_d} (y_k - \bar{y}_{U_d}) \right)^2 \right. \\
&\quad \left. + \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 \right\} \\
&= -k_c \left\{ N_c \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 \right\} \\
&= \frac{N_c}{N_c - 1} \left(\frac{N_c}{n_c} \right) \left(1 - \frac{n_c}{N_c} \right) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 \\
&= \frac{N_c}{N_c - 1} \left(\frac{N_c}{n_c} \right) \left(1 - \frac{n_c}{N_c} \right) (N_d - 1) \sigma_{y_d}^2 \\
&= N_c^2 \left(\frac{1 - f_c}{n_c} \right) \left(\frac{N_d - 1}{N_c - 1} \right) \sigma_{y_d}^2 \\
&\approx N_c^2 \left(\frac{1 - f_c}{n_c} \right) P_d \sigma_{y_d}^2, \quad \text{onde } P_d = \frac{N_d}{N_c}
\end{aligned}$$

Voltando ao caso da variância do estimador, tem-se que

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr1}) &= N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2 + p^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}^{'2} \\
&+ (R'_a)^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2 + p^2 (R'_{ab})^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}^{'2} \\
&- 2(R'_a)^2 N_a^2 A\text{Cov}(\tilde{y}_a, \tilde{x}_a) - 2p^2 R'_{ab} N_{ab}^2 A\text{Cov}(\tilde{y}'_{ab}, \tilde{x}'_{ab}) \\
&+ N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 + (1-p)^2 N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}^{'2}
\end{aligned}$$

Portanto,

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr1}) &= N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2 + p^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}^{'2} \\
&+ (R'_a)^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2 + p^2 (R'_{ab})^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}^{'2} \\
&- 2(R'_a)^2 \left(\frac{N_a}{n_a} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{ya} \sigma_{xa} \\
&- 2p^2 R'_{ab} \left(\frac{N_{ab}}{n'_{ab}} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab} \\
&+ N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 + (1-p)^2 N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}^{'2}
\end{aligned}$$

Estimador razão 2:

$$\hat{t}_{yr2} = \left[\frac{N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}}{N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}} \right] t_{X_{\mathcal{A}}} + N_b \tilde{y}_b + (1-p) N_{ab} \tilde{y}''_{ab} = \hat{G}_{\mathcal{A}} t_{X_{\mathcal{A}}} + \hat{G}_{\mathcal{B}},$$

Seja

$$\hat{t}_{y\mathcal{A}}^* = N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab} \quad \text{e} \quad \hat{t}_{x\mathcal{A}}^* = N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab} \implies \hat{G}_{\mathcal{A}} = \frac{\hat{t}_{y\mathcal{A}}^*}{\hat{t}_{x\mathcal{A}}^*}$$

$$\hat{G}_{\mathcal{B}} = N_b \tilde{y}_b + (1-p) N_{ab} \tilde{y}''_{ab} = \tilde{t}_{yb} + (1-p) \tilde{t}''_{ab}$$

Como $\hat{G}_{\mathcal{A}} = f(\hat{t}_{y\mathcal{A}}^*, \hat{t}_{x\mathcal{A}}^*)$, a idéia é obter uma aproximação linear para $\hat{G}_{\mathcal{A}}$. Utilizando o método de linearização de Taylor, tem-se que

$$\hat{G}_{\mathcal{A}} \approx \hat{G}_{\mathcal{AL}} = G_{\mathcal{A}} + \sum_{k=1}^2 a_k (\tilde{t}_{yd} - t_{yd}),$$

onde

$$G_{\mathcal{A}} = \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} = \frac{N_a \mu_{ya} + p N_{ab} \mu'_{yab}}{N_a \mu_{xa} + p N_{ab} \mu'_{xab}} = \frac{t_{ya} + p t'_{yab}}{t_{xa} + p t'_{xab}},$$

$$\begin{aligned} a_1 &= \left. \frac{\partial f}{\partial t_{y\mathcal{A}}^*} \right|_{(\hat{t}_{y\mathcal{A}}^*, \hat{t}_{x\mathcal{A}}^*) = (t_{y\mathcal{A}}^*, t_{x\mathcal{A}}^*)} = \frac{1}{t_{x\mathcal{A}}^*} = \frac{1}{t_{xa} + p t'_{xab}}, \\ a_2 &= \left. \frac{\partial f}{\partial t_{x\mathcal{A}}^*} \right|_{(\hat{t}_{y\mathcal{A}}^*, \hat{t}_{x\mathcal{A}}^*) = (t_{y\mathcal{A}}^*, t_{x\mathcal{A}}^*)} = -\frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^{*2}} = -\frac{t_{ya} + p t'_{yab}}{(t_{xa} + p t'_{xab})^2}. \end{aligned}$$

Assim,

$$\begin{aligned} \hat{G}_{\mathcal{AL}} &= G_{\mathcal{A}} + \frac{1}{t_{x\mathcal{A}}^*} (\hat{t}_{y\mathcal{A}}^* - t_{y\mathcal{A}}^*) - \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^{*2}} (\hat{t}_{x\mathcal{A}}^* - t_{x\mathcal{A}}^*) \\ &= G_{\mathcal{A}} + \frac{\hat{t}_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} - \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} - \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^{*2}} \hat{t}_{x\mathcal{A}}^* + \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} = G_{\mathcal{A}} + \frac{\hat{t}_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} - \frac{t_{y\mathcal{A}}^*}{\hat{t}_{y\mathcal{A}}^{*2}} \hat{t}_{x\mathcal{A}}^* \\ &= G_{\mathcal{A}} + \left(\frac{1}{t_{x\mathcal{A}}^*} \right) \left(\hat{t}_{y\mathcal{A}}^* - \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} \hat{t}_{x\mathcal{A}}^* \right). \end{aligned}$$

Dessa forma, tem-se que

$$A\text{Var}(\hat{t}_{yr2}) = t_{X_{\mathcal{A}}}^2 A\text{Var}(\hat{G}_{\mathcal{A}}) + \text{Var}(\hat{G}_{\mathcal{B}}) + 2 t_{X_{\mathcal{A}}} A\text{Cov}(\hat{G}_{\mathcal{A}}, \hat{G}_{\mathcal{B}})$$

Por conta dos planos amostrais terem sido aplicados independentemente em cada cadastro, $A\text{Cov}(\hat{G}_{\mathcal{A}}, \hat{G}_{\mathcal{B}}) = 0$. Portanto,

$$A\text{Var}(\hat{G}_{\mathcal{A}}) = \text{Var} \left\{ \left(\frac{1}{t_{x\mathcal{A}}^*} \right) \left(\hat{t}_{y\mathcal{A}}^* - \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} \hat{t}_{x\mathcal{A}}^* \right) \right\}$$

$$\begin{aligned}
&= \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 \text{Var} (N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab} - G_{\mathcal{A}} N_a \tilde{x}_a - p G_{\mathcal{A}} N_{ab} \tilde{x}'_{ab}) \\
&\approx \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 \{ A \text{Var}(N_a \tilde{y}_a) + p^2 A \text{Var}(N_{ab} \tilde{y}'_{ab}) + G_{\mathcal{A}}^2 A \text{Var}(N_a \tilde{x}_a) \\
&\quad + p^2 G_{\mathcal{A}}^2 \text{Var}(N_{ab} \tilde{x}'_{ab}) + 2p N_a N_{ab} A \text{Cov}(\tilde{y}_a, \tilde{y}'_{ab}) \\
&\quad - 2G_{\mathcal{A}} N_a^2 A \text{Cov}(\tilde{y}_a, \tilde{x}_a) - 2p G_{\mathcal{A}} N_a N_{ab} A \text{Cov}(\tilde{y}_a, \tilde{x}'_{ab}) \\
&\quad - 2p N_{ab} N_a A \text{Cov}(\tilde{y}'_{ab}, \tilde{x}_a) - 2p^2 N_{ab}^2 G_{\mathcal{A}} A \text{Cov}(\tilde{y}'_{ab}, \tilde{x}'_{ab}) \\
&\quad + 2p G_{\mathcal{A}} N_a N_{ab} A \text{Cov}(\tilde{x}_a, \tilde{x}'_{ab}) \}
\end{aligned}$$

$$\text{Var}(\hat{G}_{\mathcal{B}}) = \text{Var}(\tilde{t}_{yb}) + (1-p)^2 \text{Var}(\tilde{t}_{yab}) + 2(1-p) \text{Cov}(\tilde{t}_{yb}, \tilde{t}_{yab}'')$$

Como já visto anteriormente, algumas covariâncias serão nulas. Dessa forma, tem-se que

$$\begin{aligned}
A \text{Var}(\hat{G}_{\mathcal{A}}) &= \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 \left\{ N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2 + p^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}'^2 \right. \\
&\quad + G_{\mathcal{A}}^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2 + p^2 G_{\mathcal{A}}^2 N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}'^2 \\
&\quad - 2G_{\mathcal{A}} N_a^2 \frac{n_{\mathcal{A}}}{n_a^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{ya} \sigma_{xa} \\
&\quad \left. - 2p^2 G_{\mathcal{A}} N_{ab}^2 \frac{n_{\mathcal{A}}}{n_{ab}'^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab} \right\}
\end{aligned}$$

Verifica-se ainda que $\hat{G}_{\mathcal{B}}$ é uma função linear de estimadores referentes a domínios presentes no cadastro $\mathcal{B}$. Logo,

$$\text{Var}(\hat{G}_{\mathcal{B}}) = N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 + (1-p)^2 N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}''^2$$

Portanto,

$$A\text{Var}(\hat{t}_{yr2}) = t_{X_A}^2 A\text{Var}(\hat{G}_A) + \text{Var}(\hat{G}_B)$$

Sob a situação em que $U_B \subset U_A$, $N_b = 0$ e que $N_{ab} = N_B$. Logo, $N_a = N_A - N_{ab}$. A forma do estimador é dada por:

$$\begin{aligned} \hat{t}_{yr2} &= \left(\frac{N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}}{N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}} \right) t_{X_A} + (1-p) N_{ab} \tilde{y}''_{ab} \\ &= \hat{G}_A t_{X_A} + \hat{G}_B \end{aligned}$$

A partir do método de linearização de Taylor, tem-se então que uma aproximação linear para G_A pode ser dada por

$$\begin{aligned} \hat{G}_{AL} &= G_A + \frac{1}{t_{xA}^*} (\hat{t}_{yA}^* - t_{yA}^*) - \frac{t_{yA}^*}{t_{xA}^{*2}} (\hat{t}_{xA}^* - t_{xA}^*) \\ &= G_A + \frac{\hat{t}_{yA}^*}{t_{xA}^*} - \frac{t_{yA}^*}{t_{xA}^*} - \frac{t_{yA}^*}{t_{xA}^{*2}} \hat{t}_{xA}^* + \frac{t_{yA}^*}{t_{xA}^*} = G_A + \frac{\hat{t}_{yA}^*}{t_{xA}^*} - \frac{t_{yA}^*}{t_{xA}^{*2}} \hat{t}_{xA}^* \\ &= G_A + \left(\frac{1}{t_{xA}^*} \right) \left(\hat{t}_{yA}^* - \frac{t_{yA}^*}{t_{xA}^*} \hat{t}_{xA}^* \right) \end{aligned}$$

Tem-se que $\hat{G}_B = (1-p) N_{ab} \tilde{y}''_{ab}$. Logo,

$$A\text{Var}(\hat{t}_{yr2}) = t_{X_A}^2 A\text{Var}(\hat{G}_A) + \text{Var}(\hat{G}_B), \quad \text{pois } ACov(\hat{G}_A, \hat{G}_B) = 0$$

$$\begin{aligned} A\text{Var}(\hat{G}_A) &= \left(\frac{1}{t_{xA}^*} \right)^2 \left\{ N_A^2 \left(\frac{1-f_A}{n_A} \right) P_a \sigma_{ya}^2 + p^2 N_A^2 \left(\frac{1-f_A}{n_A} \right) P'_{ab} \sigma_{yab}^{'2} \right. \\ &\quad \left. + G_A^2 N_A^2 \left(\frac{1-f_A}{n_A} \right) P_a \sigma_{xa}^2 + p^2 G_A^2 N_A^2 \left(\frac{1-f_A}{n_A} \right) P'_{ab} \sigma_{xab}^{'2} \right\} \end{aligned}$$

$$\begin{aligned}
& - 2G_{\mathcal{A}}N_a^2\frac{n_{\mathcal{A}}}{n_a^2}\left(1-\frac{n_{\mathcal{A}}}{N_{\mathcal{A}}}\right)\rho_{xya}\sigma_{ya}\sigma_{xa} \\
& - 2p^2G_{\mathcal{A}}N_{ab}^2\frac{n_{\mathcal{A}}}{n_{ab}^{'2}}\left(1-\frac{n_{\mathcal{A}}}{N_{\mathcal{A}}}\right)\rho'_{xyab}\sigma'_{yab}\sigma'_{xab}\Big\}
\end{aligned}$$

Logo, a variância de $\hat{G}_{\mathcal{B}}$ é dada da seguinte forma:

$$\text{Var}(\hat{G}_{\mathcal{B}}) = (1-p)^2\text{Var}(N_{ab}\tilde{y}_{ab}'') = (1-p)^2N_{\mathcal{B}}^2\left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}}\right)P_{ab}''\sigma_{yab}''^2$$

Portanto, a variância de $\hat{t}_{yr2}$ será dada por

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr2}) &= t_{X\mathcal{A}}^2\left\{\left(\frac{1}{t_{x\mathcal{A}}^*}\right)^2N_{\mathcal{A}}^2\left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}}\right)P_a\sigma_{ya}^2 + \left(\frac{p}{t_{x\mathcal{A}}^*}\right)^2N_{\mathcal{A}}^2\left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}}\right)P'_{ab}\sigma_{yab}^{'2}\right. \\
&+ \left.\left(\frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*}\right)^2N_{\mathcal{A}}^2\left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}}\right)P_a\sigma_{xa}^2 + p^2\left(\frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*}\right)^2N_{\mathcal{A}}^2\left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}}\right)P'_{ab}\sigma_{xab}^{'2}\right. \\
&- 2\frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^{'2}}N_a^2\frac{n_{\mathcal{A}}}{n_a^2}\left(1-\frac{n_{\mathcal{A}}}{N_{\mathcal{A}}}\right)\rho_{xya}\sigma_{ya}\sigma_{xa} \\
&- 2\left(\frac{p}{t_{x\mathcal{A}}^*}\right)^2G_{\mathcal{A}}N_{ab}^2\frac{n_{\mathcal{A}}}{n_{ab}^{'2}}\left(1-\frac{n_{\mathcal{A}}}{N_{\mathcal{A}}}\right)\rho'_{xyab}\sigma'_{yab}\sigma'_{xab}\Big\} \\
&+ (1-p)^2N_{\mathcal{B}}^2\left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}}\right)P_{ab}''\sigma_{yab}''^2
\end{aligned}$$

Estimador razão 3:

$$\begin{aligned}
\hat{t}_{yr3} &= \frac{\tilde{y}_a}{\tilde{x}_a}t_{xa} + p\frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}}t'_{xab} + (1-p)\hat{N}_{ab,s}\tilde{y}_{ab}'' + \left(N_B - \hat{N}_{ab,s}\right)\tilde{y}_b \\
&= \hat{G}_{\mathcal{A}} + \hat{G}_{\mathcal{B}}
\end{aligned}$$

onde

$$\hat{G}_{\mathcal{A}} = \frac{\tilde{y}_a}{\tilde{x}_a}t_{xa} + p\frac{\tilde{y}'_{ab}}{\tilde{x}'_{ab}}t'_{xab} \quad \text{e}$$

$$\hat{G}_{\mathcal{B}} = (1-p)\hat{N}_{ab,s}\tilde{y}_{ab}'' + \left(N_B - \hat{N}_{ab,s}\right)\tilde{y}_b.$$

$\hat{G}_{\mathcal{A}}$ possui termos da forma:

$$\hat{R}_i = \frac{\tilde{y}_i}{\tilde{x}_i} \approx \hat{R}_{iL} = \frac{\mu_{yi}}{\mu_{xi}} + \frac{1}{\mu_{xi}} \left(\tilde{y}_i - \frac{\mu_{yi}}{\mu_{xi}} \tilde{x}_i \right)$$

Além disso, verifica-se que

$$\hat{G}_{\mathcal{A}} = f(\hat{N}_{ab,s}; \tilde{y}_a; \tilde{x}_a; \tilde{y}_{ab}', \tilde{x}_{ab}') \rightarrow \text{função não linear}$$

$$\hat{G}_{\mathcal{B}} = f(\hat{N}_{ab,s}; \tilde{y}_b; \tilde{y}_{ab}'') \rightarrow \text{função não linear}$$

Portanto,

$$\begin{aligned} \hat{G}_{\mathcal{A}} \approx \hat{G}_{\mathcal{AL}} &= G_{\mathcal{A}} + \left[\frac{\mu_{ya}}{\mu_{xa}} + \frac{1}{\mu_{xa}} \left(\tilde{y}_a - \frac{\mu_{ya}}{\mu_{xa}} \tilde{x}_a \right) \right] t_{xa} + p \left[\frac{\mu'_{yab}}{\mu'_{xab}} + \frac{1}{\mu'_{xab}} \left(\tilde{y}_{ab}' - \frac{\mu'_{yab}}{\mu'_{xab}} \tilde{x}_{ab}' \right) \right] t'_{xab} \\ &= G_{\mathcal{A}} + t_{ya} + N_a \tilde{y}_a - t_{ya} \tilde{x}_a + p t'_{yab} + p N_{ab} \tilde{y}_{ab}' - p t'_{yab} \tilde{x}_{ab}', \end{aligned}$$

$$\text{onde } G_{\mathcal{A}} = \frac{\mu_{ya}}{\mu_{xa}} t_{xa} + p \frac{\mu_{ya}}{\mu_{xa}} t'_{xab} = t_{ya} + p t_{yab}.$$

Seja $\hat{P}_i = \hat{N}_i \tilde{y}_i = \hat{t}_{N_i} \hat{\mu}_i$. Através do método de linearização de Taylor, verifica-se que

$$\hat{P}_i \approx \hat{P}_L = P_i + \sum_{i=1}^2 b_i (\hat{\theta}_i - \theta_i), \text{ onde:}$$

$$P_i = N_i \mu_i,$$

$$b_1 = \left. \frac{\partial \hat{P}_i}{\partial \hat{t}_{N_i}} \right|_{(\hat{t}_{N_i}, \hat{\mu}_i) = (t_{N_i}, \mu_i)} = \mu_i \quad \text{e}$$

$$b_2 = \left. \frac{\partial \hat{P}_i}{\partial \hat{t}_{N_i}} \right|_{(\hat{t}_{N_i}, \hat{\mu}_i) = (t_{N_i}, \mu_i)} = t_{N_i}.$$

Assim,

$$\begin{aligned} \hat{P}_L &= P_i + \mu_i(\hat{t}_{N_i} - t_{N_i}) + t_{N_i}(\hat{\mu}_i - \mu_i) \\ &= P_i + \mu_i \hat{t}_{N_i} - \mu_i t_{N_i} + t_{N_i} \hat{\mu}_i - t_{N_i} \mu_i \\ &= \mu_i \hat{t}_{N_i} + t_{N_i} \hat{\mu}_i - t_{N_i} \mu_i \\ &= \mu_i \hat{N}_i + N_i \tilde{y}_i - N_i \mu_i. \end{aligned}$$

Dessa forma, para $\hat{G}_{\mathcal{B}}$ tem-se que

$$\begin{aligned} \hat{G}_{\mathcal{B}} \approx \hat{G}_{BL} &= G_{\mathcal{B}} + (1-p) \left(\mu''_{yab} \hat{N}_{ab,s} + N_{ab} \tilde{y}''_{ab} - N_{ab} \mu''_{yab} \right) + N_b \tilde{y}_b - N_{ab} \tilde{y}_b \\ &\quad - \mu_{yb} \hat{N}_{ab,s} + N_{ab} \mu_{yb} \\ &= G_{\mathcal{B}} + (1-p) \left(\mu''_{yab} \hat{N}_{ab,s} + N_{ab} \tilde{y}''_{ab} - N_{ab} \mu''_{yab} \right) \\ &\quad + N_b \tilde{y}_b - \mu_{yb} \left(\hat{N}_{ab,s} - N_{ab} \right), \end{aligned}$$

onde $G_{\mathcal{B}} = (1-p)N_{ab}\mu_{yab} + N_b\mu_{yb}$, pois $N_b = N_B - N_{ab}$. Portanto,

$$\begin{aligned} \hat{t}_{yr3} \approx \hat{t}_{yr3L} &= G_{\mathcal{A}} + G_{\mathcal{B}} + t_{ya} - t_{ya} \tilde{x}_a + p t'_{yab} - p t'_{xab} \tilde{x}'_{ab} - t'_{yab} - p t'_{yab} + N_a \tilde{y}_a \\ &\quad + p N_{ab} \tilde{y}'_{ab} + N_b \tilde{y}_b + N_{ab} \tilde{y}''_{ab} - p N_{ab} \tilde{y}''_{ab} + \mu''_{yab} \hat{N}_{ab,s} - p \mu''_{yab} \hat{N}_{ab,s} \\ &\quad - \mu_{yb} \hat{N}_{ab,s} + N_{ab} \mu_{yb} \end{aligned}$$

Logo,

$$\begin{aligned} A\text{Var}(\hat{t}_{yr3}) &= t_{ya}^2 \text{Var}(\tilde{x}_a) + p^2 t_{xab}^{'2} \text{Var}(\tilde{x}_{ab}) + N_a^2 \text{Var}(\tilde{y}_a) + p^2 N_{ab}^2 \text{Var}(\tilde{y}_{ab}) \\ &\quad + N_b^2 \text{Var}(\tilde{y}_b) + (1-p)^2 N_{ab}^2 \text{Var}(\tilde{y}_{ab}) + (1-p)^2 \mu_{yab}^{'2} \text{Var}(\hat{N}_{ab,s}) \\ &\quad + \mu_{yb}^2 \text{Var}(\hat{N}_{ab,s}) + 2p t_{ya} t'_{yab} \text{ACov}(\tilde{x}_a, \tilde{x}'_{ab}) - 2N_a t_{ya} \text{ACov}(\tilde{x}_a, \tilde{y}_a) \\ &\quad - 2t_{ya} p N_{ab} \text{ACov}(\tilde{x}_a, \tilde{y}'_{ab}) + 2(1-p) N_b N_{ab} \text{ACov}(\tilde{y}_b, \tilde{y}''_{ab}) \end{aligned}$$

$$\begin{aligned}
& - 2pt'_{xab}N_a\text{ACov}(\tilde{x}'_{ab}, \tilde{y}_a) - 2p^2t'_{xab}N_{ab}\text{ACov}(\tilde{x}'_{ab}, \tilde{y}'_{ab}) \\
& + 2pN_aN_{ab}\text{ACov}(\tilde{y}_a, \tilde{y}'_{ab})
\end{aligned}$$

Como N_a , N_b e N_{ab} são desconhecidos, idéia será então obter a variância aproximada de cada média $\tilde{y}_i$, ou seja:

$$\begin{aligned}
A\text{Var}(\tilde{y}_d) &= \frac{1}{N_d^2} \sum_{k \in U_d} \sum_{l \in U_d} \Delta_{kl} \left(\frac{y_k - \bar{y}_{U_d}}{\pi_k} \right) \left(\frac{y_l - \bar{y}_{U_d}}{\pi_l} \right) \\
&= \frac{1}{N_d^2} \sum_{k \in U_d} \Delta_{kk} \left(\frac{y_k - \bar{y}_{U_d}}{\pi_k} \right)^2 + \frac{1}{N_d^2} \sum_{\substack{k \in U_d \\ k \neq l}} \sum_{l \in U_d} \Delta_{kl} \left(\frac{y_k - \bar{y}_{U_d}}{\pi_k} \right) \left(\frac{y_l - \bar{y}_{U_d}}{\pi_l} \right) \\
&= -\frac{k_C}{N_d^2} \left\{ (N_C - 1) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 - \sum_{\substack{k \in U_d \\ k \neq l}} \sum_{l \in U_d} (y_k - \bar{y}_{U_d})(y_l - \bar{y}_{U_d}) \right\} \\
&= -\frac{k_C}{N_d^2} \left\{ (N_C - 1) \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 - \left(\sum_{k \in U_d} (y_k - \bar{y}_{U_d}) \right)^2 + \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 \right\} \\
&= -\frac{k_C}{N_d^2} \left\{ N_C \sum_{k \in U_d} (y_k - \bar{y}_{U_d})^2 - (N_d \bar{y}_{U_d} - N_d \bar{y}_{U_d})^2 \right\} \\
&= \frac{1}{N_d^2} \left(1 - \frac{n_C}{N_C} \right) \left(\frac{N_C^2}{n_C} \right) \left(\frac{N_d - 1}{N_C - 1} \right) \sigma_{U_d}^2 \\
&\approx \frac{1}{N_d^2} \left(1 - \frac{n_C}{N_C} \right) \left(\frac{N_C^2}{n_C} \right) \left(\frac{N_d}{N_C} \right) \sigma_{U_d}^2 \\
&= \frac{1}{N_d} \left(1 - \frac{n_C}{N_C} \right) \left(\frac{N_C}{n_C} \right) \sigma_{U_d}^2 = \left(\frac{1 - f_C}{n_C P_d} \right) \sigma_{U_d}^2.
\end{aligned}$$

Dessa forma, tem-se que

$$\begin{aligned}
A\text{Var}(\hat{t}_{yr3}) &= t_{ya}^2 \left(\frac{1 - f_A}{n_A P_a} \right) \sigma_{xa}^2 + p^2 t_{xab}'^2 \left(\frac{1 - f_A}{n_A P_{ab}'} \right) \sigma_{xab}'^2 + N_a^2 \left(\frac{1 - f_A}{n_A P_a} \right) \sigma_{ya}^2 \\
&+ p^2 N_{ab}^2 \left(\frac{1 - f_A}{n_A P_{ab}'} \right) \sigma_{yab}'^2 + N_b^2 \left(\frac{1 - f_B}{n_B P_b} \right) \sigma_{yb}^2 + N_{ab}^2 (1 - p)^2 \left(\frac{1 - f_B}{n_B P_{ab}''} \right) \sigma_{yab}''^2
\end{aligned}$$

$$\begin{aligned}
& + \left[(1-p)^2 (\mu_{yab}''^2) + \mu_{yb}^2 \right] \frac{N_a N_b N_{ab} g_A g_B}{n_A N_b g_B + n_B N_a g_A} \\
& - 2N_a t_{ya} \frac{n_A}{n_a^2} \left(1 - \frac{n_A}{N_A} \right) \rho_{xya} \sigma_{xa} \sigma_{ya} - 2p^2 N_{ab} t'_{xab} \frac{n_A}{n_{ab}^{'2}} \left(1 - \frac{n_A}{N_A} \right) \rho'_{xyab} \sigma'_{xab} \sigma'_{yab}.
\end{aligned}$$

Estimador razão 4:

$$\begin{aligned}
\hat{t}_{yr4} &= \left[\frac{\hat{N}_a \tilde{y}_a + p \hat{N}_{ab,s} \tilde{y}'_{ab}}{\hat{N}_a \tilde{x}_a + p \hat{N}_{ab,s} \tilde{x}'_{ab}} \right] t_{X_A} + \hat{N}_{ab,s} (1-p) \tilde{y}''_{ab} + \hat{N}_b \tilde{y}_b \\
&= \hat{G}_A t_{X_A} + \hat{G}_B
\end{aligned}$$

onde $\hat{N}_a = N_A - \hat{N}_{ab,s}$; $\hat{N}_b = N_B - \hat{N}_{ab,s}$;

$$\hat{G}_A = \left[\frac{\hat{N}_a \tilde{y}_a + p \hat{N}_{ab,s} \tilde{y}'_{ab}}{\hat{N}_a \tilde{x}_a + p \hat{N}_{ab,s} \tilde{x}'_{ab}} \right]; \quad \hat{G}_B = (1-p) \hat{N}_{ab,s} \tilde{y}''_{ab} + \hat{N}_b \tilde{y}_b$$

$\hat{G}_A$ e $\hat{G}_B$ são estimadores não-lineares. Idéia é utilizar o método de linearização de Taylor de modo a obter uma aproximação linear desses estimadores.

Seja $\hat{G}_A = \frac{\hat{t}_{yA}^*}{\hat{t}_{xA}^*}$, com $\hat{t}_{yA}^* = N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}$ e $\hat{t}_{xA}^* = N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}$. Dessa forma,

$$\hat{G}_A \approx \hat{G}_{AL} = G_A + \frac{1}{\hat{t}_{xA}^*} (\hat{t}_{yA}^* - G_A \hat{t}_{xA}^*),$$

$$\text{onde } G_A = \frac{t_{yA}^*}{t_{xA}^*} = \left[\frac{N_a \mu_{ya} + p N_{ab} \mu'_{yab}}{N_a \mu_{xa} + p N_{ab} \mu'_{xab}} \right].$$

Portanto,

$$\hat{G}_{AL} = G_A + \frac{1}{N_a \mu_{xa} + p N_{ab} \mu'_{xab}} \left\{ \hat{N}_a \tilde{y}_a + p \hat{N}_{ab,s} \tilde{y}'_{ab} - G_A (\hat{N}_a \tilde{x}_a + p \hat{N}_{ab,s} \tilde{x}'_{ab}) \right\}.$$

Próxima etapa é linearizar os estimadores do tipo $\hat{N}_i \tilde{y}_i$, do mesmo modo que o visto para o estimador razão 3. Logo,

$$\begin{aligned}\hat{G}_{AL} &= G_{\mathcal{A}} + \frac{1}{N_a \mu_{xa} + p N_{ab} \mu'_{xab}} \left\{ \hat{N}_a \mu_{ya} + N_a \tilde{y}_a - N_a \mu_{ya} + p \mu'_{yab} \hat{N}_{ab,s} + p N_{ab} \tilde{y}'_{ab} \right. \\ &\quad \left. - p N_{ab} \mu'_{yab} - G_{\mathcal{A}} \left(\hat{N}_a \mu_{xa} + N_a \tilde{x}_a - N_a \mu_{xa} + p \mu'_{xab} \hat{N}_{ab,s} + p N_{ab} \tilde{x}'_{ab} - p N_{ab} \mu'_{xab} \right) \right\}.\end{aligned}$$

Dessa forma,

$$\begin{aligned}A\text{Var}(\hat{G}_{\mathcal{A}}) &= \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 \left\{ \mu_{ya}^2 \text{Var}(\hat{N}_a) + N_a^2 \text{Var}(\tilde{y}_a) + p^2 \mu_{yab}'^2 \text{Var}(\hat{N}_{ab,s}) + p^2 N_{ab}^2 \text{Var}(\tilde{y}'_{ab}) \right. \\ &\quad \left. + G_{\mathcal{A}}^2 \left(\mu_{xa}^2 \text{Var}(\hat{N}_a) + N_a^2 \text{Var}(\tilde{x}_a) + p^2 \mu_{xab}'^2 \text{Var}(\hat{N}_{ab,s}) + p^2 N_{ab}^2 \text{Var}(\tilde{x}'_{ab}) \right) \right\} \\ &\quad + 2p \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 N_a N_{ab} \text{ACov}(\tilde{y}_a, \tilde{y}'_{ab}) - 2 \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 G_{\mathcal{A}} N_a^2 \text{ACov}(\tilde{y}_a, \tilde{x}_a) \\ &\quad - 2p \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 G_{\mathcal{A}} N_a N_{ab} \text{ACov}(\tilde{y}_a, \tilde{x}'_{ab}) - 2p \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 G_{\mathcal{A}} N_a N_{ab} \text{ACov}(\tilde{y}'_{ab}, \tilde{x}_a) \\ &\quad - 2p^2 \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 G_{\mathcal{A}} N_{ab}^2 \text{ACov}(\tilde{y}'_{ab}, \tilde{x}'_{ab}) + 2p \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 G_{\mathcal{A}}^2 N_a N_{ab} \text{ACov}(\tilde{x}_a, \tilde{x}'_{ab}).\end{aligned}$$

Tem-se ainda que $\text{Var}(\hat{N}_a) = \text{Var}(N_{\mathcal{A}} - \hat{N}_{ab,s}) = \text{Var}(\hat{N}_{ab,s}) \Rightarrow \text{Var}(\hat{N}_b) = \text{Var}(\hat{N}_{ab,s})$.

