

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

UM NOVO RESÍDUO PARA CLASSES DO
MODELO DE REGRESSÃO BETA -
LINEAR E NÃO LINEAR

EVELYNE GUIMARÃES DOS SANTOS

DISSERTAÇÃO DE MESTRADO



Recife
2015

EVELYNE GUIMARÃES DOS SANTOS

UM NOVO RESÍDUO PARA CLASSES DO
MODELO DE REGRESSÃO BETA -
LINEAR E NÃO LINEAR

ORIENTADORA: PROFA. DRA. PATRÍCIA LEONE ESPINHEIRA OSPINA
CO-ORIENTADOR: PROF. DR. FRANCISCO CRIBARI NETO

Área de Concentração: Estatística Aplicada

Dissertação apresentada ao programa de Pós-graduação em
Estatística da Universidade Federal de Pernambuco como
requisito parcial para obtenção do grau de **Mestre em
Estatística**.

Recife
2015

Catálogo na fonte
Bibliotecária Joana D'Arc Leão Salvador CRB4-532

S237n Santos, Evelyne Guimarães dos.
Um novo resíduo para classes do modelo de regressão beta - linear e não linear / Evelyne Guimarães dos Santos. – Recife: O Autor, 2015.
90 f.: fig., tab.

Orientadora: Patrícia Leone Espinheira Ospina.
Dissertação (Mestrado) – Universidade Federal de Pernambuco.
CCEN, Estatística, 2015.
Inclui referências.

1. Estatística aplicada. 2. Análise de regressão. I. Ospina, Patrícia Leone Espinheira (Orientadora). II. Título.

519.5 CDD (22. ed.) UFPE-MEI 2015-098

EVELYNE GUIMARÃES DOS SANTOS

**UM NOVO RESÍDUO PARA CLASSES DO MODELO DE REGRESSÃO BETA-
LINEAR E NÃO LINEAR**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovada em: 24 de julho de 2015.

BANCA EXAMINADORA

Prof.^a Dra. Patrícia Leone Espinheira Ospina
UFPE

Prof. PhD. Getúlio José Amorim do Amaral (Examinador Interno)
UFPE

Prof.^a Dra. Michelli Karinne Barros da Silva (Examinador Externo)
UFCG

Aos meus pais, Eduardo e Emília.

Agradecimentos

Aos meus pais, que mesmo distantes sempre me apoiaram e procuraram estar presentes.

À Professora Patrícia Ospina, pela orientação e paciência durante o desenvolvimento deste trabalho e ao Professor Francisco Cribari, pela co-orientação.

À todos os professores do Programa de Pós Graduação em Estatística da UFPE, por todo conhecimento transmitido.

Aos colegas da minha turma do mestrado, Laura, Raphaela, Renan e Wanessa, por terem compartilhado muitas horas de estudo, momentos de tensão e descontração.

Aos colegas da pós-graduação, em especial a Alisson, Danielle, Diego, Fábio, Fernanda, Hildemar, Jeniffer, Jonas, Luana, Raquel, Rodrigo, Sébastien, Telma, Terezinha, por todos os momentos compartilhados.

À Valéria Bittencourt, pela atenção e eficiência.

Aos membros da banca examinadora, pelas sugestões.

À CAPES, pelo apoio financeiro.

*“Our world, our life, our destiny, are dominated by uncertainty; this
is perhaps the only statement we may assert without uncertainty”.*

de Finetti

Em situações em que o objetivo é analisar o comportamento de uma variável em função de outras, os modelos de regressão são muito utilizados. A classe de modelos de regressão beta é utilizada quando se deseja fazer esse tipo de análise e a variável resposta assume valores no intervalo $(0, 1)$, como é o caso de taxas e proporções. Ferrari e Cribari-Neto (2004) propuseram o modelo de regressão beta que utiliza uma parametrização diferente para a distribuição beta, que é indexada pela média e pelo parâmetro de precisão. Foram desenvolvidas duas extensões para este modelo, uma destas extensões foi proposta por Smithson e Verkulien (2006) e considera a precisão variável, neste caso a média e a precisão são modeladas simultaneamente. Outra extensão, proposta por Simas *et al.* (2010), considera que a média e/ou a precisão podem ser relacionadas a preditores não lineares. No processo de escolha do modelo que melhor se adequa aos dados há várias etapas envolvidas, uma delas é a análise de resíduos. Entre os objetivos desta etapa estão: detectar a presença de pontos aberrantes, que poderão ser influentes ou não, e por isso devem ser investigados; verificar se a distribuição proposta para a variável resposta e se a função de ligação estão adequadas. O objetivo desta dissertação é propor e avaliar um novo resíduo para o modelo de regressão beta e suas extensões.

Palavras-chave: Análise de resíduos. Modelo de regressão beta. Modelo não linear.

Abstract

When the interest lies in analyzing the behavior of a given variable as a function of other variables regression models are widely used. The class of beta regression models is used when the response variable assumes values in the interval $(0, 1)$, such as rates and proportions. Ferrari e Cribari-Neto (2004) proposed the beta regression model that uses a different parametrization for the beta distribution, which is indexed by the mean and by a precision parameter. Two extensions have been developed for this model. One of them was proposed by Smithson e Verkulien (2006). In this extension, the mean and precision are modeled simultaneously. Another extension, proposed by Simas *et al.* (2010), considers that the mean and/or the precision may be related to nonlinear predictors. There are several steps involved in the process of choice of the model that best fits the data, one of them being residuals analysis. Among the objectives of this stage are: to detect the presence of atypical points, which may or may not be influential, and thus should be investigated; to verify if the proposed distribution for the variable response and to determine whether the link functions are appropriate. The aim of this thesis is to propose and to evaluate a new residual which was developed for the beta regression model and its extensions.

Keywords: Residuals analysis. Beta regression model. Nonlinear regression.

1	Introdução	11
1.1	Apresentação dos capítulos	12
1.2	Suporte computacional	13
2	Modelo de regressão beta	14
2.1	Resíduos	19
2.1.1	Resíduo ponderado	19
2.1.2	Resíduo ponderado padronizado	20
3	Modelo de regressão beta com dispersão variável	22
3.1	Resíduos	27
3.1.1	Resíduo ponderado	27
3.1.2	Resíduo ponderado padronizado	28
3.1.3	Resíduo combinado	28
4	Avaliação numérica e aplicação	30
4.1	Modelo de regressão beta com dispersão constante	30
4.2	Modelo de regressão beta com dispersão variável	35
4.2.1	Modelos mal especificados	37
4.3	Aplicação	46
5	Modelo de regressão beta não linear	53
5.1	Resíduos	57
5.1.1	Resíduo ponderado	58
5.1.2	Resíduo ponderado padronizado	58
5.1.3	Resíduo combinado	59

5.2	Avaliação numérica	61
5.2.1	Modelos mal especificados	70
5.3	Aplicação	82
6	Considerações finais	88
	Referências	89

CAPÍTULO 1

Introdução

Em situações em que o objetivo é analisar o comportamento de uma variável em função de outras, os modelos de regressão são muito utilizados. A classe de modelos de regressão beta pode ser utilizada quando se deseja fazer esse tipo de análise e a variável resposta assume valores no intervalo $(0, 1)$, como é o caso de taxas e proporções. No entanto, esta classe também pode ser utilizada quando a resposta assume valores diferentes deste intervalo unitário, se seus valores pertencerem ao intervalo (a, b) , onde a e b são conhecidos, utiliza-se a transformação $(y - a)/(b - a)$ ao invés de y diretamente.

No processo de escolha do modelo que melhor se adequa aos dados há várias etapas envolvidas, uma delas é a análise de resíduos. Entre os objetivos desta etapa estão: detectar a presença de pontos aberrantes, que poderão ser influentes ou não, e por isso devem ser investigados; verificar se a distribuição proposta para a variável resposta e se a função de ligação estão adequadas. Os resíduos também são decisivos na implementação de métodos de reamostragem em modelos de regressão, processos com bootstrap. A capacidade do bootstrap em reproduzir adequadamente o verdadeiro processo de geração dos dados em um modelo de regressão pode estar diretamente relacionada com algumas propriedades dos resíduos, como no caso do “Bootstrap Selvagem” em modelos lineares clássicos. O método “Bootstrap Selvagem” introduzido por Liu (1988) se baseia na reamostragem de resíduos que consigam captar heterocedasticidade em tais modelos.

O resíduo para a t -ésima observação pode ser definido como

$$R_t = r_t(y_t, \hat{\mu}_t) \tag{1.1}$$

e mede a discrepância entre o valor observado (y_t) e o valor ajustado ($\hat{\mu}_t$) da t -ésima observação (COX; SNELL, 1968). Aqui, r_t é uma função escolhida para estabilizar a variância e/ou induzir a simetria da distribuição amostral de R_t (CORDEIRO, 1986). Entre

as escolhas mais comuns estão:

$$R_t = \frac{y_t - \hat{\mu}_t}{\sqrt{\widehat{\text{Var}}(y_t)}} \quad \text{e} \quad R_t = \frac{y_t - \hat{\mu}_t}{\sqrt{\widehat{\text{Var}}(y_t - \hat{\mu}_t)}}.$$

É interessante que a definição de resíduos através da Equação (1.1) satisfaça algumas propriedades, tais como $E(R_t) = 0$, variância constante e $\text{Cov}(R_t, R_k) = 0$, $t \neq k$, pois em muitos casos estas condições facilitam a determinação aproximada da distribuição de R_t , o que torna a interpretação dos gráficos dos resíduos, com determinação dos limites de confiança, por exemplo, mais fácil e até mesmo mais confiável. A análise dos resíduos é realizada através de gráficos, como dos resíduos *versus* índices das observações, resíduos *versus* valores ajustados, entre outros.

Na literatura sobre modelos de regressões beta já há algumas definições de resíduos. Ferrari e Cribari-Neto (2004) e Espinheira *et al.* (2008) definiram alguns resíduos para o modelo de regressão beta com dispersão constante. Ferrari *et al.* (2011) estenderam um dos resíduos propostos por Espinheira *et al.* (2008) para o modelo que considera a dispersão variável. Rocha e Simas (2011) definiram um novo resíduo para o modelo de regressão beta não linear e generalizaram alguns dos resíduos definidos por Espinheira *et al.* (2008).

O objetivo dessa dissertação é propor e avaliar um novo resíduo para a classe de modelos de regressão beta, observar o seu comportamento e compará-lo com outros já existentes na literatura. Isso será feito para os modelos de regressão beta e suas extensões: o modelo com dispersão variável e os modelos não lineares. Esse novo resíduo pode ser particularmente útil quando os dados apresentam dispersão variável, uma vez que no processo de construção desse novo resíduo é levada em consideração, além da modelagem da média, a modelagem da dispersão. Além disso, no cenário de omissão da modelagem da dispersão, este resíduo pode ser útil para a construção de estatísticas baseadas em resíduos, como por exemplo a estatística PRESS (ALLEN, 1974).

1.1 Apresentação dos capítulos

Esta dissertação está dividida em seis capítulos. No Capítulo 2 é apresentado o modelo de regressão beta com dispersão constante. No Capítulo 3 é apresentado o modelo de regressão beta com dispersão variável. No Capítulo 4 são apresentadas simulações de Monte Carlo realizadas com o objetivo de investigar a distribuição empírica do resíduo combinado e dos demais resíduos, tanto para o modelo de dispersão constante quanto para o modelo de dispersão variável. Também há simulações com modelos mal especificados. Além disso, há uma aplicação a dados reais. No Capítulo 5 são apresentados o modelo de regressão beta não linear, simulações de Monte Carlo para investigar a distribuição

empírica dos resíduos e uma aplicação a dados reais. Por último, o Capítulo 6 será de conclusões.