Logo,

$$\begin{aligned}A\text{Var}(\hat{G}_{\mathcal{A}}) &= \left(\frac{1}{t_{x_{\mathcal{A}}}^*} \right)^2 \left\{ \left(\mu_{ya}^2 + p^2 \mu_{yab}'^2 + G_{\mathcal{A}}^2 \mu_{xa}^2 + p^2 G_{\mathcal{A}}^2 \mu_{xab}'^2 \right) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \right. \\ &\quad + N_a^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}} P_a} \right) \sigma_{ya}^2 + p^2 N_{ab}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}} P'_{ab}} \right) \sigma_{yab}'^2 + G_{\mathcal{A}}^2 N_a^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}} P_a} \right) \sigma_{xa}^2 \\ &\quad + p^2 G_{\mathcal{A}}^2 N_{ab}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}} P'_{ab}} \right) \sigma_{xab}'^2 - 2 G_{\mathcal{A}} N_a^2 \frac{n_{\mathcal{A}}}{n_a^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{xa} \sigma_{ya} \\ &\quad \left. - 2 p^2 G_{\mathcal{A}} N_{ab}^2 \frac{n_{\mathcal{A}}}{n_{ab}'^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{xab} \sigma'_{yab} \right\}.\end{aligned}$$

Para $\hat{G}_{\mathcal{B}}$, tem-se que

$$\hat{G}_{\mathcal{B}} \approx \hat{G}_{\mathcal{BL}} = G_{\mathcal{B}} + (1-p) \left(\mu''_{yab} \hat{N}_{ab,s} + N_{ab} \tilde{y}''_{ab} - N_{ab} \mu''_{yab} \right) + \left(\mu_{yb} \hat{N}_b + N_b \tilde{y}_b - N_b \mu_{yb} \right).$$

Logo,

$$\begin{aligned} A\text{Var}(\hat{G}_{\mathcal{B}}) &= (1-p)^2 \left\{ \mu''_{yab}{}^2 \text{Var}(\hat{N}_{ab,s}) + N_{ab}^2 \text{Var}(\tilde{y}'_{ab}) \right\} + \mu_{yb}^2 \text{Var}(\hat{N}_b) + N_b^2 \text{Var}(\tilde{y}_b) \\ &+ 2(1-p) N_b N_{ab} A\text{Cov}(\tilde{y}''_{ab}, \tilde{y}_b) \\ &= \left\{ (1-p)^2 (\mu''_{yab}{}^2) + (\mu_{yb}^2) \right\} \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} + (1-p)^2 N_{ab}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}} P''_{ab}} \right) \sigma_{yab}''^2 \\ &+ N_b^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}} P_b} \right) \sigma_{yb}^2. \end{aligned}$$

Portanto,

$$A\text{Var}(\hat{t}_{yr4}) = t_{X_{\mathcal{A}}}^2 A\text{Var}(\hat{G}_{\mathcal{A}}) + A\text{Var}(\hat{G}_{\mathcal{B}}).$$

Apêndice B

Obtenção de p para os estimadores propostos

Estimador razão 1: Variância aproximada é dada da seguinte forma:

$$A\text{Var}(\hat{t}_{yr1}) = A_1 + p^2 A_2 + A_3 + p^2 A_4 - A_5 - p^2 A_6 + B_1 + (1 - p)^2 B_2$$

onde

$$\begin{aligned} A_1 &= N_{\mathcal{A}}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2; \\ A_2 &= N_{\mathcal{A}}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}^{'2}; \\ A_3 &= (R'_a)^2 N_{\mathcal{A}}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2; \\ A_4 &= (R'_{ab})^2 N_{\mathcal{A}}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}^{'2}; \\ A_5 &= 2(R'_a)^2 \left(\frac{N_a}{n_a} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{ya} \sigma_{xa}; \\ A_6 &= 2R'_{ab} \left(\frac{N_{ab}}{n'_{ab}} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab}; \\ B_1 &= N_{\mathcal{B}}^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 \quad \text{e} \\ B_2 &= N_{\mathcal{B}}^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}^{'2}. \end{aligned}$$

Derivando a expressão da variância aproximada em relação a p , tem-se a seguinte expressão:

$$\frac{\partial A\text{Var}(\hat{t}_{yr1})}{\partial p} = 2pA_2 + 2pA_4 - 2pA_6 - 2(1 - p)B_2.$$

Fazendo $\frac{\partial A\text{Var}(\hat{t}_{yr2})}{\partial p} = 0$, o valor de p será dado por

$$p = \frac{B_2}{A_2 + A_4 - A_6 + B_2}$$

Verifica-se ainda p é ponto de mínimo, pois

$$\frac{\partial^2 A\text{Var}(\hat{t}_{yr1})}{\partial^2 p} = 2(A_2 + A_4 - A_6 + B_2) > 0$$

Estimador razão 2: Variância aproximada é dada da seguinte forma:

$$\begin{aligned} A\text{Var}(\hat{t}_{yr2}) &= t_{x\mathcal{A}}^2 \left\{ \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}}^2 A_4 - 2G_{\mathcal{A}} A_5 - 2p^2 G_{\mathcal{A}} A_6) \right\} \\ &+ B_1 + (1-p)^2 B_2 \end{aligned}$$

onde

$$\begin{aligned} t_{x\mathcal{A}}^* &= t_{ya} + p t'_{yab} = N_a \mu_{xa} + p N_{ab} \mu_{xab} & G_{\mathcal{A}} &= \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} = \frac{N_a \mu_{ya} + p N_{ab} \mu_{yab}}{N_a \mu_{xa} + p N_{ab} \mu_{xab}} \\ A_1 &= N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{ya}^2 & A_2 &= N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{yab}^{'2} \\ A_3 &= N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P_a \sigma_{xa}^2 & A_4 &= N_{\mathcal{A}}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}}} \right) P'_{ab} \sigma_{xab}^{'2} \\ A_5 &= \left(\frac{N_a}{n_a} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{ya} \sigma_{xa} & A_6 &= \left(\frac{N_{ab}}{n'_{ab}} \right)^2 n_{\mathcal{A}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{yab} \sigma'_{xab} \\ B_1 &= N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P_b \sigma_{yb}^2 & B_2 &= N_{\mathcal{B}}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}}} \right) P''_{ab} \sigma_{yab}^{'2} \end{aligned}$$

Derivando em relação a p , tem-se a seguinte expressão:

$$\frac{\partial A\text{Var}(\hat{t}_{yr2})}{\partial p} = -\frac{1}{t_{x\mathcal{A}}^{*3}} \left\{ 2t_{x\mathcal{A}}^2 [A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}}^2 A_4 - 2G_{\mathcal{A}} A_5 - 2p^2 G_{\mathcal{A}} A_6] N_{ab} \mu'_{xab} \right\}$$

$$\begin{aligned}
& + \frac{1}{t_{x\mathcal{A}}^2} \left\{ t_{X\mathcal{A}}^2 \left[2pA_2 + 2\frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_3 N_{ab} \mu'_{yab} - 2\frac{G_{\mathcal{A}}^2}{t_{x\mathcal{A}}^*} A_3 N_{ab} \mu'_{xab} + 2pG_{\mathcal{A}}^2 A_4 \right. \right. \\
& + 2p^2 \frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_4 N_{ab} \mu'_{yab} - 2p^2 \frac{G_{\mathcal{A}}^2}{t_{x\mathcal{A}}^*} A_4 N_{ab} \mu'_{xab} - 2\frac{N_{ab} \mu'_{yab} A_5}{t_{x\mathcal{A}}^*} + 2\frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_5 N_{ab} \mu'_{xab} \\
& \left. \left. - 4pG_{\mathcal{A}} A_6 - 2p^2 \frac{N_{ab} \mu'_{yab} A_6}{t_{x\mathcal{A}}^*} + 2p^2 \frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_6 N_{ab} \mu'_{xab} \right] \right\} - 2(1-p)B_2
\end{aligned}$$

Fazendo $\frac{\partial A\text{Var}(\hat{t}_{yr2})}{\partial p} = 0$, o valor de p que minimiza a a variância aproximada de $\hat{t}_{yr2}$ é a menor raiz em valor absoluto do polinômio:

$$k_6 p^6 + k_5 p^5 + k_4 p^4 + k_3 p^3 + k_2 p^2 + k_1 p + k_0 = 0$$

onde

$$k_6 = B_2 N_{ab}^5 \mu_{xab}'^5$$

$$k_5 = 5B_2 N_a \mu_{xa} N_{ab}^4 \mu_{xab}'^4 - B_2 N_{ab}^5 \mu_{xab}'^5$$

$$k_4 = -5B_2 N_a \mu_{xa} N_{ab}^4 \mu_{xab}'^4 + 10B_2 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu_{xab}'^3$$

$$\begin{aligned}
k_3 = & -t_{X\mathcal{A}}^2 N_{ab}^2 \mu_{xab}' A_4 N_a \mu_{ya} \mu'_{yab} + t_{X\mathcal{A}}^2 N_{ab}^2 \mu_{xab}'^2 A_6 N_a \mu_{ya} - 3t_{X\mathcal{A}}^2 N_{ab}^2 \mu_{xab}' A_6 \mu'_{yab} N_a \mu_{xa} \\
& + 2t_{X\mathcal{A}}^2 A_4 N_{ab}^2 \mu_{yab}'^2 N_a \mu_{xa} + 10B_2 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu_{xab}'^2 - 10B_2 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu_{xab}'^3 \\
& + t_{X\mathcal{A}}^2 N_{ab}^2 \mu_{xab}'^2 A_2 N_a \mu_{xa}
\end{aligned}$$

$$\begin{aligned}
k_2 = & t_{X\mathcal{A}}^2 N_{ab} \mu'_{xab} A_4 N_a^2 \mu_{ya}^2 + 5B_2 N_a^4 \mu_{xa}^4 N_{ab} \mu'_{xab} - t_{X\mathcal{A}}^2 N_{ab}^3 \mu_{xab}'^3 A_1 + 2t_{X\mathcal{A}}^2 N_{ab} \mu'_{xab} A_2 N_a^2 \mu_{xa}^2 \\
& - 3t_{X\mathcal{A}}^2 A_6 N_{ab} \mu'_{yab} N_a^2 \mu_{xa}^2 + 3t_{X\mathcal{A}}^2 A_4 N_a^2 \mu_{ya} N_{ab} \mu'_{yab} \mu_{xa} - t_{X\mathcal{A}}^2 N_{ab} \mu'_{xab} A_6 N_a^2 \mu_{ya} \mu_{xa} \\
& - 10B_2 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu_{xab}'^2 - t_{X\mathcal{A}}^2 N_{ab}^3 \mu_{xab}'^2 A_3 \mu_{yab}' + 2t_{X\mathcal{A}}^2 N_{ab}^3 \mu_{xab}'^2 A_5 \mu_{yab}'
\end{aligned}$$

$$\begin{aligned}
k_1 = & t_{X\mathcal{A}}^2 N_{ab}^2 \mu'_{xab} A_5 \mu'_{yab} N_a \mu_{xa} + t_{X\mathcal{A}}^2 A_4 N_a^3 \mu_{ya}^2 \mu_{xa} + t_{X\mathcal{A}}^2 A_2 N_a^3 \mu_{xa}^3 - 2t_{X\mathcal{A}}^2 N_{ab}^2 \mu_{xab}'^2 A_1 N_a \mu_{xa} \\
& + B_2 N_a^5 \mu_{xa}^5 + 3t_{X\mathcal{A}}^2 N_{ab}^2 \mu_{xab}'^2 A_5 N_a \mu_{ya} - 2t_{X\mathcal{A}}^2 A_6 N_a^3 \mu_{ya} \mu_{xa}^2 - 5B_2 N_a^4 \mu_{xa}^4 N_{ab} \mu'_{xab} \\
& - 3t_{X\mathcal{A}}^2 N_{ab}^2 \mu'_{xab} A_3 N_a \mu_{ya} \mu'_{yab} + t_{X\mathcal{A}}^2 A_3 N_{ab}^2 \mu_{yab}'^2 N_a \mu_{xa}
\end{aligned}$$

$$\begin{aligned}
k_0 &= -B_2 N_a^5 \mu_{xa}^5 + 3t_{X_A}^2 N_{ab} \mu'_{xab} A_5 N_a^2 \mu_{ya} \mu_{xa} - t_{X_A}^2 N_{ab} \mu'_{yab} A_5 N_a^2 \mu_{xa}^2 - t_{X_A}^2 N_{ab} \mu'_{xab} A_1 N_a^2 \mu_{xa}^2 \\
&- 2t_{X_A}^2 N_{ab} \mu'_{xab} A_3 N_a^2 \mu_{ya}^2 + t_{X_A}^2 A_3 N_{ab} \mu'_{yab} N_a^2 \mu_{ya} \mu_{xa}
\end{aligned}$$

Verifica-se ainda que

$$\begin{aligned}
\frac{\partial^2 A \text{Var}(\hat{t}_{yr2})}{\partial^2 p} &= \frac{1}{t_{xA}^{*4}} \left\{ 6t_{X_A}^2 [A_1 + p^2 A_2 + G_A^2 A_3 + p^2 G_A^2 A_4 - 2G_A A_5 - 2p^2 G_A A_6] N_{ab}^2 \mu_{xab}'^2 \right\} \\
&- \frac{1}{t_{xA}^{*3}} \left\{ 4t_{X_A}^2 \left[2pA_2 + 2\frac{G_A}{t_{xA}^*} A_3 N_{ab} \mu'_{yab} - 2\frac{G_A^2}{t_{xA}^*} A_3 N_{ab} \mu'_{xab} + 2pG_A^2 A_4 \right. \right. \\
&+ 2p^2 \frac{G_A}{t_{xA}^*} A_4 N_{ab} \mu'_{yab} - 2p^2 \frac{G_A^2}{t_{xA}^*} A_4 N_{ab} \mu'_{xab} - 2\frac{N_{ab} \mu'_{yab} A_5}{t_{xA}^*} + 2\frac{G_A}{t_{xA}^*} A_5 N_{ab} \mu'_{xab} \\
&- 4p\frac{G_A}{t_{xA}^*} A_6 - 2p^2 \frac{N_{ab} \mu'_{yab} A_6}{t_{xA}^*} + 2p^2 \frac{G_A}{t_{xA}^*} A_6 N_{ab} \mu'_{xab} \left. \right] N_{ab} \mu'_{xab} \left. \right\} \\
&+ \frac{1}{t_{xA}^*} \left\{ t_{X_A}^2 [2A_2 + 2G_A^2 A_4 - 4G_A A_6 + 2\frac{N_{ab} \mu_{yab}'^2 A_3}{t_{xA}^{*2}} - 8\frac{G_A A_3 N_{ab} \mu'_{yab} \mu'_{xab}}{t_{xA}^{*2}} \right. \\
&+ 6\frac{G_A^2 A_3 N_{ab}^2 \mu_{xab}'^2}{t_{xA}^{*2}} + 8p\frac{G_A}{t_{xA}^*} A_4 N_{ab} \mu'_{yab} - 8p\frac{G_A^2}{t_{xA}^*} A_4 N_{ab} \mu'_{xab} + 2p^2 \frac{N_{ab}^2 \mu_{yab}'^2 A_4}{t_{xA}^{*2}} \\
&- 8p^2 \frac{G_A}{t_{xA}^{*2}} A_4 N_{ab}^2 \mu'_{yab} \mu'_{xab} + 6p^2 \frac{G_A^2}{t_{xA}^{*2}} A_4 N_{ab}^2 \mu_{xab} + 4\frac{N_{ab} \mu'_{yab} A_5 \mu'_{xab}}{t_{xA}^{*2}} \\
&- 4\frac{G_A}{t_{xA}^{*2}} A_5 N_{ab}^2 \mu_{xab}'^2 - 8p\frac{N_{ab} \mu'_{yab} A_6}{t_{xA}^*} + 8p\frac{G_A}{t_{xA}^*} A_6 N_{ab} \mu'_{xab} + 4p^2 \frac{N_{ab}^2 \mu'_{yab} A_6 \mu'_{xab}}{t_{xA}^{*2}} \\
&\left. \left. - 4p^2 \frac{G_A}{t_{xA}^{*2}} A_6 N_{ab}^2 \mu_{xab}'^2 \right] \right\} + 2B_2
\end{aligned}$$

É possível mostrar que de fato $\frac{\partial^2 A \text{Var}(\hat{t}_{yr2})}{\partial^2 p} > 0$, quando avaliada na menor raiz do polinômio, ou seja que a menor raiz do polinômio apresentado é ponto de mínimo. Uma alternativa é substituir o valor de p na expressão da segunda derivada.

Sob a situação em que $U_B \subset U_A$, tem-se que

$$\begin{aligned}
\hat{t}_{yr2} &= \left(\frac{N_a \tilde{y}_a + p N_{ab} \tilde{y}'_{ab}}{N_a \tilde{x}_a + p N_{ab} \tilde{x}'_{ab}} \right) t_{X_A} + (1-p) N_{ab} \tilde{y}''_{ab} \\
&= \hat{G}_A t_{X_A} + \hat{G}_B
\end{aligned}$$

$\hat{G}_{\mathcal{A}}$ e $\hat{G}_{\mathcal{B}}$ são quantidades similares às do estimador 2. Tem-se neste caso, que $N_{\mathcal{A}}$ e $N_{\mathcal{B}}$ são conhecidos, e que $N_b = 0$ e que $N_{ab} = N_B$. Logo, $N_a = N_A - N_{ab}$.

A variância aproximada de $\hat{t}_{yr2}$ é dada por:

$$\begin{aligned} A\text{Var}(\hat{t}_{yr2}) &= t_{X_{\mathcal{A}}}^2 \left\{ \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}}^2 A_4 - 2G_{\mathcal{A}} A_5 - 2p^2 G_{\mathcal{A}} A_6) \right\} \\ &+ (1-p)^2 B_1 \end{aligned}$$

onde

$$\begin{aligned} t_{x\mathcal{A}}^* &= t_{ya} + pt'_{yab} = N_a \mu_{ya} + p N_{ab} \mu'_{yab} \\ G_{\mathcal{A}} &= \frac{t_{y\mathcal{A}}^*}{t_{x\mathcal{A}}^*} = \frac{N_a \mu_{ya} + p N_{ab} \mu'_{yab}}{N_a \mu_{xa} + p N_{ab} \mu'_{xab}} \end{aligned}$$

Derivando a expressão da variância aproximada de $\hat{t}_{yr2}$ em relação a p , tem-se a seguinte expressão:

$$\begin{aligned} \frac{\partial A\text{Var}(\hat{t}_{yr2})}{\partial p} &= -\frac{1}{t_{x\mathcal{A}}^{*3}} \{ 2t_{X_{\mathcal{A}}}^2 [A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}}^2 A_4 - 2G_{\mathcal{A}} A_5 - 2p^2 G_{\mathcal{A}} A_6] N_{ab} \mu'_{xab} \} \\ &+ \frac{1}{t_{x\mathcal{A}}^{*2}} \left\{ t_{X_{\mathcal{A}}}^2 \left[2p A_2 + 2 \frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_3 N_{ab} \mu'_{yab} - 2 \frac{G_{\mathcal{A}}^2}{t_{x\mathcal{A}}^*} A_3 N_{ab} \mu'_{xab} + 2p G_{\mathcal{A}}^2 A_4 \right. \right. \\ &+ 2p^2 \frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_4 N_{ab} \mu'_{yab} - 2p^2 \frac{G_{\mathcal{A}}^2}{t_{x\mathcal{A}}^*} A_4 N_{ab} \mu'_{xab} - 2 \frac{N_{ab} \mu'_{yab} A_5}{t_{x\mathcal{A}}^*} + 2 \frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_5 N_{ab} \mu'_{xab} \\ &\left. \left. - 4p G_{\mathcal{A}} A_6 - 2p^2 \frac{N_{ab} \mu'_{yab} A_6}{t_{x\mathcal{A}}^*} + 2p^2 \frac{G_{\mathcal{A}}}{t_{x\mathcal{A}}^*} A_6 N_{ab} \mu'_{xab} \right] \right\} - 2(1-p) B_1 \end{aligned}$$

Fazendo $\frac{\partial A\text{Var}(\hat{t}_{yr2})}{\partial p} = 0$, o valor de p que minimiza a expressão da variância aproximada de $\hat{t}_{yr2}$ é dado como sendo a menor raiz do seguinte polinômio de grau 6:

$$k_6 p^6 + k_5 p^5 + k_4 p^4 + k_3 p^3 + k_2 p^2 + k_1 p + k_0 = 0,$$

onde:

$$k_6 = B_1 N_{ab}^5 \mu_{xab}'^5$$

$$k_5 = 5B_1 N_a \mu_{xa} N_{ab}^4 \mu_{xab}'^4 - B_1 N_{ab}^5 \mu_{xab}'^5$$

$$k_4 = -5B_1 N_a \mu_{xa} N_{ab}^4 \mu_{xab}'^4 + 10B_1 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu_{xab}'^3$$

$$\begin{aligned} k_3 = & -t_{X_A}^2 N_{ab}^2 \mu_{xab}' A_4 N_a \mu_{ya} \mu_{yab}' + t_{X_A}^2 N_{ab}^2 \mu_{xab}'^2 A_6 N_a \mu_{ya} - 3t_{X_A}^2 N_{ab}^2 \mu_{xab}' A_6 \mu_{yab}' N_a \mu_{xa} \\ & + 2t_{X_A}^2 A_4 N_{ab}^2 \mu_{yab}'^2 N_a \mu_{xa} + 10B_1 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu_{xab}'^2 - 10B_1 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu_{xab}'^3 \\ & + t_{X_A}^2 N_{ab}^2 \mu_{xab}'^2 A_2 N_a \mu_{xa} \end{aligned}$$

$$\begin{aligned} k_2 = & t_{X_A}^2 N_{ab} \mu_{xab}' A_4 N_a^2 \mu_{ya}^2 + 5B_1 N_a^4 \mu_{xa}^4 N_{ab} \mu_{xab}' - t_{X_A}^2 N_{ab}^3 \mu_{xab}'^3 A_1 + 2t_{X_A}^2 N_{ab} \mu_{xab}' A_2 N_a^2 \mu_{xa}^2 \\ & - 3t_{X_A}^2 A_6 N_{ab} \mu_{yab}' N_a^2 \mu_{xa}^2 + 3t_{X_A}^2 A_4 N_a^2 \mu_{ya} N_{ab} \mu_{yab}' \mu_{xa} - t_{X_A}^2 N_{ab} \mu_{xab}' A_6 N_a^2 \mu_{ya} \mu_{xa} \\ & - 10B_1 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu_{xab}'^2 - t_{X_A}^2 N_{ab}^3 \mu_{xab}'^2 A_3 \mu_{yab}' + 2t_{X_A}^2 N_{ab}^3 \mu_{xab}'^2 A_5 \mu_{yab}' \end{aligned}$$

$$\begin{aligned} k_1 = & t_{X_A}^2 N_{ab}^2 \mu_{xab}' A_5 \mu_{yab}' N_a \mu_{xa} + t_{X_A}^2 A_4 N_a^3 \mu_{ya}^2 \mu_{xa} + t_{X_A}^2 A_2 N_a^3 \mu_{xa}^3 - 2t_{X_A}^2 N_{ab}^2 \mu_{xab}'^2 A_1 N_a \mu_{xa} \\ & + B_1 N_a^5 \mu_{xa}^5 + 3t_{X_A}^2 N_{ab}^2 \mu_{xab}'^2 A_5 N_a \mu_{ya} - 2t_{X_A}^2 A_6 N_a^3 \mu_{ya} \mu_{xa}^2 - 5B_1 N_a^4 \mu_{xa}^4 N_{ab} \mu_{xab}' \\ & - 3t_{X_A}^2 N_{ab}^2 \mu_{xab}' A_3 N_a \mu_{ya} \mu_{yab}' + t_{X_A}^2 A_3 N_{ab}^2 \mu_{yab}'^2 N_a \mu_{xa} \end{aligned}$$

$$\begin{aligned} k_0 = & -B_1 N_a^5 \mu_{xa}^5 + 3t_{X_A}^2 N_{ab} \mu_{xab}' A_5 N_a^2 \mu_{ya} \mu_{xa} - t_{X_A}^2 N_{ab} \mu_{yab}' A_5 N_a^2 \mu_{xa}^2 - t_{X_A}^2 N_{ab} \mu_{xab}' A_1 N_a^2 \mu_{xa}^2 \\ & - 2t_{X_A}^2 N_{ab} \mu_{xab}' A_3 N_a^2 \mu_{ya}^2 + t_{X_A}^2 A_3 N_{ab} \mu_{yab}' N_a^2 \mu_{ya} \mu_{xa} \end{aligned}$$

Além disso,

$$\begin{aligned}
\frac{\partial^2 A\text{Var}(\hat{t}_{yr2})}{\partial^2 p} &= \frac{1}{t_{xA}^{*4}} \left\{ 6t_{xA}^2 [A_1 + p^2 A_2 + G_A^2 A_3 + p^2 G_A^2 A_4 - 2G_A A_5 - 2p^2 G_A A_6] N_{ab}^2 \mu_{xab}'^2 \right\} \\
&- \frac{1}{t_{xA}^{*3}} \left\{ 4t_{xA}^2 \left[2pA_2 + 2\frac{G_A}{t_{xA}^*} A_3 N_{ab} \mu_{yab}' - 2\frac{G_A^2}{t_{xA}^*} A_3 N_{ab} \mu_{xab}' + 2pG_A^2 A_4 \right. \right. \\
&+ 2p^2 \frac{G_A}{t_{xA}^*} A_4 N_{ab} \mu_{yab}' - 2p^2 \frac{G_A^2}{t_{xA}^*} A_4 N_{ab} \mu_{xab}' - 2\frac{N_{ab} \mu_{yab}' A_5}{t_{xA}^*} + 2\frac{G_A}{t_{xA}^*} A_5 N_{ab} \mu_{xab}' \\
&- 4p\frac{G_A}{t_{xA}^*} A_6 - 2p^2 \frac{N_{ab} \mu_{yab}' A_6}{t_{xA}^*} + 2p^2 \frac{G_A}{t_{xA}^*} A_6 N_{ab} \mu_{xab}' \left. \right] N_{ab} \mu_{xab}' \left. \right\} \\
&+ \frac{1}{t_{xA}^*} \left\{ t_{xA}^2 [2A_2 + 2G_A^2 A_4 - 4G_A A_6 + 2\frac{N_{ab}^2 \mu_{yab}'^2 A_3}{t_{xA}^{*2}} - 8\frac{G_A A_3 N_{ab}^2 \mu_{yab}' \mu_{xab}'}{t_{xA}^{*2}} \right. \\
&+ 6\frac{G_A^2 A_3 N_{ab}^2 \mu_{xab}'^2}{t_{xA}^{*2}} + 8p\frac{G_A}{t_{xA}^*} A_4 N_{ab} \mu_{yab}' - 8p\frac{G_A^2}{t_{xA}^*} A_4 N_{ab} \mu_{xab}' + 2p^2 \frac{N_{ab}^2 \mu_{yab}'^2 A_4}{t_{xA}^{*2}} \\
&- 8p^2 \frac{G_A}{t_{xA}^*} A_4 N_{ab}^2 \mu_{yab}' \mu_{xab}' + 6p^2 \frac{G_A^2}{t_{xA}^*} A_4 N_{ab}^2 \mu_{xab}' + 4\frac{N_{ab}^2 \mu_{yab}' A_5 \mu_{xab}'}{t_{xA}^{*2}} \\
&- 4\frac{G_A}{t_{xA}^{*2}} A_5 N_{ab}^2 \mu_{xab}'^2 - 8p\frac{N_{ab} \mu_{yab}' A_6}{t_{xA}^*} + 8p\frac{G_A}{t_{xA}^*} A_6 N_{ab} \mu_{xab}' + 4p^2 \frac{N_{ab}^2 \mu_{yab}' A_6 \mu_{xab}'}{t_{xA}^{*2}} \\
&\left. \left. - 4p^2 \frac{G_A}{t_{xA}^{*2}} A_6 N_{ab}^2 \mu_{xab}'^2 \right] \right\} + 2B_1
\end{aligned}$$

É possível mostrar que de fato $\frac{\partial^2 A\text{Var}(\hat{t}_{yr2})}{\partial^2 p} > 0$, quando avaliada na menor raiz do polinômio, ou seja que a menor raiz do polinômio apresentado é ponto de mínimo. Uma alternativa é substituir o valor de p na expressão da segunda derivada.

Estimador razão 3: Variância aproximada é dada da seguinte forma:

$$A\text{Var}(\hat{t}_{yr3}) = A_1 + p^2 A_2 + A_3 + p^2 A_4 - A_5 - p^2 A_6 + B_1 + (1-p)^2 B_2 + (1-p)^2 M_1 + M_2$$

onde

$$\begin{aligned}
A_1 &= t_{ya}^2 \left(\frac{1-f_A}{n_A P_a} \right) \sigma_{xa}^2 & A_2 &= t_{xab}'^2 \left(\frac{1-f_A}{n_A P_{ab}'} \right) \sigma_{xab}'^2 \\
A_3 &= N_a^2 \left(\frac{1-f_A}{n_A P_a} \right) \sigma_{ya}^2 & A_4 &= N_{ab}^2 \left(\frac{1-f_A}{n_A P_{ab}'} \right) \sigma_{yab}'^2
\end{aligned}$$

$$\begin{aligned}
A_5 &= 2N_a t_{ya} \frac{n_{\mathcal{A}}}{n_a^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}}\right) \rho_{xya} \sigma_{xa} \sigma_{ya} & A_6 &= N_{ab} t'_{xab} \frac{n_{\mathcal{A}}}{n_{ab}^{'2}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}}\right) \rho'_{xyab} \sigma'_{xab} \sigma'_{yab} \\
B_1 &= N_b^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}} P_b}\right) \sigma_{yb}^2 & B_2 &= N_{ab}^2 (1 - p)^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}} P_{ab}''}\right) \sigma_{yab}^{'2} \\
M_1 &= (\mu_{yab}^{'2}) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} & M_2 &= \mu_{yb}^2 \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}}
\end{aligned}$$

Derivando com relação a p , tem-se a seguinte expressão:

$$\frac{\partial A\text{Var}(\hat{t}_{yr3})}{\partial p} = 2pA_2 + 2pA_4 - 2pA_6 - 2(1 - p)B_2 - 2(1 - p)M_1.$$

Fazendo $\frac{\partial A\text{Var}(\hat{t}_{yr3})}{\partial p} = 0$, o valor de p é dado por

$$p = \frac{M_1 + B_2}{A_2 + A_4 - A_6 + B_2 + M_1}.$$

Neste caso, p também é ponto de mínimo, pois

$$\frac{\partial^2 A\text{Var}(\hat{t}_{yr3})}{\partial^2 p} = 2(A_2 + A_4 - A_6 + B_2 + M_1) > 0.$$

Estimador razão 4: Variância aproximada é dada da seguinte forma:

$$\begin{aligned} A\text{Var}(\hat{t}_{yr4}) &= \left(\frac{t_{X_{\mathcal{A}}}}{t_{x_{\mathcal{A}}}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}}^2 A_4 - G_{\mathcal{A}} A_5 - p^2 G_{\mathcal{A}} A_6 \\ &+ M_1 + p^2 M_2 + G_{\mathcal{A}}^2 M_3 + p^2 G_{\mathcal{A}}^2 M_4) + B_1 + (1-p)^2 B_2 + M_5 + (1-p)^2 M_6. \end{aligned}$$

onde

$$\begin{aligned} A_1 &= N_a^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P_a} \right) \sigma_{ya}^2 & A_2 &= N_{ab}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P'_{ab}} \right) \sigma_{yab}^{'2} \\ A_3 &= N_a^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P_a} \right) \sigma_{xa}^2 & A_4 &= N_{ab}^2 \left(\frac{1-f_{\mathcal{A}}}{n_{\mathcal{A}} P'_{ab}} \right) \sigma_{xab}^{'2} \\ A_5 &= 2N_a^2 \frac{n_{\mathcal{A}}}{n_a^2} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho_{xya} \sigma_{xa} \sigma_{ya} & A_6 &= 2N_{ab}^2 \frac{n_{\mathcal{A}}}{n_{ab}^{'2}} \left(1 - \frac{n_{\mathcal{A}}}{N_{\mathcal{A}}} \right) \rho'_{xyab} \sigma'_{xab} \sigma'_{yab} \\ B_1 &= N_b^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}} P_b} \right) \sigma_{yb}^2 & B_2 &= N_{ab}^2 \left(\frac{1-f_{\mathcal{B}}}{n_{\mathcal{B}} P''_{ab}} \right) \sigma_{yab}^{''2} \end{aligned}$$

$$\begin{aligned} M_1 &= (\mu_{ya}^2) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \\ M_2 &= (\mu_{yab}^{'2}) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \\ M_3 &= (\mu_{xa}^2) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \\ M_4 &= (\mu_{xab}^{'2}) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \\ M_5 &= (\mu_{yb}^2) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \\ M_6 &= (\mu_{yab}^{''2}) \frac{N_a N_b N_{ab} g_{\mathcal{A}} g_{\mathcal{B}}}{n_{\mathcal{A}} N_b g_{\mathcal{B}} + n_{\mathcal{B}} N_a g_{\mathcal{A}}} \end{aligned}$$

Derivando com relação a p , tem-se a seguinte expressão:

$$\begin{aligned}
\frac{\partial A\text{Var}(\hat{t}_{yr4})}{\partial p} = & -2t_{XA}^2 \left\{ A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}} A_4 - G_{\mathcal{A}} A_5 - p^2 G_{\mathcal{A}} A_6 \right. \\
& + M_1 + p^2 M_2 + G_{\mathcal{A}}^2 M_3 + p^2 G_{\mathcal{A}} M_4 \left. \right\} N_{ab} \mu'_{xab} / t_{xA}^{*3} \\
& + t_{XA}^2 \left\{ 2p A_2 + 2 \frac{G_{\mathcal{A}}}{t_{xA}^*} A_3 N_{ab} \mu'_{yab} - 2 \frac{G_{\mathcal{A}}^2}{t_{xA}^*} A_3 N_{ab} \mu'_{xab} + 2p G_{\mathcal{A}}^2 A_4 \right. \\
& + 2p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} A_4 N_{ab} \mu'_{yab} - 2p^2 \frac{G_{\mathcal{A}}^2}{t_{xA}^*} A_4 N_{ab} \mu'_{xab} - \frac{N_{ab} \mu'_{yab} A_5}{t_{xA}^*} \\
& + \frac{G_{\mathcal{A}}}{t_{xA}^*} A_5 N_{ab} \mu'_{xab} - 2p G_{\mathcal{A}} A_6 - p^2 \frac{N_{ab} \mu'_{yab} A_6}{t_{xA}^*} + p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} A_6 N_{ab} \mu'_{xab} \\
& + 2p M_2 + 2 \frac{G_{\mathcal{A}}}{t_{xA}^*} M_3 N_{ab} \mu'_{yab} - 2 \frac{G_{\mathcal{A}}^2}{t_{xA}^*} M_3 N_{ab} \mu'_{xab} + 2p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} M_4 N_{ab} \mu'_{yab} \\
& \left. - 2p^2 \frac{G_{\mathcal{A}}^2 M_4 N_{ab} \mu'_{xab}}{t_{xA}^*} \right\} / t_{xA}^{*2} - 2(1-p) B_2 - 2(1-p) M_6.
\end{aligned}$$

Fazendo $\frac{\partial A\text{Var}(\hat{t}_{yr4})}{\partial p} = 0$, o valor ótimo de p será a menor raiz em valor absoluto do seguinte polinômio de grau 6:

$$k_6 p^6 + k_5 p^5 + k_4 p^4 + k_3 p^3 + k_2 p^2 + k_1 p + k_0 = 0,$$

onde:

$$k_6 = 2B_2 N_{ab}^5 \mu'_{xab}{}^5 + 2M_6 N_{ab}^5 \mu'_{xab}{}^5$$

$$k_5 = -2B_2 N_{ab}^5 \mu'_{xab}{}^5 - 2M_6 N_{ab}^5 \mu'_{xab}{}^5 + 10B_2 N_a \mu_{xa} N_{ab}^4 \mu'_{xab}{}^4 + 10M_6 N_a \mu_{xa} N_{ab}^4 \mu'_{xab}{}^4$$

$$\begin{aligned}
k_4 = & -10B_2 N_a \mu_{xa} N_{ab}^4 \mu'_{xab}{}^4 + 20B_2 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu'_{xab}{}^3 - 10M_6 N_a \mu_{xa} N_{ab}^4 \mu'_{xab}{}^4 \\
& + 20M_6 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu'_{xab}{}^3
\end{aligned}$$

$$\begin{aligned}
k_3 = & 4t_{XA}^2 A_4 N_{ab}^2 \mu'_{yab}{}^2 N_a \mu_{xa} - 2t_{XA}^2 N_{ab}^2 \mu'_{xab}{}^2 A_4 N_a \mu_{ya} \mu'_{yab} - 20B_2 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu'_{xab}{}^3 \\
& + 2t_{XA}^2 N_{ab}^2 \mu'_{xab}{}^2 A_2 N_a \mu_{xa} + 2t_{XA}^2 N_{ab}^2 \mu'_{xab}{}^2 M_2 N_a \mu_{xa} - 2t_{XA}^2 N_{ab}^2 \mu'_{xab}{}^2 M_4 N_a \mu_{ya} \mu'_{yab} \\
& + 20M_6 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu'_{xab}{}^2 - 3t_{XA}^2 N_{ab}^2 \mu'_{xab}{}^2 A_6 \mu'_{yab} N_a \mu_{xa} - 20M_6 N_a^2 \mu_{xa}^2 N_{ab}^3 \mu'_{xab}{}^3 \\
& + 20B_2 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu'_{xab}{}^2 + t_{XA}^2 N_{ab}^2 \mu'_{xab}{}^2 A_6 N_a \mu_{ya} + 4t_{XA}^2 M_4 N_{ab}^2 \mu'_{yab}{}^2 N_a \mu_{xa}
\end{aligned}$$

$$\begin{aligned}
k_2 &= 4t_{XA}^2 N_{ab}\mu'_{xab} A_2 N_a^2 \mu_{xa}^2 - 20B_2 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu_{xab}'^2 - 2t_{XA}^2 N_{ab}^3 \mu_{xab}'^3 A_1 \\
&- 2t_{XA}^2 N_{ab}^3 \mu_{xab}'^2 M_3 \mu_{yab}'^2 - 2t_{XA}^2 N_{ab}\mu'_{xab} A_4 N_a^2 \mu_{ya}^2 - 3t_{XA}^2 A_6 N_{ab}\mu'_{yab} N_a^2 \mu_{xa}^2 \\
&- 2t_{XA}^2 N_{ab}^3 \mu_{xab}'^3 M_1 + 2t_{XA}^2 N_{ab}^3 \mu_{xab}'^2 A_5 \mu_{yab}' + 6t_{XA}^2 M_4 N_a^2 \mu_{ya} N_{ab}\mu'_{yab} \mu_{xa} \\
&- 2t_{XA}^2 N_{ab}\mu'_{xab} M_4 N_a^2 \mu_{ya}^2 + 10B_2 N_a^4 \mu_{xa}^4 N_{ab}\mu'_{xab} + 6t_{XA}^2 A_4 N_a^2 \mu_{ya} N_{ab}\mu'_{yab} \mu_{xa} \\
&+ 4t_{XA}^2 N_{ab}\mu'_{xab} M_2 N_a^2 \mu_{xa}^2 - t_{XA}^2 N_{ab}\mu'_{xab} A_6 N_a^2 \mu_{ya} \mu_{xa} - 2t_{XA}^2 N_{ab}^3 \mu_{xab}'^3 A_3 \mu_{yab}'^2 \\
&- 20M_6 N_a^3 \mu_{xa}^3 N_{ab}^2 \mu_{xab}'^2 + 10M_6 N_a^4 \mu_{xa}^4 N_{ab}\mu'_{xab} \\
\\
k_1 &= -10B_2 N_a^4 \mu_{xa}^4 N_{ab}\mu'_{xab} + 2B_2 N_a^5 \mu_{xa}^5 + 2t_{XA}^2 A_2 N_a^3 \mu_{xa}^3 \\
&- 6t_{XA}^2 N_{ab}^2 \mu_{xab}'^2 A_3 N_a \mu_{ya} \mu_{yab}' + 2M_6 N_a^5 \mu_{xa}^5 + 3t_{XA}^2 N_{ab}^2 \mu_{xab}'^2 A_5 N_a \mu_{ya} \\
&+ t_{XA}^2 N_{ab}^2 \mu_{xab}'^2 A_5 \mu_{yab}' N_a \mu_{xa} - 4t_{XA}^2 N_{ab}^2 \mu_{xab}'^2 M_1 N_a \mu_{xa} \\
&- 10M_6 N_a^4 \mu_{xa}^4 N_{ab}\mu'_{xab} - 6t_{XA}^2 N_{ab}^2 \mu_{xab}'^2 M_3 N_a \mu_{ya} \mu_{yab}' + 2t_{XA}^2 M_3 N_{ab}^2 \mu_{yab}'^2 N_a \mu_{xa} \\
&+ 2t_{XA}^2 M_2 N_a^3 \mu_{xa}^3 + 2t_{XA}^2 A_3 N_{ab}^2 \mu_{yab}'^2 N_a \mu_{xa} - 4t_{XA}^2 N_{ab}^2 \mu_{xab}'^2 A_1 N_a \mu_{xa} \\
&+ 2t_{XA}^2 M_4 N_a^3 \mu_{ya}^2 \mu_{xa} - 2t_{XA}^2 A_6 N_a^3 \mu_{ya} \mu_{xa}^2 + 2t_{XA}^2 A_4 N_a^3 \mu_{ya}^2 \mu_{xa} \\
\\
k_0 &= -2B_2 N_a^5 \mu_{xa}^5 - 2M_6 N_a^5 \mu_{xa}^5 + 3t_{XA}^2 N_{ab}\mu'_{xab} A_5 N_a^2 \mu_{ya} \mu_{xa} \\
&- 2t_{XA}^2 N_{ab}\mu'_{xab} A_1 N_a^2 \mu_{xa}^2 - 4t_{XA}^2 N_{ab}\mu'_{xab} A_3 N_a^2 \mu_{ya}^2 - 2t_{XA}^2 N_{ab}\mu'_{xab} M_1 N_a^2 \mu_{xa}^2 \\
&- 4t_{XA}^2 N_{ab}\mu'_{xab} M_3 N_a^2 \mu_{ya}^2 + 2t_{XA}^2 A_3 N_{ab}\mu'_{yab} N_a^2 \mu_{ya} \mu_{xa} - t_{XA}^2 N_{ab}\mu'_{yab} A_5 N_a^2 \mu_{xa}^2 \\
&+ 2t_{XA}^2 M_3 N_{ab}\mu'_{yab} N_a^2 \mu_{ya} \mu_{xa}
\end{aligned}$$

Além disso, tem-se ainda que

$$\begin{aligned}
\frac{\partial^2 A \text{Var}(\hat{t}_{yr4})}{\partial^2 p} &= 6t_{XA}^2 \left\{ A_1 + p^2 A_2 + G_{\mathcal{A}}^2 A_3 + p^2 G_{\mathcal{A}}^2 A_4 - G_{\mathcal{A}}^2 A_5 - p^2 G_{\mathcal{A}}^2 A_6 + M_1 \right. \\
&+ \left. p^2 M_2 + G_{\mathcal{A}}^2 M_3 + p^2 G_{\mathcal{A}}^2 M_4 \right\} N_{ab}^2 \mu_{xab}'^2 / t_{xA}^{*4} \\
&- 4t_{XA}^2 \left\{ 2pA_2 + 2 \frac{G_{\mathcal{A}} A_3 N_{ab}\mu'_{yab}}{t_{xA}^{*2}} - 2 \frac{G_{\mathcal{A}}^2}{t_{xA}^*} A_3 N_{ab}\mu'_{xab} + 2pG_{\mathcal{A}}^2 A_4 \right. \\
&+ \left. 2p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} A_4 N_{ab}\mu'_{yab} - 2p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} A_4 N_{ab}\mu'_{xab} - \frac{N_{ab}\mu'_{yab} A_5}{t_{xA}^*} + \frac{G_{\mathcal{A}}}{t_{xA}^*} A_5 N_{ab}\mu'_{xab} \right. \\
&- \left. 2pG_{\mathcal{A}} A_6 - p^2 \frac{N_{ab}\mu'_{yab} A_6}{t_{xA}^*} + p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} A_6 N_{ab}\mu'_{xab} + 2pM_2 + 2 \frac{G_{\mathcal{A}}}{t_{xA}^*} M_3 N_{ab}\mu'_{yab} \right. \\
&- \left. 2 \frac{G_{\mathcal{A}}}{t_{xA}^*} M_3 N_{ab}\mu'_{xab} + 2pG_{\mathcal{A}}^2 M_4 + 2p^2 \frac{G_{\mathcal{A}}}{t_{xA}^*} M_4 N_{ab}\mu'_{yab} - \frac{N_{ab}\mu'_{yab} A_5}{t_{xA}^*} \right\}
\end{aligned}$$

$$\begin{aligned}
& + \frac{G_A}{t_{xA}^*} A_5 N_{ab} \mu'_{xab} - 2p G_A A_6 - p^2 \frac{N_{ab} \mu'_{yab} A_6}{t_{xA}^*} + p^2 \frac{G_A}{t_{xA}^*} A_6 N_{ab} \mu'_{xab} + 2p M_2 \\
& + 2 \frac{G_A}{t_{xA}^*} M_3 N_{ab} \mu'_{yab} - 2 \frac{G_A}{t_{xA}^*} M_3 N_{ab} \mu'_{xab} + 2p G_A^2 M_4 + 2p^2 \frac{G_A}{t_{xA}^*} M_4 N_{ab} \mu'_{yab} \\
& - 2p^2 \frac{G_A}{t_{xA}^*} M_4 N_{ab} \mu'_{xab} \Big\} N_{ab} \mu'_{xab} / t_{xA}^{*3} \\
& + t_{xA}^2 \left\{ 2A_2 + 2M_2 + 2p^2 \frac{N_{ab}^2 \mu_{yab}'^2 M_4}{t_{xA}^{*2}} - 8p^2 \frac{G_A}{t_{xA}^{*2}} M_4 N_{ab}^2 \mu'_{yab} \mu'_{xab} \right. \\
& + 6p^2 \frac{G_A^2}{t_{xA}^*} M_4 N_{ab}^2 \mu_{xab}'^2 - 8p^2 \frac{G_A}{t_{xA}^{*2}} A_4 M_{ab}^2 \mu'_{yab} \mu'_{xab} + 6p^2 \frac{G_A^2}{t_{xA}^{*2}} A_4 N_{ab}^2 \mu_{xab}'^2 \\
& - 2G_A A_6 + 2G_A^2 M_4 + 2 \frac{N_{ab}^2 \mu_{yab}'^2 A_3}{t_{xA}^{*2}} - 8 \frac{G_A}{t_{xA}^*} N_{ab}^2 \mu'_{yab} \mu'_{xab} + 2 \frac{N_{ab}^2 \mu'_{yab} A_5 \mu'_{xab}}{t_{xA}^{*2}} \\
& - 2 \frac{G_A}{t_{xA}^{*2}} A_5 N_{ab}^2 \mu_{xab}'^2 - 4p \frac{N_{ab} \mu'_{yab} A_6}{t_{xA}^*} + 4p \frac{G_A}{t_{xA}^*} A_6 N_{ab} \mu'_{xab} + 2p^2 \frac{N_{ab}^2 \mu'_{yab} A_6 \mu'_{xab}}{t_{xA}^*} \\
& - 2p^2 \frac{G_A}{t_{xA}^{*3}} + 6 \frac{G_A^2}{t_{xA}^{*2}} A_3 N_{ab}^2 \mu_{xab}'^2 + 8p \frac{G_A^2}{t_{xA}^*} A_4 N_{ab} \mu'_{yab} - 8p \frac{G_A^2}{t_{xA}^*} N_{ab} \mu'_{xab} \\
& + 2p^2 \frac{N_{ab}^2 \mu_{yab}'^2 A_4}{t_{xA}^{*2}} + 2G_A^2 A_4 + 2 \frac{N_{ab}^2 \mu'_{yab} M_3}{t_{xA}^*} - 8 \frac{G_A}{t_{xA}^{*2}} M_3 N_{ab}^2 \mu'_{yab} \mu'_{xab} \\
& + 6 \frac{G_A^2}{t_{xA}^{*2}} M_3 N_{ab}^2 \mu_{xab}'^2 + 8p \frac{G_A}{t_{xA}^{*2}} M_4 N_{ab} \mu'_{yab} - 8p \frac{G_A^2}{t_{xA}^*} M_4 N_{ab} \mu'_{xab} \Big\} / t_{xA}^{*2} \\
& + 2B_2 + 2M_6
\end{aligned}$$

É possível mostrar que de fato $\frac{\partial^2 AVar(\hat{t}_{yr4})}{\partial^2 p} > 0$, quando avaliada na menor raiz do polinômio, ou seja que a menor raiz do polinômio apresentado é ponto de mínimo. Uma alternativa é substituir o valor de p na expressão da segunda derivada.

Apêndice C

Variâncias aproximadas dos estimadores sob o plano AAE

Estimador razão separado 1

$$\hat{t}_{yrs1} = \sum_{j=1}^{\mathcal{J}} \tilde{\delta}_{a1(j)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} p_{(jw)} \tilde{\delta}'_{ab1(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b1(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab1(jw)},$$

Através do método de linearização de Taylor, temos que, para $j \in U_a$, $w \in U_b$, $jw \in U_{ab} \subset U_A$ e $jw \in U_{ab} \subset U_B$, o estimador pode ser reescrito da seguinte forma:

$$\begin{aligned} \hat{t}_{yrs1} &= \sum_{j=1}^{\mathcal{J}} (t_{ya(j)} + \tilde{t}_{ya(j)} - R_{a(j)} \tilde{t}_{xa(j)}) \\ &+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (p_{(jw)} t'_{yab(jw)} + p_{(jw)} \tilde{t}'_{yab(jw)} - p_{(jw)} R_{ab(jw)} \tilde{t}'_{xab(jw)}) \\ &+ \sum_{w=1}^{\mathcal{W}} \tilde{t}_{yb(jw)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{t}''_{yab(jw)}, \end{aligned} \quad (7.1)$$

onde:

$$\tilde{t}_{ya(j)} = N_{a(j)} \tilde{y}_{a(j)}; \quad \tilde{t}_{yb(w)} = N_{b(w)} \tilde{y}_{b(w)}; \quad \tilde{t}'_{yab(jw)} = N_{ab(jw)} \tilde{y}'_{ab(jw)}; \quad \tilde{t}''_{yab(jw)} = N_{ab(jw)} \tilde{y}''_{ab(jw)}.$$

Seja

$$\tilde{R}_{a(j)} = \frac{\tilde{t}_{ya}}{\tilde{t}_{xa}} \quad \tilde{R}_{ab(jw)} = \frac{\tilde{t}_{yab(jw)}}{\tilde{t}_{xab(jw)}}$$

A variância aproximada do estimador é dada por:

$$\begin{aligned}
A\text{Var}(\hat{t}_{yrs1}) &= \sum_{j=1}^{\mathcal{J}} [\text{Var}(\tilde{t}_{ya(j)}) + R_{a(j)}^2 \text{Var}(\tilde{t}_{xa(j)}) - 2R_{a(j)} \text{ACov}(\tilde{t}_{ya(j)}; \tilde{t}_{xa(j)})] \\
&+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} [p_{(jw)}^2 \text{Var}(\tilde{t}'_{yab(jw)}) + p_{(jw)}^2 R_{ab(jw)}^2 \text{Var}(\tilde{t}'_{xab(jw)}) \\
&- 2R_{ab(jw)} p_{(jw)}^2 \text{ACov}(\tilde{t}_{yab(jw)}; \tilde{t}_{xab(jw)})] \\
&+ \sum_{w=1}^{\mathcal{W}} \text{Var}(\tilde{t}_{yb(w)}) + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 \text{Var}(\tilde{t}''_{yab(jw)}) \\
&= \sum_{j=1}^{\mathcal{J}} \left[N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 + R_{a(j)}^2 N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \right. \\
&- 2R_{a(j)}^2 \left(\frac{N_{a(j)}}{n_{a(j)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \left. \right] \\
&+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left[p_{(jw)}^2 N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \right. \\
&+ p_{(jw)}^2 R_{ab(jw)}^2 N_{\mathcal{A}(j)} \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\
&- 2R_{ab(jw)} p_{(jw)}^2 \left(\frac{N_{ab(jw)}}{n'_{ab(jw)}} \right) n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \left. \right] \\
&+ \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\
&+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2.
\end{aligned}$$

Os termos que compõem as somas em j e em w são similares aos termos do estimador razão 1, sob o plano de Amostragem Aleatória Simples. Neste caso, as constantes de ponderação $p_{(jw)}$ são os pontos que minimizam a variância aproximada do estimador para dois estratos j e w , e são obtidas através do mesmo procedimento descrito para o estimador razão 1. $R_{a(j)}$ e $R_{ab(jw)}$ são valores populacionais referentes a cada estrato.

Estimador razão separado 2

Para $j \in U_a$, $w \in U_b$, $jw \in U_{ab} \subset U_{\mathcal{A}}$ e $jw \in U_{ab} \subset U_{\mathcal{B}}$, o estimador pode ser reescrito da seguinte forma:

$$\hat{t}_{yrs2} = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab2(jw)} t_{\mathcal{I}(j)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b2(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab2(jw)}$$

A componente $\tilde{\delta}'_{ab2(jw)}$ tem forma similar à forma do termo $G_{\mathcal{A}}$ definida para o estimador 2, e os demais componentes da soma que são referentes ao cadastro $\mathcal{B}$ são os mesmos que os definidos para o estimador razão separado 1. Logo, através do método de linearização de Taylor, temos que a variância do estimador razão separado 2 será dada por:

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs2}) = & \sum_{j=1}^J \sum_{w=1}^W t_{\mathcal{I}(j)}^2 \left(\frac{1}{t_{x\mathcal{A}(jw)}^*} \right)^2 \left\{ N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ & + p_{(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma'_{yab(jw)}{}^2 \\ & + G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ & + p^2 G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma'_{xab(jw)}{}^2 \\ & - 2G_{\mathcal{A}(jw)} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \\ & \left. - 2p^2 G_{\mathcal{A}(jw)} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ & + \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\ & + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma''_{yab(jw)}{}^2, \end{aligned}$$

onde

$$t_{x\mathcal{A}(jw)}^* = t_{xa(j)} + p t'_{xab(jw)} \quad G_{\mathcal{A}(jw)} = \frac{t_{ya(j)} + p t'_{yab(jw)}}{t_{xa(j)} + p t'_{xab(jw)}}.$$

Os valores $p_{(jw)}$ são obtidos individualmente da mesma forma que a descrita para o estimador razão separado 2. Cada $p_{(jw)}$ é o valor que minimiza a variância do estimador combinado para dois cadastros j e w . Como a forma da variância é dada para o estimador sem correção, é possível incorporar uma correção na forma da variância do estimador, e assim evitar que termos repetidos aumentem o valor da variância.

Sob a situação em que $U_B \subset U_A$, $N_{b(w)} = 0$, $\forall w$ e a expressão de $\hat{t}_{yrs2}$ reduz-se à:

$$\hat{t}_{yrs2} = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \tilde{\delta}'_{ab2(jw)} t_{\mathcal{I}(j)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab2(jw)}.$$

Logo, a forma da variância do estimador é dada por

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs2}) &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} t_{\mathcal{I}(j)}^2 \left(\frac{1}{t_{x\mathcal{A}(jw)}^*} \right)^2 \left\{ N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ &+ p_{(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 G_{\mathcal{A}(jw)}^2 N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2G_{\mathcal{A}(jw)} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \\ &- \left. 2p^2 G_{\mathcal{A}(jw)} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ &+ \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)})^2 N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2. \end{aligned}$$

Estimador razão separado 3

$$\hat{t}_{yrs3} = \sum_{j=1}^{\mathcal{J}} \tilde{\delta}_{a3(j)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} p_{(jw)} \tilde{\delta}'_{ab3(jw)} + \sum_{w=1}^{\mathcal{W}} \tilde{\delta}_{b3(w)} + \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (1 - p_{(jw)}) \tilde{\delta}''_{ab3(jw)}.$$

Observa-se que cada termo em j e w tem a mesma forma que os termos utilizados no estimador razão 3. Assim, através do método de linearização de Taylor, é possível obter um estimador aproximado para $\hat{t}_{yrs3}$. Dessa forma, a variância aproximada do estimador é dada por:

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs3}) = & \sum_{j=1}^J \left[t_{ya(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{xa(j)}^2 + N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{ya(j)}^2 \right. \\ & \left. - 2N_{a(j)} t_{ya(j)} \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{xa(j)} \sigma_{ya(j)} \right] \\ & + \sum_{j=1}^J \sum_{w=1}^W \left[p_{(jw)}^2 t_{xab(jw)}'^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{xab(jw)}'^2 \right. \\ & + p_{(jw)}^2 N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{yab(jw)}'^2 \\ & \left. - 2p_{(jw)}^2 N_{ab(jw)} t_{xab(jw)}' \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{xab(jw)} \sigma'_{yab(jw)} \right] \\ & + \sum_{w=1}^W N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P_b} \right) \sigma_{yb}^2 \\ & + \sum_{j=1}^J \sum_{w=1}^W \mu_{yb(w)}^2 \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ & + \sum_{j=1}^J \sum_{w=1}^W \left[N_{ab(jw)}^2 (1 - p_{(jw)})^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P''_{ab(jw)}} \right) \sigma_{yab(jw)}''^2 \right. \\ & \left. + (1 - p_{(jw)})^2 (\mu_{yab(jw)}''^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \right]. \end{aligned}$$

Estimador razão separado 4

Podemos reescrever o estimador razão separado 4 na forma:

$$\hat{t}_{yrs4} = \sum_{j=1}^J \sum_{w=1}^W \tilde{\delta}'_{ab4(jw)} t_{\mathcal{I}(j)} + \sum_{w=1}^W \tilde{\delta}_{b4(w)} + \sum_{j=1}^J \sum_{w=1}^W (1 - p_{(jw)}) \tilde{\delta}''_{ab4(jw)}.$$

Os termos em j e w que estão nos estimadores obtidos dos cadastros $\mathcal{A}$ e $\mathcal{B}$ são idênticos aos termos $G_{\mathcal{A}}$ e $G_{\mathcal{B}}$ do estimador razão 4. Dessa forma, através do método de linearização de Taylor, é possível obter a variância do estimador razão separado 4, a qual é dada pela expressão:

$$\begin{aligned} A\text{Var}(\hat{t}_{yrs4}) &= \sum_{j=1}^J \sum_{w=1}^W t_{\mathcal{I}(j)}^2 \left(\frac{1}{t_{x\mathcal{A}(jw)}^*} \right)^2 \left\{ \left(\mu_{ya(j)}^2 + p^2 \mu_{yab(jw)}'^2 + G_{\mathcal{A}(j)}^2 \mu_{xa(j)}^2 \right. \right. \\ &+ \left. p^2 G_{\mathcal{A}(jw)}^2 \mu_{xab(jw)}'^2 \right) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ &+ N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{ya(j)}^2 + p^2 N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{yab(jw)}'^2 \\ &+ G_{\mathcal{A}(jw)}^2 N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{xa(j)}^2 + p^2 G_{\mathcal{A}(j)}^2 N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{xab(jw)}'^2 \\ &- 2 G_{\mathcal{A}(jw)} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{xa(j)} \sigma_{ya(j)} \\ &- \left. 2 p^2 G_{\mathcal{A}(jw)} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{xab(jw)} \sigma'_{yab(jw)} \right\} \\ &+ \sum_{w=1}^W N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P_b} \right) \sigma_{yb}^2 + \sum_{j=1}^J \sum_{w=1}^W \mu_{yab(jw)}^2 \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \\ &+ \sum_{j=1}^J \sum_{w=1}^W \left[N_{ab(jw)}^2 (1 - p_{(jw)})^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)} P''_{ab(jw)}} \right) \sigma_{yab(jw)}''^2 \right. \\ &+ \left. (1 - p_{(jw)})^2 (\mu_{yab(jw)}''^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \right]. \end{aligned}$$

Os valores $p_{(jw)}$ são obtidos separadamente, e cada $p_{(jw)}$ é o valor que minimiza a variância do estimador para o total de dois estratos j e w .

Estimador razão combinado 1

Podemos reescrever o estimador razão combinado 1 da seguinte forma:

$$\begin{aligned}\hat{t}_{yrc1} &= \frac{\tilde{t}_{ya(j)}}{\tilde{t}_{xa(j)}} t_{xa(j)} + p \frac{\tilde{t}'_{yab(jw)}}{\tilde{t}'_{xab(jw)}} t_{xab(j)} + \tilde{t}_{yb(w)} + (1-p)\tilde{t}_{yab(jw)} \\ &= \hat{R}_{a(j)} t_{xa(j)} + p \hat{R}_{ab(jw)} t_{xab(jw)} + \tilde{t}_{yb(w)} + (1-p)\tilde{t}_{yab(jw)}.\end{aligned}$$

onde:

$$\begin{aligned}\hat{t}_{ya(j)} &= \sum_{j=1}^{\mathcal{J}} N_{a(j)} \tilde{y}_{a(j)}; & \hat{t}_{xa(j)} &= \sum_{j=1}^{\mathcal{J}} N_{a(j)} \tilde{x}_{a(j)}; \\ \hat{t}'_{yab(jw)} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)} \tilde{y}'_{ab(jw)}; & \hat{t}'_{xab(jw)} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)} \tilde{x}'_{ab(jw)}; \\ \hat{t}_{yb(w)} &= \sum_{w=1}^{\mathcal{W}} N_{b(w)} \tilde{y}_{b(w)}; & \hat{t}''_{yab(jw)} &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)} \tilde{y}''_{ab(jw)}; \\ \hat{R}_{a(j)} &= \frac{\hat{t}_{ya(j)}}{\hat{t}_{xa(j)}}; & \hat{R}_{ab(jw)} &= \frac{\hat{t}_{yab(jw)}}{\hat{t}_{xab(jw)}}.\end{aligned}$$

Utilizando o método de linearização de Taylor e tomando $\hat{t}_{ya(j)}$, $\hat{t}_{xa(j)}$, $\hat{t}'_{yab(jw)}$, $\hat{t}'_{xab(jw)}$, $\hat{t}_{yb(w)}$ e $\hat{t}''_{yab(jw)}$ como variáveis de referência para linearização, o procedimento será o mesmo que o estimador razão 1. Dessa forma, a variância do estimador será dada por

$$\begin{aligned}A\text{Var}(\hat{t}_{yrc1}) &= \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \\ &+ p^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ R_{a(j)}^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 R_{ab(jw)}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2R_{a(j)}^2 \sum_{j=1}^{\mathcal{J}} \left(\frac{N_{a(j)}}{n_{a(j)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)}\end{aligned}$$

$$\begin{aligned}
& - 2p^2 R_{ab(jw)} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left(\frac{N_{ab(jw)}}{n'_{ab(jw)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \\
& + \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\
& + (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma''_{yab(jw)}{}^2.
\end{aligned}$$

Através do mesmo procedimento utilizado para o estimador razão 1, temos que o valor de p que minimiza a variância do estimador razão combinado 1 é dado por:

$$p = \frac{T_4}{T_1 + T_2 - 2T_3 + T_4},$$

onde

$$\begin{aligned}
T_1 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}^2; \\
T_2 &= R_{ab(jw)}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}^2; \\
T_3 &= 2R_{ab(jw)} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left(\frac{N_{ab(jw)}}{n'_{ab(jw)}} \right)^2 n_{\mathcal{A}(j)} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} e \\
T_4 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma''_{yab(jw)}{}^2.
\end{aligned}$$

Estimador razão combinado 2

É possível reescrever o estimador 2 da seguinte forma:

$$\hat{t}_{yrc2} = \left(\frac{\hat{t}_{ya(j)} + p\hat{t}'_{yab(jw)}}{\hat{t}_{xa(j)} + p\hat{t}'_{xab(jw)}} \right) t_{X_{\mathcal{A}}} + \hat{t}_{yb(w)} + (1-p)\hat{t}_{yab(jw)}.$$

Seja

$$\hat{G}_{\mathcal{A}_C} = \frac{\hat{t}_{yA}^*}{\hat{t}_{xA}^*} = \frac{\hat{t}_{ya(j)} + p\hat{t}'_{yab(jw)}}{\hat{t}_{xa(j)} + p\hat{t}'_{xab(jw)}},$$

com $\hat{t}_{ya(j)}$, $\hat{t}_{xa(j)}$, $\hat{t}'_{yab(jw)}$, $\hat{t}'_{xab(jw)}$, $\hat{t}_{yb(w)}$ e $\hat{t}''_{yab(jw)}$ definidos da mesma forma que as mostradas para o estimador razão combinado 1.

Tomando $\hat{t}_{yA}^*$ e $\hat{t}_{xA}^*$ como referência para linearização, da mesma forma que a vista para o estimador razão 2, tem-se que

$$\hat{G}_{\mathcal{A}_C} \approx \hat{G}_{\mathcal{A}_C L} = G_{\mathcal{A}_C} + \left\{ \left(\frac{1}{t_{xA}^*} \right) \left(\hat{t}_{yA}^* - \frac{t_{yA}^*}{t_{xA}^*} \hat{t}_{xA}^* \right) \right\}$$

onde $G_{\mathcal{A}_C} = \frac{t_{yA}^*}{t_{xA}^*}$ é o valor de referência para a aproximação (parâmetro de interesse).

Dessa forma, através do mesmo procedimento utilizado para obtenção da variância do estimador razão 2, temos que

$$\begin{aligned} A\text{Var}(\hat{t}_{yrc2}) &= t_{X_{\mathcal{A}}}^2 \left(\frac{1}{t_{xA}^*} \right)^2 \left\{ \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ &+ p^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya} \sigma_{ya(j)} \sigma_{xa(j)} \Big\} \end{aligned}$$

$$\begin{aligned}
& - 2p^2 G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \Big\} \\
& + \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\
& + (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{ab(jw)}'' \sigma_{yab(jw)}''^2 \\
& = t_{X_{\mathcal{A}}}^2 \left\{ \left(\frac{1}{t_{x\mathcal{A}}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}_C}^2 A_3 + p^2 G_{\mathcal{A}_C}^2 A_4 - 2G_{\mathcal{A}_C} A_5 \right. \\
& \quad \left. - 2p^2 G_{\mathcal{A}_C} A_6) \right\} + B_1 + (1-p)^2 B_2
\end{aligned}$$

Derivando a expressão da variância e igualando a expressão da derivada a zero, temos que o valor de p que minimiza a variância do estimador razão combinado 2 será dado pela menor solução do seguinte polinômio de grau 6: $k_6 p^6 + k_5 p^5 + k_4 p^4 + k_3 p^3 + k_2 p^2 + k_1 p + k_0 = 0$, onde:

$$\begin{aligned}
k_6 &= B_2 t_{xab(jw)}^5 \\
k_5 &= -B_2 t_{xab(jw)}^5 + 5B_2 t_{xa(j)} t_{xab(jw)}^4 \\
k_4 &= -5B_2 t_{xa(j)} t_{xab(jw)}^4 + 10B_2 t_{xa(j)}^2 t_{xab(jw)}^3 \\
k_3 &= t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_6 t_{ya(j)} - 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_6 t_{yab(jw)} t_{xa(j)} + t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_2 t_{xa(j)} \\
& - 10B_2 t_{xa(j)}^2 t_{xab(jw)}^3 + 10B_2 t_{xa(j)}^3 t_{xab(jw)}^2 - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_4 t_{ya(j)} t_{yab(jw)} + 2t_{X_{\mathcal{A}}}^2 A_4 t_{yab(jw)}^2 t_{xa(j)} \\
k_2 &= -t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_6 t_{ya(j)} t_{xa(j)} + 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_2 t_{xa(j)}^2 + 5B_2 t_{xa(j)}^4 t_{xab(jw)} \\
& - t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^3 A_1 - 10B_2 t_{xa(j)}^3 t_{xab(jw)}^2 - 3t_{X_{\mathcal{A}}}^2 A_6 t_{yab(jw)} t_{xa(j)}^2 \\
& - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_4 t_{ya(j)}^2 + 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_5 t_{yab(jw)} - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_3 t_{yab(jw)}^2 \\
& + 3t_{X_{\mathcal{A}}}^2 A_4 t_{ya(j)} t_{yab(jw)} t_{xa(j)} \\
k_1 &= t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_5 t_{yab(jw)} t_{xa(j)} + B_2 t_{xa(j)}^5 + 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_5 t_{ya(j)} \\
& + t_{X_{\mathcal{A}}}^2 A_2 t_{xa(j)}^3 - 5B_2 t_{xa(j)}^4 t_{xab(jw)} - 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_1 t_{xa(j)} - 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_3 t_{ya(j)} t_{yab(jw)} \\
& + t_{X_{\mathcal{A}}}^2 A_4 t_{ya(j)}^2 t_{xa(j)} - 2t_{X_{\mathcal{A}}}^2 A_6 t_{ya(j)} t_{xa(j)}^2 + t_{X_{\mathcal{A}}}^2 A_3 t_{yab(jw)}^2 t_{xa(j)} \\
k_0 &= t_{X_{\mathcal{A}}}^2 t_{yab(jw)} A_5 t_{xa(j)}^2 - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_1 t_{xa(j)}^2 + t_{X_{\mathcal{A}}}^2 A_3 t_{yab(jw)} t_{ya(j)} t_{xa(j)}
\end{aligned}$$

$$- 2t_{X_A}^2 t_{xab(jw)} A_3 t_{ya(j)}^2 - B_2 t_{xa(j)}^5 + 3t_{X_A}^2 t_{xab(jw)} A_5 t_{ya(j)} t_{xa(j)}$$

Os totais populacionais nas expressões das constantes k_i são referentes aos estratos, domínios e variáveis envolvidas no processo de estimação.