1.2 Suporte computacional

As simulações computacionais deste trabalho foram realizadas através da linguagem de programação matricial Ox, que é gratuita para fins acadêmicos e se encontra disponível em <http://www.doornik.com>. Detalhes sobre esta linguagem de programação podem ser encontrados em Doornik (2009). Já os gráficos foram produzidos utilizando o software R em sua versão 3.1.2, que é um ambiente computacional para o desenvolvimento de gráficos e cálculos, principalmente estatísticos. O R encontra-se gratuitamente disponível em <http://www.R-project.org/>.

Modelo de regressão beta

Nos modelos de regressão beta há a suposição de que a variável resposta segue distribuição beta, que é muito flexível, pois ao variar seus dois parâmetros sua densidade pode apresentar diferentes formas. A variável aleatória y tem distribuição beta se a sua função de densidade é dada por

$$f(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1}, \quad 0 < y < 1,$$

onde $p, q > 0$ são parâmetros de forma e $\Gamma(\cdot)$ é a função gama, isto é, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. A média e a variância de y são

$$\begin{aligned} E(y) &= \frac{p}{p+q} \quad \text{e} \\ \text{Var}(y) &= \frac{pq}{(p+q)^2(p+q+1)}. \end{aligned}$$

Nos modelos de regressão o usual é modelar a média e definir um modelo que contenha o parâmetro de precisão (ou dispersão). Neste sentido, o modelo de regressão beta proposto por Ferrari e Cribari-Neto (2004) utiliza uma parametrização diferente para a distribuição beta, a saber: $\mu = p/(p+q)$ e $\phi = p+q$, ou seja, $p = \mu\phi$ e $q = (1-\mu)\phi$. Neste caso a função de densidade beta é expressa por

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, \quad (2.1)$$

em que $0 < \mu < 1$ e $\phi > 0$. Assim,

$$\begin{aligned} \mu &= E(y), \\ \frac{V(\mu)}{1+\phi} &= \text{Var}(y), \end{aligned}$$

em que $V(\mu) = \mu(1 - \mu)$ é a função de variância.

Note que uma vez fixada a média da variável resposta, quanto maior o valor de ϕ , menor será a variância de y . Assim, ϕ pode ser interpretado como um parâmetro de precisão do modelo e ϕ^{-1} é o parâmetro de dispersão.

Sejam y_1, \dots, y_n variáveis aleatórias independentes, em que cada y_t , $t = 1, \dots, n$, possui densidade (2.1) com média μ_t e parâmetro de precisão ϕ , que é desconhecido. O modelo de regressão beta pode ser escrito como

$$g(\mu_t) = \sum_{i=1}^k x_{ti}\beta_i = \eta_t,$$

em que $\beta = (\beta_1, \dots, \beta_k)^\top$ é um vetor de parâmetros desconhecidos ($\beta \in \mathbb{R}^k$) e x_{t1}, \dots, x_{tk} são observações das k covariáveis fixas e conhecidas ($k < n$). A função $g(\cdot)$ é a função de ligação que é estritamente monótona e duas vezes diferenciável, tem domínio no intervalo $(0, 1)$ e a imagem compreende todos os reais. Como exemplos de funções de ligação pode-se citar: a logito, $g(\mu) = \log\{\mu/(1 - \mu)\}$, a probito, $g(\mu) = \Phi^{-1}(\mu)$, em que $\Phi(\cdot)$ é a função de distribuição acumulada de uma variável aleatória normal padrão, a log-log, $g(\mu) = -\log\{-\log(\mu)\}$, entre outras.

Neste modelo, $\mu_t = g^{-1}(\eta_t)$ e $\text{Var}(y_t) = V(g^{-1}(\eta_t))/(1 + \phi)$. Portanto, como a variância da variável resposta depende de μ_t , mesmo nos casos em que o parâmetro de precisão é constante para todas as observações, as variâncias não são. Isto é, o modelo clássico de regressão beta definido por Ferrari e Cribari-Neto (2004) é naturalmente heterocedástico.

A estimação dos parâmetros pode ser realizada a partir do método de máxima verossimilhança. Assim, baseando-se em uma amostra de n observações independentes, o logaritmo da função de verossimilhança é dado por

$$\ell(\beta, \phi) = \sum_{t=1}^n \ell_t(\mu_t, \phi),$$

em que

$$\begin{aligned} \ell_t(\mu_t, \phi) &= \log \Gamma(\phi) - \log \Gamma(\mu_t \phi) - \log \Gamma((1 - \mu_t)\phi) + (\mu_t \phi - 1) \log y_t \\ &\quad + \{(1 - \mu_t)\phi - 1\} \log(1 - y_t). \end{aligned}$$

A função escore, obtida através da diferenciação do logaritmo da função de verossimilhança com relação aos seus parâmetros desconhecidos, é dada por $(U_\beta(\beta, \phi)^\top, U_\phi(\beta, \phi)^\top)^\top$, um vetor de dimensão $(k + 1) \times 1$. Assim, os componentes do vetor escore de β são dados por

$$\frac{\partial \ell(\beta, \phi)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi)}{\partial \mu_t} \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_i}, \quad i = 1, \dots, k, \quad (2.2)$$

com

$$\frac{\partial \ell_t(\mu_t, \phi)}{\partial \mu_t} = \phi \left[\log \left(\frac{y_t}{1 - y_t} \right) - \{\psi(\mu_t \phi) - \psi((1 - \mu_t)\phi)\} \right], \quad (2.3)$$

em que $\psi(\cdot)$ é a função digama, isto é, $\psi(z) = d \log \Gamma(z)/dz$. Além disso, $d\mu_t/d\eta_t = 1/g'(\mu_t)$ e $\partial\eta_t/\partial\beta_i = x_{ti}$. Assim, a Equação (2.2) reduz-se a

$$\frac{\partial\ell(\beta, \phi)}{\partial\beta_i} = \sum_{t=1}^n \phi(y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti},$$

em que

$$y_t^* = \log\left(\frac{y_t}{1-y_t}\right) \quad \text{e} \quad \mu_t^* = \psi(\mu_t\phi) - \psi((1-\mu_t)\phi). \quad (2.4)$$

Adicionalmente, sob certas condições de regularidade, tem-se que

$$E\left(\frac{\partial\ell_t(\mu_t, \phi)}{\partial\mu_t}\right) = 0 \Rightarrow E(\phi(y_t^* - \mu_t^*)) = 0 \Rightarrow E(y_t^*) = \mu_t^*. \quad (2.5)$$

Em forma matricial, a função escore para β , vetor de dimensão k , é dada por

$$U_\beta(\beta, \phi) = \phi X^\top T(y^* - \mu^*),$$

em que X é a matriz de covariadas de ordem $n \times k$ cuja t -ésima linha é x_t^\top , $T = \text{diag}\{1/g'(\mu_1), \dots, 1/g'(\mu_n)\}$, $y^* = (y_1^*, \dots, y_n^*)^\top$, $\mu^* = (\mu_1^*, \dots, \mu_n^*)^\top$.

A função escore para o parâmetro de precisão, ϕ , é dada por

$$\frac{\partial\ell(\beta, \phi)}{\partial\phi} = \sum_{t=1}^n \frac{\partial\ell_t(\mu_t, \phi)}{\partial\phi},$$

com

$$\begin{aligned} \frac{\partial\ell_t(\mu_t, \phi)}{\partial\phi} &= \mu_t \left[\log\left(\frac{y_t}{1-y_t}\right) - \psi(\mu_t\phi) + \psi((1-\mu_t)\phi) \right] \\ &+ \log(1-y_t) - \psi((1-\mu_t)\phi) + \psi(\phi), \end{aligned} \quad (2.6)$$

ou seja,

$$U_\phi(\beta, \phi) = \sum_{t=1}^n \mu_t (y_t^* - \mu_t^*) + \log(1-y_t) - \psi((1-\mu_t)\phi) + \psi(\phi),$$

que é um escalar.

Para a obtenção da matriz de informação de Fisher para β e ϕ são necessárias as derivadas de segunda ordem do logaritmo da função de verossimilhança em relação aos parâmetros. A partir de (2.2), as derivadas de segunda ordem de $\ell(\beta, \phi)$ em relação à β_i e β_j , com $i, j = 1, \dots, k$, são

$$\begin{aligned} \frac{\partial^2\ell(\beta, \phi)}{\partial\beta_i\partial\beta_j} &= \sum_{t=1}^n \frac{\partial}{\partial\mu_t} \left(\frac{\partial\ell_t(\mu_t, \phi)}{\partial\mu_t} \frac{d\mu_t}{d\eta_t} \frac{\partial\eta_t}{\partial\beta_i} \right) \frac{d\mu_t}{d\eta_t} \frac{\partial\eta_t}{\partial\beta_j} \\ &= \sum_{t=1}^n \left\{ \frac{\partial^2\ell_t(\mu_t, \phi)}{\partial\mu_t^2} \frac{d\mu_t}{d\eta_t} + \frac{\partial\ell_t(\mu_t, \phi)}{\partial\mu_t} \frac{\partial}{\partial\mu_t} \left(\frac{d\mu_t}{d\eta_t} \right) \right\} \frac{d\mu_t}{d\eta_t} x_{ti} x_{tj}. \end{aligned}$$

Baseado em (2.5), segue que

$$\mathbb{E} \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_i \partial \beta_j} \right) = \sum_{t=1}^n \mathbb{E} \left(\frac{\partial^2 \ell_t(\mu_t, \phi)}{\partial \mu_t^2} \right) \left(\frac{d\mu_t}{d\eta_t} \right)^2 x_{ti} x_{tj}.$$

De (2.3) obtem-se que

$$\frac{\partial^2 \ell_t(\mu_t, \phi)}{\partial \mu_t^2} = -\phi^2 \{ \psi'(\mu_t \phi) + \psi'((1 - \mu_t)\phi) \},$$

em que $\psi'(\cdot)$ é a função trigama, isto é, $\psi'(z) = d^2 \log \Gamma(z)/dz^2$. Assim,

$$\mathbb{E} \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_i \partial \beta_j} \right) = -\phi \sum_{t=1}^n w_t x_{ti} x_{tj},$$

com

$$w_t = \phi [\psi'(\mu_t \phi) + \psi'((1 - \mu_t)\phi)] \frac{1}{\{g'(\mu_t)\}^2} = \phi v_t \frac{1}{\{g'(\mu_t)\}^2} \quad e$$

$$v_t = \psi'(\mu_t \phi) + \psi'((1 - \mu_t)\phi).$$

Sua forma matricial é dada por

$$\mathbb{E} \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta \partial \beta^\top} \right) = -\phi X^\top W X,$$

em que $W = \text{diag}\{w_1, \dots, w_n\}$.

A derivada de (2.2) com relação a ϕ é dada por

$$\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_i \partial \phi} = \sum_{t=1}^n \left[(y_t^* - \mu_t^*) - \phi \frac{\partial \mu_t^*}{\partial \phi} \right] \frac{1}{g'(\mu_t)} x_{ti},$$

em que $\partial \mu_t^*/\partial \phi = \mu_t \psi'(\mu_t \phi) - \psi'((1 - \mu_t)\phi)(1 - \mu_t)$ e com base em (2.5), segue que

$$\mathbb{E} \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_i \partial \phi} \right) = -\sum_{t=1}^n c_t \frac{1}{g'(\mu_t)} x_{ti},$$

em que $c_t = \phi \{ \mu_t \psi'(\mu_t \phi) - (1 - \mu_t) \psi'((1 - \mu_t)\phi) \}$. Em notação matricial,

$$\mathbb{E} \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta \partial \phi} \right) = -X^\top T c,$$

com $c = (c_1, \dots, c_n)^\top$.