Sob a situação em que $U_B \subset U_A$, a forma do estimador razão combinado 2 reduz-se à:

$$\hat{t}_{yrc2} = \left(\frac{\hat{t}_{ya(j)} + p\hat{t}'_{yab(jw)}}{\hat{t}_{xa(j)} + p\hat{t}'_{xab(jw)}} \right) t_{X_A} + (1-p)\hat{t}''_{yab(jw)}.$$

Com isso, a forma da variância do estimador é dada por

$$\begin{aligned} AVar(\hat{t}_{yrc2}) &= t_{X_A}^2 \left(\frac{1}{t_{x_A}^*} \right)^2 \left\{ \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{ya(j)}^2 \right. \\ &+ p^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{yab(jw)}'^2 \\ &+ G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P_{a(j)} \sigma_{xa(j)}^2 \\ &+ p^2 G_{\mathcal{A}_C}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{A}(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) P'_{ab(jw)} \sigma_{xab(jw)}'^2 \\ &- 2G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya} \sigma_{ya(j)} \sigma_{xa(j)} \\ &- \left. 2p^2 G_{\mathcal{A}_C} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\ &+ \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1-f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P_{b(w)} \sigma_{yb(w)}^2 \\ &+ (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{\mathcal{B}(w)} \left(\frac{1-f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) P''_{ab(jw)} \sigma_{yab(jw)}''^2 \\ &= t_{X_A}^2 \left\{ \left(\frac{1}{t_{x_A}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}_C}^2 A_3 + p^2 G_{\mathcal{A}_C}^2 A_4 - 2G_{\mathcal{A}_C} A_5 \right. \end{aligned}$$

$$- 2p^2 G_{\mathcal{A}_C} A_6) \Big\} + (1-p)^2 B_1$$

Derivando a expressão da variância e igualando a expressão da derivada a zero, temos que o valor de p que minimiza a variância do estimador é a menor solução (parte real) do seguinte polinômio de grau 6: $k_6 p^6 + k_5 p^5 + k_4 p^4 + k_3 p^3 + k_2 p^2 + k_1 p + k_0 = 0$, onde:

$$\begin{aligned} k_6 &= B_1 t_{xab(jw)}^5 \\ k_5 &= -B_1 t_{xab(jw)}^5 + 5B_1 t_{xa(j)} t_{xab(jw)}^4 \\ k_4 &= -5B_1 t_{xa(j)} t_{xab(jw)}^4 + 10B_1 t_{xa(j)}^2 t_{xab(jw)}^3 \\ k_3 &= t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_6 t_{ya(j)} - 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_6 t_{yab(jw)} t_{xa(j)} + t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_2 t_{xa(j)} \\ &\quad - 10B_1 t_{xa(j)}^2 t_{xab(jw)}^3 + 10B_1 t_{xa(j)}^3 t_{xab(jw)}^2 - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_4 t_{ya(j)} t_{yab(jw)} + 2t_{X_{\mathcal{A}}}^2 A_4 t_{yab(jw)}^2 t_{xa(j)} \\ k_2 &= -t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_6 t_{ya(j)} t_{xa(j)} + 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_2 t_{xa(j)}^2 + 5B_1 t_{xa(j)}^4 t_{xab(jw)} \\ &\quad - t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^3 A_1 - 10B_1 t_{xa(j)}^3 t_{xab(jw)}^2 - 3t_{X_{\mathcal{A}}}^2 A_6 t_{yab(jw)} t_{xa(j)}^2 \\ &\quad - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_4 t_{ya(j)}^2 + 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_5 t_{yab(jw)} - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_3 t_{yab(jw)}^2 \\ &\quad + 3t_{X_{\mathcal{A}}}^2 A_4 t_{ya(j)} t_{yab(jw)} t_{xa(j)} \\ k_1 &= t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_5 t_{yab(jw)} t_{xa(j)} + B_1 t_{xa(j)}^5 + 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_5 t_{ya(j)} \\ &\quad + t_{X_{\mathcal{A}}}^2 A_2 t_{xa(j)}^3 - 5B_1 t_{xa(j)}^4 t_{xab(jw)} - 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_1 t_{xa(j)} - 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_3 t_{ya(j)} t_{yab(jw)} \\ &\quad + t_{X_{\mathcal{A}}}^2 A_4 t_{ya(j)}^2 t_{xa(j)} - 2t_{X_{\mathcal{A}}}^2 A_6 t_{ya(j)} t_{xa(j)}^2 + t_{X_{\mathcal{A}}}^2 A_3 t_{yab(jw)}^2 t_{xa(j)} \\ k_0 &= t_{X_{\mathcal{A}}}^2 t_{yab(jw)} A_5 t_{xa(j)}^2 - t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_1 t_{xa(j)}^2 + t_{X_{\mathcal{A}}}^2 A_3 t_{yab(jw)} t_{ya(j)} t_{xa(j)} \\ &\quad - 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_3 t_{ya(j)}^2 - B_1 t_{xa(j)}^5 + 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_5 t_{ya(j)} t_{xa(j)} \end{aligned}$$

Estimador razão combinado 3

Podemos reescrever o estimador 3 da seguinte forma:

$$\begin{aligned}\hat{t}_{yrc3} &= \frac{\hat{t}_{ya(j)}}{\hat{t}_{xa(j)}}t_{xa(j)} + p\frac{\hat{t}'_{yab(jw)}}{\hat{t}'_{xab(jw)}}t_{xab(j)} + \hat{t}_{yb(w)} + (1-p)\hat{t}''_{yab(jw)} \\ &= \hat{G}_{\mathcal{A}} + \hat{G}_{\mathcal{B}}\end{aligned}\tag{7.2}$$

Os componentes $\hat{G}_{\mathcal{A}}$ e $\hat{G}_{\mathcal{B}}$ são funções não-lineares que estão em termos dos estimadores para os domínios dentro de cada cada estrato j e w , onde

$$\begin{aligned}\hat{G}_{\mathcal{A}} &= \frac{\hat{t}_{ya(j)}}{\hat{t}_{xa(j)}}t_{xa(j)} + p\frac{\hat{t}'_{yab(jw)}}{\hat{t}'_{xab(jw)}}t_{xab(j)} = \hat{R}_a t_{xa(j)} + p\hat{R}_{ab}t_{xab(j)} \\ \hat{G}_{\mathcal{B}} &= \hat{t}_{yb(w)} + (1-p)\hat{t}_{yab(jw)}.\end{aligned}$$

Através do método de linearização de Taylor, temos que

$$\hat{G}_{\mathcal{A}} \approx G_{\mathcal{AL}} = G_{\mathcal{A}} + t_{ya(j)} + pt_{yab(jw)} + \hat{t}_{ya(j)} + p\hat{t}'_{yab(jw)} - R_a \hat{t}_{xa(j)} - pR_{ab}\hat{t}'_{xab(jw)}.$$

$\hat{G}_{\mathcal{B}}$ está em função de termos não-lineares do tipo $\hat{N}_i \tilde{y}_i$, para cada estrato j e w na população. Logo, através do método de linearização de Taylor verifica-se que

$$\begin{aligned}\hat{G}_{\mathcal{B}} &= G_{\mathcal{B}} + (1-p) \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yab(jw)} \hat{N}_{ab,s(jw)} + N_{ab(jw)} \tilde{y}''_{ab(jw)} - N_{ab(jw)} \mu''_{yab(jw)}) \\ &\quad + \sum_{w=1}^{\mathcal{W}} \left(N_{b(w)} \tilde{y}_{b(w)} - \mu_{yb(w)} (\hat{N}_{ab,s(w)} - N_{ab(jw)}) \right).\end{aligned}$$

Verifica-se ainda que a linearização de $G_{\mathcal{A}}$ ainda possui termos não-lineares da forma $\hat{N}_i \tilde{y}_i$. Aplicando novamente o método de linearização de Taylor e calculando a variância aproximada de $\hat{G}_{\mathcal{A}} + \hat{G}_{\mathcal{B}}$, temos que

$$\begin{aligned}
A\text{Var}(\hat{t}_{yrc3}) &= \sum_{j=1}^{\mathcal{J}} \text{Var}(\mu_{ya(j)}\hat{N}_{a(j)} + N_{a(j)}\tilde{y}_{a(j)} - N_{a(j)}\mu_{ya(j)}) \\
&+ p^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \text{Var}(\mu_{yab(jw)}\hat{N}_{ab,s(jw)} + N_{ab(jw)}\tilde{y}_{ab(jw)} - N_{ab(jw)}\mu_{yab(jw)}) \\
&+ R_a^2 \sum_{j=1}^{\mathcal{J}} \text{Var}(\mu_{xa(j)}\hat{N}_{a(j)} + N_{a(j)}\tilde{x}_{a(j)} - N_{a(j)}\mu_{xa(j)}) \\
&+ p^2 R_{ab}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \text{Var}(\mu_{xab(jw)}\hat{N}_{ab,s(jw)} + N_{ab(jw)}\tilde{x}'_{ab(jw)} - N_{ab(jw)}\mu_{xab(jw)}) \\
&- 2R_a \sum_{j=1}^{\mathcal{J}} \text{ACov}(\mu_{ya(j)}\hat{N}_{a(j)} + N_{a(j)}\tilde{y}_{a(j)}) \\
&- N_{a(j)}\mu_{ya(j)} ; \mu_{xa(j)}\hat{N}_{a(j)} + N_{a(j)}\tilde{x}_{a(j)} - N_{a(j)}\mu_{xa(j)}) \\
&- 2p^2 R_{ab} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \text{ACov}(\mu_{yab(jw)}\hat{N}_{ab,s(jw)} + N_{ab(jw)}\tilde{y}'_{ab(jw)}) \\
&- N_{ab(jw)}\mu_{yab(jw)} ; \mu_{xab(jw)}\hat{N}_{ab,s(jw)} + N_{ab(jw)}\tilde{x}_{ab(jw)} - N_{ab(jw)}\mu_{xab(jw)}) \\
&+ \sum_{w=1}^{\mathcal{W}} \text{Var}(\mu_{yb(w)}\hat{N}_{b(w)} + N_{b(w)}\tilde{y}_{b(w)} - N_{b(w)}\mu_{yb(w)}) \\
&+ (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \text{Var}(\mu_{yab(jw)}\hat{N}_{ab,s(jw)} + N_{ab(jw)}\tilde{y}''_{ab(jw)} - N_{ab(jw)}\mu_{yab(jw)}) \\
&= \sum_{j=1}^{\mathcal{J}} \left\{ \mu_{ya(j)}^2 \text{Var}(\hat{N}_{a(j)}) + N_{a(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{ya(j)}^2 \right\} \\
&+ p^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}^2 \text{Var}(\hat{N}_{ab,s(jw)}) + N_{ab(jw)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma'_{yab(jw)}{}^2 \right\} \\
&+ R_a^2 \sum_{j=1}^{\mathcal{J}} \left\{ \mu_{xa(j)}^2 \text{Var}(\hat{N}_{a(j)}) + N_{a(j)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{xa(j)}^2 \right\} \\
&+ p^2 R_{ab}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{xab(jw)}^2 \text{Var}(\hat{N}_{ab,s(jw)}) + N_{ab(jw)}^2 \left(\frac{1-f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma'_{xab(jw)}{}^2 \right\} \\
&- 2R_a \sum_{j=1}^{\mathcal{J}} \left\{ \mu_{ya(j)}\mu_{xa(j)} \text{Var}(\hat{N}_{a(j)}) + N_{a(j)}^2 \text{ACov}(\tilde{y}_{a(j)} ; \tilde{x}_{a(j)}) \right\} \\
&- 2p^2 R_{ab} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}\mu_{xab(jw)} \text{Var}(\hat{N}_{ab,s(jw)}) + N_{ab(jw)}^2 \text{ACov}(\tilde{y}'_{ab(jw)} ; \tilde{x}'_{ab(jw)}) \right\}
\end{aligned}$$

$$\begin{aligned}
& + \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yb(w)}^2 \text{Var}(\hat{N}_{b(w)}) + N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) \sigma_{yb(w)}^2 \right\} \\
& + (1-p)^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}^2 \text{Var}(\hat{N}_{ab,s(jw)}) + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) \sigma_{yab(jw)}^{\prime\prime 2} \right\}
\end{aligned}$$

Tem-se ainda que

$$\text{Var}(\hat{N}_{a(j)}) = \text{Var}(\hat{N}_{b(w)}) = \text{Var}(\hat{N}_{ab,s(jw)}) = \frac{N_{a(j)}N_{b(w)}N_{ab(jw)}g_{\mathcal{A}(j)}g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)}N_{b(w)}g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)}N_{a(j)}g_{\mathcal{A}(j)}},$$

$$\text{ACov}(\tilde{y}_{a(j)} ; \tilde{x}_{a(j)}) = \left(\frac{n_{\mathcal{A}(j)}}{n_{a(j)}} \right)^2 \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya(j)} \sigma_{ya(j)} \sigma_{xa(j)} \quad \text{e}$$

$$\text{ACov}(\tilde{y}'_{ab(jw)} ; \tilde{x}'_{ab(jw)}) = \left(\frac{n_{\mathcal{A}(j)}}{n'_{ab(jw)}} \right)^2 \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)}.$$

Logo, o valor de p que minimiza a expressão da variância aproximada do estimador razão combinado 3 é dado por:

$$p = \frac{D}{A + B - 2C + D},$$

onde

$$\begin{aligned}
A &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}^2 \frac{N_{a(j)}N_{b(w)}N_{ab(jw)}g_{\mathcal{A}(j)}g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)}N_{b(w)}g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)}N_{a(j)}g_{\mathcal{A}(j)}} + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{yab(jw)}^{\prime 2} \right\} \\
B &= R_{ab}^2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{xab(jw)}^2 \frac{N_{a(j)}N_{b(w)}N_{ab(jw)}g_{\mathcal{A}(j)}g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)}N_{b(w)}g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)}N_{a(j)}g_{\mathcal{A}(j)}} + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)}} \right) \sigma_{xab(jw)}^{\prime 2} \right\} \\
C &= R_{ab} \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)} \mu_{xab(jw)} \frac{N_{a(j)}N_{b(w)}N_{ab(jw)}g_{\mathcal{A}(j)}g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)}N_{b(w)}g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)}N_{a(j)}g_{\mathcal{A}(j)}} \right. \\
& \quad \left. + N_{ab(jw)}^2 \left(\frac{n_{\mathcal{A}(j)}}{n'_{ab(jw)}} \right)^2 \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{yab(jw)} \sigma'_{xab(jw)} \right\} \\
D &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} \left\{ \mu_{yab(jw)}^2 \frac{N_{a(j)}N_{b(w)}N_{ab(jw)}g_{\mathcal{A}(j)}g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)}N_{b(w)}g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)}N_{a(j)}g_{\mathcal{A}(j)}} + N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{B}(w)}}{n_{\mathcal{B}(w)}} \right) \sigma_{yab(jw)}^{\prime\prime 2} \right\}
\end{aligned}$$

Estimador razão combinado 4

Conforme visto para o estimador razão combinado 2, é possível reescrever o estimador 4 da seguinte forma:

$$\hat{t}_{yrc4} = \left(\frac{\hat{t}_{ya(j)} + p\hat{t}'_{yab(jw)}}{\hat{t}_{xa(j)} + p\hat{t}'_{xab(jw)}} \right) t_{X_{\mathcal{A}}} + \hat{t}_{yb(w)} + (1-p)\hat{t}_{yab(jw)},$$

onde

$$\hat{G}_{\mathcal{A}_C} = \frac{\hat{t}_{yA}^*}{\hat{t}_{xA}^*} = \frac{\hat{t}_{ya(j)} + p\hat{t}_{yab(jw)}}{\hat{t}_{xa(j)} + p\hat{t}_{xab(jw)}}.$$

Porém, da mesma forma que o estimador razão combinado 3, os estimadores em $\hat{G}_A$ estão em função de estimadores da forma $\hat{N}_i \tilde{y}_i$. Logo, utilizando o mesmo procedimento adotado para o estimador razão 4, obtém-se a seguinte expressão para a variância do estimador razão combinado 4:

$$\begin{aligned} AVar(\hat{t}_{yrc4}) &= \left(\frac{t_{X_{\mathcal{A}_C}}}{t_{x_{\mathcal{A}_C}}^*} \right)^2 (A_1 + p^2 A_2 + G_{\mathcal{A}_C}^2 A_3 + p^2 G_A^2 A_4 - G_{\mathcal{A}_C} A_5 - p^2 G_{\mathcal{A}_C} A_6 \\ &+ M_1 + p^2 M_2 + G_{\mathcal{A}_C}^2 M_3 + p^2 G_{\mathcal{A}_C}^2 M_4) + B_1 + (1-p)^2 B_2 + M_5 + (1-p)^2 M_6, \end{aligned}$$

onde

$$\begin{aligned} A_1 &= \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{ya(j)}^2; \quad A_2 = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{yab(jw)}'^2; \\ A_3 &= \sum_{j=1}^{\mathcal{J}} N_{a(j)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P_{a(j)}} \right) \sigma_{xa(j)}^2; \quad A_4 = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{A}(j)}}{n_{\mathcal{A}(j)} P'_{ab(jw)}} \right) \sigma_{xab(jw)}'^2; \\ A_5 &= \sum_{j=1}^{\mathcal{J}} 2N_{a(j)}^2 \frac{n_{\mathcal{A}(j)}}{n_{a(j)}^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho_{xya} \sigma_{xa(j)} \sigma_{ya(j)}; \\ A_6 &= 2 \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \frac{n_{\mathcal{A}(j)}}{n_{ab(jw)}'^2} \left(1 - \frac{n_{\mathcal{A}(j)}}{N_{\mathcal{A}(j)}} \right) \rho'_{xyab(jw)} \sigma'_{xab(jw)} \sigma'_{yab(jw)}; \end{aligned}$$

$$B_1 = \sum_{w=1}^{\mathcal{W}} N_{b(w)}^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}} P_{b(w)}} \right) \sigma_{yb(w)}^2 \quad B_2 = \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} N_{ab(jw)}^2 \left(\frac{1 - f_{\mathcal{B}}}{n_{\mathcal{B}} P_{ab(jw)}''} \right) \sigma_{yab(jw)}'';$$

$$\begin{aligned} M_1 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{ya(j)}^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}}; \\ M_2 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yab(jw)}'^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}}; \\ M_3 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{xa(j)}^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}}; \\ M_4 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{xab(jw)}'^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}}; \\ M_5 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yb(w)}^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \quad \text{e} \\ M_6 &= \sum_{j=1}^{\mathcal{J}} \sum_{w=1}^{\mathcal{W}} (\mu_{yab(jw)}''^2) \frac{N_{a(j)} N_{b(w)} N_{ab(jw)} g_{\mathcal{A}(j)} g_{\mathcal{B}(w)}}{n_{\mathcal{A}(j)} N_{b(w)} g_{\mathcal{B}(w)} + n_{\mathcal{B}(w)} N_{a(j)} g_{\mathcal{A}(j)}} \end{aligned}$$

Para M_1 , M_3 e M_5 , verifica-se que $N_{ab(jw)} = 0$ quando não há interseção entre os estratos j e w . Com isso, o valor de p que minimiza a variância do estimador razão combinado 4 é dada pela menor raiz real de um polinômio de grau 6, onde os coeficientes k_i , $i = 0, \dots, 6$ são dados por:

$$k_6 = 2B_2 t_{xab(jw)}^5 + 2M_6 t_{xab(jw)}^5$$

$$k_5 = -2B_2 t_{xab(jw)}^5 - 2M_6 t_{xab(jw)}^5 + 10B_2 t_{xa(j)} t_{xab(jw)}^4 + 10M_6 t_{xa(j)} t_{xab(jw)}^4$$

$$k_4 = 20M_6 t_{xa(j)}^2 t_{xab(jw)}^3 - 10M_6 t_{xa(j)} t_{xab(jw)}^4 - 10B_2 t_{xa(j)} t_{xab(jw)}^4 + 20B_2 t_{xa(j)}^2 t_{xab(jw)}^3$$

$$\begin{aligned} k_3 &= -20M_6 t_{xa(j)}^2 t_{xab(jw)}^3 + t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 A_6 t_{ya(j)} - 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)} M_4 t_{ya(j)} t_{yab(jw)} \\ &+ 4t_{X_{\mathcal{A}}}^2 A_4 t_{yab(jw)}^2 t_{xa(j)} + 4t_{X_{\mathcal{A}}}^2 M_4 t_{yab(jw)}^2 t_{xa(j)} - 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_4 t_{ya(j)} t_{yab(jw)} \\ &+ 2t_{X_{\mathcal{A}}}^2 t_{xab(jw)}^2 M_2 t_{xa(j)} - 3t_{X_{\mathcal{A}}}^2 t_{xab(jw)} A_6 t_{yab(jw)} t_{xa(j)} - 20B_2 t_{xa(j)}^2 t_{xab(jw)}^3 \end{aligned}$$

$$\begin{aligned}
& + 20B_2t_{xa(j)}^3t_{xab(jw)}^2 + 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}^2A_2t_{xa(j)} + 20M_6t_{xa(j)}^3t_{xab(jw)}^2 \\
k_2 & = 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}^2A_5t_{yab(jw)} + 4t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_2t_{xa(j)}^2 - 20M_6t_{xa(j)}^3t_{xab(jw)}^2 \\
& - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_4t_{ya(j)}^2 + 6t_{X_{\mathcal{A}}}^2A_4t_{ya(j)}t_{yab(jw)}t_{xa(j)} - t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_6t_{ya(j)}t_{xa(j)} \\
& - 20B_2t_{xa(j)}^3t_{xab(jw)}^2 - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}M_3t_{yab(jw)}^2 - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}M_4t_{ya(j)}^2 \\
& + 6t_{X_{\mathcal{A}}}^2M_4t_{ya(j)}t_{yab(jw)}t_{xa(j)} + 10B_2t_{xa(j)}^4t_{xab(jw)} + 10M_6t_{xa(j)}^4t_{xab(jw)} \\
& - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_3t_{yab(jw)}^2 + 4t_{X_{\mathcal{A}}}^2t_{xab(jw)}M_2t_{xa(j)}^2 \\
& - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}^3A_1 - 3t_{X_{\mathcal{A}}}^2A_6t_{yab(jw)}t_{xa(j)}^2 - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}^3M_1 \\
k_1 & = 3t_{X_{\mathcal{A}}}^2t_{xab(jw)}^2A_5t_{ya(j)} + 2B_2t_{xa(j)}^5 - 6t_{X_{\mathcal{A}}}^2t_{xab(jw)}M_3t_{ya(j)}t_{yab(jw)} + 2t_{X_{\mathcal{A}}}^2A_2t_{xa(j)}^3 \\
& - 10M_6t_{xa(j)}^4t_{xab(jw)} - 4t_{X_{\mathcal{A}}}^2t_{xab(jw)}^2A_1t_{xa(j)} - 6t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_3t_{ya(j)}t_{yab(jw)} \\
& + 2t_{X_{\mathcal{A}}}^2A_4t_{ya(j)}^2t_{xa(j)} - 4t_{X_{\mathcal{A}}}^2t_{xab(jw)}^2M_1t_{xa(j)} + 2t_{X_{\mathcal{A}}}^2M_2t_{xa(j)}^3 - 10B_2t_{xa(j)}^4t_{xab(jw)} \\
& + 2M_6t_{xa(j)}^5 + t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_5t_{yab(jw)}t_{xa(j)} - 2t_{X_{\mathcal{A}}}^2A_6t_{ya(j)}t_{xa(j)}^2 + 2t_{X_{\mathcal{A}}}^2M_3t_{yab(jw)}^2t_{xa(j)} \\
& + 2t_{X_{\mathcal{A}}}^2A_3t_{yab(jw)}^2t_{xa(j)} + 2t_{X_{\mathcal{A}}}^2M_4t_{ya(j)}^2t_{xa(j)} \\
k_0 & = -4t_{X_{\mathcal{A}}}^2t_{xab(jw)}M_3t_{ya(j)}^2 - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_1t_{xa(j)}^2 + 2t_{X_{\mathcal{A}}}^2A_3t_{yab(jw)}t_{ya(j)}t_{xa(j)} \\
& - 4t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_3t_{ya(j)}^2 - t_{X_{\mathcal{A}}}^2t_{yab(jw)}A_5t_{xa(j)}^2 - 2t_{X_{\mathcal{A}}}^2t_{xab(jw)}M_1t_{xa(j)}^2 - 2M_6t_{xa(j)}^5 \\
& + 3t_{X_{\mathcal{A}}}^2t_{xab(jw)}A_5t_{ya(j)}t_{xa(j)} - 2B_2t_{xa(j)}^5 + 2t_{X_{\mathcal{A}}}^2M_3t_{yab(jw)}t_{ya(j)}t_{xa(j)}
\end{aligned}$$

Apêndice D

Programas utilizados na avaliação numérica

PROGRAMA 1

```
//#####  
//##                                     #  
//## Mestrado Em Estatística           #  
//## Dissertação de Mestrado           #  
//## Autor: Hemílio Fernandes           #  
//## Orientador: Cristiano Ferraz       #  
//##                                     #  
//## Descrição: Esse programa avalia os estimadores propostos   #  
//##                para o caso de existência de variável auxiliar #  
//##                através do método de simulação de Monte Carlo. #  
//##                Para o cenário 2: NA, NB, Na, Nb e Nab são conhecidos #  
//##                                     #  
//#####  
#include<oxstd.h>  
#include<oxprob.h>  
#import<maximize>  
#import<solvenle>  
#include <oxfloat.h>  
  
//NÚMERO DE RÉPLICAS  
static decl NREP = 1000;  
  
//INÍCIO DO PROGRAMA  
main()  
{  
  ranseed("GM");  
  ranseed({1965,2001});  
  decl N=<2000,1750,1500,1250,1000,750>;  
  decl NAB=<500,1000,1500,2000,2500,3000>;  
  decl amostraA=<125,250,500,750,1000, 1250, 1000,1000,1000,1000,1000>;  
  decl amostraB=<125,250,500,750,1000, 1250, 200,300,400,500,700,900>;  
  decl tempo,arq;  
  decl amo, i, k, l, m, n, o, x, j, ind1, ind2,vec1,vec2,vech;  
  decl uni1, uni2, ord1, ord2;  
  decl mya, mxa, myb, myab, mxab;  
  decl YA, YB, XA, Xa, Xab, Ya, Yab, Yb, cadA, cadB, cadcruz;  
  decl tyA,txA, tyB, txa, txab, tya, tyab, tyb, tygeral,param;  
  decl NA, NB, Na, Nb, Nab, NTOT;  
  decl yA, yB, ya, yb, yab1, yab2;  
  decl nA, nB, na, nab1, nab2, nb;  
  decl pa, pb, pab1, pab2;  
  decl est1, est2, est5, esthart;  
  decl dif1, dif2, dif5, difhart;  
  decl vies1, vies2, vies5, vieshart;  
  decl exp1, exp2, exp5, exphart;
```

```

decl amcadA, amcadB;
decl fA, fB, fAinv, fBinv, gA, gB;
decl estya, estxa, estyab1, estxab1, estyb, estyab2, varya, varyb, varyab1, varyab2, varxa, varxab1;
decl p1, p2, p5, phart;
decl a1e1, a2e1, a3e1, a4e1, a5e1, a6e1, b1e1, b2e1, ra, rab1, cor_a, cor_ab;
decl GA, tespxa, tespya;
decl a1e2, a2e2, a3e2, a4e2, a5e2, a6e2, b1e2, b2e2;
decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
decl medest1, medest2, medhart;
decl varest1, varest2, varhart, dphart, dp1, dp2, eqmfb;
decl eqm1, eqm2, eqmhart, vrhart, vr1, vr2;

//MÉDIAS DOS DOMÍNIOS
mya=15;
mx=12;
myb=8;
myab=6;
mxab=6;
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\programas\\finais\\estimadores_1_2.txt", "a");

//CONSTRUINDO OS VALORES DA VARIÁVEL DE INTERESSE PARA OS DOMÍNIOS

//CONSTRUINDO A ESTRUTURA DAS OBSERVAÇÕES NA POPULAÇÃO PARA OS DOIS CADASTROS

//INÍCIO DO LOOP PARA AS ESTRUTURAS DE POPULAÇÃO ONDE O TAMANHO DE Nab VARIA
fprintf(arq, "RESULTADOS DA SIMULACAO");
//CONTADOR DE TEMPO
tempo=timer();
for(j=0; j<columns(N); j++)
{
println("#####");
println("VALOR DE Na, Nb e Nab: ", N[j]~N[j]~NAB[j]);
println("#####");
fprintf(arq, "#####");
println(arq, "VALOR DE Na, Nb e Nab: ", N[j]~N[j]~NAB[j]);
fprintf(arq, "#####");
ranseed(-1);
Xa = mx+2*rann(N[j],1);
Xab = mxab+rann(NAB[j],1);
XA=Xa|Xab;
Ya = mya*Xa+sqrt(Xa).*rann(N[j],1);
Yab = myab*Xab+sqrt(Xab).*rann(NAB[j],1);
cadcruz=Yab~Xab;
YA = Ya|Yab;
Yb = myb+3*rann(N[j],1);
YB = Yb|Yab;

//CADASTROS FORMADOS
cadA=YA~XA;
cadB=YB;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
NA = rows(YA);
NB = rows(YB);
Na = rows(Ya);
Nb = rows(Yb);
Nab = rows(cadcruz);
NTOT = Na+Nab+Nb;

//TOTAIS POPULACIONAIS
txA=sumc(XA);
tyA = sumc(YA);
tyB = sumc(YB);

```



```

for(l = 0; l < rows(amcadB); l++)
{
ord2 = uni2[l];
amcadB[l]=cadB[ord2];
}

// SEPARANDO AS OBSERVAÇÕES PARA CADA DOMÍNIO

//Domínio a
ya = zeros(NA,2);
for(m = 0; m < Na; m++)
{
for(n = 0; n < nA; n++)
{
if(amcadA[n][]==cadA[m][])
{
ya[n][]=amcadA[n][];
}
}
}
ya = selectifr(ya,ya.!=0);
ya = selectifr(ya,ya.!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(NA,2);
for(m = Na; m < NA; m++)
{
for(n = 0; n < nA; n++)
{
if(amcadA[n][]==cadA[m][])
{
yab1[n][]=amcadA[n][];
}
}
}
yab1 = selectifr(yab1,yab1.!=0);
yab1 = selectifr(yab1,yab1.!=.NaN);
nab1 = rows(yab1);

//Domínio b
yb = zeros(NB,1);
for(m = 0; m < Nb; m++)
{
for(n = 0; n < nB; n++)
{
if(amcadB[n]==cadB[m])
{
yb[n]=amcadB[n];
}
}
}
yb = selectifr(yb,yb.!=0);
yb = selectifr(yb,yb.!=.NaN);
nb = rows(yb);

//Domínio ab, CADASTRO B
yab2 = zeros(NA,1);
for(m = Nb; m < NB; m++)
{
for(n = 0; n < nB; n++)
{

```

```

if (amcadB[n]==cadB[m])
{
yab2[n]=amcadB[n];
}
}
}
yab2 = selectifr(yab2,yab2.!=0);
yab2 = selectifr(yab2,yab2.!=.NaN);
nab2 = rows(yab2);
//fprintln(arq,"na: ",na);
//fprintln(arq,"nab1: ",nab1);
//fprintln(arq,"nb: ",nb);
//fprintln(arq,"nab2: ",nab2);

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya=meanc(ya[] [0]);
estxa=meanc(xa[] [1]);
estyab1=meanc(yab1[] [0]);
estxab1=meanc(xab1[] [1]);
estyb=meanc(yb);
estyab2=meanc(yab2);
varya=varc(ya[] [0]);
varyb=varc(yb);
varyab1=varc(yab1[] [0]);
varyab2=varc(yab2);
varxa=varc(ya[] [1]);
varxab1=varc(yab1[] [1]);

//ESTIMADOR PROPOSTO POR HARTLEY

phart=(2*Nb*fBinv*pab2*varyab2)/(2*NA*fAinv*pab1*varyab1+2*Nb*fBinv*pab2*varyab2);
esthart[i]=(Na*estya + Nab*(phart*estyab1+(1-phart)*estyab2) + Nb*estyb)/NTOT;

//CÁLCULOS - ESTIMADOR 1
ra=estya/estxa;
rab1=estyab1/estxab1;
cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e1=NA^2*((1-fA)/(nA))*pa*varya;
a2e1=NA^2*((1-fA)/(nA))*pab1*varyab1;
a3e1=ra^2*NA^2*((1-fA)/(nA))*pa*varxa;
a4e1=rab1^2*NA^2*((1-fA)/(nA))*pab1*varxab1;
a5e1=2*ra^2*(Na/na)^2*nA*(1-fA)*cor_a[0] [0]*sqrt(varya)*sqrt(varxa);
a6e1=2*rab1*(Nab/nab1)^2*nA*(1-fA)*cor_ab[0] [0]*sqrt(varyab1)*sqrt(varxab1);
b1e1=NB^2*((1-fB)/(nB))*pb*varyb;
b2e1=NB^2*((1-fB)/(nB))*pab2*varyab2;

//VALOR DE P1 e ESTIMADOR
p1 = b2e1/(a2e1+a4e1-a6e1+b2e1);
p1 = fabs(p1);
est1[i]=(estya/estxa)*txa +Nab*(p1*(estyab1/estxab1)*(txab/Nab)+(1-p1)*estyab2) + Nb*estyb)/NTOT;

//CÁLCULOS - ESTIMADOR 2
a1e2=NA^2*((1-fA)/(nA))*pa*varya;
a2e2=NA^2*((1-fA)/(nA))*pab1*varyab1;
a3e2=NA^2*((1-fA)/(nA))*varxa;
a4e2=NA^2*((1-fA)/(nA))*pab1*varxab1;
a5e2=2*(Na/na)^2*nA*(1-fA)*cor_a[0] [0]*sqrt(varya)*sqrt(varxa);
a6e2=2*(Nab/nab1)^2*nA*(1-fA)*cor_ab[0] [0]*sqrt(varyab1)*sqrt(varxab1);
b1e2=NB^2*((1-fB)/(nB))*pb*varyb;
b2e2=NB^2*((1-fB)/(nB))*pab2*varyab2;

//COEFICIENTES DO POLINÔMIO DE GRAU 6
k6 = b2e2*Nab^5*estxab1^5;

```

```

k5 = 5*b2e2*Na*estxa*Nab^4*estxab1^4-b2e2*Nab^5*estxab1^5;
k4 = 10*b2e2*Na^2*estxa^2*Nab^3*estxab1^3-5*b2e2*Na*estxa*Nab^4*estxab1^4;
k3 = -txA^2*Nab^2*estxab1*a4e2*Na*estya*estyab1+txA^2*Nab^2*estxab1^2*a6e2*Na*estya
-3*txA^2*Nab^2*estxab1*a6e2*estyab1*Na*estxa+2*txA^2*a4e2*Nab^2*estyab1^2*Na*estxa
+10*b2e2*Na^3*estxa^3*Nab^2*estxab1^2-10*b2e2*Na^2*estxa^2*Nab^3*estxab1^3+txA^2*Nab^2*estxab1^2*a2e2*Na*estxa;
k2 = -txA^2*Nab*estxab1*a4e2*Na^2*estya^2+3*txA^2*a4e2*Na^2*estya*Nab*estyab1*estxa
+2*txA^2*Nab*estxab1*a2e2*Na^2*estxa^2+5*b2e2*Na^4*estxa^4*Nab*estxab1
-txA^2*Nab^3*estxab1^3*a1e2-3*txA^2*a6e2*Nab*estyab1*Na^2*estxa^2-txA^2*Nab*estxab1*a6e2*Na^2*estya*estxa
-10*b2e2*Na^3*estxa^3*Nab^2*estxab1^2-txA^2*Nab^3*estxab1*a3e2*estyab1^2+2*txA^2*Nab^3*estxab1^2*a5e2*estyab1;
k1 = txA^2*Nab^2*estxab1*a5e2*estyab1*Na*estxa+txA^2*a4e2*Na^3*estya^2*estxa+txA^2*a2e2*Na^3*estxa^3
-2*txA^2*Nab^2*estxab1^2*a1e2*Na*estxa+b2e2*Na^5*estxa^5+3*txA^2*Nab^2*estxab1^2*a5e2*Na*estya
-2*txA^2*a6e2*Na^3*estya*estxa^2-5*b2e2*Na^4*estxa^4*Nab*estxab1-3*txA^2*Nab^2*estxab1*a3e2*Na*estya*estyab1
+txA^2*a3e2*Nab^2*estyab1^2*Na*estxa;
k0 = -b2e2*Na^5*estxa^5+3*txA^2*Nab*estxab1*a5e2*Na^2*estya*estxa-txA^2*Nab*estyab1*a5e2*Na^2*estxa^2
-txA^2*Nab*estxab1*a1e2*Na^2*estxa^2-2*txA^2*Nab*estxab1*a3e2*Na^2*estya^2+txA^2*a3e2*Nab*estyab1*Na^2*estya*estxa;
vk=k6~k5~k4~k3~k2~k1~k0;
polyroots(vk,&rootk);
p2=min(fabs(rootk[0] []));

//ESTIMADOR 2
est2[i] = ((Na*estya + p2*Nab*estyab1)/(Na*estxa + p2*Nab*estxab1))*txA + Nb*estyb + (1-p2)*Nab*estyab2)/NTOT;
}
//FIM DO LOOP DE MONTE CARLO

//OBTENÇÃO DE ESTIMATIVAS VIA MONTE CARLO
medhart = meanc(esthart);
medest1=meanc(est1);
medest2=meanc(est2);
varhart=varc(esthart);
varest1=varc(est1);
varest2=varc(est2);

//CÁLCULO DE VIÉS, VIÉS RELATIVO E ERRO QUADRÁTICO MÉDIO PARA CADA ESTIMADOR
vieshart = medhart - param;
vrhart = fabs(vieshart)/param;
vies1 = medest1 - param;
vr1 = fabs(vies1)/param;
vies2 = medest2 - param;
vr2 = fabs(vies2)/param;
dphart = sqrt(varhart);
dp1 = sqrt(varest1);
dp2 = sqrt(varest2);
eqmhart = varhart + vieshart^2;
eqm1 = varest1 + vies1^2;
eqm2 = varest2 + vies2^2;
vec1= medest1~vies1~vr1~dp1~eqm1~varest1;
vech = medhart~vieshart~vrhart~dphart~eqmhart~varhart;
vec2 = medest2~vies2~vr2~dp2~eqm2~varest2;

fprintf(arq, "VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintf(arq, "HARTLEY ", "%.5f ", vech);
fprintf(arq, "");
fprintf(arq, "EST 1 ", "%.5f ", vec1);
fprintf(arq, "");
fprintf(arq, "EST 2 ", "%.5f ", vec2);
fprintf(arq, "-----");
fprintf(arq, "-----");
}
//FIM DO LOOP PARA OS TAMANHOS DE AMOSTRA
println("FIM DO PASSO ", j);
println("#####");
println("-----");
}
//FIM DO LOOP PARA AS ESTRUTURAS DE POPULAÇÃO ONDE O TAMANHO DE Nab VARIA
fprintf(arq, "-----");

```



```
fprintln(arq, "Tempo de execucao:", timespan(tempo));
fprintln(arq, "-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}
//FIM DO PROGRAMA
```