De (2.6), a segunda derivada de $\ell(\beta, \phi)$ com relação a ϕ é dada por

$$\frac{\partial^2 \ell(\beta, \phi)}{\partial \phi^2} = -\sum_{t=1}^n [\mu_t^2 \psi'(\mu_t \phi) + (1 - \mu_t)^2 \psi'((1 - \mu_t)\phi) - \psi'(\phi)] = -\sum_{t=1}^n d_t,$$

em que $d_t = \mu_t^2 \psi'(\mu_t \phi) + (1 - \mu_t)^2 \psi'((1 - \mu_t)\phi) - \psi'(\phi)$. Portanto,

$$\mathbb{E} \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \phi^2} \right) = -\sum_{t=1}^n d_t.$$

Em notação matricial,

$$E\left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \phi^2}\right) = -\text{tr}(D),$$

em que $D = \text{diag}\{d_1, \dots, d_n\}$ e $\text{tr}(D)$ indica traço da matriz D .

Finalmente, a matriz de informação de Fisher é dada por

$$K = K(\beta, \phi) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\phi} \\ K_{\phi\beta} & K_{\phi\phi} \end{pmatrix},$$

onde $K_{\beta\beta} = \phi X^\top W X$, $K_{\beta\phi} = (K_{\phi\beta})^\top = X^\top T c$ e $K_{\phi\phi} = \text{tr}(D)$.

Sob certas condições de regularidade (SEN; SINGER, 1993) e em grandes amostras,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\phi} \end{pmatrix} \sim N_{k+1} \left(\begin{pmatrix} \beta \\ \phi \end{pmatrix}, K^{-1} \right),$$

aproximadamente, em que $\hat{\beta}$ e $\hat{\phi}$ são os estimadores de máxima verossimilhança de β e ϕ , respectivamente. Diferente do que ocorre em modelos lineares generalizados, aqui os parâmetros β e ϕ não são ortogonais e, portanto, $\hat{\beta}$ e $\hat{\phi}$ não são assintoticamente independentes. Estes estimadores são obtidos através do sistema

$$\begin{cases} U_\beta(\beta, \phi) = 0 \\ U_\phi(\beta, \phi) = 0. \end{cases}$$

Eles não podem ser expressos através de uma forma fechada. Eles devem ser obtidos numericamente utilizando algum algoritmo de otimização não linear, como por exemplo, Newton-Raphson, Escore de Fisher, BHHH, BFGS. Esse tipo de otimização exige a especificação de valores iniciais para serem utilizados no esquema iterativo. Ferrari e Cribari-Neto (2004) sugeriram utilizar como um ponto inicial para β a estimativa de mínimos quadrados ordinários desse vetor de parâmetros, que é obtida de uma regressão linear em que a variável resposta é a variável transformada através da função de ligação $g(\cdot)$. Assim, o estimador de mínimos quadrados ordinários é $(X^\top X)^{-1} X^\top z$, em que $z = (g(y_1), \dots, g(y_n))^\top$.

Em relação ao parâmetro de precisão, a sugestão do valor inicial é baseado em $\text{Var}(y_t) = \mu_t(1 - \mu_t)/(1 + \phi)$, que pode ser reescrito como $\phi = [\mu_t(1 - \mu_t)/\text{Var}(y_t)] - 1$. O valor inicial sugerido é

$$\phi^{(0)} = \frac{1}{n} \sum_{t=1}^n \frac{\check{\mu}_t(1 - \check{\mu}_t)}{\check{\sigma}_t^2} - 1,$$

em que $\check{\mu}_t$ é obtido aplicando-se $g^{-1}(\cdot)$ para o t -ésimo valor ajustado da regressão linear de $g(y_1), \dots, g(y_n)$ em X , isto é,

$$\check{\mu}_t = g^{-1}(x_t^\top (X^\top X)^{-1} X^\top z) \quad \text{e} \quad \check{\sigma}_t^2 = \frac{\check{e}^\top \check{e}}{(n - k)g'(\check{\mu}_t)^2}.$$

Aqui, $\check{e} = z - X(X^\top X)^{-1}X^\top z$ é vetor de resíduos de mínimos quadrados ordinários de uma regressão linear com a variável resposta transformada. Para se obter $\check{\sigma}_t^2$ utiliza-se a expansão até a primeira ordem da função $g(y_t)$ em série de Taylor ao redor do ponto μ_t e aplica-se a variância, tal que

$$\text{Var}[g(y_t)] \approx \text{Var}[g(\mu_t) + (y_t - \mu_t)g'(\mu_t)] = \text{Var}(y_t)[g'(\mu_t)]^2,$$

isto é,

$$\text{Var}(y_t) \approx \frac{\text{Var}[g(y_t)]}{[g'(\mu_t)]^2} \Rightarrow \check{\sigma}^2 = \widetilde{\text{Var}}(y_t) \approx \frac{\widehat{\text{Var}}[g(y_t)]}{[g'(\check{\mu}_t)]^2},$$

em que $\widehat{\text{Var}}[g(y_t)] = \check{e}^\top \check{e}/(n - k)$.

2.1 Resíduos

A definição de resíduo para o modelo de regressão beta é feita de forma análoga à que é feita em modelos lineares generalizados. No entanto, as propriedades dos resíduos não são as mesmas, por isso torna-se importante definir resíduos cujas propriedades sejam conhecidas. A seguir serão apresentados alguns dos resíduos que já existem na literatura, os resíduos ponderado padronizado 1 e o resíduo ponderado padronizado 2, propostos por Espinheira *et al.* (2008). Aqui os chamaremos de resíduo ponderado e resíduo ponderado padronizado, respectivamente.

2.1.1 Resíduo ponderado

Uma proposta de Espinheira *et al.* (2008) de resíduo para o modelo de regressão beta com dispersão constante é baseada no processo iterativo Scoring de Fisher para β , que é dado por

$$\beta^{(m+1)} = \beta^{(m)} + \left(K_{\beta\beta}^{(m)}\right)^{-1} U_{\beta}^{(m)}(\beta, \gamma),$$

em que m representa os passos necessários até a convergência do processo. Portanto, o processo iterativo de interesse é dado por

$$\beta^{(m+1)} = \beta^{(m)} + \left(\phi^{(m)} X^\top W^{(m)} X\right)^{-1} \phi^{(m)} X^\top T^{(m)} (y^* - \mu^{*(m)}),$$

que pode ser expresso como

$$\beta^{(m+1)} = \beta^{(m)} + \left(X^\top \Phi^{(m)} W^{(m)} X\right)^{-1} X^\top \Phi^{(m)} T^{(m)} (y^* - \mu^{*(m)}),$$

em que $\Phi = \text{diag}\{\phi, \dots, \phi\}$.

O processo acima pode ser representado como um processo iterativo de mínimos quadrados reponderados, ou seja,

$$\beta^{(m+1)} = \left(X^\top \Phi^{(m)} W^{(m)} X\right)^{-1} X^\top \Phi^{(m)} W^{(m)} z^{(m)},$$

em que $z^{(m)} = \eta^{(m)} + W^{-1(m)}T^{(m)}(y^* - \mu^{*(m)})$ com $\eta = (\eta_1, \dots, \eta_n)^\top = X\beta$. Após a convergência, temos que

$$\hat{\beta} = (X^\top \hat{\Phi} \hat{W} X)^{-1} X^\top \hat{\Phi} \hat{W} z, \quad (2.7)$$

com

$$z = \hat{\eta} + \hat{W}^{-1} \hat{T}(y^* - \hat{\mu}^*),$$

que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de $\Phi^{1/2}W^{1/2}z$ em $\Phi^{1/2}W^{1/2}X$. O resíduo de mínimos quadrados dessa regressão é dado por

$$\begin{aligned} r^\beta &= \hat{\Phi}^{1/2} \hat{W}^{1/2} z - \hat{\Phi}^{1/2} \hat{W}^{1/2} \hat{\eta} = \hat{\Phi}^{1/2} \hat{W}^{1/2} (z - \hat{\eta}) \\ &= \hat{\Phi}^{1/2} \hat{W}^{1/2} (\hat{\eta} + \hat{W}^{-1} \hat{T}(y^* - \hat{\mu}^*) - \hat{\eta}) \\ &= \Phi^{1/2} \hat{W}^{-1/2} \hat{T}(y^* - \hat{\mu}^*). \end{aligned} \quad (2.8)$$

Assim, o resíduo da t -ésima observação é dado por

$$r_t^\beta = \hat{\phi}^{1/2} \frac{1}{\hat{w}_t^{1/2}} \frac{1}{g'(\hat{\mu}_t)} (y_t^* - \hat{\mu}_t^*),$$

com $w_t = \phi_t v_t [1/g'(\mu_t)]^2$. Logo, este resíduo é definido como

$$r_t^\beta = \frac{(y_t^* - \hat{\mu}_t^*)}{\sqrt{\hat{v}_t}}, \quad (2.9)$$

sendo

$$v_t = \psi'(\mu_t \phi) + \psi'((1 - \mu_t) \phi). \quad (2.10)$$

2.1.2 Resíduo ponderado padronizado

Uma possível padronização de (2.9), também proposta por Espinheira *et al.* (2008), baseia-se na variância de z . Para isso, a equação (2.7) pode ser reescrita como

$$(X^\top \hat{\Phi} \hat{W} X) \hat{\beta} = X^\top \hat{\Phi} \hat{W} z. \quad (2.11)$$

Dado que $\text{Cov}(\hat{\beta}) \approx \phi^{-1} (X^\top W X)^{-1} = (X^\top \Phi W X)^{-1}$ e considerando $\hat{W} \approx W$ e $\hat{\Phi} \approx \Phi$ na Equação (2.11), segue que

$$\begin{aligned} (X^\top \Phi W X) \text{Cov}(\hat{\beta}) (X^\top \Phi W X)^\top &\approx (X^\top \Phi W) \text{Cov}(z) (X^\top \Phi W)^\top \\ (X^\top \Phi W X) (X^\top \Phi W X)^{-1} (X^\top \Phi W X) &\approx (X^\top \Phi W) \text{Cov}(z) (W \Phi X) \\ \text{Cov}(z) &\approx W^{-1} \Phi^{-1}. \end{aligned}$$

Assim, $\widehat{\text{Cov}}(z) \approx \hat{W}^{-1} \hat{\Phi}^{-1}$.

Baseando-se em (2.7), o resíduo em (2.8) também pode ser expresso por

$$r^\beta = \hat{\Phi}^{1/2} \hat{W}^{1/2} (I - X (X^\top \hat{\Phi} \hat{W} X)^{-1} X^\top \hat{\Phi} \hat{W}) z.$$

Assim, mais uma vez considerando $\widehat{W} \approx W$ e $\widehat{\Phi} \approx \Phi$, segue que

$$\begin{aligned} \text{Cov}(r^\beta) &\approx (\Phi^{1/2}W^{1/2} - \Phi^{1/2}W^{1/2}X(X^\top\Phi WX)^{-1}X^\top\Phi W)\text{Cov}(z) \\ &\quad \times (\Phi^{1/2}W^{1/2} - \Phi^{1/2}W^{1/2}X(X^\top\Phi WX)^{-1}X^\top\Phi W)^\top \\ &\approx (\text{I} - \Phi^{1/2}W^{1/2}X(X^\top\Phi^{1/2}WX)^{-1}X^\top\Phi^{1/2}W^{1/2}) \\ &\approx (\text{I} - H^*). \end{aligned}$$

Logo, $\widehat{\text{Cov}}(r^\beta) \approx \text{I} - \widehat{H}^*$, com $\widehat{H}^* = \widehat{\Phi}^{1/2}\widehat{W}^{1/2}X(X^\top\widehat{\Phi}^{1/2}\widehat{W}X)^{-1}X^\top\widehat{\Phi}^{1/2}\widehat{W}^{1/2}$, em que H^* é a matriz de projeção de $\Phi^{1/2}W^{1/2}z$ contra $\Phi^{1/2}W^{1/2}X$ e é uma matriz simétrica e idempotente.