PROGRAMA 2

```

#####
//##                                     #
//## Mestrado Em Estatística           #
//## Dissertação de Mestrado           #
//## Autor: Hemílio Fernandes           #
//## Orientador: Cristiano Ferraz       #
//##                                     #
//## Descrição: Esse programa avalia os estimadores propostos   #
//##                para o caso de existência de variável auxiliar #
//##                através do método de simulação de Monte Carlo. #
//##                Para o cenário 3: NA, NB, são conhecidos, porém #
//##                Na, Nb e Nab são desconhecidos                #
//##                                     #
#####

#include<oxstd.h>
#include<oxprob.h>
#import<maximize>
#import<solvenle>
#include <oxfloat.h>

//NÚMERO DE RÉPLICAS
static decl NREP = 1000;

// INÍCIO DO PROGRAMA
main()
{
  ranseed("GM");
  ranseed({1965,2001});

  decl N=<2000,1750,1500>;
  decl NAB=<500,1000,1500>;
  decl amostraA=<125,250,500,750,1000,1250>;
  decl amostraB=<125,250,500,750,1000,1250>;
  //decl amostraA=<1000,1000,1000,1000,1000,1000>;
  //decl amostraB=<200,300,400,500,700,900>;
  decl tempo,arq;
  decl amo, i, k, l, m, n, o , x, j, ind1;
  decl uni1, uni2, ord1, ord2;
  decl mya, mxa, myb, myab, mxab;
  decl YA, YB, XA, Xa, Xab, Ya, Yab, Yb, cadA, cadB, cadcruz;
  decl tyA,txA, tyB, txA, txab, tya, tyab, tyb, tygeral, param;
  decl NA, NB, Na, Nb, Nab, NTOT;
  decl yA, yB, ya, yb, yab1, yab2;
  decl nA, nB, na, nab1, nab2, nb;
  decl pa, pb, pab1, pab2;
  decl est3, est4, estfuller;
  decl dif3, dif4, dif_fuller;
  decl vies3, vies4 , viesfuller;
  decl exp3, exp4, expfuller;
  decl amcadA, amcadB;
  decl fA, fB, fAinv, fBinv, gA, gB;
  decl estya,estxa,estya1,estxab1,estyb,estya2,varya,varyb,varyab1,varyab2,varxa,varxab1;
  decl p3, p4, pfuller;
  decl a1e3, a2e3, a3e3, a4e3, a5e3, a6e3, b1e3, b2e3, m1e3, m2e3, cor_a, cor_ab;
  decl GA, texpa, texpa;
  decl a1e4, a2e4, a3e4, a4e4, a5e4, a6e4, b1e4, b2e4, m1e4, m2e4, m3e4, m4e4, m5e4, m6e4;
  decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
  decl medest3, medest4, medfuller, viesfb, vrfb, vr3, vr4;
  decl varest3, varest4, varfuller, dpfb, dp3, dp4;
  decl eqm3, eqm4, mv3, mv4, mvfuller, vareqm3, vareqm4, vareqmfuller, eqmfb;
  decl Naest, Nbest, Nabest, ntotest, paest, pbest, pab1est, pab2est;

```

```

decl c1, c2, c3, c, rootc, wfuller, vecfb, vec3, vec4;

//MÉDIAS DOS DOMÍNIOS
mya=15;
mxa=12;
myb=8;
myab=6;
mxab=6;
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\programas\\resultados_est3_4_finalissimo", "a");

//INÍCIO DO LOOP PARA AS ESTRUTURAS DE POPULAÇÃO ONDE O TAMANHO DE Nab VARIA
fprintf(arq,"RESULTADOS DA SIMULACAO");
tempo=timer();
for(j=0;j<columns(N);j++)
{
println("#####");
println("VALOR DE Na, Nb e Nab: ", N[j]~N[j]~NAB[j]);
println("#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, Nb e Nab: ", N[j]~N[j]~NAB[j]);
fprintf(arq,"#####");
ranseed(-1);

//CONSTRUÍDO A ESTRUTURA DAS OBSERVAÇÕES NA POPULAÇÃO PARA OS DOIS CADASTROS
Xa = mxa+2*rann(N[j],1);
Xab= mxab+rann(NAB[j],1);
XA=Xa|Xab;
Ya = mya*Xa+sqrt(Xa).*rann(N[j],1);
Yab = myab*Xab+sqrt(Xab).*rann(NAB[j],1);
cadcruz=Yab~Xab;
YA = Ya|Yab;
Yb = myb+3*rann(N[j],1);
YB = Yb|Yab;

//CADASTROS FORMADOS
cadA=YA~XA;
cadB=YB;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
NA = rows(YA);
NB = rows(YB);
Na = rows(Ya);
Nb = rows(Yb);
Nab = rows(cadcruz);
NTOT = Na+Nab+Nb;

//TOTAIS POPULACIONAIS
txA=sumc(XA);
tyA = sumc(YA);
tyB = sumc(YB);
txa = sumc(Xa);
txab=sumc(Xab);
tya=sumc(Ya);
tyab=sumc(Yab);
tyb=sumc(Yb);
tygeral = tyA + tyab + tyb;

//PARÂMETRO DE INTERESSE
param=tygeral/NTOT;

//PROPORÇÃO POPULACIONAL PARA CADA DOMÍNIO - QUANTIDADES CONHECIDAS

```

187

```

{
for(n = 0; n < nA; n++)
{
if(amcadA[n][]==cadA[m][])
{
ya[n][]=amcadA[n][];
}
}
}
ya = selectifr(ya,ya!=0);
ya = selectifr(ya,ya!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(NA,2);
for(m = Na; m < NA; m++)
{
for(n = 0; n < nA; n++)
{
if(amcadA[n][]==cadA[m][])
{
yab1[n][]=amcadA[n][];
}
}
}
yab1 = selectifr(yab1,yab1!=0);
yab1 = selectifr(yab1,yab1!=.NaN);
nab1 = rows(yab1);

//Domínio a
yb = zeros(NB,1);
for(m = 0; m < Nb; m++)
{
for(n = 0; n < nB; n++)
{
if(amcadB[n]==cadB[m])
{
yb[n]=amcadB[n];
}
}
}
yb = selectifr(yb,yb!=0);
yb = selectifr(yb,yb!=.NaN);
nb = rows(yb);

//Domínio ab, CADASTRO B
yab2 = zeros(NA,1);
for(m = Nb; m < NB; m++)
{
for(n = 0; n < nB; n++)
{
if(amcadB[n]==cadB[m])
{
yab2[n]=amcadB[n];
}
}
}
yab2 = selectifr(yab2,yab2!=0);
yab2 = selectifr(yab2,yab2!=.NaN);
nab2 = rows(yab2);

//PROPORÇÃO POPULACIONAL ESTIMADA PARA CADA DOMÍNIO
paest=na/nA;

```

```

pbest=nb/nB;
pab1est=nab1/nA;
pab2est=nab2/nB;

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya=meanc(ya[] [0]);
estxa=meanc(xa[] [1]);
estyab1=meanc(yab1[] [0]);
estxab1=meanc(xab1[] [1]);
estyb=meanc(yb);
estyab2=meanc(yab2);
varya=varc(ya[] [0]);
varyb=varc(yb);
varyab1=varc(yab1[] [0]);
varyab2=varc(yab2);
varxa=varc(ya[] [1]);
varxab1=varc(yab1[] [1]);

//ESTIMADOR PROPOSTO POR FULLER

c1 = nA*gB + nB*gA;
c2 = (-1)*(nA*nB*gB + nB*NA*gA + nab1*NA*gB + nab2*NB*gA);
c3= (nab1*gB + nab2*gA)*NA*NB;
c=c1~c2~c3;
polyroots(c,&rootc);
Nabest=min(rootc[0] []);
Naest = NA - Nabest;
Nbest = NB - Nabest;
ntotest = Naest+Nbest+Nabest;
wfuller=(nab1*(1-fB))/(nab1*(1-fB) + nab2*(1-fA));
medfuller=wfuller*estyab1 + (1-wfuller)*estyab2;
estfuller[i]=(Naest*estya + Nabest*medfuller + Nbest*estyb)/ntotest;

//CÁLCULOS - ESTIMADOR 3

cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e3=(Naest*estya)^2*((1-fA)/(nA*paest))*varxa;
a2e3=(Nabest*estxab1)^2*((1-fA)/(nA*pab1est))*varxab1;
a3e3=Naest^2*((1-fA)/(nA*paest))*varya;
a4e3=Nabest^2*((1-fA)/(nA*pab1est))*varyab1;
a5e3=2*Naest*Naest*estya*(nA/(na^2))*((1-fA)*cor_a[0] [0]*sqrt(varya)*sqrt(varxa);
a6e3=2*Nabest*Nabest*estxab1*(nA/(nab1^2))*((1-fA)*cor_ab[0] [0]*sqrt(varyab1)*sqrt(varxab1);
b1e3=(Nbest^2)*((1-fB)/(nB*pbest))*varyb;
b2e3=(Nabest^2)*((1-fB)/(nB*pab2est))*varyab2;
m1e3=(estyab2)^2*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));
m2e3=(estyb)^2*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));

//VALOR DE P1 e ESTIMADOR
p3 = (m1e3+b2e3)/(a2e3+a4e3-a6e3+b2e3+m1e3);
est3[i]=((estya/estxa)*txa +p3*(estyab1/estxab1)*txab + (1-p3)*Nabest*estyab2 + Nbest*estyb)/ntotest;

//CÁLCULOS - ESTIMADOR 4
a1e4=Naest^2*((1-fA)/(nA*paest))*varya;
a2e4=Nabest^2*((1-fA)/(nA*pab1est))*varyab1;
a3e4=Naest^2*((1-fA)/(nA*paest))*varxa;
a4e4=Nabest^2*((1-fA)/(nA*pab1est))*varxab1;
a5e4=2*Naest^2*(nA/(na^2))*((1-fA)*cor_a[0] [0]*sqrt(varya)*sqrt(varxa);
a6e4=2*Nabest^2*(nA/(nab1^2))*((1-fA)*cor_ab[0] [0]*sqrt(varyab1)*sqrt(varxab1);
b1e4=Nbest^2*((1-fB)/(nB*pbest))*varyb;

```

```

b2e4=Nabest^2*((1-fB)/(nB*pab2est))*varyab2;
m1e4=(estya^2)*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));
m2e4=(estxab2^2)*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));
m3e4=(estxa^2)*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));
m4e4=(estxab1^2)*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));
m5e4=(estyb^2)*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));
m6e4=(estxab2^2)*((Naest*Nbest*Nabest*gA*gB)/(nA*Nbest*gB+nB*Naest*gA));

//COEFICIENTES DO POLINÔMIO DE GRAU 6
k6 = 2*b2e4*Nabest^5*estxab1^5+2*m6e4*Nabest^5*estxab1^5;
k5 = -2*b2e4*Nabest^5*estxab1^5-2*m6e4*Nabest^5*estxab1^5
+10*b2e4*Naest*estxa*Nabest^4*estxab1^4+10*m6e4*Naest*estxa*Nabest^4*estxab1^4;
k4 = -10*b2e4*Naest*estxa*Nabest^4*estxab1^4
+ 20*b2e4*Naest^2*estxa^2*Nabest^3*estxab1^3-10*m6e4*Naest*estxa*Nabest^4*estxab1^4
+20*m6e4*Naest^2*estxa^2*Nabest^3*estxab1^3;
k3 = 4*txA^2*a4e4*Nabest^2*estxab1^2*Naest*estxa-2*txA^2*Nabest^2*estxab1*a4e4*Naest*estya*estxab1
-20*b2e4*Naest^2*estxa^2*Nabest^3*estxab1^3+2*txA^2*Nabest^2*estxab1^2*a2e4*Naest*estxa
+2*txA^2*Nabest^2*estxab1^2*m2e4*Naest*estxa-2*txA^2*Nabest^2*estxab1*m4e4*Naest*estya*estxab1
+20*m6e4*Naest^3*estxa^3*Nabest^2*estxab1^2-3*txA^2*Nabest^2*estxab1*a6e4*estya*Naest*estxa
-20*m6e4*Naest^2*estxa^2*Nabest^3*estxab1^3+20*b2e4*Naest^3*estxa^3*Nabest^2*estxab1^2
+txA^2*Nabest^2*estxab1^2*a6e4*Naest*estya+4*txA^2*m4e4*Naest^2*estxab1^2*Naest*estxa;
k2 = 4*txA^2*Nabest*estxab1*a2e4*Naest^2*estxa^2-20*b2e4*Naest^3*estxa^3*Nabest^2*estxab1^2
-2*txA^2*Nabest^3*estxab1^3*a1e4-2*txA^2*Nabest^3*estxab1*m3e4*estya*estxab1^2
-2*txA^2*Nabest*estxab1*a4e4*Naest^2*estya^2-3*txA^2*a6e4*Naest*estxab1*Naest^2*estxa^2
-2*txA^2*Nabest^3*estxab1^3*m1e4+2*txA^2*Nabest^3*estxab1^2*a5e4*estya*estxab1
+6*txA^2*m4e4*Naest^2*estya*Naest*estya*estxab1*estxa-2*txA^2*Nabest*estxab1*m4e4*Naest^2*estya^2
+10*b2e4*Naest^4*estxa^4*Naest*estxab1+6*txA^2*a4e4*Naest^2*estya*Naest*estya*estxab1*estxa
+4*txA^2*Nabest*estxab1*m2e4*Naest^2*estxa^2-txA^2*Nabest*estxab1*a6e4*Naest^2*estya*estxa
-2*txA^2*Nabest^3*estxab1*a3e4*estya*estxab1^2-20*m6e4*Naest^3*estxa^3*Nabest^2*estxab1^2
+10*m6e4*Naest^4*estxa^4*Naest*estxab1;
k1 = -10*b2e4*Naest^4*estxa^4*Naest*estxab1+2*b2e4*Naest^5*estxa^5+2*txA^2*a2e4*Naest^3*estxa^3
-6*txA^2*Nabest^2*estxab1*a3e4*Naest*estya*estya*estxab1+2*m6e4*Naest^5*estxa^5
+3*txA^2*Nabest^2*estxab1^2*a5e4*Naest*estya+txA^2*Nabest^2*estxab1*a5e4*estya*Naest*estxa
-4*txA^2*Nabest^2*estxab1^2*m1e4*Naest*estxa-10*m6e4*Naest^4*estxa^4*Naest*estxab1
-6*txA^2*Nabest^2*estxab1*m3e4*Naest*estya*estya*estxab1+2*txA^2*m3e4*Naest^2*estya*estxab1^2*Naest*estxa
+2*txA^2*m2e4*Naest^3*estxa^3+2*txA^2*a3e4*Naest^2*estya*estxab1^2*Naest*estxa
-4*txA^2*Nabest^2*estxab1^2*a1e4*Naest*estxa+2*txA^2*m4e4*Naest^3*estya^2*estxa
-2*txA^2*a6e4*Naest^3*estya*estxa^2+2*txA^2*a4e4*Naest^3*estya^2*estxa;
k0 = -2*b2e4*Naest^5*estxa^5-2*m6e4*Naest^5*estxa^5+3*txA^2*Nabest*estxab1*a5e4*Naest^2*estya*estxa
-2*txA^2*Nabest*estxab1*a1e4*Naest^2*estxa^2-4*txA^2*Nabest*estxab1*a3e4*Naest^2*estya^2
-2*txA^2*Nabest*estxab1*m1e4*Naest^2*estxa^2-4*txA^2*Nabest*estxab1*m3e4*Naest^2*estya^2
+2*txA^2*a3e4*Naest*estya*estxab1*Naest^2*estya*estxa-txA^2*Nabest*estya*estya*estxab1*a5e4*Naest^2*estxa^2
+2*txA^2*m3e4*Naest*estya*estxab1*Naest^2*estya*estxa;
vk=k6~k5~k4~k3~k2~k1~k0;
polyroots(vk,&rootk);
p4=min(fabs(rootk[0] []));
if (p4>1)
{
p4=fabs(1-p4);
}

//ESTIMADOR 4
est4[i] = (((Naest*estya + p4*Nabest*estya*estxab1)/(Naest*estxa + p4*Nabest*estxab1))*txA
+ Nbest*estya + (1-p4)*Nabest*estya*estxab2)/ntotest;
}

//FIM DO LOOP DE MONTE CARLO

//OBTENÇÃO DE ESTIMATIVAS VIA MONTE CARLO
medfuller = meanc(estfuller);
medest3=meanc(est3);
medest4=meanc(est4);
varfuller=varc(estfuller);
varest3=varc(est3);

```

```

varest4=varc(est4);

//CÁLCULO DE VIÉS, VIÉS RELATIVO E ERRO QUADRÁTICO MÉDIO PARA CADA ESTIMADOR
viesfuller = medfuller - param;
vrfb = fabs(viesfuller)/param;
vies3 = medest3 - param;
vr3 = fabs(vies3)/param;
vies4 = medest4 - param;
vr4 = fabs(vies4)/param;
dpfb = sqrt(varfuller);
dp3 = sqrt(varest3);
dp4 = sqrt(varest4);
eqmfb = varfuller + viesfuller^2;
eqm3 = varest3 + vies3^2;
eqm4 = varest4 + vies4^2;
vecfb = medfuller~viesfuller~vrfb~dpfb~eqmfb~varfuller;
vec3= medest3~vies3~vr3~dp3~eqm3~varest3;
vec4 = medest4~vies4~vr4~dp4~eqm4~varest4;

fprintf(arq, "VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintf(arq, "FULLER ", "%.5f ", vecfb);
fprintf(arq, "");
fprintf(arq, "EST 3 ", "%.5f ", vec3);
fprintf(arq, "");
fprintf(arq, "EST 4 ", "%.5f ", vec4);
fprintf(arq, "-----");
fprintf(arq, "-----");
//
}
//FIM DO LOOP PARA OS TAMANHOS DE AMOSTRA
println("FIM DO PASSO ", j);
println("#####");
println("-----");

}
//FIM DO LOOP PARA AS ESTRUTURAS DE POPULAÇÃO ONDE O TAMANHO DE Nab VARIA
fprintf(arq, "PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
fprintf(arq, "-----");
fprintf(arq, "Tempo de execucao:", timespan(tempo));
fprintf(arq, "-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}
//FIM DO PROGRAMA

```


PROGRAMA 3

```

#####
//##                                     #
//## Mestrado Em Estatística           #
//## Dissertação de Mestrado           #
//## Autor: Hemílio Fernandes           #
//## Orientador: Cristiano Ferraz      #
//##                                     #
//## Descrição: Esse programa avalia o estimador de Hartley           #
//## e o estimador 2 através do método de simulação                   #
//## de Monte Carlo.                                                  #
//## Caso onde U_B pertence a U_A                                       #
//##                                     #
#####

#include<oxstd.h>
#include<oxprob.h>
#import<maximize>
#import<solvenle>
#include <oxfloat.h>

//NÚMERO DE RÉPLICAS
static decl NREP = 1000;

//INÍCIO DO PROGRAMA
main()
{
  ranseed("GM");
  ranseed({1965,2001});
  decl N=<2000,1750,1500>;
  decl NAB=<500,1000,1500>;
  decl amostraA=<125,250,500,750,1000,1250,1750,2000>;
  decl amostraB=<25,50,100,150,200,250,300,350>;
  decl tempo,arq;
  decl amo, i, k, l, m, n, o , x, j, ind1;
  decl uni1, uni2, ord1, ord2;
  decl mya, mxa, myb, myab, mxab;
  decl YA, YB, XA, Xa, Xab, Ya, Yab, cadA, cadB, cadcruz;
  decl tyA,txA, tyB, txa, txab, tya, tyab, tygeral, param;
  decl NA, NB, Na, Nab, NTOT;
  decl yA, yB, ya, yab1, yab2;
  decl nA, nB, na, nab1, nab2;
  decl pa, pab1, pab2;
  decl esthart, est5;
  decl dif5, difhart;
  decl vies5, vieshart;
  decl exp5, exphart;
  decl amcadA, amcadB;
  decl fA, fB, fAinv, fBinv, gA, gB;
  decl estya, estxa, estyab1, estxab1, estyab2, varya, varyab1, varyab2, varxa, varxab1;
  decl p5, phart;
  decl a1e5, a2e5, a3e5, a4e5, a5e5, a6e5, b1e5, ra, rab, cor_a, cor_ab;
  decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
  decl medest5, medhart;
  decl varest5, varhart;
  decl eqm5, eqmhart, mv5, mvhart, vareqm5, vareqmhart,vrhart, vr5, dp hart, dp5,vec5,vech;
  decl uniA1,uniA2,uniA3,uniB1,uniB2,uniB3,amcadA1,amcadA2,amcadA3,amcadB1,amcadB2,amcadB3;
  decl rab1,a1e1,a2e1,a3e1,a4e1,a5e1,a6e1,b2e1;

//MÉDIAS DOS DOMÍNIOS
mya=15;

```

```

mx=12;
myb=8;
myab=6;
mxab=6;
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\programas\\resultados_screen_est1_ROCHEDO.txt", "a");

//INÍCIO DO LOOP PARA AS ESTRUTURAS DE POPULAÇÃO ONDE O TAMANHO DE Nab VARIA
fprintf(arq,"RESULTADOS DA SIMULACAO");
//CONTADOR DE TEMPO
tempo=timer();
for(j=0;j<columns(N);j++)
{
println("#####");
println("VALOR DE Na, e Nab=NB: ", N[j]~NAB[j]);
println("#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, e Nab=NB: ", N[j]~NAB[j]);
fprintf(arq,"#####");
ranseed(-1);

//CONSTRUINDO A ESTRUTURA DAS OBSERVAÇÕES NA POPULAÇÃO PARA OS DOIS CADASTROS

//CENÁRIO 1
Xa = mx+2*rann(N[j],1);
Xab= mxab+rann(NAB[j],1);
XA=Xa|Xab;
Ya = mya*Xa+sqrt(Xa).*rann(N[j],1);
Yab = myab*Xab+sqrt(Xab).*rann(NAB[j],1);
cadacruz=Yab~Xab;
YA = Ya|Yab;
YB = Yab;

//CENÁRIO 2
//Xa = mx+2*rann(N[j],1);
//Xab= mxab+rann(NAB[j],1);
//XA=Xa|Xab;
//Ya = mya+Xa+rann(N[j],1);
//Yab = myab+Xab+rann(NAB[j],1);
//cadacruz=Yab~Xab;
//YA = Ya|Yab;
//YB = Yab;

//CENÁRIO 3
//Xa = mx+2*rann(N[j],1);
//Xab= mxab+rann(NAB[j],1);
//XA=Xa|Xab;
//Ya = mya+rann(N[j],1);
//Yab = myab+rann(NAB[j],1);
//cadacruz=Yab~Xab;
//YA = Ya|Yab;
//YB = Yab;

//CADASTROS FORMADOS
cadA=YA~XA;
cadB=YB;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
NA = rows(YA);
NB = rows(YB);
Na = rows(Ya);
Nab = NB;

```

194

```

}

//RETIRANDO A AMOSTRA DO CADASTRO B
for(l = 0; l < rows(amcadB); l++)
{
ord2 = uni2[l];
amcadB[l]=cadB[ord2];
}

// SEPARANDO AS OBSERVAÇÕES PARA CADA DOMÍNIO
//Domínio a
ya = zeros(NA,2);
for(m = 0; m < Na; m++)
{
for(n = 0; n < nA; n++)
{
if(amcadA[n] []==cadA[m] [])
{
ya[n] []=amcadA[n] [];
}
}
}
ya = selectifr(ya,ya.!=0);
ya = selectifr(ya,ya.!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(NA,2);
for(m = Na; m < NA; m++)
{
for(n = 0; n < nA; n++)
{
if(amcadA[n] []==cadA[m] [])
{
yab1[n] []=amcadA[n] [];
}
}
}
yab1 = selectifr(yab1,yab1.!=0);
yab1 = selectifr(yab1,yab1.!=.NaN);
nab1 = rows(yab1);

yab2 = amcadB;
nab2 = rows(yab2);
nB = nab2;

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya=meanc(ya[] [0]);
estxa=meanc(ya[] [1]);
estyab1=meanc(yab1[] [0]);
estxab1=meanc(yab1[] [1]);
estyab2=meanc(yab2);
varya=varc(ya[] [0]);
varyab1=varc(yab1[] [0]);
varyab2=varc(yab2);
varxa=varc(ya[] [1]);
varxab1=varc(yab1[] [1]);

//ESTIMADOR PROPOSTO POR HARTLEY
phart=(2*Nb*fBinvpab2*varyab2)/(2*Na*fAinvpab1*varyab1+2*Nb*fBinvpab2*varyab2);

```

```

esthart[i]=(Na*estya + Nab*(phart*estyab1+(1-phart)*estyab2))/NTOT;

CÁLCULOS - ESTIMADOR 5

cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e5=NA^2*((1-fA)/(nA))*pa*varya;
a2e5=NA^2*((1-fA)/(nA))*pab1*varyab1;
a3e5=NA^2*((1-fA)/(nA))*pa*varxa;
a4e5=NA^2*((1-fA)/(nA))*pab1*varxab1;
a5e5=2*(Na/na)^2*nA*(1-fA)*cor_a[0][0]*sqrt(varya)*sqrt(varxa);
a6e5=2*(Nab/nab1)^2*nA*(1-fA)*cor_ab[0][0]*sqrt(varyab1)*sqrt(varxab1);
b1e5=NB^2*((1-fB)/(nB))*pab2*varyab2;

COEFICIENTES DO POLINÔMIO DE GRAU 6
k6 = b1e5*Nab^5*estxab1^5;
k5 = 5*b1e5*Na*estxa*Nab^4*estxab1^4-b1e5*Nab^5*estxab1^5;
k4 = 10*b1e5*Na^2*estxa^2*Nab^3*estxab1^3-5*b1e5*Na*estxa*Nab^4*estxab1^4;
k3 = -txA^2*Nab^2*estxab1*a4e5*Na*estya*estyab1+txA^2*Nab^2*estxab1^2*a6e5*Na*estya
-3*txA^2*Nab^2*estxab1*a6e5*estyab1*Na*estxa+2*txA^2*a4e5*Nab^2*estyab1^2*Na*estxa
+10*b1e5*Na^3*estxa^3*Nab^2*estxab1^2-10*b1e5*Na^2*estxa^2*Nab^3*estxab1^3
+txA^2*Nab^2*estxab1^2*a2e5*Na*estxa;
k2 = -txA^2*Nab*estxab1*a4e5*Na^2*estya^2+3*txA^2*a4e5*Na^2*estya*Nab*estyab1*estxa
+2*txA^2*Nab*estxab1*a2e5*Na^2*estxa^2+5*b1e5*Na^4*estxa^4*Nab*estxab1
-txA^2*Nab^3*estxab1^3*a1e5-3*txA^2*a6e5*Nab*estyab1*Na^2*estxa^2
-txA^2*Nab*estxab1*a6e5*Na^2*estya*estxa-10*b1e5*Na^3*estxa^3*Nab^2*estxab1^2
-txA^2*Nab^3*estxab1*a3e5*estyab1^2+2*txA^2*Nab^3*estxab1^2*a5e5*estyab1;
k1 = txA^2*Nab^2*estxab1*a5e5*estyab1*Na*estxa+txA^2*a4e5*Na^3*estya^2*estxa
+txA^2*a2e5*Na^3*estxa^3-2*txA^2*Nab^2*estxab1^2*a1e5*Na*estxa+b1e5*Na^5*estxa^5
+3*txA^2*Nab^2*estxab1^2*a5e5*Na*estya-2*txA^2*a6e5*Na^3*estya*estxa^2
-5*b1e5*Na^4*estxa^4*Nab*estxab1-3*txA^2*Nab^2*estxab1*a3e5*Na*estya*estyab1
+txA^2*a3e5*Nab^2*estyab1^2*Na*estxa;
k0 = -b1e5*Na^5*estxa^5+3*txA^2*Nab*estxab1*a5e5*Na^2*estya*estxa
-txA^2*Nab*estyab1*a5e5*Na^2*estxa^2-txA^2*Nab*estxab1*a1e5*Na^2*estxa^2
-2*txA^2*Nab*estxab1*a3e5*Na^2*estya^2+txA^2*a3e5*Nab*estyab1*Na^2*estya*estxa;
vk=k6~k5~k4~k3~k2~k1~k0;
polyroots(vk,&rootk);
p5=min(fabs(rootk[0][]));

ESTIMADOR 5
est5[i] = (((Na*estya + p5*Nab*estyab1)/(Na*estxa + p5*Nab*estxab1))*txA
+ (1-p5)*Nab*estyab2)/NTOT;
}
//FIM DO LOOP DE MONTE CARLO

//OBTENÇÃO DE ESTIMATIVAS VIA MONTE CARLO
medhart = meanc(esthart);
medest5=meanc(est5);
varhart=varc(esthart);
varest5=varc(est5);

//CÁLCULO DE VIÉS, VIÉS RELATIVO E ERRO QUADRÁTICO MÉDIO PARA CADA ESTIMADOR
vieshart = medhart - param;
vrhart = fabs(vieshart)/param;
vies5 = medest5 - param;
vr5 = fabs(vies5)/param;
dphart = sqrt(varhart);
dp5 = sqrt(varest5);
eqmhart = varhart + vieshart^2;
eqm5 = varest5 + vies5^2;
vec5 = medest5~vies5~vr5~dp5~eqm5~varest5;
vech = medhart~vieshart~vrhart~dphart~eqmhart~varhart;

fprintf(arq, "VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintf(arq, "HARTLEY ", "%.5f ", vech);

```

```

fprintln(arq,"");
fprintln(arq, "EST 5 ", "%.5f ", vec5);
fprintln(arq,"-----");
fprintln(arq,"-----");
}
//FIM DO LOOP PARA OS TAMANHOS DE AMOSTRA
println("FIM DO PASSO ", j);
println("#####");
println("-----");
}
//FIM DO LOOP PARA AS ESTRUTURAS DE POPULAÇÃO ONDE O TAMANHO DE Nab VARIA
fprintln(arq, "-----");
fprintln(arq, "Tempo de execucao:", timespan(tempo));
fprintln(arq, "-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}
//FIM DO PROGRAMA

```

PROGRAMA 4

```

#####
/**
/**# Mestrado Em Estatística #
/**# Dissertação de Mestrado #
/**# Autor: Hemílio Fernandes #
/**# Orientador: Cristiano Ferraz #
/**# #
/**# Descrição: Esse programa avalia os estimadores propostos #
/**# para o caso de existência de variável auxiliar #
/**# através do método de simulação de Monte Carlo. #
/**# Para o caso em que é considerado o plano de #
/**# Amostragem Estratificada (Estimadores Separados) #
/**# Estimadores avaliados: H, 1, 2, FB, 3 e 4 #
/**# #
/**# #
#####
#include<oxstd.h>
#include<oxprob.h>
#import<maximize>
#import<solvenle>
#include <oxfloat.h>

//NÚMERO DE RÉPLICAS
static decl NREP = 1000;

// INÍCIO DO PROGRAMA
main()
{
ranseed("GM");
ranseed({1965,2001});
//decl N=<2000,1750,1500>;
//decl NAB=<500,1000,1500>;
decl Na1=<3500,3000>;
decl NAB1=<500,1500>;
decl Na2=<1500,1250>;
decl NAB2=<500,1000>;
decl Na3=<1100,900>;
decl NAB3=<500,700>;
decl Nb1=<1500,1000>;
decl Nb2=<700,450>;
decl Nb3=<300,100>;
decl amostraA=<125,250,500,750,1000,1250>;
decl amostraB=<125,250,500,750,1000,1250>;
decl amcadA, amcadB, gA, gB, fA, fB, uni1, uni2, ord1, ord2;
decl h, i, j, k, l, m, n, o, p, q, r, s, t, u, amo, mc,ind1, fAinv, fBinv;
decl st, cadA1, cadA2, cadA3, cadA, cadB1, cadB2, cadB3, cadB;
decl VNA, VNB, VNa, VNb, VNab, na, nb, nab1, nab2, nA, nB;
decl ya, yb, yab1, yab2;
decl es1, es2, es3, es4, esh, esfb;
decl ec1, ec2, ec3, ec4, ech, ecfb;
decl esth, est1, est2, est3, est4, estfb;
decl a1e3,a2e3,a3e3,a4e3,a5e3,a6e3,b1e3,b2e3,m1e3,m2e3;
decl Xa, Xab, Ya, Yab, Yb, XA, YA, YB;
decl txA, tya, tyb, tyab, txa, txab, tyA, tyB, tygeral_st;
decl vs1, vs2, vs3, vs4, vsh, vsfb, eqms1, eqms2, eqms3, eqms4, eqmsh, eqmsfb;
decl vc1, vc2, vc3, vc4, vch, vcfb, eqmc1, eqmc2, eqmc3, eqmc4, eqmch, eqmcfb;
decl vrs1, vrs2, vrs3, vrs4, vrsh, vrsfb;
decl vars1, vars2, vars3, vars4, varsh, varsfb;
decl varc1, varc2, varc3, varc4, varch, varcfb;
decl dps1, dps2, dps3, dps4, dpsh, dpsfb;
decl dpc1, dpc2, dpc3, dpc4, dpch, dpcfb;
decl mya, myab, mxa, mxab, myb;

```

```

decl a1e1, a2e1, a3e1, a4e1, a5e1, a6e1, b1e1, b2e1, ra, rab1, cor_a, cor_ab;
decl a1e2, a2e2, a3e2, a4e2, a5e2, a6e2, b1e2, b2e2;
decl a1e4,a2e4,a3e4,a4e4,a5e4,a6e4,b1e4,b2e4,m1e4,m2e4,m3e4,m4e4,m5e4,m6e4;
decl estya,estxa,estyab1,estxab1,estyb,estyab2,varya,varyb,varyab1,varyab2,varxa,varxab1;
decl p1, p2, p3, p4, ph, pfb;
decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
decl u0, u1, u2, u3, u4, u5, u6, vu, rootu;
decl c0, c1, c2, c3, c4, c5, c6, vc, rootc;
decl NTOT, medpar, pargeral, valor;
decl pa, pb, pab1, pab2;
decl paest, pbest, pab1est, pab2est;
decl atual1, atual2, m1, m2, m3, m4, mh, mfb, vech, vec1, vec2, vec3, vec4, vecfb;
decl ntotest, Nabest,Naest,Nbest;
decl wfuller, medfuller, arq;
decl tempo;

//MÉDIAS DOS ESTRATOS
h = 3;
mya = <16,24,32>;
mxs = <8,12,16>;
myb = <7,9,2>;
myab = <10,12,14>;
mxab = <5,6,7>;

//INÍCIO DO LOOP(i) PARA AS ESTRUTURAS DE POPULAÇÃO ONDE Nab VARIA EM CADA ESTRATO
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\caso_estratificado_separado.txt", "a");
tempo=timer();
fprintf(arq,"RESULTADOS DA SIMULACAO: ");
for(i=0;i<columns(N);i++)
{
println("#####");
println("VALOR DE Na, Nb e Nab estrato 1: ", Na1[i]~Nb1[i]~NAB1[i]);
println("#####");
println("VALOR DE Na, Nb e Nab estrato 2: ", Na2[i]~Nb2[i]~NAB2[i]);
println("#####");
println("VALOR DE Na, Nb e Nab estrato 3: ", Na3[i]~Nb3[i]~NAB3[i]);
println("#####");
println("#####");
fprintf(arq,"VALOR DE Na, Nb e Nab estrato 1: ", Na1[i]~Nb1[i]~NAB1[i]);
fprintf(arq,"#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, Nb e Nab estrato 2: ", Na2[i]~Nb2[i]~NAB2[i]);
fprintf(arq,"#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, Nb e Nab estrato 3: ", Na3[i]~Nb3[i]~NAB3[i]);
fprintf(arq,"#####");
Xa1 = zeros(Na1[i],h);
Xa2 = zeros(Na2[i],h);
Xa3 = zeros(Na3[i],h);
Ya1 = zeros(Na1[i],h);
Ya2 = zeros(Na2[i],h);
Ya3 = zeros(Na3[i],h);
Xab1 = zeros(NAB1[i],h);
Xab2 = zeros(NAB2[i],h);
Xab3 = zeros(NAB3[i],h);
Yab = zeros(NAB[i],h);
Yb = zeros(N[i],h);
YA = zeros(N[i]+NAB[i],h);
XA = zeros(N[i]+NAB[i],h);
YB = zeros(N[i]+NAB[i],h);
NTOT=zeros(h,1);
VNA = zeros(h,1);