Portanto, a padronização do resíduo é dada por

$$r_{pt}^\beta = \frac{r_t^\beta}{\sqrt{\widehat{\text{Cov}}(r_t^\beta)}} = \frac{y_t^* - \widehat{\mu}_t^*}{\sqrt{\widehat{v}_t(1 - \widehat{h}_{tt}^*)}}, \quad (2.12)$$

em que \widehat{h}_{tt}^* é o t -ésimo elemento da diagonal principal de H^* e v_t está definido em (2.10).

No Capítulo 4 algumas propriedades dos resíduos ponderado e ponderado padronizado serão avaliadas numericamente.

Modelo de regressão beta com dispersão variável

No modelo de regressão beta proposto por Ferrari e Cribari-Neto (2004) o parâmetro de dispersão é assumido ser constante ao longo das observações. Similar ao que ocorre em modelos lineares generalizados, o conceito de heteroscedasticidade e dispersão não constante difere do utilizado nos modelos lineares normais, onde ambos significam que as variâncias não são constantes. No caso dos modelos normais o parâmetro de dispersão é a variância.

Aqui, quando se considera dispersão variável, a média e a dispersão são modeladas simultaneamente (SMITHSON; VERKULIEN, 2006). Assim, também é definida uma estrutura de regressão para o parâmetro de precisão. A estimação conjunta dos parâmetros pode ser realizada pelo método da máxima verossimilhança.

Considerando y_1, \dots, y_n variáveis aleatórias independentes, onde cada y_t , $t = 1, \dots, n$, segue a densidade (2.1), a média e o parâmetro de precisão satisfazem as seguintes relações funcionais:

$$g(\mu_t) = \eta_{1t} = \sum_{i=1}^k x_{ti}\beta_i \quad \text{e} \quad h(\phi_t) = \eta_{2t} = \sum_{j=1}^q z_{tj}\gamma_j,$$

em que $\beta = (\beta_1, \dots, \beta_k)^\top$ e $\gamma = (\gamma_1, \dots, \gamma_q)^\top$, são vetores de parâmetros desconhecidos ($\beta \in \mathbb{R}^k$ e $\gamma \in \mathbb{R}^q$), x_{t1}, \dots, x_{tk} e z_{t1}, \dots, z_{tq} são observações de k e q covariáveis fixas e conhecidas ($k < n$ e $q < n$), $g(\cdot)$ e $h(\cdot)$ são funções de ligação, funções estritamente monótonas e duas vezes diferenciáveis. O logaritmo da função de verossimilhança é dado por

$$\ell(\beta, \gamma) = \sum_{t=1}^n \ell_t(\mu_t, \phi_t), \tag{3.1}$$

em que

$$\begin{aligned} \ell_t(\mu_t, \phi_t) &= \log \Gamma(\phi_t) - \log \Gamma(\mu_t \phi_t) - \log \Gamma((1 - \mu_t)\phi_t) + (\mu_t \phi_t - 1) \log y_t \\ &\quad + \{(1 - \mu_t)\phi_t - 1\} \log(1 - y_t). \end{aligned}$$

O vetor escore é dado por $(U_\beta(\beta, \gamma)^\top, U_\gamma(\beta, \gamma)^\top)^\top$. Assim, os componentes do vetor escore de β são da forma

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i}, \quad i = 1, \dots, k, \quad (3.2)$$

com

$$\frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} = \phi_t \left[\log \left(\frac{y_t}{1 - y_t} \right) - \{\psi(\mu_t \phi_t) - \psi((1 - \mu_t)\phi_t)\} \right], \quad (3.3)$$

em que $\psi(\cdot)$ é a função digama. Portanto, a Equação (3.2) reduz-se a

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti}, \quad (3.4)$$

em que

$$y_t^* = \log \left(\frac{y_t}{1 - y_t} \right) \quad \text{e} \quad \mu_t^* = \psi(\mu_t \phi_t) - \psi((1 - \mu_t)\phi_t). \quad (3.5)$$

Em forma matricial a função escore de $\beta = (\beta_1, \dots, \beta_k)$ é dada por

$$U_\beta(\beta, \phi) = X^\top \Phi T (y^* - \mu^*),$$

em que X é a matriz de covariadas de ordem $n \times k$ cuja t -ésima linha é x_t^\top , $\Phi = \text{diag}\{\phi_1, \dots, \phi_n\}$, $T = \text{diag}\{1/g'(\mu_1), \dots, 1/g'(\mu_n)\}$, $y^* = (y_1^*, \dots, y_n^*)^\top$ e $\mu^* = (\mu_1^*, \dots, \mu_n^*)^\top$.

Adicionalmente, a função de densidade beta pode ser escrita na forma exponencial biparamétrica canônica dada por

$$f(y_t; \mu_t, \phi) = \exp\{\tau_1 T_1 + \tau_2 T_2 - A(\tau)\} \frac{1}{y_t(1 - y_t)},$$

em que $\tau = (\tau_1, \tau_2) = (\mu_t \phi_t, \phi_t)$,

$$(T_1, T_2) = \left(\log \left(\frac{y_t}{1 - y_t} \right), \log(1 - y_t) \right) \quad \text{e}$$

$$A(\tau) = \{-\log \Gamma(\phi_t) + \log \Gamma(\mu_t \phi_t) + \log \Gamma((1 - \mu_t)\phi_t)\}.$$

Assim, tem-se que

$$E(T_1) = E(y_t^*) = \frac{\partial A(\tau)}{\partial \tau_1} = \psi(\mu_t \phi_t) - \psi((1 - \mu_t)\phi_t) = \mu_t^*, \quad (3.6)$$

$$\text{Var}(T_1) = \text{Var}(y_t^*) = \frac{\partial^2 A(\tau)}{\partial \tau_1^2} = \psi'(\mu_t \phi_t) + \psi'((1 - \mu_t)\phi_t) = v_t, \quad (3.7)$$

$$E(T_2) = E(\log(1 - y_t)) = \frac{\partial A(\tau)}{\partial \tau_2} = \psi((1 - \mu_t)\phi_t) - \psi(\phi_t), \quad (3.8)$$

$$\text{Var}(T_2) = \text{Var}(\log(1 - y_t)) = \frac{\partial^2 A(\tau)}{\partial \tau_2^2} = \psi'((1 - \mu_t)\phi_t) - \psi'(\phi_t), \quad (3.9)$$

$$\text{Cov}(T_1, T_2) = \frac{\partial^2 A(\tau)}{\partial \tau_1 \partial \tau_2} = -\psi'((1 - \mu_t)\phi_t), \quad (3.10)$$

(LEHMANN; CASELLA, 1998). Essas quantidades serão importantes para a definição dos resíduos.

A função escore para γ_j , $j = 1, \dots, n$, é dada por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j}, \quad (3.11)$$

em que

$$\begin{aligned} \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} = \mu_t \left[\log \left(\frac{y_t}{1 - y_t} \right) - \{ \psi(\mu_t \phi_t) - \psi((1 - \mu_t) \phi_t) \} \right] \\ + \log(1 - y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t). \end{aligned} \quad (3.12)$$

Além disso, $d\phi_t/d\eta_{2t} = 1/h'(\mu_t)$ e $\partial \eta_{2t}/\partial \gamma_j = z_{tj}$.

Utilizando as definições de y^* e μ^* em (3.5), a função escore para cada um dos parâmetros γ_j é dada por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n [\mu_t(y_t^* - \mu_t^*) + \log(1 - y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t)] \frac{1}{h'(\phi_t)} z_{tj}, \quad (3.13)$$

que pode ser expressa por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n a_t \frac{1}{h'(\phi_t)} z_{tj}, \quad (3.14)$$

em que $a_t = \mu_t(y_t^* - \mu_t^*) + \log(1 - y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t)$.

Na forma matricial temos que

$$U_\gamma(\beta, \gamma) = Z^\top H a,$$

em que Z é uma matriz $n \times q$ cuja t -ésima linha é z_t^\top , $H = \text{diag}\{1/h'(\phi_1), \dots, 1/h'(\phi_n)\}$ e $a = (a_1, \dots, a_n)^\top$.

Para construir a matriz de informação de Fisher para β e γ são necessárias as segundas derivadas do logaritmo da função de verossimilhança com relação aos parâmetros. Assim, a partir de (3.2) temos que, para $i, p = 1, \dots, k$,

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \beta_p} &= \sum_{t=1}^n \frac{\partial}{\partial \mu_t} \left(\frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right) \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_p} \\ &= \sum_{t=1}^n \left\{ \frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \mu_t^2} \frac{d\mu_t}{d\eta_{1t}} + \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial}{\partial \mu_t} \left(\frac{d\mu_t}{d\eta_{1t}} \right) \right\} \frac{d\mu_t}{d\eta_{1t}} x_{ti} x_{tp}. \end{aligned}$$

A partir de (3.6), temos que $E(\partial \ell(\mu_t, \phi_t)/\partial \mu_t) = 0$. Então,

$$E \left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_i \partial \beta_p} \right) = \sum_{t=1}^n E \left(\frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \mu_t^2} \right) \left(\frac{d\mu_t}{d\eta_{1t}} \right)^2 x_{ti} x_{tp}.$$

De (3.3) segue que

$$\frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \mu_t^2} = -\phi_t^2 \{ \psi'(\mu_t \phi_t) + \psi'((1 - \mu_t) \phi_t) \},$$

em que $\psi'(\cdot)$ é a função trigama. Assim,

$$E\left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_i \partial \beta_p}\right) = -\sum_{t=1}^n \phi_t w_t x_{ti} x_{tp},$$

com

$$w_t = \phi_t [\psi'(\mu_t \phi_t) + \psi'((1 - \mu_t) \phi_t)] \frac{1}{\{g'(\mu_t)\}^2}.$$

Sua forma matricial é dada por

$$E\left(\frac{\partial^2 \ell(\beta, \phi)}{\partial \beta \partial \beta^\top}\right) = -X^\top \Phi W X,$$

em que $W = \text{diag}\{w_1, \dots, w_n\}$.

De (3.4), as segundas derivadas de $\ell(\beta, \gamma)$ em relação à β_i e γ_j , para $i = 1, \dots, k$ e $j = 1, \dots, q$, são dadas por

$$\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} = \sum_{t=1}^n \frac{\partial}{\partial \phi_t} \left(\phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti} \right) \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j},$$

em que

$$\frac{\partial}{\partial \phi_t} \left(\phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti} \right) = \{(y_t^* - \mu_t^*) - \phi_t [\psi'(\mu_t \phi_t) \mu_t - \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)]\} \frac{1}{g'(\mu_t)} x_{ti},$$

$d\phi_t/d\eta_{2t} = 1/h'(\phi_t)$ e $\partial \eta_{2t}/\partial \gamma_j = z_{tj}$.

Como $E(y_t^* - \mu_t^*) = 0$, segue que

$$E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_j}\right) = -\sum_{t=1}^n c_t \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} x_{ti} z_{tj},$$

em que $c_t = \phi_t \{\mu_t \psi'(\mu_t \phi_t) - (1 - \mu_t) \psi'((1 - \mu_t) \phi_t)\}$. Em notação matricial pode ser escrito como

$$E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta \partial \gamma^\top}\right) = -X^\top C T H Z,$$

com $C = \text{diag}\{c_1, \dots, c_n\}$.

A partir de (3.11), as derivadas de $\ell(\beta, \gamma)$ com relação a γ_j e γ_l , para $j, l = 1, \dots, q$, são dadas por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \gamma_l} &= \sum_{t=1}^n \frac{\partial}{\partial \phi_t} \left(\frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \right) \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_l} \\ &= \sum_{t=1}^n \left(\frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \phi_t^2} \frac{d\phi_t}{d\eta_{2t}} + \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial}{\partial \phi_t} \left(\frac{d\phi_t}{d\eta_{2t}} \right) \right) \frac{d\phi_t}{d\eta_{2t}} z_{tl} z_{tj}. \end{aligned}$$

Como $E(\partial \ell(\mu_t, \phi_t)/\partial \mu_t) = 0$, temos que

$$E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \gamma_l}\right) = \sum_{t=1}^n E\left(\frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \phi_t^2}\right) \left(\frac{d\phi_t}{d\eta_{2t}}\right)^2 z_{tl} z_{tj}.$$

Utilizando (3.12), segue que

$$E \left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_i \partial \gamma_j} \right) = - \sum_{t=1}^n [\psi'(\mu_t \phi_t) \mu_t^2 + \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)^2 - \psi'(\phi_t)] \frac{1}{\{h'(\phi_t)\}^2} z_{it} z_{tj}.$$

Em forma matricial,

$$E \left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma \partial \gamma^\top} \right) = -Z^\top D^* Z,$$

em que $D^* = \text{diag}\{d_1^*, \dots, d_n^*\}$, com

$$d_t^* = [\psi'(\mu_t \phi_t) \mu_t^2 + \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)^2 - \psi'(\phi_t)] \frac{1}{\{h'(\phi_t)\}^2}.$$

Finalmente, a matriz informação de Fisher para β e γ é dada por

$$K^* = K^*(\beta, \gamma) = \begin{pmatrix} K_{\beta\beta}^* & K_{\beta\gamma}^* \\ K_{\gamma\beta}^* & K_{\gamma\gamma}^* \end{pmatrix},$$

em que $K_{\beta\beta}^* = X^\top \Phi W X$, $K_{\beta\gamma}^* = (K_{\gamma\beta}^*)^\top = X^\top C T H Z$ e $K_{\gamma\gamma}^* = Z^\top D^* Z$.

Sob certas condições de regularidade (SEN; SINGER, 1993) e em grandes amostras tem-se que

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \sim N_{k+q} \left(\begin{pmatrix} \beta \\ \gamma \end{pmatrix}, K^{*-1} \right),$$

aproximadamente. Aqui, $\hat{\beta}$ e $\hat{\gamma}$ são os estimadores de máxima verossimilhança de β e γ , que são soluções do sistema

$$\begin{cases} U_\beta(\beta, \gamma) = 0 \\ U_\gamma(\beta, \gamma) = 0. \end{cases}$$

Os estimadores de máxima verossimilhança não podem ser expressos através de uma forma fechada, por isso são obtidos numericamente utilizando algum algoritmo de otimização não linear. Portanto, há a necessidade de especificar valores iniciais para serem utilizados no esquema iterativo. Ferrari *et al.* (2011) sugeriram que o valor inicial para β seja o mesmo utilizado no modelo de regressão beta com dispersão constante, isto é, $\beta^{(0)} = (X^\top X)^{-1} X^\top g(y_t)$. Com relação a ϕ_t , $t = 1, \dots, n$, o chute inicial é da forma

$$\phi_t^{(0)} = \frac{\check{\mu}_t(1 - \check{\mu}_t)}{\check{\sigma}^2} - 1,$$

em que $\check{\mu}_t = g^{-1}(x_t^\top (X^\top X)^{-1} X^\top z)$ e $\check{\sigma}_t^2 = \check{e}^\top \check{e} / [(n-t)g'(\check{\mu}_t)^2]$, com $\check{e} = z - X(X^\top X)^{-1} X^\top z$ e $z = g(y_t)$.

Foi considerado que $h(\phi_t) = \sum_{j=1}^q z_{tj} \gamma_j$ e a estimativa inicial de γ é dada por $(Z^\top Z)^{-1} Z^\top k$, em que $k = (h(\phi_1^{(0)}), \dots, h(\phi_n^{(0)}))^\top$.

3.1 Resíduos

O resíduo ponderado padronizado proposto por Espinheira *et al.* (2008) para o modelo de regressão beta com dispersão constante foi adaptado para os casos em que a dispersão varia ao longo das observações, a extensão desse resíduo é apresentada em Ferrari *et al.* (2011). Nas seções seguintes além da extensão do resíduo ponderado padronizado, apresentaremos a extensão para o resíduo ponderado e o novo resíduo que estamos propondo, o resíduo combinado.

3.1.1 Resíduo ponderado

No modelo de regressão beta com dispersão variável o processo iterativo Scoring de Fisher para estimar β é dado por

$$\beta^{(m+1)} = \beta^{(m)} + (X^\top \Phi^{(m)} W^{(m)} X)^{-1} X^\top \Phi^{(m)} T^{(m)}(y^* - \mu^{*(m)}).$$

Na forma de um processo iterativo de mínimos quadrados reponderados o processo é dado por

$$\beta^{(m+1)} = (X^\top \Phi^{(m)} W^{(m)} X)^{-1} X^\top \Phi^{(m)} W^{(m)} u_1^{(m)},$$

em que $u_1^{(m)} = \eta^{(m)} + W^{-1(m)} T^{(m)}(y^* - \mu^{*(m)})$.

Após a convergência, temos que

$$\hat{\beta} = (X^\top \hat{\Phi} \hat{W} X)^{-1} X^\top \hat{\Phi} \hat{W} u_1, \quad (3.15)$$

com

$$u_1 = \hat{\eta} + \hat{W}^{-1} \hat{T}(y^* - \hat{\mu}^*),$$

que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de $\Phi^{1/2} W^{1/2} u_1$ em $\Phi^{1/2} W^{1/2} X$. O resíduo de mínimos quadrados dessa regressão é dado por

$$r^\beta = \hat{\Phi}^{1/2} \hat{W}^{1/2} u_1 - \hat{\Phi}^{1/2} \hat{W}^{1/2} X \hat{\beta} = \hat{\Phi}^{1/2} \hat{W}^{1/2} (u_1 - X \hat{\beta}) = \hat{\Phi}^{1/2} \hat{W}^{-1/2} \hat{T}(y^* - \hat{\mu}^*), \quad (3.16)$$

então, o resíduo da t -ésima observação é dado por

$$r_t^\beta = \frac{y_t^* - \hat{\mu}_t^*}{\sqrt{\hat{v}_t}}, \quad (3.17)$$

em que $v_t = \text{Var}(y_t^*)$ está definido em (3.7).

3.1.2 Resíduo ponderado padronizado

Uma possível padronização do resíduo (3.17) baseia-se na variância de u_1 . Considerando que $\text{Cov}(\hat{\beta}) \approx (X^\top \Phi W X)^{-1}$, $\widehat{W} \approx W$, $\widehat{\Phi} \approx \Phi$ e reescrevendo a Equação (3.15) como

$$(X^\top \widehat{\Phi} \widehat{W} X) \hat{\beta} = X^\top \widehat{\Phi} \widehat{W} u_1,$$

tem-se que

$$\begin{aligned} (X^\top \Phi W X) \text{Cov}(\hat{\beta}) (X^\top \Phi W X)^\top &\approx (X^\top \Phi W) \text{Cov}(u_1) (X^\top \Phi W)^\top \\ (X^\top \Phi W X) (X^\top \Phi W X)^{-1} (X^\top \Phi W X) &\approx X^\top \Phi W \text{Cov}(u_1) W \Phi X \\ \text{Cov}(u_1) &\approx W^{-1} \Phi^{-1}. \end{aligned}$$

Logo, $\widehat{\text{Cov}}(u_1) \approx \widehat{W}^{-1} \widehat{\Phi}^{-1}$. Em seguida expressamos o resíduo em (3.16) como

$$\begin{aligned} r^\beta &= \widehat{\Phi}^{1/2} \widehat{W}^{1/2} (u_1 - X (X^\top \widehat{\Phi} \widehat{W} X)^{-1} X^\top \widehat{\Phi} \widehat{W} u_1) \\ &= (\widehat{\Phi}^{1/2} \widehat{W}^{1/2} - \widehat{\Phi}^{1/2} \widehat{W}^{1/2} X (X^\top \widehat{\Phi} \widehat{W} X)^{-1} X^\top \widehat{\Phi} \widehat{W}) u_1. \end{aligned}$$

Considerando que $\text{Cov}(u_1) \approx W^{-1} \phi^{-1}$, $\widehat{W} \approx W$ e $\widehat{\Phi} \approx \Phi$, segue que

$$\begin{aligned} \text{Cov}(r^\beta) &\approx (\Phi^{1/2} W^{1/2} - \Phi^{1/2} W^{1/2} X (X^\top \Phi W X)^{-1} X^\top \Phi W) \text{Cov}(u_1) \\ &\quad \times (\Phi^{1/2} W^{1/2} - \Phi^{1/2} W^{1/2} X (X^\top \Phi W X)^{-1} X^\top \Phi W)^\top \\ &\approx (I - \Phi^{1/2} W^{1/2} X (X^\top \Phi W X)^{-1} X^\top \Phi^{1/2} W^{1/2}) \\ &\approx (I - G). \end{aligned}$$

Logo, $\widehat{\text{Cov}}(r^\beta) \approx (I - \widehat{G})$, com $\widehat{G} = \widehat{\Phi}^{1/2} \widehat{W}^{1/2} X (X^\top \widehat{\Phi} \widehat{W} X)^{-1} X^\top \widehat{\Phi}^{1/2} \widehat{W}^{1/2}$, em que G é a matriz de projeção de $\Phi^{1/2} W^{1/2} u_1$ em $\Phi^{1/2} W^{1/2} X$ e é uma matriz simétrica e idempotente. Note que $\Phi = \text{diag}\{\phi_1, \dots, \phi_n\}$.

Assim, o resíduo ponderado padronizado é definido por

$$r_{pt}^\beta = \frac{r_t^\beta}{\sqrt{\widehat{\text{Cov}}(r_t^\beta)}} = \frac{y_t^* - \widehat{\mu}_t^*}{\sqrt{\widehat{v}_t(1 - \widehat{g}_{tt})}},$$

em que g_{tt} é o t -ésimo elemento da matriz diagonal G e v_t está definido em (3.7).

3.1.3 Resíduo combinado

O resíduo combinado é o que estamos propondo nesta dissertação. A construção desse resíduo é em parte similar à construção do resíduo ponderado, que é baseado no processo iterativo Scoring de Fisher para estimar β . Agora considera-se adicionalmente o processo iterativo para γ que é dado por

$$\gamma^{(m+1)} = \gamma^{(m)} + (K_{\gamma\gamma}^{*(m)})^{-1} U_\gamma^{(m)}(\beta, \gamma),$$

em que m representa os passos necessários até a convergência do processo. Assim, o processo de interesse é dado por

$$\gamma^{(m+1)} = \gamma^{(m)} + (Z^\top D^{*(m)} Z)^{-1} Z^\top H^{(m)} a^{(m)}, \quad (3.18)$$

em que $a = (a_1, \dots, a_n)$ com

$$a_t = \mu_t(y_t^* - \mu_t^*) + \log(1 - y_t) - \psi((1 - \mu_t)\phi_t) + \psi(\phi_t). \quad (3.19)$$

O processo em (3.18) também pode ser representado como um processo iterativo de mínimos quadrados ponderados, ou seja,

$$\gamma^{(m+1)} = (Z^\top D^{*(m)} Z)^{-1} Z^\top D^{*(m)} u_2^{(m)},$$

em que $u_2^{(m)} = \eta_2^{(m)} + D^{*(m)} H^{(m)} a^{(m)}$, com $\eta_2 = (\eta_{21}, \dots, \eta_{2n}) = Z\gamma$. Na convergência do processo temos que

$$\hat{\gamma} = (Z^\top \hat{D}^* Z)^\top Z^\top \hat{D}^* \hat{u}_2,$$

com $\hat{u}_2 = \hat{\eta}_2 + \hat{D}^{*-1} \hat{H} \hat{a}$, que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de $D^{*1/2} Z$ em $D^{*1/2} u_2$, onde o resíduo ordinário é dado por

$$r^\gamma = D^{*1/2} u_2 - D^{*1/2} Z\gamma = D^{*1/2} (\hat{\eta}_2 + \hat{D}^* \hat{H} \hat{a} - \hat{\eta}_2) = (\hat{D}^*)^{-1/2} \hat{H} \hat{a}.$$

Assim, o resíduo da t -ésima observação é

$$r_t^\gamma = \frac{a_t}{\sqrt{\hat{\xi}_t}}, \quad (3.20)$$

em que $\xi_t = \mu_t^2 \psi'(\mu_t \phi_t) + (1 - \mu_t)^2 \psi'((1 - \mu_t)\phi_t) - \psi'(\phi_t) = \text{Var}(a_t)$ e a_t está definido em (3.19).