```



```

VNB = zeros(h,1);
VNa = zeros(h,1);
VNB = zeros(h,1);
VNab = zeros(h,1);
txA=zeros(h,1);
tyA = zeros(h,1);
tyB = zeros(h,1);
txa = zeros(h,1);
txab=zeros(h,1);
tya=zeros(h,1);
tyab=zeros(h,1);
tyb=zeros(h,1);
medpar=zeros(h,1);
pa =zeros(h,1);
pb =zeros(h,1);
pab1 =zeros(h,1);
pab2 =zeros(h,1);
nA=zeros(h,1);
nB=zeros(h,1);
esth=zeros(h,1);
estfb=zeros(h,1);
est1=zeros(h,1);
est2=zeros(h,1);
est3=zeros(h,1);
est4=zeros(h,1);
ntotest=zeros(h,1);
Nabest=zeros(h,1);

//LOOP PARA CADA ESTRATO
for(j=0;j<h;j++)
{
Xa[] [j]=mxa[j]+rann(N[i],1);
txa[j]=sumc(Xa[] [j]);
Xab[] [j]=mxab[j]+rann(NAB[i],1);
txab[j]=sumc(Xab[] [j]);
Yb[] [j]=myb[j]+rann(N[i],1);
Ya[] [j]=mya[j].*Xa[] [j]+sqrt(Xa[] [j]).*rann(N[i],1);
Yab[] [j]=myab[j].*Xab[] [j]+sqrt(Xab[] [j]).*rann(NAB[i],1);
YB[] [j]=Yb[] [j]|Yab[] [j];
YA[] [j]=Ya[] [j]|Yab[] [j];
XA[] [j]=Xa[] [j]|Xab[] [j];
}

//COMPOSIÇÃO DOS CADASTROS
cadA1=YA[] [0]~XA[] [0];
cadA2=YA[] [1]~XA[] [1];
cadA3=YA[] [2]~XA[] [2];
cadA=cadA1~cadA2~cadA3;
cadB1=YB[] [0];
cadB2=YB[] [1];
cadB3=YB[] [2];
cadB=cadB1~cadB2~cadB3;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
VNA = rows(YA);
VNB = rows(YB);
VNa = rows(Ya[] [0])~rows(Ya[] [1])~rows(Ya[] [2]);
VNa = VNa';
VNB = rows(Yb[] [0])~rows(Yb[] [1])~rows(Yb[] [2]);
VNB = VNB';
VNab = rows(Yab[] [0])~rows(Yab[] [1])~rows(Yab[] [2]);
VNab = VNab';
NTOT = (VNa[0]+VNab[0]+VNB[0])~(VNa[1]+VNab[1]+VNB[1])~(VNa[2]+VNab[2]+VNB[2]);

```

```

NTOT = NTOT';
txA=sumc(XA[] [0])~sumc(XA[] [1])~sumc(XA[] [2]);
txA=txA';
tyA = sumc(YA[] [0])~sumc(YA[] [1])~sumc(YA[] [2]);
tyA = tyA';
tyB = sumc(YB[] [0])~sumc(YB[] [1])~sumc(YB[] [2]);
tyB = tyB';
txa = sumc(Xa[] [0])~sumc(Xa[] [1])~sumc(Xa[] [2]);
txa = txa';
txab= sumc(Xab[] [0])~sumc(Xab[] [1])~sumc(Xab[] [2]);
txab = txab';
tya = sumc(Ya[] [0])~sumc(Ya[] [1])~sumc(Ya[] [2]);
tya = tya';
tyab= sumc(Yab[] [0])~sumc(Yab[] [1])~sumc(Yab[] [2]);
tyab=tyab';
tyb = sumc(Yb[] [0])~sumc(Yb[] [1])~sumc(Yb[] [2]);
tyb=tyb';
tygeral_st=(tya[0]+tyab[0]+tyb[0])~(tya[1]+tyab[1]+tyb[1])~(tya[2]+tyab[2]+tyb[2]);

//PARÂMETRO DE INTERESSE EM CADA ESTRATO
for(j=0;j<h;j++)
{
valor=tygeral_st[j]/NTOT[j];
medpar[j]=valor;
}

//PARÂMETRO TOTAL
pargerall=sumc(medpar);

//PROPORÇÃO POPULACIONAL - ESTIMADOR 1 e 2
pa[0]=VNa[0]/VNA;
pa[1]=VNa[1]/VNA;
pa[2]=VNa[2]/VNA;
pa=pa[0]~pa[1]~pa[2];
pa=pa';
pb[0]=VNB[0]/VNB;
pb[1]=VNB[1]/VNB;
pb[2]=VNB[2]/VNB;
pb=pb[0]~pb[1]~pb[2];
pb=pb';
pab1[0]=VNab[0]/VNA;
pab1[1]=VNab[1]/VNA;
pab1[2]=VNab[2]/VNA;
pab1=pab1[0]~pab1[1]~pab1[2];
pab1=pab1';
pab2[0]=VNab[0]/VNB;
pab2[1]=VNab[1]/VNB;
pab2[2]=VNab[2]/VNB;
pab2=pab2[0]~pab2[1]~pab2[2];
pab2=pab2';

println("-----");
println("VALOR POPULACIONAL DA MEDIA NO ESTRATO ", h-2, ": ", tygeral_st[0]/NTOT[0]);
println("VALOR POPULACIONAL DA MEDIA NO ESTRATO ", h-1, ": ", tygeral_st[1]/NTOT[1]);
println("VALOR POPULACIONAL DA MEDIA NO ESTRATO ", h, ": ", tygeral_st[2]/NTOT[2]);
println("VALOR POPULACIONAL DA MEDIA NA POPULACAO (PARAMETRO DE INTERESSE): ", pargerall/3);
println("-----");
fprintf(arq,"-----");
fprintf(arq,"VALOR POPULACIONAL DA MEDIA NO ESTRATO ", h-2, ": ", tygeral_st[0]/NTOT[0]);
fprintf(arq,"VALOR POPULACIONAL DA MEDIA NO ESTRATO ", h-1, ": ", tygeral_st[1]/NTOT[1]);
fprintf(arq,"VALOR POPULACIONAL DA MEDIA NO ESTRATO ", h, ": ", tygeral_st[2]/NTOT[2]);
fprintf(arq,"VALOR POPULACIONAL DA MEDIA NA POPULACAO (PARAMETRO DE INTERESSE): ", pargerall/3);
fprintf(arq,"-----");

```



```

ya = zeros(VNA,2);
for(m = 0; m < VNa[k]; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n] []==atual1[m] [])
{
ya[n] []=amcadA[n] [];
}
}
}
ya = selectifr(ya,ya.!=0);
ya = selectifr(ya,ya.!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(VNA,2);
for(m = VNa[k]; m < VNA; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n] []==atual1[m] [])
{
yab1[n] []=amcadA[n] [];
}
}
}
yab1 = selectifr(yab1,yab1.!=0);
yab1 = selectifr(yab1,yab1.!=.NaN);
nab1 = rows(yab1);

//Domínio b
yb = zeros(VNB,1);
for(m = 0; m < VNb[k]; m++)
{
for(n = 0; n < nB[k]; n++)
{
if(amcadB[n]==atual2[m])
{
yb[n]=amcadB[n];
}
}
}
yb = selectifr(yb,yb.!=0);
yb = selectifr(yb,yb.!=.NaN);
nb = rows(yb);

//Domínio ab, CADASTRO B
yab2 = zeros(VNA,1);
for(m = VNb[k]; m < VNB; m++)
{
for(n = 0; n < nB[k]; n++)
{
if(amcadB[n]==atual2[m])
{
yab2[n]=amcadB[n];
}
}
}
yab2 = selectifr(yab2,yab2.!=0);
yab2 = selectifr(yab2,yab2.!=.NaN);
nab2 = rows(yab2);

//PROPORÇÃO POPULACIONAL - ESTIMADOR 3 e 4
paest=na/nA[k];

```

```

pbest=nb/nB[k];
pab1est=nab1/nA[k];
pab2est=nab2/nB[k];

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya=meanc(ya[] [0]);
estxa=meanc(ya[] [1]);
estyab1=meanc(yab1[] [0]);
estxab1=meanc(yab1[] [1]);
estyb=meanc(yb);
estyab2=meanc(yab2);
varya=varc(ya[] [0]);
varyb=varc(yb);
varyab1=varc(yab1[] [0]);
varyab2=varc(yab2);
varxa=varc(ya[] [1]);
varxab1=varc(yab1[] [1]);

//ESTIMADOR DE HARTLEY
fAinv=VNA/nA[k];
fBinv=VNB/nB[k];
ph=(2*VNB*fBinv*pab2[k]*varyab2)/(2*VNA*fAinv*pab1[k]*varyab1+2*VNB*fBinv*pab2[k]*varyab2);
esth[k]=(VNa[k]*estya + VNab[k]*(ph*estyab1+(1-ph)*estyab2) + VNb[k]*estyb)/NTOT[k];

//ESTIMADOR 1
//CÁLCULOS - ESTIMADOR 1
ra=estya/estxa;
rab1=estyab1/estxab1;
cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e1=VNA^2*((1-fA)/(nA[k]))*pa[k]*varya;
a2e1=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varyab1;
a3e1=ra^2*VNA^2*((1-fA)/(nA[k]))*pa[k]*varxa;
a4e1=rab1^2*VNA^2*((1-fA)/(nA[k]))*pab1[k]*varxab1;
a5e1=2*ra^2*(VNa[k]/na)^2*nA[k]*(1-fA)*cor_a[0][0]*sqrt(varya)*sqrt(varxa);
a6e1=2*rab1*(VNab[k]/nab1)^2*nA[k]*(1-fA)*cor_ab[0][0]*sqrt(varyab1)*sqrt(varxab1);
b1e1=VNB^2*((1-fB)/(nB[k]))*pb[k]*varyb;
b2e1=VNB^2*((1-fB)/(nB[k]))*pab2[k]*varyab2;

//VALOR DE P1 e ESTIMADOR
p1 = b2e1/(a2e1+a4e1-a6e1+b2e1);
p1 = fabs(p1);
est1[k]=((estya/estxa)*txa[k] +VNab[k]*(p1*(estyab1/estxab1)
*(txab[k]/VNab[k])+(1-p1)*estyab2) + VNb[k]*estyb)/NTOT[k];

//ESTIMADOR 2
//CÁLCULOS - ESTIMADOR 2
a1e2=VNA^2*((1-fA)/(nA[k]))*pa[k]*varya;
a2e2=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varyab1;
a3e2=pa[k]*VNA^2*((1-fA)/(nA[k]))*varxa;
a4e2=pab1[k]*VNA^2*((1-fA)/(nA[k]))*varxab1;
a5e2=2*(VNa[k]/na)^2*nA[k]*(1-fA)*cor_a[0][0]*sqrt(varya)*sqrt(varxa);
a6e2=2*(VNab[k]/nab1)^2*nA[k]*(1-fA)*cor_ab[0][0]*sqrt(varyab1)*sqrt(varxab1);
b1e2=VNB^2*((1-fB)/(nB[k]))*pb[k]*varyb;
b2e2=VNB^2*((1-fB)/(nB[k]))*pab2[k]*varyab2;

//COEFICIENTES DO POLINÔMIO DE GRAU 6
k6 = b2e2*VNab[k]^5*estxab1^5;
k5 = 5*b2e2*VNa[k]*estxa*VNab[k]^4*estxab1^4-b2e2*VNab[k]^5*estxab1^5;
k4 = 10*b2e2*VNa[k]^2*estxa^2*VNab[k]^3*estxab1^3-5*b2e2*VNa[k]*estxa*VNab[k]^4*estxab1^4;
k3 = -txA[k]^2*VNab[k]^2*estxab1*a4e2*VNa[k]*estya*estyab1
+txA[k]^2*VNab[k]^2*estxab1^2*a6e2*VNa[k]*estya
-3*txA[k]^2*VNab[k]^2*estxab1*a6e2*estyab1*VNa[k]*estxa

```

```

+2*txA[k]^2*a4e2*VNab[k]^2*estyab1^2*VNa[k]*estxa+10*b2e2*VNa[k]^3*estxa^3*VNab[k]^2*estxab1^2
-10*b2e2*VNa[k]^2*estxa^2*VNab[k]^3*estxab1^3+txA[k]^2*VNab[k]^2*estxab1^2*a2e2*VNa[k]*estxa;
k2 = -txA[k]^2*VNab[k]*estxab1*a4e2*VNa[k]^2*estya^2+3*txA[k]^2*a4e2*VNa[k]^2*estya*VNab[k]*estyab1*estxa
+2*txA[k]^2*VNab[k]*estxab1*a2e2*VNa[k]^2*estxa^2+5*b2e2*VNa[k]^4*estxa^4*VNab[k]*estxab1
-txA[k]^2*VNab[k]^3*estxab1^3*a1e2-3*txA[k]^2*a6e2*VNab[k]*estyab1*VNa[k]^2*estxa^2
-txA[k]^2*VNab[k]*estxab1*a6e2*VNa[k]^2*estya*estxa-10*b2e2*VNa[k]^3*estxa^3*VNab[k]^2*estxab1^2
-txA[k]^2*VNab[k]^3*estxab1*a3e2*estyab1^2+2*txA[k]^2*VNab[k]^3*estxab1^2*a5e2*estyab1;
k1 = txA[k]^2*VNab[k]^2*estxab1*a5e2*estyab1*VNa[k]*estxa+txA[k]^2*a4e2*VNa[k]^3*estya^2*estxa
+txA[k]^2*a2e2*VNa[k]^3*estxa^3-2*txA[k]^2*VNab[k]^2*estxab1^2*a1e2*VNa[k]*estxa+b2e2*VNa[k]^5*estxa^5
+3*txA[k]^2*VNab[k]^2*estxab1^2*a5e2*VNa[k]*estya-2*txA[k]^2*a6e2*VNa[k]^3*estya*estxa^2
-5*b2e2*VNa[k]^4*estxa^4*VNab[k]*estxab1-3*txA[k]^2*VNab[k]^2*estxab1*a3e2*VNa[k]*estya*estyab1
+txA[k]^2*a3e2*VNab[k]^2*estyab1^2*VNa[k]*estxa;
k0 = -b2e2*VNa[k]^5*estxa^5+3*txA[k]^2*VNab[k]*estxab1*a5e2*VNa[k]^2*estya*estxa
-txA[k]^2*VNab[k]*estyab1*a5e2*VNa[k]^2*estxa^2-txA[k]^2*VNab[k]*estxab1*a1e2*VNa[k]^2*estxa^2
-2*txA[k]^2*VNab[k]*estxab1*a3e2*VNa[k]^2*estya^2+txA[k]^2*a3e2*VNab[k]*estyab1*VNa[k]^2*estya*estxa;
vk=k6^k5^k4^k3^k2^k1^k0;
polyroots(vk,&rootk);
p2=min(fabs(rootk[0]));

//ESTIMADOR 2
est2[k] = (((VNa[k]*estya + p2*VNab[k]*estyab1)/(VNa[k]*estxa + p2*VNab[k]*estxab1))*txA[k]
+ VNab[k]*estyb + (1-p2)*VNab[k]*estyab2)/NTOT[k];

//ESTIMADOR PROPOSTO POR FULLER

c1 = nA[k]*gB + nB[k]*gA;
c2 = (-1)*(nA[k]*VNB*gB + nB[k]*VNA*gA + nab1*VNA*gB + nab2*VNB*gA);
c3= (nab1*gB + nab2*gA)*VNA*VNB;
vc=c1^c2^c3;
polyroots(vc,&rootc);
Nabest[k]=min(rootc[0]);
Naest = VNA - Nabest[k];
Nbest = VNB - Nabest[k];
ntotest[k] = Naest+Nbest+Nabest[k];
wfuller=(nab1*(1-fB))/(nab1*(1-fB) + nab2*(1-fA));
medfuller=wfuller*estyab1 + (1-wfuller)*estyab2;
estfb[k]=(Naest*estya + Nabest[k]*medfuller + Nbest*estyb)/ntotest[k];

cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e3=(Naest*estya)^2*((1-fA)/(nA[k]*paest))*varxa;
a2e3=(Nabest[k]*estxab1)^2*((1-fA)/(nA[k]*pab1est))*varxab1;
a3e3=Naest^2*((1-fA)/(nA[k]*paest))*varya;
a4e3=Nabest[k]^2*((1-fA)/(nA[k]*pab1est))*varyab1;
a5e3=2*Naest*Naest*estya*(nA[k]/(na^2))*(1-fA)*cor_a[0][0]*sqrt(varya)*sqrt(varxa);
a6e3=2*Nabest[k]*Nabest[k]*estxab1*(nA[k]/(nab1^2))*(1-fA)*cor_ab[0][0]*sqrt(varyab1)*sqrt(varxab1);
b1e3=(Nbest^2)*((1-fB)/(nB[k]*pbest))*varyb;
b2e3=(Nabest[k]^2)*((1-fB)/(nB[k]*pab2est))*varyab2;
m1e3=(estyab2)^2*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));
m2e3=(estyb)^2*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));

//VALOR DE P1 e ESTIMADOR
p3 = (m1e3+b2e3)/(a2e3+a4e3-a6e3+b2e3+m1e3);
est3[k]=((estya/estxa)*txA[k] +p3*(estyab1/estxab1)*txab[k] + (1-p3)*Nabest[k]*estyab2 + Nbest*estyb)/ntotest[k];

//CÁLCULOS - ESTIMADOR 4
a1e4=Naest^2*((1-fA)/(nA[k]*paest))*varya;
a2e4=Nabest[k]^2*((1-fA)/(nA[k]*pab1est))*varyab1;
a3e4=Naest^2*((1-fA)/(nA[k]*paest))*varxa;
a4e4=Nabest[k]^2*((1-fA)/(nA[k]*pab1est))*varxab1;
a5e4=2*Naest^2*(nA[k]/(na^2))*(1-fA)*cor_a[0][0]*sqrt(varya)*sqrt(varxa);
a6e4=2*Nabest[k]^2*(nA[k]/(nab1^2))*(1-fA)*cor_ab[0][0]*sqrt(varyab1)*sqrt(varxab1);
b1e4=Nbest^2*((1-fB)/(nB[k]*pbest))*varyb;
b2e4=Nabest[k]^2*((1-fB)/(nB[k]*pab2est))*varyab2;

```

```

m1e4=(estya^2)*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));
m2e4=(estyab2^2)*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));
m3e4=(estxa^2)*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));
m4e4=(estxab1^2)*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));
m5e4=(estyb^2)*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));
m6e4=(estyab2^2)*((Naest*Nbest*Nabest[k]*gA*gB)/(nA[k]*Nbest*gB+nB[k]*Naest*gA));

//COEFICIENTES DO POLINÔMIO DE GRAU 6
u6 = 2*b2e4*Nabest[k]^5*estxab1^5+2*m6e4*Nabest[k]^5*estxab1^5;
u5 = -2*b2e4*Nabest[k]^5*estxab1^5-2*m6e4*Nabest[k]^5*estxab1^5+10*b2e4*Naest*estxa*Nabest[k]^4*estxab1^4
+10*m6e4*Naest*estxa*Nabest[k]^4*estxab1^4;
u4 = -10*b2e4*Naest*estxa*Nabest[k]^4*estxab1^4 + 20*b2e4*Naest^2*estxa^2*Nabest[k]^3*estxab1^3
-10*m6e4*Naest*estxa*Nabest[k]^4*estxab1^4+20*m6e4*Naest^2*estxa^2*Nabest[k]^3*estxab1^3;
u3 = 4*txA[k]^2*a4e4*Nabest[k]^2*estyab1^2*Naest*estxa-2*txA[k]^2*Nabest[k]^2*estxab1*a4e4*Naest*estya*estyab1
-20*b2e4*Naest^2*estxa^2*Nabest[k]^3*estxab1^3+2*txA[k]^2*Nabest[k]^2*estxab1*a2e4*Naest*estxa
+2*txA[k]^2*Nabest[k]^2*estxab1^2*m2e4*Naest*estxa-2*txA[k]^2*Nabest[k]^2*estxab1*m4e4*Naest*estya*estyab1
+20*m6e4*Naest^3*estxa^3*Nabest[k]^2*estxab1^2-3*txA[k]^2*Nabest[k]^2*estxab1*a6e4*estyab1*Naest*estxa
-20*m6e4*Naest^2*estxa^2*Nabest[k]^3*estxab1^3+20*b2e4*Naest^3*estxa^3*Nabest[k]^2*estxab1^2
+txA[k]^2*Nabest[k]^2*estxab1^2*a6e4*Naest*estya+4*txA[k]^2*m4e4*Nabest[k]^2*estyab1^2*Naest*estxa;
u2 = 4*txA[k]^2*Nabest[k]*estxab1*a2e4*Naest^2*estxa^2-20*b2e4*Naest^3*estxa^3*Nabest[k]^2*estxab1^2
-2*txA[k]^2*Nabest[k]^3*estxab1^3*a1e4-2*txA[k]^2*Nabest[k]^3*estxab1*m3e4*estyab1^2
-2*txA[k]^2*Nabest[k]*estxab1*a4e4*Naest^2*estya^2-3*txA[k]^2*a6e4*Nabest[k]*estyab1*Naest^2*estxa^2
-2*txA[k]^2*Nabest[k]^3*estxab1^3*m1e4+2*txA[k]^2*Nabest[k]^3*estxab1^2*a5e4*estyab1
+6*txA[k]^2*m4e4*Naest^2*estya*Nabest[k]*estyab1*estxa-2*txA[k]^2*Nabest[k]*estxab1*m4e4*Naest^2*estya^2
+10*b2e4*Naest^4*estxa^4*Nabest[k]*estxab1+6*txA[k]^2*a4e4*Naest^2*estya*Nabest[k]*estyab1*estxa
+4*txA[k]^2*Nabest[k]*estxab1*m2e4*Naest^2*estxa^2-txA[k]^2*Nabest[k]*estxab1*a6e4*Naest^2*estya*estxa
-2*txA[k]^2*Nabest[k]^3*estxab1*a3e4*estyab1^2-20*m6e4*Naest^3*estxa^3*Nabest[k]^2*estxab1^2
+10*m6e4*Naest^4*estxa^4*Nabest[k]*estxab1;
u1 = -10*b2e4*Naest^4*estxa^4*Nabest[k]*estxab1+2*b2e4*Naest^5*estxa^5+2*txA[k]^2*a2e4*Naest^3*estxa^3
-6*txA[k]^2*Nabest[k]^2*estxab1*a3e4*Naest*estya*estyab1+2*m6e4*Naest^5*estxa^5
+3*txA[k]^2*Nabest[k]^2*estxab1^2*a5e4*Naest*estya*txA[k]^2*Nabest[k]^2*estxab1*a5e4*estyab1*Naest*estxa
-4*txA[k]^2*Nabest[k]^2*estxab1^2*m1e4*Naest*estxa-10*m6e4*Naest^4*estxa^4*Nabest[k]*estxab1
-6*txA[k]^2*Nabest[k]^2*estxab1*m3e4*Naest*estya*estyab1+2*txA[k]^2*m3e4*Nabest[k]^2*estyab1^2*Naest*estxa
+2*txA[k]^2*m2e4*Naest^3*estxa^3+2*txA[k]^2*a3e4*Nabest[k]^2*estyab1^2*Naest*estxa
-4*txA[k]^2*Nabest[k]^2*estxab1^2*a1e4*Naest*estxa+2*txA[k]^2*m4e4*Naest^3*estya^2*estxa
-2*txA[k]^2*a6e4*Naest^3*estya*estxa^2+2*txA[k]^2*a4e4*Naest^3*estya^2*estxa;
u0 = -2*b2e4*Naest^5*estxa^5-2*m6e4*Naest^5*estxa^5+3*txA[k]^2*Nabest[k]*estxab1*a5e4*Naest^2*estya*estxa
-2*txA[k]^2*Nabest[k]*estxab1*a1e4*Naest^2*estxa^2-4*txA[k]^2*Nabest[k]*estxab1*a3e4*Naest^2*estya^2
-2*txA[k]^2*Nabest[k]*estxab1*m1e4*Naest^2*estxa^2-4*txA[k]^2*Nabest[k]*estxab1*m3e4*Naest^2*estya^2
+2*txA[k]^2*a3e4*Nabest[k]*estyab1*Naest^2*estya*estxa-txA[k]^2*Nabest[k]*estyab1*a5e4*Naest^2*estxa^2
+2*txA[k]^2*m3e4*Nabest[k]*estyab1*Naest^2*estya*estxa;
vu=u6~u5~u4~u3~u2~u1~u0;
polyroots(vu,&rootu);
p4=min(fabs(rootu[0] []));

//ESTIMADOR 4
est4[k] = (((Naest*estya + p4*Nabest[k]*estyab1)/(Naest*estxa + p4*Nabest[k]*estxab1))*txA[k]
+ Nbest*estyb + (1-p4)*Nabest[k]*estyab2)/ntotest[k];

}
//FIM DO LOOP(k) PARA OS ESTRATOS

//ESTIMADOR SEPARADO 1
esh[mc]=sumc(esth)/3;
esfb[mc]=sumc(estfb)/3;
es1[mc] =sumc(est1)/3;
es2[mc]=sumc(est2)/3;
es3[mc]=sumc(est3)/3;
es4[mc]=sumc(est4)/3;
}
//FIM DO LOOP(mc) DE MONTE CARLO
mh = meanc(esh);
mfb = meanc(esfb);
m1 = meanc(es1);

```

```

m2 = meanc(es2);
m3 = meanc(es3);
m4 = meanc(es4);
vsh = mh - pargeral/3;
vsfb = mfb - pargeral/3;
vs1 = m1 - pargeral/3;
vs2 = m2 - pargeral/3;
vs3 = m3 - pargeral/3;
vs4 = m4 - pargeral/3;
vrsh = fabs(vsh)/(pargeral/3);
vrsfb = fabs(vsfb)/(pargeral/3);
vrs1 = fabs(vs1)/(pargeral/3);
vrs2 = fabs(vs2)/(pargeral/3);
vrs3 = fabs(vs3)/(pargeral/3);
vrs4 = fabs(vs4)/(pargeral/3);
varsh = varc(esh);
varsfb = varc(esfb);
vars1 = varc(es1);
vars2 = varc(es2);
vars3 = varc(es3);
vars4 = varc(es4);
dpsh = sqrt(varsh);
dpsfb = sqrt(varsfb);
dps1 = sqrt(vars1);
dps2 = sqrt(vars2);
dps3 = sqrt(vars3);
dps4 = sqrt(vars4);
eqmsh = varsh + vsh^2;
eqmsfb = varsfb + vsfb^2;
eqms1 = vars1 + vs1^2;
eqms2 = vars2 + vs2^2;
eqms3 = vars3 + vs3^2;
eqms4 = vars4 + vs4^2;
vech = mh~vsh~vrsh~dpsh~eqmsh~varsh;
vecfb = mfb~vsfb~vrsfb~dpsfb~eqmsfb~varsfb;
vec1 = m1~vs1~vrs1~dps1~eqms1~vars1;
vec2 = m2~vs2~vrs2~dps2~eqms2~vars2;
vec3 = m3~vs3~vrs3~dps3~eqms3~vars3;
vec4 = m4~vs4~vrs4~dps4~eqms4~vars4;

fprintln(arq,"VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintln(arq,"HARTLEY SEPARADO", "%.5f ", vech);
fprintln(arq,"");
fprintln(arq,"FULLER SEPARADO", "%.5f ", vecfb);
fprintln(arq,"");
fprintln(arq,"ESTIMADOR SEPARADO 1 ", "%.5f ", vec1);
fprintln(arq,"");
fprintln(arq,"ESTIMADOR SEPARADO 2 ", "%.5f ", vec2);
fprintln(arq,"");
fprintln(arq,"ESTIMADOR SEPARADO 3 ", "%.5f ", vec3);
fprintln(arq,"");
fprintln(arq,"ESTIMADOR SEPARADO 4 ", "%.5f ", vec4);
fprintln(arq,"-----");
fprintln(arq,"-----");

}
//FIM DO LOOP(j) PARA OS TAMANHOS DE AMOSTRA
println("FIM DO PASSO ", j);
println("#####");
println("-----");
}
//FIM DO LOOP(i) PARA TAMANHOS DE POPULAÇÃO
fprintln(arq,"PRONTO, TERMINOU!!!!!!!!!!!!!!!!!!!!");
fprintln(arq,"-----");

```



```
fprintln(arq,"Tempo de execucao:", timespan(tempo));
fprintln(arq,"-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}
//FIM DO PROGRAMA
```

PROGRAMA 5

```

#####
/**#
/**# Mestrado Em Estatística
/**# Dissertação de Mestrado
/**# Autor: Hemílio Fernandes
/**# Orientador: Cristiano Ferraz
/**#
/**# Descrição: Esse programa avalia os estimadores propostos
/**# para o caso de existência de variável auxiliar
/**# através do método de simulação de Monte Carlo.
/**# Para o caso em que é considerado o plano de
/**# Amostragem Estratificada (Estimadores Separados)
/**# Estimadores avaliados: H e 2.
/**#
/**#
#####
#include<oxstd.h>
#include<oxprob.h>
#import<maximize>
#import<solvenle>
#include <oxfloat.h>

//NÚMERO DE RÉPLICAS
static decl NREP = 1000;

// INÍCIO DO PROGRAMA
main()
{
ranseed("GM");
ranseed({1965,2001});
static decl N=<2000,1750,1500,1250,1000,750>;
static decl NAB=<500,1000,1500,2000,2500,3000>;
static decl amostraA=<125,250,500,750,1000,1250,1500,1750,500,500,500,500,500,500,500>;
static decl amostraB=<25,50,100,150,200,250,300,350,25,50,100,150,200,250,300,350>;
decl amcadA, amcadB, gA, gB, fA, fB, uni1, uni2, ord1, ord2;
decl h, i, j, k, l, m, n, o, p, q, r, s, t, u, amo, mc, ind1, fAinv, fBinv;
decl st, cadA1, cadA2, cadA3, cadA, cadB1, cadB2, cadB3, cadB;
decl VNA, VNB, VNa, VNb, VNab, na, nb, nab1, nab2, nA, nB;
decl ya, yb, yab1, yab2;
decl es5,esh;
decl ec5, ech;
decl esth, est5;
decl a1e3,a2e3,a3e3,a4e3,a5e3,a6e3,b1e3,b2e3,m1e3,m2e3;
decl Xa, Xab, Ya, Yab, Yb, XA, YA, YB;
decl txA, tya, tyb, tyab, txab, tyA, tyB, tygeral_st;
decl vs5, vsh, eqmsh, eqms5;
decl vc5, vch, eqmch, eqmc5;
decl vrs5, vrsh;
decl vars5, varsh;
decl varc5, varch;
decl dps5, dps5;
decl dpc5, dpch;
decl mya, myab, mxa, mxab, myb;
decl a1e1, a2e1, a3e1, a4e1, a5e1, a6e1, b1e1, b2e1, ra, rab1, cor_a, cor_ab;
decl a1e2, a2e2, a3e2, a4e2, a5e2, a6e2, b1e2, b2e2;
decl a1e4,a2e4,a3e4,a4e4,a5e4,a6e4,b1e4,b2e4,m1e4,m2e4,m3e4,m4e4,m5e4,m6e4;
decl estya,estxa,estyab1,estxab1,estyb,estyab2,varya,varyb,varyab1,varyab2,varxa,varxab1;
decl p5,ph;
decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
decl u0, u1, u2, u3, u4, u5, u6, vu, rootu;
decl c0, c1, c2, c3, c4, c5, c6, vc, rootc;
decl NTOT, medpar, pargeral, valor;

```

```

decl pa, pb, pab1, pab2;
decl paest, pbest, pab1est, pab2est;
decl atual1, atual2, m5, mh, vech, vec5;
decl tempo, arq;

//MÉDIAS DOS ESTRATOS
h = 3;
mya = <16,24,32>;
mxa = <8,12,16>;
myab = <10,12,14>;
mxab = <5,6,7>;

//INÍCIO DO LOOP(i) PARA AS ESTRUTURAS DE POPULAÇÃO ONDE Nab VARIA EM CADA ESTRATO
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\programas\\finais\\caso_estratificado_screening.txt", "a");
tempo=timer();
fprintf(arq,"RESULTADOS DA SIMULACAO: ");
for(i=0;i<columns(N);i++)
{
println("#####");
println("VALOR DE Na, Nb e Nab em cada estrato: ", N[i]~N[i]~NAB[i]);
println("#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, Nb e Nab em cada estrato: ", N[i]~N[i]~NAB[i]);
fprintf(arq,"#####");
Xa = zeros(N[i],h);
Ya = zeros(N[i],h);
Xab = zeros(NAB[i],h);
Yab = zeros(NAB[i],h);
YA = zeros(N[i]+NAB[i],h);
XA = zeros(N[i]+NAB[i],h);
YB = zeros(NAB[i],h);
NTOT=zeros(h,1);
VNA = zeros(h,1);
VNB = zeros(h,1);
VNa = zeros(h,1);
VNb = zeros(h,1);
VNab = zeros(h,1);
txA=zeros(h,1);
tyA = zeros(h,1);
tyB = zeros(h,1);
txa = zeros(h,1);
txab=zeros(h,1);
tya=zeros(h,1);
tyab=zeros(h,1);
tyb=zeros(h,1);
medpar=zeros(h,1);
pa =zeros(h,1);
pb =zeros(h,1);
pab1 =zeros(h,1);
pab2 =zeros(h,1);
nA=zeros(h,1);
nB=zeros(h,1);
esth=zeros(h,1);
est5=zeros(h,1);

//LOOP PARA CADA ESTRATO
for(j=0;j<h;j++)
{
Xa[j]=mxa[j]+rann(N[i],1);
txa[j]=sumc(Xa[j]);
Xab[j]=mxab[j]+rann(NAB[i],1);
txab[j]=sumc(Xab[j]);
Ya[j]=mya[j].*Xa[j]+sqrt(Xa[j]).*rann(N[i],1);
Yab[j]=myab[j].*Xab[j]+sqrt(Xab[j]).*rann(NAB[i],1);

```

```

YB[] [j]=Yab[] [j];
YA[] [j]=Ya[] [j]|Yab[] [j];
XA[] [j]=Xa[] [j]|Xab[] [j];
}

//COMPOSIÇÃO DOS CADASTROS
cadA1=YA[] [0]~XA[] [0];
cadA2=YA[] [1]~XA[] [1];
cadA3=YA[] [2]~XA[] [2];
cadA=cadA1~cadA2~cadA3;
cadB1=YB[] [0];
cadB2=YB[] [1];
cadB3=YB[] [2];
cadB=cadB1~cadB2~cadB3;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
VNA = rows(YA);
VNB = rows(YB);
VNa = rows(Ya[] [0])~rows(Ya[] [1])~rows(Ya[] [2]);
VNa = VNa';
VNab = rows(Yab[] [0])~rows(Yab[] [1])~rows(Yab[] [2]);
VNab = VNab';
NTOT = (VNa[0]+VNab[0]+VNB[0])~(VNa[1]+VNab[1]+VNB[1])~(VNa[2]+VNab[2]+VNB[2]);
NTOT = NTOT';
txA=sumc(XA[] [0])~sumc(XA[] [1])~sumc(XA[] [2]);
txA=txA';
tyA = sumc(YA[] [0])~sumc(YA[] [1])~sumc(YA[] [2]);
tyA = tyA';
tyB = sumc(YB[] [0])~sumc(YB[] [1])~sumc(YB[] [2]);
tyB = tyB';
txa = sumc(Xa[] [0])~sumc(Xa[] [1])~sumc(Xa[] [2]);
txa = txa';
txab= sumc(Xab[] [0])~sumc(Xab[] [1])~sumc(Xab[] [2]);
txab = txab';
tya = sumc(Ya[] [0])~sumc(Ya[] [1])~sumc(Ya[] [2]);
tya = tya';
tyab= sumc(Yab[] [0])~sumc(Yab[] [1])~sumc(Yab[] [2]);
tyab=tyab';
tygeral_st=(tya[0]+tyab[0])~(tya[1]+tyab[1])~(tya[2]+tyab[2]);

//PARÂMETRO DE INTERESSE EM CADA ESTRATO
for(j=0;j<h;j++)
{
valor=tygeral_st[j]/NTOT[j];
medpar[j]=valor;
}

//PARÂMETRO TOTAL
pargerall=sumc(medpar);

//PROPORÇÃO POPULACIONAL - ESTIMADOR 1 e 2
pa[0]=VNa[0]/VNA;
pa[1]=VNa[1]/VNA;
pa[2]=VNa[2]/VNA;
pa=pa[0]~pa[1]~pa[2];
pa=pa';
pab1[0]=VNab[0]/VNA;
pab1[1]=VNab[1]/VNA;
pab1[2]=VNab[2]/VNA;
pab1=pab1[0]~pab1[1]~pab1[2];
pab1=pab1';
pab2[0]=VNab[0]/VNB;
pab2[1]=VNab[1]/VNB;