O resíduo combinado é construído baseando-se nos processos iterativos Scoring de Fisher de β e γ , dessa forma o resíduo carrega informação sobre todos os parâmetros do modelo. Assim, baseando-se na definição de resíduos em (1.1), o resíduo combinado, que tem componentes dos resíduos (3.17) e (3.20), é dado por

$$r_t^{\beta\gamma} = (y_t^* - \hat{\mu}^*) + \hat{a}_t. \quad (3.21)$$

A partir de (3.7), (3.9) e (3.10), a variância de (3.21) é dada por

$$\text{Var}(y_t^* + \mu_t^* y_t^* + \log(1 - y_t)) = (1 + \mu_t)^2 \psi'(\mu_t \phi_t) + \mu_t^2 \psi'((1 - \mu_t)\phi_t) - \psi'(\phi_t) = \zeta_t.$$

Assim, o resíduo combinado padronizado é definido por

$$r_t^{\beta\gamma} = \frac{(y_t^* - \hat{\mu}^*) + \hat{a}_t}{\sqrt{\hat{\zeta}_t}}, \quad (3.22)$$

em que $\hat{\zeta}_t$ é dado por ζ_t avaliado em $\hat{\mu}_t$ e $\hat{\phi}_t$. Para o modelo com dispersão constante considera-se que $\phi_1 = \dots = \phi_n = \phi$.

Algumas propriedades desses resíduos serão avaliadas no Capítulo 4.

Avaliação numérica e aplicação no modelo de regressão beta linear

Neste capítulo serão apresentados resultados de simulações que objetivaram investigar a distribuição empírica do resíduo combinado e dos demais resíduos. Primeiro considerou-se o modelo de regressão beta com dispersão constante e depois o modelo com dispersão variável. Além disso, foram realizadas simulações com modelos mal especificados e, por último, há uma aplicação a dados reais.

4.1 Modelo de regressão beta com dispersão constante

Para investigar a distribuição empírica do resíduo combinado e dos demais resíduos quando o modelo de regressão beta tem dispersão constante foram realizadas simulações de Monte Carlo com 5000 réplicas considerando diferentes cenários. No primeiro caso foi considerada uma amostra de tamanho 20. Assim, y_1, \dots, y_{20} são variáveis aleatórias independentes, onde $y_t \sim B(\mu_t, \phi)$, para $t = 1, \dots, 20$. A média da variável resposta satisfaz a relação funcional

$$\log\left(\frac{\mu_t}{1 - \mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, \quad t = 1, \dots, 20. \quad (4.1)$$

As covariáveis foram geradas de forma independente através da distribuição uniforme, ou seja, $X_{ti} \sim U(0, 1)$, $i = 2, \dots, 5$, e foram mantidas constantes durante todo o experimento. Os valores verdadeiros para os parâmetros foram definidos por $\beta_1 = 1.2$, $\beta_2 = -2.2$, $\beta_3 = -2.2$, $\beta_4 = 1.2$, $\beta_5 = 1.1$, porém foram considerados três cenários diferentes para o parâmetro de precisão: 15, 50 e 400. Nessa configuração a média da variável resposta assume valores no intervalo (0.2, 0.9).

As análises foram feitas através de algumas medidas que podem caracterizar uma distribuição: a média, o erro-padrão, a assimetria e a curtose. Para a distribuição normal padrão os valores dessas medidas são 0, 1, 0 e 3, respectivamente. Através das Tabelas 4.1, 4.2 e 4.3 observa-se que as médias dos três resíduos estudados são semelhantes e bem próximas de zero, a similaridade também ocorrendo com os erros-padrão, que estão próximos de um. No entanto, nota-se que o resíduo combinado possui, em geral, assimetria negativa enquanto os outros resíduos têm assimetria positiva, mas todas próximas de zero. Já para a curtose, os valores são muito parecidos para todos os resíduos. O menor valor observado foi 2.515 (resíduo ponderado quando $\phi = 400$) e o maior valor observado corresponde a $\phi = 15$ para o resíduo combinado, quando a curtose é 2.989.

Assim, a partir das Tabelas 4.1, 4.2 e 4.3, nota-se que em um cenário onde o tamanho da amostra é consideravelmente pequeno, quando $\mu \in (0.2, 0.9)$, ou seja, os valores estão bem distribuídos no intervalo $(0, 1)$, as distribuições dos três resíduos aproximam-se bem da distribuição normal padrão, em especial quando ϕ aumenta.

Tabela 4.1: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 20$, quando $\phi = 15$ e $\mu \in (0.2, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	-0.001	-0.001	-0.006	1.037	1.135	1.037	-0.015	-0.015	-0.291	2.652	2.651	2.772
2	0.041	0.046	0.039	1.011	1.138	1.015	0.361	0.359	0.215	2.872	2.869	2.819
3	0.005	0.006	0.004	1.004	1.142	1.000	-0.216	-0.215	-0.424	2.703	2.706	2.826
4	-0.001	-0.000	-0.006	1.055	1.177	1.054	-0.064	-0.063	-0.344	2.718	2.716	2.855
5	0.010	0.012	0.008	1.020	1.158	1.016	-0.073	-0.072	-0.349	2.762	2.762	2.919
6	-0.004	-0.006	0.011	0.833	1.133	0.833	-0.000	0.000	-0.219	2.622	2.623	2.673
7	-0.010	-0.012	-0.005	0.950	1.140	0.952	0.103	0.101	-0.157	2.779	2.780	2.821
8	-0.021	-0.024	-0.020	0.985	1.146	0.982	-0.039	-0.040	-0.294	2.724	2.724	2.811
9	-0.013	-0.014	-0.023	1.089	1.158	1.088	0.071	0.071	-0.231	2.745	2.745	2.777
10	-0.028	-0.031	-0.029	1.011	1.148	1.014	0.020	0.022	-0.255	2.687	2.688	2.781
11	-0.005	-0.006	0.004	0.894	1.144	0.898	0.001	0.003	-0.231	2.623	2.621	2.709
12	-0.014	-0.017	-0.009	0.898	1.113	0.903	0.226	0.221	0.058	2.824	2.818	2.804
13	0.023	0.028	0.029	0.939	1.142	0.937	0.023	0.023	-0.231	2.701	2.700	2.750
14	-0.027	-0.032	-0.026	0.990	1.155	0.991	0.089	0.089	-0.188	2.746	2.748	2.811
15	0.011	0.012	0.004	1.078	1.160	1.075	-0.204	-0.204	-0.460	2.796	2.797	2.989
16	0.008	0.009	0.004	1.035	1.166	1.034	-0.048	-0.047	-0.317	2.654	2.652	2.797
17	0.023	0.029	0.028	0.930	1.140	0.924	-0.199	-0.195	-0.397	2.819	2.806	2.934
18	0.008	0.009	0.003	1.048	1.160	1.052	0.171	0.171	-0.083	2.762	2.762	2.758
19	0.007	0.008	0.000	1.058	1.142	1.054	-0.018	-0.017	-0.320	2.797	2.796	2.910
20	-0.002	-0.002	-0.001	0.970	1.144	0.976	0.196	0.196	0.001	2.776	2.773	2.741

Uma outra forma de verificar se os resíduos apresentam distribuição aproximadamente normal ou normal é graficamente. Assim, com o intuito de comparar os quantis empíricos dos resíduos com os quantis teóricos da distribuição normal padrão, foram construídos gráficos normais de probabilidade, que são apresentados nas Figuras 4.1 e 4.2.

Aqui foram considerados seis diferentes cenários, mas ainda considerando o modelo em (4.1). Consideramos uma amostra de tamanho 80, em que os valores das covariáveis foram obtidos replicando quatro vezes os valores anteriores. Os valores verdadeiros dos

Tabela 4.2: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 20$, quando $\phi = 150$ e $\mu \in (0.2, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.024	0.026	0.022	1.047	1.146	1.046	-0.030	-0.030	-0.120	2.609	2.608	2.634
2	0.006	0.006	0.005	1.049	1.166	1.050	0.124	0.123	0.070	2.693	2.694	2.687
3	0.004	0.004	0.003	1.016	1.154	1.015	-0.017	-0.017	-0.087	2.597	2.598	2.607
4	0.011	0.013	0.010	1.038	1.159	1.036	0.019	0.019	-0.065	2.589	2.589	2.589
5	-0.004	-0.004	-0.004	1.018	1.157	1.018	-0.024	-0.024	-0.110	2.673	2.673	2.687
6	0.007	0.010	0.012	0.852	1.165	0.852	0.000	0.000	-0.072	2.641	2.642	2.645
7	-0.013	-0.016	-0.012	0.962	1.160	0.963	0.004	0.005	-0.075	2.650	2.651	2.657
8	0.001	0.001	0.002	0.972	1.134	0.971	0.008	0.008	-0.068	2.610	2.610	2.614
9	-0.002	-0.002	-0.005	1.077	1.148	1.077	0.008	0.008	-0.083	2.648	2.649	2.657
10	-0.015	-0.017	-0.016	1.016	1.156	1.017	0.009	0.009	-0.082	2.713	2.713	2.720
11	-0.002	-0.002	0.001	0.890	1.144	0.889	0.046	0.046	-0.027	2.634	2.637	2.630
12	0.006	0.007	0.007	0.950	1.161	0.950	0.069	0.068	0.010	2.724	2.721	2.726
13	-0.008	-0.010	-0.006	0.942	1.151	0.942	0.023	0.023	-0.056	2.635	2.634	2.638
14	-0.005	-0.006	-0.004	0.970	1.134	0.971	0.035	0.035	-0.049	2.665	2.665	2.672
15	0.013	0.014	0.011	1.076	1.158	1.075	-0.053	-0.052	-0.133	2.621	2.621	2.638
16	-0.011	-0.012	-0.011	1.013	1.143	1.013	-0.013	-0.013	-0.104	2.732	2.733	2.749
17	-0.003	-0.004	-0.002	0.957	1.169	0.956	-0.042	-0.041	-0.109	2.660	2.658	2.668
18	0.001	0.001	-0.000	1.047	1.160	1.047	0.050	0.050	-0.030	2.668	2.669	2.667
19	-0.009	-0.010	-0.012	1.071	1.155	1.070	0.025	0.025	-0.067	2.660	2.660	2.660
20	0.007	0.008	0.007	0.983	1.155	0.983	0.095	0.095	0.031	2.661	2.658	2.649