```

212

```

ord1 = uni1[n];
amcadA[n][]=atual1[ord1][];
}

//RETIRANDO A AMOSTRA DO CADASTRO B
for(m = 0; m < rows(amcadB); m++)
{
ord2 = uni2[m];
amcadB[m]=atual2[ord2];
}

// SEPARANDO AS OBSERVAÇÕES PARA CADA DOMÍNIO EM CADA ESTRATO

//Domínio a
ya = zeros(VNA,2);
for(m = 0; m < VNa[k]; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n][]==atual1[m][])
{
ya[n][]=amcadA[n][];
}
}
}
ya = selectifr(ya,ya.!=0);
ya = selectifr(ya,ya.!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(VNA,2);
for(m = VNa[k]; m < VNA; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n][]==atual1[m][])
{
yab1[n][]=amcadA[n][];
}
}
}
yab1 = selectifr(yab1,yab1.!=0);
yab1 = selectifr(yab1,yab1.!=.NaN);
nab1 = rows(yab1);

//Domínio ab, CADASTRO B
yab2 = zeros(VNA,1);
for(m = VNb[k]; m < VNB; m++)
{
for(n = 0; n < nB[k]; n++)
{
if(amcadB[n]==atual2[m])
{
yab2[n]=amcadB[n];
}
}
}
yab2 = amcadB;
yab2 = selectifr(yab2,yab2.!=.NaN);
nab2 = rows(yab2);

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya=meanc(ya[][0]);
estxa=meanc(ya[][1]);

```

```

estyab1=mean(yab1[] [0]);
estxab1=mean(yab1[] [1]);
estyab2=mean(yab2);
varya=var(ya[] [0]);
varyab1=var(yab1[] [0]);
varyab2=var(yab2);
varxa=var(ya[] [1]);
varxab1=var(yab1[] [1]);

//ESTIMADOR DE HARTLEY
fAinv=VNA/nA[k];
fBinv=VNB/nB[k];
ph=(2*VNB*fBinv*pab2[k]*varyab2)/(2*VNA*fAinv*pab1[k]*varyab1+2*VNB*fBinv*pab2[k]*varyab2);
esth[k]=(VNA[k]*estya + VNa[k]*(ph*estyab1+(1-ph)*estyab2))/NTOT[k];

//ESTIMADOR 5
//CÁLCULOS - ESTIMADOR 5
cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e2=VNA^2*((1-fA)/(nA[k]))*pa[k]*varya;
a2e2=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varyab1;
a3e2=pa[k]*VNA^2*((1-fA)/(nA[k]))*varxa;
a4e2=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varxab1;
a5e2=2*(VNa[k]/na)^2*nA[k]*(1-fA)*cor_a[0][0]*sqrt(varya)*sqrt(varxa);
a6e2=2*(VNa[k]/nab1)^2*nA[k]*(1-fA)*cor_ab[0][0]*sqrt(varyab1)*sqrt(varxab1);
b1e2=VNB^2*((1-fB)/(nB[k]))*pab2[k]*varyab2;

//COEFICIENTES DO POLINÔMIO DE GRAU 6
k6 = b1e2*VNa[k]^5*estxab1^5;
k5 = 5*b1e2*VNa[k]*estxa*VNa[k]^4*estxab1^4-b1e2*VNa[k]^5*estxab1^5;
k4 = 10*b1e2*VNa[k]^2*estxa^2*VNa[k]^3*estxab1^3
-5*b1e2*VNa[k]*estxa*VNa[k]^4*estxab1^4;
k3 = -txA[k]^2*VNa[k]^2*estxab1*a4e2*VNa[k]*estya*estyab1
+txA[k]^2*VNa[k]^2*estxab1^2*a6e2*VNa[k]*estya
-3*txA[k]^2*VNa[k]^2*estxab1*a6e2*estyab1*VNa[k]*estxa+2*txA[k]^2*a4e2*VNa[k]^2*estyab1^2*VNa[k]*estxa
+10*b1e2*VNa[k]^3*estxa^3*VNa[k]^2*estxab1^2-10*b1e2*VNa[k]^2*estxa^2*VNa[k]^3*estxab1^3
+txA[k]^2*VNa[k]^2*estxab1^2*a2e2*VNa[k]*estxa;
k2 = -txA[k]^2*VNa[k]*estxab1*a4e2*VNa[k]^2*estya^2
+3*txA[k]^2*a4e2*VNa[k]^2*estya*VNa[k]*estyab1*estxa
+2*txA[k]^2*VNa[k]*estxab1*a2e2*VNa[k]^2*estxa^2+5*b1e2*VNa[k]^4*estxa^4*VNa[k]*estxab1
-txA[k]^2*VNa[k]^3*estxab1^3*a1e2-3*txA[k]^2*a6e2*VNa[k]*estyab1*VNa[k]^2*estxa^2
-txA[k]^2*VNa[k]*estxab1*a6e2*VNa[k]^2*estya*estxa-10*b1e2*VNa[k]^3*estxa^3*VNa[k]^2*estxab1^2
-txA[k]^2*VNa[k]^3*estxab1*a3e2*estyab1^2+2*txA[k]^2*VNa[k]^3*estxab1^2*a5e2*estyab1;
k1 = txA[k]^2*VNa[k]^2*estxab1*a5e2*estyab1*VNa[k]*estxa+txA[k]^2*a4e2*VNa[k]^3*estya^2*estxa
+txA[k]^2*a2e2*VNa[k]^3*estxa^3-2*txA[k]^2*VNa[k]^2*estxab1^2*a1e2*VNa[k]*estxa+b1e2*VNa[k]^5*estxa^5
+3*txA[k]^2*VNa[k]^2*estxab1^2*a5e2*VNa[k]*estya-2*txA[k]^2*a6e2*VNa[k]^3*estya*estxa^2
-5*b1e2*VNa[k]^4*estxa^4*VNa[k]*estxab1-3*txA[k]^2*VNa[k]^2*estxab1*a3e2*VNa[k]*estya*estyab1
+txA[k]^2*a3e2*VNa[k]^2*estyab1^2*VNa[k]*estxa;
k0 = -b1e2*VNa[k]^5*estxa^5+3*txA[k]^2*VNa[k]*estxab1*a5e2*VNa[k]^2*estya*estxa
-txA[k]^2*VNa[k]*estyab1*a5e2*VNa[k]^2*estxa^2-txA[k]^2*VNa[k]*estxab1*a1e2*VNa[k]^2*estxa^2
-2*txA[k]^2*VNa[k]*estxab1*a3e2*VNa[k]^2*estya^2+txA[k]^2*a3e2*VNa[k]*estyab1*VNa[k]^2*estya*estxa;
vk=k6~k5~k4~k3~k2~k1~k0;
polyroots(vk,&rootk);
p5=min(fabs(rootk[0] []));

//ESTIMADOR 5
est5[k] = (((VNa[k]*estya + p5*VNa[k]*estyab1)/(VNa[k]*estxa + p5*VNa[k]*estxab1))*txA[k]
+ (1-p5)*VNa[k]*estyab2)/NTOT[k];

}
//FIM DO LOOP(k) PARA OS ESTRATOS

//ESTIMADOR SEPARADO 1

```

```

esh[mc]=sumc(esth)/3;
es5[mc]=sumc(est5)/3;
}
//FIM DO LOOP(mc) DE MONTE CARLO
mh = meanc(esh);
m5 = meanc(es5);
vsh = mh - pargeral/3;
vs5 = m5 - pargeral/3;
vrsh = fabs(vsh)/(pargeral/3);
vrs5 = fabs(vs5)/(pargeral/3);
varsh = varc(esh);
vars5 = varc(es5);
dpsh = sqrt(varsh);
dps5 = sqrt(vars5);
eqmsh = varsh + vsh^2;
eqms5 = vars5 + vs5^2;
vech = mh~vsh~vrsh~dpsh~eqmsh~varsh;
vec5 = m5~vs5~vrs5~dps5~eqms5~vars5;

fprintf(arq,"VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintf(arq,"HARTLEY SEPARADO", "%.5f ", vech);
fprintf(arq,"");
fprintf(arq,"ESTIMADOR SEPARADO 5 ", "%.5f ", vec5);
fprintf(arq,"-----");
fprintf(arq,"-----");

}
//FIM DO LOOP(j) PARA OS TAMANHOS DE AMOSTRA
println("FIM DO PASSO ", j);
println("#####");
println("-----");
}
//FIM DO LOOP(i) PARA TAMANHOS DE POPULAÇÃO
fprintf(arq,"PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
fprintf(arq,"-----");
fprintf(arq,"Tempo de execucao:", timespan(tempo));
fprintf(arq,"-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}

```


PROGRAMA 6

```

#####
//##                                     #
//## Mestrado Em Estatística           #
//## Dissertação de Mestrado           #
//## Autor: Hemílio Fernandes           #
//## Orientador: Cristiano Ferraz      #
//##                                     #
//## Descrição: Esse programa avalia os estimadores propostos   #
//##                para o caso de existência de variável auxiliar #
//##                através do método de simulação de Monte Carlo. #
//##                Para o caso em que é considerado o plano de   #
//##                Amostragem Estratificada (Estimadores Combinados) #
//##                Estimadores avaliados: H, 1, 2, FB, 3 e 4.     #
//##                                     #
//##                                     #
#####
#include<oxstd.h>
#include<oxprob.h>
#import<maximize>
#import<solvenle>
#include <oxfloat.h>

//NÚMERO DE RÉPLICAS
static decl NREP = 1000;

// INÍCIO DO PROGRAMA
main()
{
  ranseed("GM");
  ranseed({1965,2001});
  decl N=<2000,1750,1500>;
  decl NAB=<500,1000,1500>;
  decl amostraA=<125,250,500,750,1000,1250>;
  decl amostraB=<125,250,500,750,1000,1250>;
  decl amcadA, amcadB, gA, gB, fA, fB, uni1, uni2, ord1, ord2;
  decl h, i, j, k, l, m, n, o, p, q, r, s, t, u, amo, mc, ind1, fAinv, fBinv;
  decl st, cadA1, cadA2, cadA3, cadA, cadB1, cadB2, cadB3, cadB;
  decl VNA, VNB, VNa, VNb, VNab, na, nb, nab1, nab2, nA, nB;
  decl ya, yb, yab1, yab2;
  decl ec1, ec2, ec3, ec4, ech, ecfb;
  decl esth, est1, est2, est3, est4, estfb;
  decl a1e3,a2e3,a3e3,a4e3,a5e3,a6e3,b1e3,b2e3,m1e3,m2e3;
  decl Xa, Xab, Ya, Yab, Yb, XA, YA, YB;
  decl txA, tya, tyb, tyab, txa, txab, tyA, tyB, tygeral_st;
  decl vc1, vc2, vc3, vc4, vch, vcfb, eqmc1, eqmc2, eqmc3, eqmc4, eqmch, eqmcfb;
  decl vrc1, vrc2, vrc3, vrc4, vrch, vrcfb;
  decl varc1, varc2, varc3, varc4, varch, varcfb;
  decl dpc1, dpc2, dpc3, dpc4, dpch, dpcfb;
  decl mya, myab, mxa, mxab, myb;
  decl a1e1, a2e1, a3e1, a4e1, a5e1, a6e1, b1e1, b2e1, ra, rab1, cor_a, cor_ab;
  decl a1e2, a2e2, a3e2, a4e2, a5e2, a6e2, b1e2, b2e2;
  decl a1e4,a2e4,a3e4,a4e4,a5e4,a6e4,b1e4,b2e4,m1e4,m2e4,m3e4,m4e4,m5e4,m6e4;
  decl estya,estxa,estyab1,estxab1,estyb,estyab2,varya,varyb,varyab1,varyab2,varxa,varxab1;
  decl p1, p2, p3, p4, ph, pfb;
  decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
  decl u0, u1, u2, u3, u4, u5, u6, vu, rootu;
  decl c0, c1, c2, c3, c4, c5, c6, vc, rootc;
  decl NTOT, medpar, pargeral, valor;
  decl pa, pb, pab1, pab2;
  decl paest, pbest, pablest, pab2est;
  decl atual1, atual2, m1, m2, m3, m4, mh, mfb, vech, vec1, vec2, vec3, vec4, vecfb;
  decl ntotest, Nabest,Naest,Nbest;

```

```

decl wfuller, medfuller, arq;
decl tempo;
decl ngeral, ngeralest, txageral, txAgeral, txabgeral;

//MÉDIAS DOS ESTRATOS
h = 3;
mya = <10,10,10>;
mya = <8,12,16>;
myb = <7,9,2>;
myab = <5,5,5>;
mxab = <5,6,7>;

//INÍCIO DO LOOP(i) PARA AS ESTRUTURAS DE POPULAÇÃO ONDE Nab VARIA EM CADA ESTRATO
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\caso_estratificado_combinados.txt", "a");
tempo=timer();
fprintf(arq,"RESULTADOS DA SIMULACAO: ");
for(i=0;i<columns(N);i++)
{
println("#####");
println("VALOR DE Na, Nb e Nab em cada estrato: ", N[i]~N[i]~NAB[i]);
println("#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, Nb e Nab em cada estrato: ", N[i]~N[i]~NAB[i]);
fprintf(arq,"#####");
Xa = zeros(N[i],h);
Ya = zeros(N[i],h);
Xab = zeros(NAB[i],h);
Yab = zeros(NAB[i],h);
Yb = zeros(N[i],h);
YA = zeros(N[i]+NAB[i],h);
XA = zeros(N[i]+NAB[i],h);
YB = zeros(N[i]+NAB[i],h);
NTOT=zeros(h,1);
VNA = zeros(h,1);
VNB = zeros(h,1);
VNa = zeros(h,1);
VNb = zeros(h,1);
VNab = zeros(h,1);
txA=zeros(h,1);
tyA = zeros(h,1);
tyB = zeros(h,1);
txa = zeros(h,1);
txab=zeros(h,1);
tya=zeros(h,1);
tyab=zeros(h,1);
tyb=zeros(h,1);
medpar=zeros(h,1);
pa =zeros(h,1);
pb =zeros(h,1);
pab1 =zeros(h,1);
pab2 =zeros(h,1);
nA=zeros(h,1);
nB=zeros(h,1);
esth=zeros(h,1);
estfb=zeros(h,1);
est1=zeros(h,1);
est2=zeros(h,1);
est3=zeros(h,1);
est4=zeros(h,1);
ntotest=zeros(h,1);
Nabest=zeros(h,1);
estya=zeros(h,1);
estxa=zeros(h,1);
estyab1=zeros(h,1);
estxab1=zeros(h,1);

```

```

estyb=zeros(h,1);
estyab2=zeros(h,1);
varya=zeros(h,1);
varyb=zeros(h,1);
varyab1=zeros(h,1);
varyab2=zeros(h,1);
varxa=zeros(h,1);
varxab1=zeros(h,1);
ra=zeros(h,1);
rab1=zeros(h,1);
cor_a=zeros(h,1);
cor_ab=zeros(h,1);
a1e1=zeros(h,1);
a2e1=zeros(h,1);
a3e1=zeros(h,1);
a4e1=zeros(h,1);
a5e1=zeros(h,1);
a6e1=zeros(h,1);
b1e1=zeros(h,1);
b2e1=zeros(h,1);
a1e2=zeros(h,1);
a2e2=zeros(h,1);
a3e2=zeros(h,1);
a4e2=zeros(h,1);
a5e2=zeros(h,1);
a6e2=zeros(h,1);
b1e2=zeros(h,1);
b2e2=zeros(h,1);
a1e3=zeros(h,1);
a2e3=zeros(h,1);
a3e3=zeros(h,1);
a4e3=zeros(h,1);
a5e3=zeros(h,1);
a6e3=zeros(h,1);
b1e3=zeros(h,1);
b2e3=zeros(h,1);
m1e3=zeros(h,1);
m2e3=zeros(h,1);
a1e4=zeros(h,1);
a2e4=zeros(h,1);
a3e4=zeros(h,1);
a4e4=zeros(h,1);
a5e4=zeros(h,1);
a6e4=zeros(h,1);
b1e4=zeros(h,1);
b2e4=zeros(h,1);
m1e4=zeros(h,1);
m2e4=zeros(h,1);
m3e4=zeros(h,1);
m4e4=zeros(h,1);
m5e4=zeros(h,1);
m6e4=zeros(h,1);
Naest=zeros(h,1);
Nbest=zeros(h,1);
medfuller=zeros(h,1);

//LOOP PARA CADA ESTRATO
for(j=0;j<h;j++)
{
Xa[] [j]=mxa[j]+rann(N[i],1);
txa[j]=sumc(Xa[] [j]);
Xab[] [j]=mxab[j]+rann(NAB[i],1);
txab[j]=sumc(Xab[] [j]);
Yb[] [j]=myb[j]+rann(N[i],1);

```

```

Ya[] [j]=mya[j].*Xa[] [j]+sqrt(Xa[] [j]).*rann(N[i],1);
Yab[] [j]=myab[j].*Xab[] [j]+sqrt(Xab[] [j]).*rann(NAB[i],1);
YB[] [j]=Yb[] [j]|Yab[] [j];
YA[] [j]=Ya[] [j]|Yab[] [j];
XA[] [j]=Xa[] [j]|Xab[] [j];
}

//COMPOSIÇÃO DOS CADASTROS
cadA1=YA[] [0]~XA[] [0];
cadA2=YA[] [1]~XA[] [1];
cadA3=YA[] [2]~XA[] [2];
cadA=cadA1~cadA2~cadA3;
cadB1=YB[] [0];
cadB2=YB[] [1];
cadB3=YB[] [2];
cadB=cadB1~cadB2~cadB3;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
VNA = rows(YA);
VNB = rows(YB);
VNa = rows(Ya[] [0])~rows(Ya[] [1])~rows(Ya[] [2]);
VNa = VNa';
VNb = rows(Yb[] [0])~rows(Yb[] [1])~rows(Yb[] [2]);
VNB = VNB';
VNab = rows(Yab[] [0])~rows(Yab[] [1])~rows(Yab[] [2]);
VNab = VNab';
NTOT = (VNa[0]+VNab[0]+VNB[0])~(VNa[1]+VNab[1]+VNB[1])~(VNa[2]+VNab[2]+VNB[2]);
NTOT = NTOT';
txA=sumc(XA[] [0])~sumc(XA[] [1])~sumc(XA[] [2]);
txA=txA';
tyA = sumc(YA[] [0])~sumc(YA[] [1])~sumc(YA[] [2]);
tyA = tyA';
tyB = sumc(YB[] [0])~sumc(YB[] [1])~sumc(YB[] [2]);
tyB = tyB';
txa = sumc(Xa[] [0])~sumc(Xa[] [1])~sumc(Xa[] [2]);
txa = txa';
txab= sumc(Xab[] [0])~sumc(Xab[] [1])~sumc(Xab[] [2]);
txab = txab';
tya = sumc(Ya[] [0])~sumc(Ya[] [1])~sumc(Ya[] [2]);
tya = tyab';
tyab= sumc(Yab[] [0])~sumc(Yab[] [1])~sumc(Yab[] [2]);
tyab=tyab';
tyb = sumc(Yb[] [0])~sumc(Yb[] [1])~sumc(Yb[] [2]);
tyb=tyb';
tygeral_st=(tya[0]+tyab[0]+tyb[0])~(tya[1]+tyab[1]+tyb[1])~(tya[2]+tyab[2]+tyb[2]);

//PARÂMETRO DE INTERESSE EM CADA ESTRATO
for(j=0;j<h;j++)
{
valor=tygeral_st[j]/NTOT[j];
medpar[j]=valor;
}

//PARÂMETRO TOTAL
pargerall=sumc(medpar);

//PROPORÇÃO POPULACIONAL - ESTIMADOR 1 e 2
pa[0]=VNa[0]/VNA;
pa[1]=VNa[1]/VNA;
pa[2]=VNa[2]/VNA;
pa=pa[0]~pa[1]~pa[2];
pa=pa';
pb[0]=VNB[0]/VNB;

```



```

uni2 = sortcindex(ranu(rows(atual2),1));

//RETIRANDO AS AMOSTRAS DOS ESTRATOS
for(n=0; n<rows(amcadA);n++)
{
ord1 = uni1[n];
amcadA[n][]=atual1[ord1][];
}

//RETIRANDO A AMOSTRA DO CADASTRO B
for(m = 0; m < rows(amcadB); m++)
{
ord2 = uni2[m];
amcadB[m]=atual2[ord2];
}

// SEPARANDO AS OBSERVAÇÕES PARA CADA DOMÍNIO EM CADA ESTRATO

//Domínio a
ya = zeros(VNA,2);
for(m = 0; m < VNa[k]; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n][]==atual1[m][])
{
ya[n][]=amcadA[n][];
}
}
}
ya = selectifr(ya,ya.!=0);
ya = selectifr(ya,ya.!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(VNA,2);
for(m = VNa[k]; m < VNA; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n][]==atual1[m][])
{
yab1[n][]=amcadA[n][];
}
}
}
yab1 = selectifr(yab1,yab1.!=0);
yab1 = selectifr(yab1,yab1.!=.NaN);
nab1 = rows(yab1);

//Domínio b
yb = zeros(VNB,1);
for(m = 0; m < VNb[k]; m++)
{
for(n = 0; n < nB[k]; n++)
{
if(amcadB[n]==atual2[m])
{
yb[n]=amcadB[n];
}
}
}
yb = selectifr(yb,yb.!=0);
yb = selectifr(yb,yb.!=.NaN);
nb = rows(yb);

```

```

//Domínio ab, CADASTRO B
yab2 = zeros(VNA,1);
for(m = VNb[k]; m < VNB; m++)
{
for(n = 0; n < nB[k]; n++)
{
if(amcadB[n]==atual2[m])
{
yab2[n]=amcadB[n];
}
}
}
yab2 = selectifr(yab2,yab2.!=0);
yab2 = selectifr(yab2,yab2.!=.NaN);
nab2 = rows(yab2);

//PROPORÇÃO POPULACIONAL - ESTIMADOR 3 e 4
paest=na/nA[k];
pbest=nb/nB[k];
pab1est=nab1/nA[k];
pab2est=nab2/nB[k];

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya[k]=meanc(ya[] [0]);
estxa[k]=meanc(ya[] [1]);
estyab1[k]=meanc(yab1[] [0]);
estxab1[k]=meanc(yab1[] [1]);
estyb[k]=meanc(yb);
estyab2[k]=meanc(yab2);
varya[k]=varc(ya[] [0]);
varyb[k]=varc(yb);
varyab1[k]=varc(yab1[] [0]);
varyab2[k]=varc(yab2);
varxa[k]=varc(ya[] [1]);
varxab1[k]=varc(yab1[] [1]);

//QUANTIDADES PARA O ESTIMADOR DE HARTLEY
fAinv=VNA/nA[k];
fBinv=VNB/nB[k];

//ESTIMADOR 1
//CÁLCULOS - ESTIMADOR 1
ra[k]=estya[k]/estxa[k];
rab1[k]=estyab1[k]/estxab1[k];
cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e1[k]=VNA^2*((1-fA)/(nA[k]))*pa[k]*varya[k];
a2e1[k]=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varyab1[k];
a3e1[k]=ra[k]^2*VNA^2*((1-fA)/(nA[k]))*pa[k]*varxa[k];
a4e1[k]=rab1[k]^2*VNA^2*((1-fA)/(nA[k]))*pab1[k]*varxab1[k];
a5e1[k]=2*ra[k]^2*(VNa[k]/na)^2*nA[k]*(1-fA)*cor_a[0][0]*sqrt(varya[k])*sqrt(varxa[k]);
a6e1[k]=2*rab1[k]*(VNab[k]/nab1)^2*nA[k]*(1-fA)*cor_ab[0][0]*sqrt(varyab1[k])*sqrt(varxab1[k]);
b1e1[k]=VNB^2*((1-fB)/(nB[k]))*pb[k]*varyb[k];
b2e1[k]=VNB^2*((1-fB)/(nB[k]))*pab2[k]*varyab2[k];

//ESTIMADOR 2
//CÁLCULOS - ESTIMADOR 2
a1e2[k]=VNA^2*((1-fA)/(nA[k]))*pa[k]*varya[k];
a2e2[k]=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varyab1[k];
a3e2[k]=pa[k]*VNA^2*((1-fA)/(nA[k]))*varxa[k];

```

```

a4e2[k]=pab1[k]*VNA^2*((1-fA)/(nA[k]))*varxab1[k];
a5e2[k]=2*(VNa[k]/na)^2*nA[k]*(1-fA)*cor_a[0][0]*sqrt(varya[k])*sqrt(varxa[k]);
a6e2[k]=2*(VNab[k]/nab1)^2*nA[k]*(1-fA)*cor_ab[0][0]*sqrt(varyab1[k])*sqrt(varxab1[k]);
b1e2[k]=VNB^2*((1-fB)/(nB[k]))*pb[k]*varyb[k];
b2e2[k]=VNB^2*((1-fB)/(nB[k]))*pab2[k]*varyab2[k];

//CÁLCULOS PARA ESTIMADOR DE FULLER

c1 = nA[k]*gB + nB[k]*gA;
c2 = (-1)*(nA[k]*VNB*gB + nB[k]*VNA*gA + nab1*VNA*gB + nab2*VNB*gA);
c3= (nab1*gB + nab2*gA)*VNA*VNB;
vc=c1^2-c2^2-c3;
polyroots(vc,&rootc);
Nabest[k]=min(rootc[0]);
Naest[k] = VNA - Nabest[k];
Nbest[k] = VNB - Nabest[k];
ntotest[k] = Naest[k]+Nbest[k]+Nabest[k];
wfuller=(nab1*(1-fB))/(nab1*(1-fB) + nab2*(1-fA));
medfuller[k]=wfuller*estyab1[k] + (1-wfuller)*estyab2[k];

//CÁLCULOS - ESTIMADOR 3
cor_a=correlation(ya);
cor_ab=correlation(yab1);
a1e3[k]=(Naest[k]*estya[k])^2*((1-fA)/(nA[k]*paest))*varxa[k];
a2e3[k]=(Nabest[k]*estxab1[k])^2*((1-fA)/(nA[k]*pab1est))*varxab1[k];
a3e3[k]=Naest[k]^2*((1-fA)/(nA[k]*paest))*varya[k];
a4e3[k]=Nabest[k]^2*((1-fA)/(nA[k]*pab1est))*varyab1[k];
a5e3[k]=2*Naest[k]*Naest[k]*estya[k]*(nA[k]/(na^2))*(1-fA)*cor_a[0][0]*sqrt(varya[k])*sqrt(varxa[k]);
a6e3[k]=2*Nabest[k]*Nabest[k]*estxab1[k]*(nA[k]/(nab1^2))*(1-fA)*cor_ab[0][0]*sqrt(varyab1[k])*sqrt(varxab1[k]);
b1e3[k]=(Nbest[k]^2)*((1-fB)/(nB[k]*pbest))*varyb[k];
b2e3[k]=(Nabest[k]^2)*((1-fB)/(nB[k]*pab2est))*varyab2[k];
m1e3[k]=(estya[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));
m2e3[k]=(estyb[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));

//CÁLCULOS - ESTIMADOR 4
a1e4[k]=Naest[k]^2*((1-fA)/(nA[k]*paest))*varya[k];
a2e4[k]=Nabest[k]^2*((1-fA)/(nA[k]*pab1est))*varyab1[k];
a3e4[k]=Naest[k]^2*((1-fA)/(nA[k]*paest))*varxa[k];
a4e4[k]=Nabest[k]^2*((1-fA)/(nA[k]*pab1est))*varxab1[k];
a5e4[k]=2*Naest[k]^2*(nA[k]/(na^2))*(1-fA)*cor_a[0][0]*sqrt(varya[k])*sqrt(varxa[k]);
a6e4[k]=2*Nabest[k]^2*(nA[k]/(nab1^2))*(1-fA)*cor_ab[0][0]*sqrt(varyab1[k])*sqrt(varxab1[k]);
b1e4[k]=Nbest[k]^2*((1-fB)/(nB[k]*pbest))*varyb[k];
b2e4[k]=Nabest[k]^2*((1-fB)/(nB[k]*pab2est))*varyab2[k];
m1e4[k]=(estya[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));
m2e4[k]=(estyab2[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));
m3e4[k]=(estxa[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));
m4e4[k]=(estxab1[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));
m5e4[k]=(estyb[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));
m6e4[k]=(estyab2[k]^2)*((Naest[k]*Nbest[k]*Nabest[k]*gA*gB)/(nA[k]*Nbest[k]*gB+nB[k]*Naest[k]*gA));

}
//FIM DO LOOP(k) PARA OS ESTRATOS
ngeral=sumc(NTOT);
ngerales=sumc(ntotest);
txabgeral=sumc(txab);
txageral=sumc(txa);
txAgeral=sumc(txA);

//ESTIMADOR DE HARTLEY COMBINADO
ph=(2*sumc(VNB*fBinV*pab2.*varyab2))/(2*sumc(VNA*fAinV*pab1.*varyab1)+2*sumc(VNB*fBinV*pab2.*varyab2));
ech[mc]=(sumc(VNa.*estya) + ph*sumc(VNab.*estyab1)+(1-ph)*sumc(VNab.*estyab2) + sumc(VNB.*estyb))/ngeral;

```



```

//VALOR DE P1 e ESTIMADOR
p1 = sumc(b2e1)/(sumc(a2e1)+sumc(a4e1)-sumc(a6e1)+sumc(b2e1));
p1 = fabs(p1);
ec1[mc]=((sumc(estya)/sumc(estxa))*txageral +p1*(sumc(estyab1)/sumc(estxab1))
*txabgeral+(1-p1)*sumc(VNab.*estyab2) + sumc(VNb.*estyb))/ngeral;

//COEFICIENTES DO POLINÔMIO DE GRAU 6 - ESTIMADOR COMBINADO 2
k6 = sumc(b2e2)*sumc(VNab.*estxab1);
k5 = -sumc(b2e2)*sumc(VNab.*estxab1)+5*sumc(b2e2)*sumc(VNa.*estya)*sumc(VNab.*estxab1);
k4 = 10*sumc(b2e2)*sumc(VNa.*estxa)^2*sumc(VNab.*estxab1)^3
-5*sumc(b2e2)*sumc(VNa.*estxa)*sumc(VNab.*estxab1)^4;
k3 = txAgeral^2*sumc(VNab.*estyab1)^2*sumc(a6e2)*sumc(VNa.*estya)
- 3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a6e2)*sumc(VNab.*estyab1)*sumc(VNa.*estya)
+txAgeral^2*sumc(VNab.*estxab1)*sumc(a2e2)*sumc(VNa.*estxa)
-10*sumc(b2e2)*sumc(VNa.*estxa)^2*sumc(VNab.*estxab1)^3
+10*sumc(b2e2)*sumc(VNa.*estxa)^3*sumc(VNab.*estxab1)^2
-txAgeral^2*sumc(VNab.*estxab1)*sumc(a4e2)*sumc(VNa.*estya)*sumc(VNab.*estyab1)
+2*txAgeral^2*sumc(a4e2)*sumc(VNab.*estyab1)^2*sumc(VNa.*estxa);
k2 = -txAgeral^2*sumc(VNab.*estxab1)*sumc(a6e2)*sumc(VNa.*estya)*sumc(VNa.*estxa)
+2*txAgeral^2*sumc(VNab.*estxab1)*sumc(a2e2)*sumc(VNa.*estxa)^2
+5*sumc(b2e2)*sumc(VNa.*estxa)^4*sumc(VNab.*estxab1)-txAgeral^2*sumc(VNab.*estxab1)^3*sumc(a1e2)
-10*sumc(b2e2)*sumc(VNa.*estxa)^3*sumc(VNab.*estxab1)^2
-3*txAgeral^2*sumc(a6e2)*sumc(VNab.*estyab1)*sumc(VNa.*estya)^2
- txAgeral^2*sumc(VNab.*estxab1)*sumc(a4e2)*sumc(VNa.*estya)^2
+2*txAgeral^2*sumc(VNab.*estxab1)^2*sumc(a5e2)*sumc(VNab.*estyab1)
-txAgeral^2*sumc(VNab.*estyab1)*sumc(a3e2)*sumc(VNab.*estyab1)^2
+3*txAgeral^2*sumc(a4e2)*sumc(VNa.*estya)*sumc(VNab.*estyab1)*sumc(VNa.*estxa);
k1 = txAgeral^2*sumc(VNab.*estxab1)*sumc(a5e2)*sumc(VNab.*estyab1)*sumc(VNa.*estxa)
+sumc(b2e2)*sumc(VNa.*estxa)^5+3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a5e2)*sumc(VNa.*estya)
+txAgeral^2*sumc(a2e2)*sumc(VNa.*estxa)^3-sumc(b2e2)*sumc(VNa.*estxa)^4*sumc(VNab.*estxab1)
-2*txAgeral^2*sumc(VNab.*estxab1)^2*sumc(a1e2)*sumc(VNa.*estxa)
-3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a3e2)*sumc(VNa.*estya)*sumc(VNab.*estyab1)
+txAgeral^2*sumc(a4e2)*sumc(VNa.*estya)^2*sumc(VNa.*estxa)
-2*txAgeral^2*sumc(a6e2)*sumc(VNa.*estya)*sumc(VNa.*estxa)^2
+txAgeral^2*sumc(a3e2)*sumc(VNab.*estyab1)^2*sumc(VNa.*estxa);
k0 = txAgeral^2*sumc(VNab.*estyab1)*sumc(a5e2)*sumc(VNa.*estxa)^2
-txAgeral^2*sumc(VNab.*estxab1)*sumc(a1e2)*sumc(VNa.*estxa)^2
+txAgeral^2*sumc(a3e2)*sumc(VNa.*estya)*sumc(VNa.*estxa)
-2*txAgeral^2*sumc(VNab.*estxab1)*sumc(a3e2)*sumc(VNa.*estya)^2
-sumc(b2e2)*sumc(VNa.*estxa)^5+3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a5e2)*sumc(VNa.*estya)*sumc(VNa.*estxa);
vk=k6~k5~k4~k3~k2~k1~k0;
polyroots(vk,&rootk);
p2=min(fabs(rootk[0][ ]));

//ESTIMADOR COMBINADO 2
ec2[mc] = (((sumc(VNa.*estya) + p2*sumc(VNab.*estyab1))/(sumc(VNa.*estxa)
+ p2*sumc(VNab.*estxab1)))*txAgeral + sumc(VNb.*estyb) + (1-p2)*sumc(VNab.*estyab2))/ngeral;

//ESTIMADOR DE FULLER COMBINADO
ecfb[mc]=(sumc(Naest.*estya) + sumc(Nabest.*medfuller)
+ sumc(Nbest.*estyb))/ngeralest;

//ESTIMADOR COMIBNADO 3
p3 = (sumc(m1e3)+sumc(b2e3))/(sumc(a2e3)+sumc(a4e3)-sumc(a6e3)+sumc(b2e3)+sumc(m1e3));
ec3[mc]=((sumc(VNa.*estya)/sumc(VNa.*estxa))*txageral
+p3*(sumc(VNab.*estyab1)/sumc(VNab.*estxab1))*txabgeral
+ (1-p3)*sumc(Nabest.*estyab2) + sumc(Nbest.*estyb))/ngeralest;

//COEFICIENTES DO POLINÔMIO DE GRAU 6 - ESTIMADOR COMBINADO 4
u6 = 2*sumc(b2e4)*sumc(Nabest.*estxab1)^5+2*sumc(m6e4)*sumc(Nabest.*estxab1)^5;
u5 = -2*sumc(b2e4)*sumc(Nabest.*estxab1)^5-2*sumc(m6e4)*sumc(Nabest.*estxab1)^5
+10*sumc(b2e4)*sumc(Naest.*estxa)*sumc(Nabest.*estxab1)^4