Tabela 4.3: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 20$, quando $\phi = 400$ e $\mu \in (0.2, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	-0.028	-0.031	-0.029	1.068	1.170	1.068	0.042	0.042	-0.013	2.663	2.662	2.662
2	-0.019	-0.021	-0.020	1.041	1.156	1.041	0.092	0.092	0.059	2.645	2.644	2.635
3	0.014	0.016	0.014	1.000	1.135	1.000	-0.076	-0.076	-0.120	2.683	2.684	2.698
4	0.010	0.011	0.009	1.044	1.166	1.044	-0.061	-0.061	-0.113	2.595	2.595	2.605
5	-0.040	-0.046	-0.041	1.032	1.173	1.033	0.034	0.034	-0.017	2.619	2.618	2.614
6	-0.013	-0.018	-0.011	0.863	1.181	0.863	0.076	0.076	0.032	2.598	2.598	2.594
7	-0.002	-0.003	-0.001	0.945	1.139	0.945	0.014	0.014	-0.034	2.735	2.736	2.734
8	0.003	0.003	0.003	0.979	1.142	0.979	-0.061	-0.061	-0.112	2.720	2.721	2.729
9	-0.001	-0.001	-0.003	1.064	1.133	1.064	-0.003	-0.003	-0.057	2.593	2.593	2.597
10	0.000	0.000	0.000	1.017	1.157	1.018	-0.049	-0.049	-0.102	2.644	2.644	2.654
11	-0.001	-0.001	0.000	0.904	1.163	0.904	0.016	0.016	-0.028	2.600	2.600	2.602
12	0.022	0.027	0.023	0.942	1.151	0.942	0.020	0.019	-0.013	2.673	2.671	2.670
13	-0.006	-0.008	-0.006	0.953	1.164	0.953	-0.023	-0.022	-0.071	2.613	2.613	2.611
14	-0.001	-0.001	-0.001	0.986	1.152	0.985	0.012	0.012	-0.036	2.583	2.583	2.587
15	0.019	0.021	0.018	1.073	1.155	1.072	-0.030	-0.030	-0.081	2.680	2.680	2.688
16	0.020	0.023	0.019	1.037	1.170	1.037	-0.018	-0.018	-0.069	2.515	2.516	2.521
17	-0.003	-0.004	-0.002	0.934	1.141	0.934	-0.070	-0.071	-0.110	2.706	2.703	2.715
18	0.020	0.022	0.019	1.038	1.150	1.038	0.014	0.014	-0.035	2.710	2.710	2.714
19	0.008	0.008	0.007	1.061	1.144	1.061	-0.046	-0.046	-0.102	2.691	2.691	2.702
20	0.005	0.006	0.005	0.972	1.142	0.972	0.076	0.077	0.039	2.601	2.601	2.597

parâmetros de regressão são $\beta_1 = -1.5$, $\beta_2 = -1.2$, $\beta_3 = 1.0$, $\beta_4 = 1.1$ e $\beta_5 = -1.3$. Consideramos três valores diferentes para o parâmetro de precisão: 50, 150 e 400. Neste contexto, $\mu \in (0.05, 0.5)$. Outros cenários considerados foram com os valores dos parâmetros definidos como $\beta_1 = 0.8$, $\beta_2 = -1.0$, $\beta_3 = -1.0$, $\beta_4 = 1.2$ e $\beta_5 = 1.1$, novamente com diferentes valores para o parâmetro de precisão: 50, 150 e 400. Neste cenário, a média da variável resposta assume valores em $(0.5, 0.9)$.

Nos cenários considerados todos os resíduos possuem distribuição empírica consideravelmente próxima da normal padrão. No entanto, em todos os casos quando a média da variável resposta pertence ao intervalo $(0.05, 0.5)$ as distribuições destes resíduos apresentam leve assimetria, o que é mais evidente no resíduo combinado. Nota-se também que à medida que ϕ aumenta a distribuição do resíduo combinado se aproxima mais da distribuição normal. Quando a média da variável resposta pertence ao intervalo $(0.5, 0.9)$, para todos os valores da precisão considerados, o resíduo combinado possui a distribuição mais próxima da distribuição normal padrão quando comparada a dos outros resíduos.

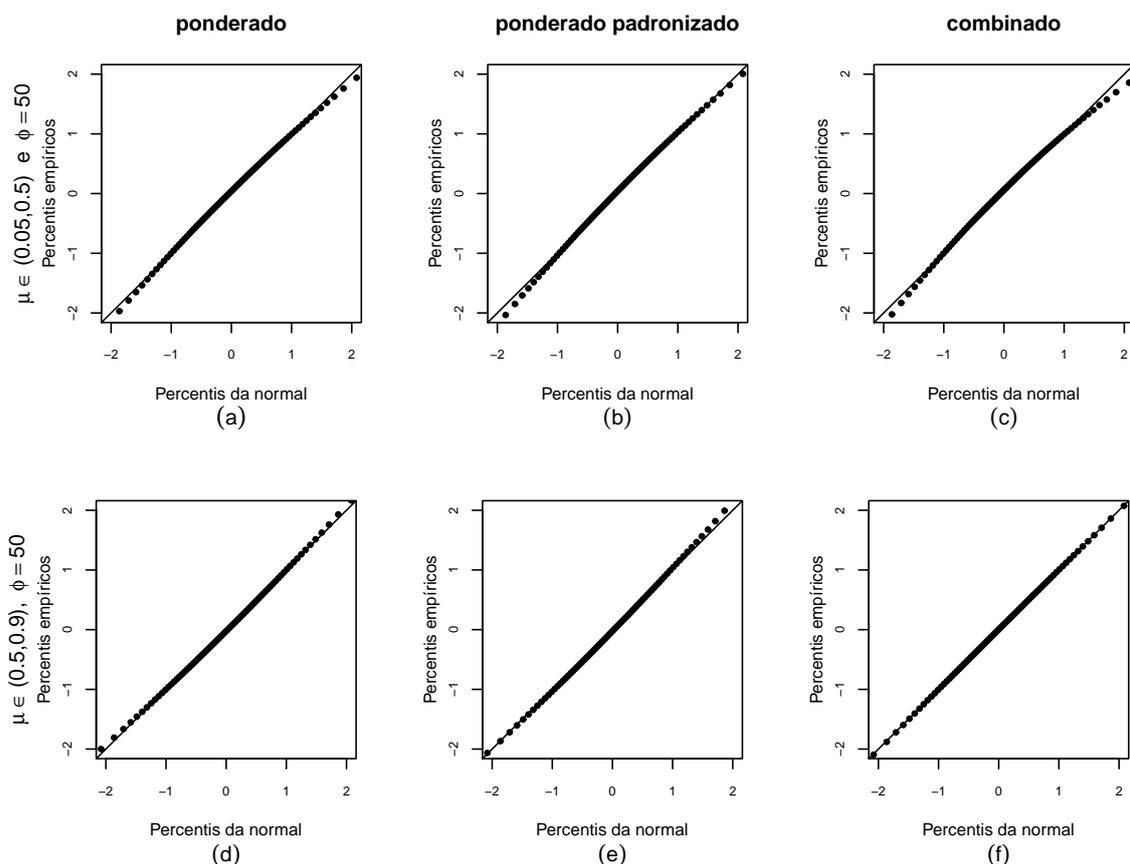


Figura 4.1: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 80$.

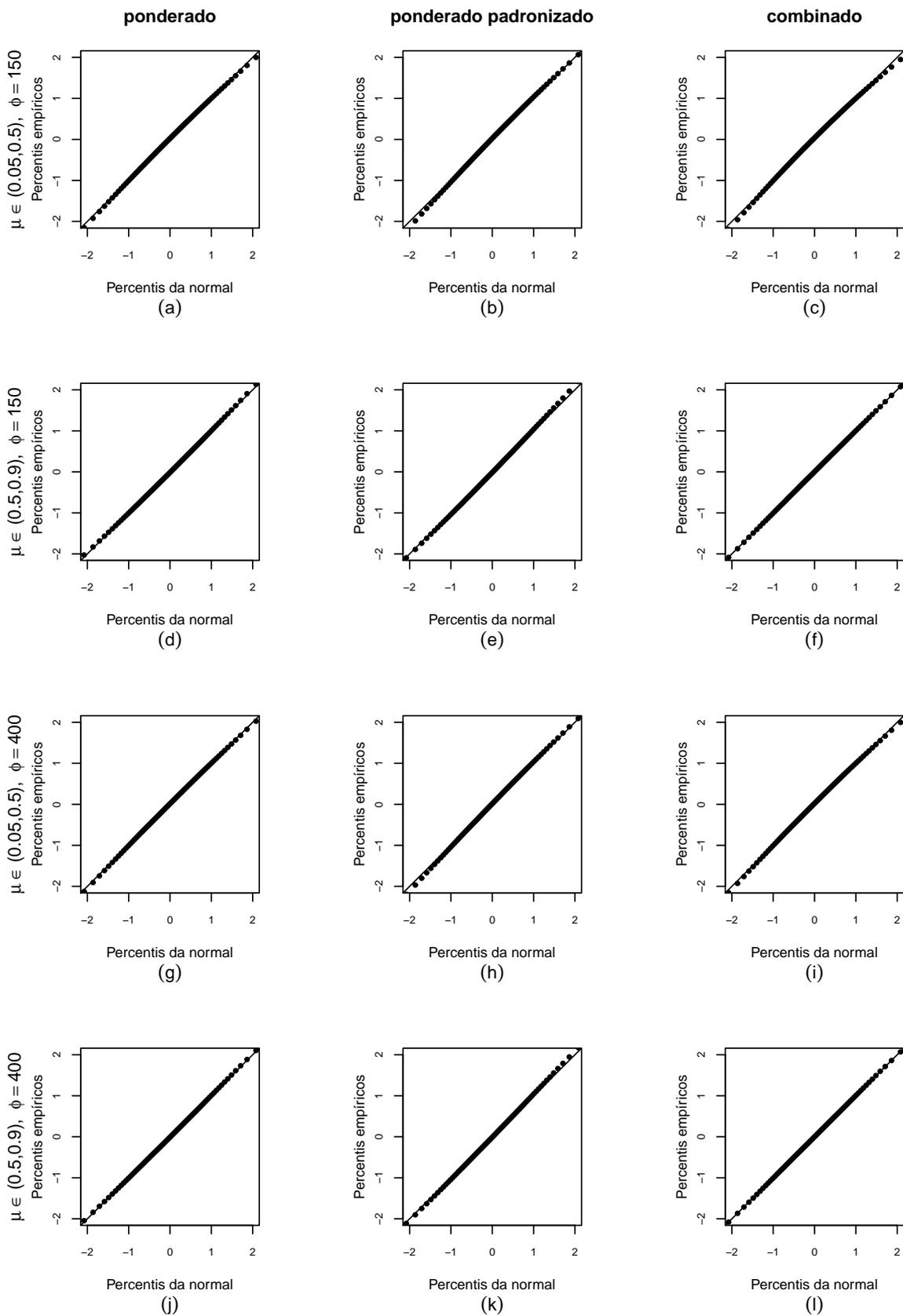


Figura 4.2: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 80$.

4.2 Modelo de regressão beta com dispersão variável

Para investigar as distribuições empíricas do resíduo combinado e dos demais resíduos no modelo de regressão beta com dispersão variável também foram realizadas simulações de Monte Carlo com 5000 réplicas em diferentes cenários. Nas primeiras análises foram consideradas amostras de tamanho 20. Assim, y_1, \dots, y_{20} são variáveis aleatórias independentes, onde $y_t \sim B(\mu_t, \phi_t)$, para $t = 1, \dots, 20$. A média da variável resposta e o parâmetro de precisão satisfazem as relações funcionais

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, \\ \log(\phi_t) &= \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}, \quad t = 1, \dots, 20. \end{aligned} \quad (4.2)$$

Os valores das covariadas foram gerados de forma independente através da distribuição uniforme, $X_{ti} \sim U(0, 1)$ e $Z_{ti} \sim U(-0.5, 0.5)$, $i = 2, \dots, 5$, e foram mantidos constantes durante todo o experimento. Temos que $\mu \in (0.2, 0.9)$, tendo os valores dos parâmetros sido definidos como $\beta = (1.2, -2.2, -2.2, 1.2, 1.1)$. Para o submodelo da precisão consideramos três cenários diferentes, que conduziram a diferentes valores do grau de heterogeneidade da dispersão dos dados, que é definido como

$$\lambda = \frac{\max(\phi_t)}{\min(\phi_t)}, \quad t = 1, \dots, n.$$

No primeiro cenário consideramos $\gamma_1 = 2.5$, $\gamma_2 = 1.4$, $\gamma_3 = 1.5$, $\gamma_4 = 1.2$ e $\gamma_5 = 1.4$, que conduz a $\lambda = 20$. No segundo cenário, $\gamma_1 = 2.5$, $\gamma_2 = 2.1$, $\gamma_3 = 1.6$, $\gamma_4 = 1.7$ e $\gamma_5 = 1.8$, que leva a $\lambda = 50$. Por último, definimos $\gamma_1 = 2.5$, $\gamma_2 = 2.4$, $\gamma_3 = 1.9$, $\gamma_4 = 2.0$ e $\gamma_5 = 2.2$, o que acarreta $\lambda = 100$.