```

```

+10*sumc(m6e4)*sumc(Naest.*estxa)*sumc(Nabest.*estxab1)^4;
u4 = -10*sumc(b2e4)*sumc(Naest.*estxa)*sumc(Nabest.*estxab1)^4
+ 20*sumc(b2e4)*sumc(Naest.*estxa)^2*sumc(Nabest.*estxab1)^3
+ 20*sumc(m6e4)*sumc(Naest.*estxa)^2*sumc(Nabest.*estxab1)^3
- 10*sumc(m6e4)*sumc(Naest.*estxa)*sumc(Nabest.*estxab1)^4;
u3 = -20*sumc(m6e4)*sumc(Naest.*estxa)^2*sumc(Nabest.*estxab1)^3
+ txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(a6e4)*sumc(Naest.*estya)
- 2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(m4e4)*sumc(Naest.*estya)*sumc(Nabest.*estya)
+4*txAgeral^2*sumc(a4e4)*sumc(Nabest.*estya)^2*sumc(Naest.*estxa)
+4*txAgeral^2*sumc(m4e4)*sumc(Nabest.*estya)^2*sumc(Naest.*estxa)
-2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a4e4)*sumc(Naest.*estya)*sumc(Nabest.*estya)
+2*txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(m2e4)*sumc(Naest.*estxa)
-3*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a6e4)*sumc(Nabest.*estya)*sumc(Naest.*estxa)
-20*sumc(b2e4)*sumc(Naest.*estxa)^2*sumc(Nabest.*estxab1)^3
+20*sumc(b2e4)*sumc(Naest.*estxa)^3*sumc(Nabest.*estxab1)^2
+2*txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(a2e4)*sumc(Naest.*estxa)
+20*sumc(m6e4)*sumc(Naest.*estxa)^3*sumc(Nabest.*estxab1)^2;
u2 = 2*txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(a5e4)*sumc(Nabest.*estya)
+4*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a2e4)*sumc(Naest.*estxa)^2
-20*sumc(m6e4)*sumc(Naest.*estxa)^3*sumc(Nabest.*estxab1)^2
-2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a4e4)*sumc(Naest.*estya)^2
+6*txAgeral^2*sumc(a4e4)*sumc(Naest.*estya)*sumc(Nabest.*estya)*sumc(Naest.*estxa)
-txAgeral^2*sumc(Nabest.*estxab1)*sumc(a6e4)*sumc(Naest.*estya)*sumc(Naest.*estxa)
-20*sumc(b2e4)*sumc(Naest.*estxa)^3*sumc(Nabest.*estxab1)^2
-2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(m3e4)*sumc(Nabest.*estya)^2
-2*txAgeral*sumc(Nabest.*estxab1)*sumc(m3e4)*sumc(Nabest.*estya)^2
-2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(m4e4)*sumc(Naest.*estya)^2
+6*txAgeral*sumc(m4e4)*sumc(Naest.*estya)*sumc(Nabest.*estya)*sumc(Naest.*estxa)
+10*sumc(b2e4)*sumc(Naest.*estxa)^4*sumc(Nabest.*estxab1)
+10*sumc(m6e4)*sumc(Naest.*estxa)^4*sumc(Nabest.*estxab1)
-2*txAgeral*sumc(Nabest.*estxab1)*sumc(a3e4)*sumc(Nabest.*estya)^2
+4*txAgeral*sumc(Nabest.*estxab1)*sumc(m2e4)*sumc(Naest.*estxa)^2
-2*txAgeral*sumc(Nabest.*estxab1)^3*sumc(a1e4)
-3*txAgeral^2*sumc(a6e4)*sumc(Nabest.*estya)*sumc(Naest.*estxa)^2
-2*txAgeral^2*sumc(Nabest.*estxab1)^3*sumc(m1e4);
u1 = 3*txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(a5e4)*sumc(Naest.*estya)
+2*sumc(b2e4)*sumc(Naest.*estxa)^5
-6*txAgeral^2*sumc(Nabest.*estxab1)*sumc(m3e4)*sumc(Naest.*estya)*sumc(Nabest.*estya)
+2*txAgeral^2*sumc(a2e4)*sumc(Naest.*estxa)^3
-10*sumc(m6e4)*sumc(Naest.*estxa)^4*sumc(Nabest.*estxab1)
-4*txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(a1e4)*sumc(Naest.*estxa)
-6*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a3e4)*sumc(Naest.*estya)*sumc(Nabest.*estya)
+2*txAgeral^2*sumc(a4e4)*sumc(Naest.*estya)^2*sumc(Naest.*estxa)
-4*txAgeral^2*sumc(Nabest.*estxab1)^2*sumc(m1e4)*sumc(Naest.*estxa)
+2*txAgeral^2*sumc(m2e4)*sumc(Naest.*estxa)^3
-10*sumc(b2e4)*sumc(Naest.*estxa)^4*sumc(Nabest.*estxab1)
+2*sumc(m6e4)*sumc(Naest.*estxa)^5
+txAgeral^2*sumc(Nabest.*estxab1)*sumc(a5e4)*sumc(Nabest.*estya)*sumc(Naest.*estxa)
-2*txAgeral^2*sumc(a6e4)*sumc(Naest.*estya)*sumc(Naest.*estxa)^2
+2*txAgeral^2*sumc(m3e4)*sumc(Nabest.*estya)^2*sumc(Naest.*estxa)
+2*txAgeral^2*sumc(a3e4)*sumc(Naest.*estya)^2*sumc(Naest.*estxa)
+2*txAgeral^2*sumc(m4e4)*sumc(Naest.*estya)^2*sumc(Naest.*estxa);
u0 = -4*txAgeral^2*sumc(Nabest.*estxab1)*sumc(m3e4)*sumc(Naest.*estya)^2
-2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a1e4)*sumc(Naest.*estxa)^2
+2*txAgeral^2*sumc(a3e4)*sumc(Nabest.*estya)*sumc(Naest.*estya)*sumc(Naest.*estxa)
-4*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a3e4)*sumc(Naest.*estya)^2
-txAgeral^2*sumc(Nabest.*estya)*sumc(a5e4)*sumc(Naest.*estxa)^2
-2*txAgeral^2*sumc(Nabest.*estxab1)*sumc(m1e4)*sumc(Naest.*estxa)^2
-2*sumc(m6e4)*sumc(Naest.*estxa)^5
+3*txAgeral^2*sumc(Nabest.*estxab1)*sumc(a5e4)*sumc(Naest.*estya)*sumc(Naest.*estxa)
-2*sumc(b2e4)*sumc(Naest.*estxa)^5
+2*txAgeral^2*sumc(m3e4)*sumc(Nabest.*estya)*sumc(Naest.*estya)*sumc(Naest.*estxa);
vu=u6~u5~u4~u3~u2~u1~u0;
polyroots(vu,&rootu);
p4=min(fabs(rootu[0] []));

```

```

if(p4>1)
{
p4=fabs(1-p4);
}

println(p2~p4);
//ESTIMADOR COMBINADO 4
ec4[mc] = (((sumc(Naest.*estya) + p4*sumc(Nabest.*estyab1))/(sumc(Naest.*estxa)
+ p4*sumc(Nabest.*estxab1)))*txAgeral + sumc(Nbest.*estyb) + (1-p4)*sumc(Nabest.*estyab2))/ngeralest;
ind1=mc+1;
if(ind1==NREP)
{
println("FIM para n = ", amostraA[amo]~amostraB[amo]);
}
}
//FIM DO LOOP(mc) DE MONTE CARLO
mh = meanc(ech);
mfb = meanc(ecfb);
m1 = meanc(ec1);
m2 = meanc(ec2);
m3 = meanc(ec3);
m4 = meanc(ec4);
vch = mh - pargeral/3;
vcfb = mfb - pargeral/3;
vc1 = m1 - pargeral/3;
vc2 = m2 - pargeral/3;
vc3 = m3 - pargeral/3;
vc4 = m4 - pargeral/3;
vrch = fabs(vch)/(pargeral/3);
vrcfb = fabs(vcfb)/(pargeral/3);
vrc1 = fabs(vc1)/(pargeral/3);
vrc2 = fabs(vc2)/(pargeral/3);
vrc3 = fabs(vc3)/(pargeral/3);
vrc4 = fabs(vc4)/(pargeral/3);
varch = varc(ech);
varcfb = varc(ecfb);
varc1 = varc(ec1);
varc2 = varc(ec2);
varc3 = varc(ec3);
varc4 = varc(ec4);
dpch = sqrt(varch);
dpcfb = sqrt(varcfb);
dpc1 = sqrt(varc1);
dpc2 = sqrt(varc2);
dpc3 = sqrt(varc3);
dpc4 = sqrt(varc4);
eqmch = varch + vch^2;
eqmcfb = varcfb + vcfb^2;
eqmc1 = varc1 + vc1^2;
eqmc2 = varc2 + vc2^2;
eqmc3 = varc3 + vc3^2;
eqmc4 = varc4 + vc4^2;
vech = mh~vch~vrch~dpch~eqmch~varch;
vecfb = mfb~vcfb~vrcfb~dpcfb~eqmcfb~varcfb;
vec1 = m1~vc1~vrc1~dpc1~eqmc1~varc1;
vec2 = m2~vc2~vrc2~dpc2~eqmc2~varc2;
vec3 = m3~vc3~vrc3~dpc3~eqmc3~varc3;
vec4 = m4~vc4~vrc4~dpc4~eqmc4~varc4;

fprintf(arq,"VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintf(arq,"HARTLEY COMBINADO", "%.5f ", vech);
fprintf(arq,"");
fprintf(arq,"ESTIMADOR COMBINADO 1 ", "%.5f ", vec1);
fprintf(arq,"");
fprintf(arq,"ESTIMADOR COMBINADO 2 ", "%.5f ", vec2);

```

```

fprintln(arq,"");
fprintln(arq,"FULLER COMBINADO", "%.5f ", vecfb);
fprintln(arq,"");
fprintln(arq,"ESTIMADOR COMBINADO 3 ", "%.5f ", vec3);
fprintln(arq,"");
fprintln(arq,"ESTIMADOR COMBINADO 4 ", "%.5f ", vec4);
fprintln(arq,"-----");
fprintln(arq,"-----");

println("FIM DO PASSO ", amo );
println("#####");
println("-----");
}
//FIM DO LOOP(j) PARA OS TAMANHOS DE AMOSTRA
}
//FIM DO LOOP(i) PARA TAMANHOS DE POPULAÇÃO
fprintln(arq,"PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
fprintln(arq,"-----");
fprintln(arq,"Tempo de execucao:", timespan(tempo));
fprintln(arq,"-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}
//FIM DO PROGRAMA

```

PROGRAMA 7

```

#####
//##                                     #
//## Mestrado Em Estatística           #
//## Dissertação de Mestrado           #
//## Autor: Hemílio Fernandes           #
//## Orientador: Cristiano Ferraz      #
//##                                     #
//## Descrição: Esse programa avalia os estimadores propostos   #
//##                para o caso de existência de variável auxiliar #
//##                através do método de simulação de Monte Carlo. #
//##                Para o caso em que é considerado o plano de   #
//##                Amostragem Estratificada (Estimadores Combinados) #
//##                Estimadores avaliados: H e 2.                 #
//##                                     #
//##                                     #
#####
#include<oxstd.h>
#include<oxprob.h>
#import<maximize>
#import<solvenle>
#include <oxfloat.h>

//NÚMERO DE RÉPLICAS
static decl NREP = 1000;

// INÍCIO DO PROGRAMA
main()
{
  ranseed("GM");
  ranseed({1965,2001});
  decl N=<2000,1750,1500,1250,1000,750>;
  decl NAB=<500,1000,1500,2000,2500,3000>;
  decl amostraA=<125,250,500,750,1000,1250,1000,1000,1000,1000,1000>;
  decl amostraB=<125,250,500,750,1000,1250,200,300,400,500,700,900>;
  decl amcadA, amcadB, gA, gB, fA, fB, uni1, uni2, ord1, ord2;
  decl h, i, j, k, l, m, n, o, p, q, r, s, t, u, amo, mc, ind1, fAinv, fBinv;
  decl st, cadA1, cadA2, cadA3, cadA, cadB1, cadB2, cadB3, cadB;
  decl VNA, VNB, VNa, VNb, VNab, na, nb, nab1, nab2, nA, nB;
  decl ya, yb, yab1, yab2;
  decl esth;
  decl Xa, Xab, Ya, Yab, Yb, XA, YA, YB;
  decl txA, tya, tyb, tyab, txa, txab, tyA, tyB, tygeral_st;
  decl atual1, atual2, ech, mh, vech, vrch, m5,vc5,vrc5,varc5,ec5,dpc5,eqmc5,vec5, vch, eqmch;
  decl varch, varcfb, dpch, dpcfb;
  decl mya, myab, mxa, mxab, myb;
  decl a1e1, a2e1, a3e1, a4e1, a5e1, a6e1, b1e1, b2e1, ra, rab1, cor_a, cor_ab;
  decl a1e5, a2e5, a3e5, a4e5, a5e5, a6e5, b1e5, b2e5;
  decl a1e4,a2e4,a3e4,a4e4,a5e4,a6e4,b1e4,b2e4,m1e4,m2e4,m3e4,m4e4,m5e4,m6e4;
  decl estya,estxa,estyab1,estxab1,estyb,estyab2,varya,varyb,varyab1,varyab2,varxa,varxab1;
  decl p1, p5, p3, p4, ph, pfb;
  decl k0, k1, k2, k3, k4, k5, k6, vk, rootk;
  decl u0, u1, u2, u3, u4, u5, u6, vu, rootu;
  decl c0, c1, c2, c3, c4, c5, c6, vc, rootc;
  decl NTOT, medpar, pargeral, valor;
  decl pa, pb, pab1, pab2;
  decl paest, pbest, pab1est, pab2est;
  decl ntotest, Nabest,Naest,Nbest;
  decl wfuller, medfuller, arq;
  decl tempo;
  decl ngeral, ngeralest, txageral, txAgeral, txabgeral;

//MÉDIAS DOS ESTRATOS

```

```

h = 3;
mya = <10,10,10>;
mxm = <8,12,16>;
myab = <5,5,5>;
mxab = <5,6,7>;

//INÍCIO DO LOOP(i) PARA AS ESTRUTURAS DE POPULAÇÃO ONDE Nab VARIA EM CADA ESTRATO
arq=fopen("C:\\Userroot\\Alunos\\Hemílio\\dissertacao\\estratificado_combinados_screen.txt", "a");
tempo=timer();
fprintf(arq,"RESULTADOS DA SIMULACAO: ");
for(i=0;i<columns(N);i++)
{
println("#####");
println("VALOR DE Na, Nb e Nab em cada estrato: ", N[i]~N[i]~NAB[i]);
println("#####");
fprintf(arq,"#####");
fprintf(arq,"VALOR DE Na, Nb e Nab em cada estrato: ", N[i]~N[i]~NAB[i]);
fprintf(arq,"#####");
Xa = zeros(N[i],h);
Ya = zeros(N[i],h);
Xab = zeros(NAB[i],h);
Yab = zeros(NAB[i],h);
YA = zeros(N[i]+NAB[i],h);
XA = zeros(N[i]+NAB[i],h);
YB = zeros(NAB[i],h);
NTOT=zeros(h,1);
VNA = zeros(h,1);
VNB = zeros(h,1);
VNa = zeros(h,1);
VNab = zeros(h,1);
txA=zeros(h,1);
tyA = zeros(h,1);
tyB = zeros(h,1);
txa = zeros(h,1);
txab=zeros(h,1);
tya=zeros(h,1);
tyab=zeros(h,1);
medpar=zeros(h,1);
pa =zeros(h,1);
pb =zeros(h,1);
pab1 =zeros(h,1);
pab2 =zeros(h,1);
nA=zeros(h,1);
nB=zeros(h,1);
esth=zeros(h,1);
ntotest=zeros(h,1);
estya=zeros(h,1);
estxa=zeros(h,1);
estyab1=zeros(h,1);
estxab1=zeros(h,1);
estyab2=zeros(h,1);
varya=zeros(h,1);
varyab1=zeros(h,1);
varyab2=zeros(h,1);
varxa=zeros(h,1);
varxab1=zeros(h,1);
ra=zeros(h,1);
rab1=zeros(h,1);
cor_a=zeros(h,1);
cor_ab=zeros(h,1);
a1e5=zeros(h,1);
a2e5=zeros(h,1);
a3e5=zeros(h,1);
a4e5=zeros(h,1);
a5e5=zeros(h,1);

```

```

a6e5=zeros(h,1);
b1e5=zeros(h,1);

//LOOP PARA CADA ESTRATO
for(j=0;j<h;j++)
{
Xa[] [j]=mxa[j]+rann(N[i],1);
txa[j]=sumc(Xa[] [j]);
Xab[] [j]=mxab[j]+rann(NAB[i],1);
txab[j]=sumc(Xab[] [j]);
Yb[] [j]=myb[j]+rann(N[i],1);
Ya[] [j]=mya[j].*Xa[] [j]+sqrt(Xa[] [j]).*rann(N[i],1);
Yab[] [j]=myab[j].*Xab[] [j]+sqrt(Xab[] [j]).*rann(NAB[i],1);
YB[] [j]=Yb[] [j]|Yab[] [j];
YA[] [j]=Ya[] [j]|Yab[] [j];
XA[] [j]=Xa[] [j]|Xab[] [j];
}

//COMPOSIÇÃO DOS CADASTROS
cadA1=YA[] [0]~XA[] [0];
cadA2=YA[] [1]~XA[] [1];
cadA3=YA[] [2]~XA[] [2];
cadA=cadA1~cadA2~cadA3;
cadB1=YB[] [0];
cadB2=YB[] [1];
cadB3=YB[] [2];
cadB=cadB1~cadB2~cadB3;

//MEDIDAS REFERENTES ÀS POPULAÇÕES GERADAS

//TAMANHO POPULACIONAL
VNA = rows(YA);
VNB = rows(YB);
Vna = rows(Ya[] [0])~rows(Ya[] [1])~rows(Ya[] [2]);
VNa = Vna';
VNab = rows(Yab[] [0])~rows(Yab[] [1])~rows(Yab[] [2]);
VNab = VNab';
NTOT = (Vna[0]+VNab[0])~(Vna[1]+VNab[1])~(Vna[2]+VNab[2]);
NTOT = NTOT';
txA=sumc(XA[] [0])~sumc(XA[] [1])~sumc(XA[] [2]);
txA=txA';
tyA = sumc(YA[] [0])~sumc(YA[] [1])~sumc(YA[] [2]);
tyA = tyA';
tyB = sumc(YB[] [0])~sumc(YB[] [1])~sumc(YB[] [2]);
tyB = tyB';
txa = sumc(Xa[] [0])~sumc(Xa[] [1])~sumc(Xa[] [2]);
txa = txa';
txab= sumc(Xab[] [0])~sumc(Xab[] [1])~sumc(Xab[] [2]);
txab = txab';
tya = sumc(Ya[] [0])~sumc(Ya[] [1])~sumc(Ya[] [2]);
tya = tya';
tyab= sumc(Yab[] [0])~sumc(Yab[] [1])~sumc(Yab[] [2]);
tyab=tyab';
tygeral_st=(tya[0]+tyab[0])~(tya[1]+tyab[1])~(tya[2]+tyab[2]);

//PARÂMETRO DE INTERESSE EM CADA ESTRATO
for(j=0;j<h;j++)
{
valor=tygeral_st[j]/NTOT[j];
medpar[j]=valor;
}

//PARÂMETRO TOTAL
pargerall=sumc(medpar);

```



```

uni1 = sortcindex(ranu(rows(atual1),1));
uni2 = sortcindex(ranu(rows(atual2),1));

//RETIRANDO AS AMOSTRAS DOS ESTRATOS
for(n=0; n<rows(amcadA);n++)
{
ord1 = uni1[n];
amcadA[n][]=atual1[ord1][];
}

//RETIRANDO A AMOSTRA DO CADASTRO B
for(m = 0; m < rows(amcadB); m++)
{
ord2 = uni2[m];
amcadB[m]=atual2[ord2];
}

// SEPARANDO AS OBSERVAÇÕES PARA CADA DOMÍNIO EM CADA ESTRATO

//Domínio a
ya = zeros(VNA,2);
for(m = 0; m < VNa[k]; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n][]==atual1[m][] )
{
ya[n][]=amcadA[n][];
}
}
}
ya = selectifr(ya,ya.!=0);
ya = selectifr(ya,ya.!=.NaN);
na = rows(ya);

//Domínio ab, CADASTRO A
yab1 = zeros(VNA,2);
for(m = VNa[k]; m < VNA; m++)
{
for(n = 0; n < nA[k]; n++)
{
if(amcadA[n][]==atual1[m][] )
{
yab1[n][]=amcadA[n][];
}
}
}
yab1 = selectifr(yab1,yab1.!=0);
yab1 = selectifr(yab1,yab1.!=.NaN);
nab1 = rows(yab1);

yab2 = amcadB;
yab2 = selectifr(yab2,yab2.!=.NaN);
nab2 = rows(yab2);

//PROPORÇÃO POPULACIONAL - ESTIMADOR 3 e 4
paest=na/nA[k];
pab1est=nab1/nA[k];
pab2est=nab2/nB[k];

//QUANTIDADES A SEREM CALCULADAS NA AMOSTRA
estya[k]=meanc(ya[][0]);
estxa[k]=meanc(ya[][1]);
estyab1[k]=meanc(yab1[][0]);

```

```

estxab1[k]=meanc(yab1[][1]);
estyab2[k]=meanc(yab2);
varya[k]=varc(ya[][0]);
varyab1[k]=varc(yab1[][0]);
varyab2[k]=varc(yab2);
varxa[k]=varc(ya[][1]);
varxab1[k]=varc(yab1[][1]);

//QUANTIDADES PARA O ESTIMADOR DE HARTLEY
fAinv=VNA/nA[k];
fBinv=VNB/nB[k];

//ESTIMADOR 5
//CÁLCULOS - ESTIMADOR 5
a1e5[k]=VNA^2*((1-fA)/(nA[k]))*pa[k]*varya[k];
a2e5[k]=VNA^2*((1-fA)/(nA[k]))*pab1[k]*varyab1[k];
a3e5[k]=pa[k]*VNA^2*((1-fA)/(nA[k]))*varxa[k];
a4e5[k]=pab1[k]*VNA^2*((1-fA)/(nA[k]))*varxab1[k];
a5e5[k]=2*(VNa[k]/na)^2*nA[k]*(1-fA)*cor_a[0][0]*sqrt(varya[k])*sqrt(varxa[k]);
a6e5[k]=2*(VNab[k]/nab1)^2*nA[k]*(1-fA)*cor_ab[0][0]*sqrt(varyab1[k])*sqrt(varxab1[k]);
b1e5[k]=VNB^2*((1-fB)/(nB[k]))*pab2[k]*varyab2[k];

}
//FIM DO LOOP(k) PARA OS ESTRATOS
ngeral=sumc(NTOT);
ngeralest=sumc(ntotest);
txabgeral=sumc(txab);
txageral=sumc(txa);
txAgeral=sumc(txA);

//ESTIMADOR DE HARTLEY COMBINADO
ph=(2*sumc(VNB*fBinv*pab2.*varyab2))/(2*sumc(VNA*fAinv*pab1.*varyab1)+2*sumc(VNB*fBinv*pab2.*varyab2));
ech[mc]=(sumc(VNa.*estya) + ph*sumc(VNab.*estyab1)+(1-ph)*sumc(VNab.*estyab2))/ngeral;

//COEFICIENTES DO POLINÔMIO DE GRAU 6 - ESTIMADOR COMBINADO 5
k6 = sumc(b1e5)*sumc(VNab.*estxab1);
k5 = -sumc(b1e5)*sumc(VNab.*estxab1)
+5*sumc(b1e5)*sumc(VNa.*estya)*sumc(VNab.*estxab1);
k4 = 10*sumc(b1e5)*sumc(VNa.*estxa)^2*sumc(VNab.*estxab1)^3
-5*sumc(b1e5)*sumc(VNa.*estxa)*sumc(VNab.*estxab1)^4;
k3 = txAgeral^2*sumc(VNab.*estyab1)^2*sumc(a6e5)*sumc(VNa.*estya)
- 3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a6e5)*sumc(VNab.*estyab1)*sumc(VNa.*estya)
+txAgeral^2*sumc(VNab.*estxab1)*sumc(a2e5)*sumc(VNa.*estxa)
-10*sumc(b1e5)*sumc(VNa.*estxa)^2*sumc(VNab.*estxab1)^3
+10*sumc(b1e5)*sumc(VNa.*estxa)^3*sumc(VNab.*estxab1)^2
-txAgeral^2*sumc(VNab.*estxab1)*sumc(a4e5)*sumc(VNa.*estya)*sumc(VNab.*estyab1)
+2*txAgeral^2*sumc(a4e5)*sumc(VNab.*estyab1)^2*sumc(VNa.*estxa);
k2 = -txAgeral^2*sumc(VNab.*estxab1)*sumc(a6e5)*sumc(VNa.*estya)*sumc(VNa.*estxa)
+2*txAgeral^2*sumc(VNab.*estxab1)*sumc(a2e5)*sumc(VNa.*estxa)^2
+5*sumc(b1e5)*sumc(VNa.*estxa)^4*sumc(VNab.*estxab1)
-txAgeral^2*sumc(VNab.*estxab1)^3*sumc(a1e5)
-10*sumc(b1e5)*sumc(VNa.*estxa)^3*sumc(VNab.*estxab1)^2
-3*txAgeral^2*sumc(a6e5)*sumc(VNab.*estyab1)*sumc(VNa.*estya)^2
- txAgeral^2*sumc(VNab.*estxab1)*sumc(a4e5)*sumc(VNa.*estya)^2
+2*txAgeral^2*sumc(VNab.*estxab1)^2*sumc(a5e5)*sumc(VNab.*estyab1)
-txAgeral^2*sumc(VNab.*estyab1)*sumc(a3e5)*sumc(VNab.*estyab1)^2
+3*txAgeral^2*sumc(a4e5)*sumc(VNa.*estya)*sumc(VNab.*estyab1)*sumc(VNa.*estxa);
k1 = txAgeral^2*sumc(VNab.*estxab1)*sumc(a5e5)*sumc(VNab.*estyab1)*sumc(VNa.*estxa)
+sumc(b1e5)*sumc(VNa.*estxa)^5+3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a5e5)*sumc(VNa.*estya)
+txAgeral^2*sumc(a2e5)*sumc(VNa.*estxa)^3-sumc(b1e5)*sumc(VNa.*estxa)^4*sumc(VNab.*estxab1)
-2*txAgeral^2*sumc(VNab.*estxab1)^2*sumc(a1e5)*sumc(VNa.*estxa)
-3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a3e5)*sumc(VNa.*estya)*sumc(VNab.*estyab1)

```

```

+txAgeral^2*sumc(a4e5)*sumc(VNa.*estya)^2*sumc(VNa.*estxa)
-2*txAgeral^2*sumc(a6e5)*sumc(VNa.*estya)*sumc(VNa.*estxa)^2
+txAgeral^2*sumc(a3e5)*sumc(VNab.*estyab1)^2*sumc(VNa.*estxa);
k0 = txAgeral^2*sumc(VNab.*estyab1)*sumc(a5e5)*sumc(VNa.*estxa)^2
-txAgeral^2*sumc(VNab.*estxab1)*sumc(a1e5)*sumc(VNa.*estxa)^2
+txAgeral^2*sumc(a3e5)*sumc(VNa.*estya)*sumc(VNa.*estxa)
-2*txAgeral^2*sumc(VNab.*estxab1)*sumc(a3e5)*sumc(VNa.*estya)^2
-sumc(b1e5)*sumc(VNa.*estxa)^5
+3*txAgeral^2*sumc(VNab.*estxab1)*sumc(a5e5)*sumc(VNa.*estya)*sumc(VNa.*estxa);
vk=k6~k5~k4~k3~k2~k1~k0;
polyroots(vk,&rootk);
p5=min(fabs(rootk[0][]));

//ESTIMADOR COMBINADO 5
ec5[mc] = (((sumc(VNa.*estya) + p5*sumc(VNab.*estyab1))/(sumc(VNa.*estxa)
+ p5*sumc(VNab.*estxab1)))*txAgeral + (1-p5)*sumc(VNab.*estyab2))/ngeral;

ind1=mc+1;
if(ind1==NREP)
{
println("FIM para n = ", amostraA[amo]~amostraB[amo]);
}
}
//FIM DO LOOP(mc) DE MONTE CARLO
mh = meanc(ech);
m5 = meanc(ec5);
vch = mh - pargeral/3;
vc5 = m5 - pargeral/3;
vrch = fabs(vch)/(pargeral/3);
vrc5 = fabs(vc5)/(pargeral/3);
varch = varc(ech);
varc5 = varc(ec5);
dpch = sqrt(varch);
dpc5 = sqrt(varc5);
eqmch = varch + vch^2;
eqmc5 = varc5 + vc5^2;
vech = mh~vch~vrch~dpch~eqmch~varch;
vec5 = m5~vc5~vrc5~dpc5~eqmc5~varc5;

fprintf(arq,"VALORES: ", "MEDIA ", "VIES ", "VIES RELATIVO ", "D.P. ", "E.Q.M. ", "VARIANCIA");
fprintf(arq,"HARTLEY COMBINADO", "%.5f ", vech);
fprintf(arq,"");
fprintf(arq,"ESTIMADOR COMBINADO 5 ", "%.5f ", vec5);
fprintf(arq,"-----");
fprintf(arq,"-----");

println("FIM DO PASSO ", amo );
println("#####");
println("-----");
}
//FIM DO LOOP(j) PARA OS TAMANHOS DE AMOSTRA
}
//FIM DO LOOP(i) PARA TAMANHOS DE POPULAÇÃO
fprintf(arq,"PRONTO, TERMINOU!!!!!!!!!!!!!!!!!!!!");
fprintf(arq,"-----");
fprintf(arq,"Tempo de execucao:", timespan(tempo));
fprintf(arq,"-----");
fclose(arq);
println("PRONTO, TERMINOU!!!!!!!!!!!!!!!!!!!!");
println("-----");
println("Tempo de execucao:", timespan(tempo));
println("-----");
}
//FIM DO PROGRAMA

```

Referências Bibliográficas

- [1] Hartley, H. O. (1962). *Multiple Frame Surveys*. Proceeding of the Social Statistics Association - ASA.
- [2] Kish, L. (1965). *Survey Sampling*. Wiley, New York.
- [3] Fuller, W. A. & Burmeister, L. F. (1972). *Estimators for Samples Selected from two overlapping frames*. Proceeding of the Social Statistics Association - ASA.
- [4] Cochran, W. G. (1977). *Sampling Techniques*. Wiley, New York, 3d ed.
- [5] Hidiroglou, M. A. & Srinath, K. P. (1981). *Some Estimators of a Population Total From Simple Random Samples Containing Large Units*. Journal of the American Statistical Association. September 1981, **81**, nº 375, Theory and Methods Section.
- [6] Biemer, P. P. (1983). *Optimal Dual Frame Design: Results of a Simulation Study*. Proceedings of the Survey Research Methods Section of the American Statistical Association. **5**, nº 630.
- [7] Biemer, P. P. (1984). *Methodology for Optimal Dual Frame Sample Design*. Bureau of the Census, Statistical Research Division Report Series. January, 1984.
- [8] Fecso, R. , Tortora, R. D. & Vogel, F. A. (1985). *Recent and Future Activities to improve sampling frames for agriculture*. Proceedings of the Survey Research Methods Section of the American Statistical Association.

- [9] Lepkowski, J. M. & Groves, R. M. (1986). *Estimators Based on Several Stratified Samples With Applications to Multiple Frame Surveys*. Journal of the American Statistical Association. December 1986, **81**, n° 396 Applications.
- [10] Lepkowski, J. M. & Groves, R. M. (1986). *A Mean Squared Errors Model for Dual Frame, Mixed Mode Survey Design*. Journal of the American Statistical Association. December 1986, **81**, n° 396 Applications.
- [11] Skinner, C. J. (1991). *On the Efficiency of Raking Ratio Estimation for Multiple Frame Surveys*. Journal of the American Statistical Association. December 1986, **81**, n° 396 Applications.
- [12] Särndal, Swensson & Wretman (1992). *Model Assisted Survey Sampling*. Springer Series in Statistics, 245-260.
- [13] Skinner, C. J. & Rao, J. N. K. (1996). *Estimation in Dual Frame Surveys With Complex Designs*. Journal of the American Statistical Association. March 1996, **81**, n° 433 Theory and Methods.
- [14] *Multiple Frame Agricultural Surveys - vol. 1 e 2*. FAO Statistical Development Series. Food and Agriculture Organization of the United Nations
- [15] Haines, D. E. & Pollock, K. H. (1998). *Estimating the number of active and successful bald eagle nests: an application of the dual frame method*. Environmental and Ecological Statistics. **5**, 245–256.
- [16] Hulliger, B. (1999). *Simple and Robust Estimators for Sampling*. Proceedings of the Survey Research Methods Section of the American Statistical Association, 1999.
- [17] Lohr, S. L. (1999). *Sampling: Design and Analysis*. Brooks/Cole Publishing Company, 81-83.
- [18] Lohr, S. L. (2000). *Youth in Transition Survey: A case study in design and development*. Proceedings of the Survey Research Methods Section of the American Statistical Association, 2000.

- [19] Carfagna, E. (2004). *List Frames, area frames and administrative data, are complementary or in competition?*. invited paper to the Meeting MEXSAI Conference organized by Eurostat, FAO, OCSE, UN/ECE, NASS/USDA, JRC, ISI, Istat, SAGARPA;. Cancun, Mexico, 2-4 November
- [20] Alpizar-Jara, R. , Pollock, K. H. & Haines, D. E. (2005). *Mark-recapture estimator for dual frame population size of prominent nesting structures: the effect of uncertain detection probability*. Environmental and Ecological Statistics. **81**, 155–168.
- [21] Bolfarine, H & Bussab, W.O. (2005). *Elementos de Amostragem*. ABE-Projeto Fisher. São Paulo: Edgard Blücher.
- [22] Ferraz, C. (2006). *A Abordagem de Dual Frame e suas Potencialidades de Aplicação em Pesquisas Agropecuárias*. Texto base para a mesa redonda “Pesquisas Agropecuárias por Amostragem Probabilística”. V CONFEST, 2º Encontro Nacional de Produtores e Usuários de informações sociais, econômicas e territoriais; IBGE, Rio de Janeiro, 21 a 25 de Agosto de 2006.