A primeira avaliação da distribuição empírica desses resíduos foi realizada através das médias, dos erros-padrão, das assimetrias e das curtoses. Nas Tabelas 4.4, 4.5 e 4.6, observam-se que as médias dos três resíduos são semelhantes e próximas de zero. Os erros-padrão encontram-se, em geral, próximos de um; entre os casos que mais distam de um estão: 0.558, o menor valor observado, no resíduo ponderado e 1.671, o maior, no resíduo ponderado padronizado, ambos valores no cenário em que $\lambda = 100$. Com relação à assimetria, observa-se que o resíduo combinado possui uma assimetria negativa, fato que não ocorre com os outros resíduos. Todos os resíduos, no entanto, possuem assimetrias próximas de zero. Já para a curtose, os resíduos possuem medidas similares, no entanto algumas se distanciam bastante de três, como por exemplo 1.362, para o resíduo ponderado padronizado e 5.034, no resíduo combinado, valores que aparecem quando $\lambda = 100$.

Para comparar os quantis empíricos dos resíduos com os quantis teóricos da distribuição normal padrão foram construídos gráficos normais de probabilidade, que são apresentados nas Figuras 4.3 e 4.4. Aqui ainda consideramos o modelo da Equação (4.2). O tamanho

Tabela 4.4: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 20$, quando $\lambda = 20$ e $\mu \in (0.2, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$
1	-0.002	0.002	0.010	0.976	1.079	0.976	-0.005	0.003	-0.258	2.291	2.312	2.338
2	-0.017	-0.016	-0.016	0.897	0.948	0.899	0.437	0.419	0.285	2.342	2.356	2.269
3	-0.030	-0.035	-0.039	1.109	1.253	1.110	-0.157	-0.135	-0.317	2.316	2.251	2.410
4	0.031	0.044	0.023	1.174	1.599	1.173	-0.056	-0.055	-0.105	1.630	1.514	1.640
5	-0.013	-0.016	-0.017	1.148	1.580	1.149	-0.006	0.012	-0.076	1.719	1.463	1.718
6	-0.005	-0.001	-0.001	0.719	1.350	0.721	0.001	0.012	-0.200	3.903	2.183	3.989
7	-0.024	-0.033	-0.008	0.785	1.122	0.779	0.179	0.104	-0.081	3.066	2.626	2.997
8	0.023	0.028	0.033	0.783	1.078	0.783	-0.119	-0.101	-0.351	3.058	2.370	3.168
9	-0.018	-0.018	-0.005	0.997	1.033	0.995	0.064	0.075	-0.236	2.186	2.205	2.241
10	0.012	0.014	-0.003	1.178	1.252	1.179	-0.009	-0.001	-0.219	2.201	2.174	2.253
11	0.026	0.023	0.003	1.100	1.182	1.099	0.045	0.029	-0.221	2.046	1.978	2.089
12	-0.094	-0.096	-0.081	0.838	0.929	0.837	0.581	0.516	0.382	2.749	2.772	2.578
13	0.005	0.031	0.009	0.691	1.350	0.691	0.010	0.048	-0.149	3.732	2.169	3.776
14	-0.005	-0.007	-0.025	1.284	1.437	1.286	0.022	0.012	-0.102	2.155	1.912	2.160
15	-0.001	-0.003	-0.013	1.153	1.209	1.153	-0.171	-0.165	-0.327	2.032	2.007	2.108
16	0.029	0.030	0.033	1.083	1.115	1.079	-0.043	-0.043	-0.386	2.425	2.417	2.570
17	-0.032	-0.031	-0.037	1.007	1.201	1.013	-0.268	-0.198	-0.476	2.675	2.415	2.897
18	-0.058	-0.060	-0.046	0.932	0.964	0.945	0.267	0.259	-0.082	2.320	2.338	2.317
19	0.007	0.001	0.024	1.004	1.064	1.007	-0.093	-0.079	-0.395	2.621	2.565	2.720
20	-0.059	-0.063	-0.053	0.898	1.029	0.899	0.311	0.265	0.122	2.529	2.578	2.444

Tabela 4.5: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 20$, quando $\lambda = 50$ e $\mu \in (0.2, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$
1	0.014	0.013	0.029	0.987	1.063	0.986	-0.030	-0.036	-0.305	2.288	2.343	2.377
2	-0.031	-0.032	-0.028	0.852	0.890	0.858	0.469	0.448	0.311	2.440	2.443	2.383
3	-0.018	-0.020	-0.025	1.097	1.256	1.097	-0.110	-0.094	-0.256	2.353	2.273	2.418
4	-0.005	-0.006	-0.012	1.153	1.642	1.154	-0.009	-0.002	-0.046	1.584	1.453	1.593
5	0.016	0.021	0.009	1.186	1.630	1.186	-0.033	-0.028	-0.090	1.662	1.402	1.668
6	0.010	0.025	0.014	0.680	1.412	0.680	-0.035	0.001	-0.215	4.077	2.113	4.191
7	-0.037	-0.053	-0.021	0.789	1.081	0.787	0.174	0.127	-0.082	2.991	2.595	2.936
8	-0.004	-0.018	0.002	0.782	1.116	0.782	-0.057	-0.063	-0.278	3.034	2.337	3.126
9	-0.020	-0.019	-0.003	0.986	1.019	0.988	0.032	0.030	-0.262	2.187	2.231	2.249
10	0.018	0.017	0.002	1.193	1.261	1.194	-0.028	-0.029	-0.217	2.170	2.150	2.225
11	-0.007	-0.006	-0.037	1.130	1.186	1.135	0.097	0.102	-0.194	1.984	1.942	2.045
12	-0.141	-0.143	-0.128	0.816	0.888	0.820	0.634	0.604	0.410	2.959	3.048	2.766
13	-0.004	0.006	-0.000	0.610	1.354	0.612	-0.100	0.031	-0.253	4.426	2.140	4.522
14	0.015	0.018	-0.004	1.317	1.454	1.317	0.010	0.003	-0.103	2.117	1.916	2.130
15	0.023	0.024	0.009	1.179	1.240	1.177	-0.141	-0.141	-0.285	2.006	1.982	2.080
16	0.012	0.012	0.025	1.043	1.063	1.042	-0.092	-0.090	-0.523	2.621	2.609	2.870
17	-0.032	-0.041	-0.037	1.007	1.201	1.012	-0.201	-0.148	-0.389	2.587	2.345	2.745
18	-0.028	-0.028	-0.016	0.917	0.939	0.929	0.297	0.288	-0.100	2.358	2.368	2.403
19	0.017	0.015	0.036	1.007	1.054	1.008	-0.069	-0.077	-0.360	2.648	2.615	2.723
20	-0.031	-0.027	-0.025	0.902	1.006	0.903	0.266	0.249	0.080	2.441	2.493	2.381

da amostra é 80 e os valores das covariáveis foram obtidos replicando-se quatro vezes os valores anteriores a fim de manter o grau de heterogeneidade constante. Consideramos dois cenários para a média da variável resposta, $\mu \in (0.05, 0.5)$, que corresponde a $\beta_1 = -1.5$, $\beta_2 = -1.2$, $\beta_3 = 1.0$, $\beta_4 = 1.1$ e $\beta_5 = -1.3$, e $\mu \in (0.5, 0.9)$, que corresponde a $\beta_1 = 0.8$,

Tabela 4.6: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 20$, quando $\lambda = 100$ e $\mu \in (0.2, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$
1	0.003	0.006	0.019	0.984	1.048	0.983	-0.010	-0.006	-0.280	2.273	2.288	2.339
2	-0.001	-0.002	-0.000	0.848	0.880	0.851	0.481	0.466	0.310	2.484	2.478	2.425
3	-0.005	-0.008	-0.014	1.098	1.265	1.098	-0.152	-0.139	-0.294	2.307	2.206	2.385
4	0.000	-0.000	-0.004	1.126	1.671	1.127	-0.010	0.012	-0.042	1.599	1.439	1.613
5	0.016	0.022	0.010	1.217	1.660	1.216	-0.017	-0.024	-0.064	1.604	1.362	1.604
6	-0.005	-0.006	-0.001	0.637	1.423	0.638	-0.000	0.043	-0.195	4.831	2.070	4.945
7	-0.022	-0.041	-0.007	0.801	1.095	0.799	0.157	0.097	-0.078	2.761	2.528	2.749
8	-0.015	-0.032	-0.009	0.783	1.166	0.786	-0.078	-0.056	-0.274	3.025	2.279	3.121
9	-0.005	-0.004	0.011	0.984	1.006	0.983	0.034	0.040	-0.290	2.192	2.211	2.266
10	0.008	0.007	-0.008	1.214	1.270	1.215	-0.011	-0.014	-0.192	2.145	2.120	2.185
11	0.036	0.036	0.000	1.149	1.192	1.148	0.090	0.080	-0.221	1.983	1.946	2.058
12	-0.162	-0.166	-0.147	0.785	0.846	0.788	0.687	0.616	0.470	2.934	3.099	2.773
13	0.009	0.041	0.012	0.558	1.381	0.559	-0.011	0.020	-0.152	4.981	2.082	5.034
14	-0.029	-0.034	-0.048	1.337	1.474	1.339	0.040	0.031	-0.058	2.097	1.890	2.094
15	0.024	0.027	0.011	1.177	1.235	1.174	-0.131	-0.126	-0.274	2.005	1.984	2.071
16	0.029	0.028	0.050	0.998	1.013	0.992	-0.061	-0.061	-0.555	2.780	2.771	3.068
17	-0.043	-0.050	-0.050	1.029	1.211	1.033	-0.178	-0.134	-0.373	2.555	2.313	2.716
18	-0.035	-0.035	-0.020	0.889	0.904	0.900	0.362	0.357	-0.097	2.470	2.472	2.497
19	0.031	0.035	0.048	1.013	1.052	1.010	-0.041	-0.038	-0.332	2.587	2.563	2.665
20	-0.036	-0.034	-0.030	0.915	1.020	0.915	0.303	0.289	0.125	2.477	2.562	2.405

$\beta_2 = -1.0$, $\beta_3 = -1.0$, $\beta_4 = 1.2$ e $\beta_5 = -1.1$. Para o submodelo da precisão consideramos três cenários: $\lambda = 20$, dado que $\gamma_1 = 3.0$, $\gamma_2 = 1.7$, $\gamma_3 = 1.5$, $\gamma_4 = 1.7$, $\gamma_5 = 1.4$; $\lambda = 50$, dado que $\gamma_1 = 3.0$, $\gamma_2 = 2.5$, $\gamma_3 = 1.8$, $\gamma_4 = 1.7$ e $\gamma_5 = 2.1$, e $\lambda = 100$, dado que $\gamma_1 = 3.2$, $\gamma_2 = 2.7$, $\gamma_3 = -0.9$, $\gamma_4 = 2.5$, $\gamma_5 = 3.8$.

Nos cenários considerados todos os resíduos possuem distribuição empírica consideravelmente próximas da distribuição normal padrão. No entanto, em todos os casos quando a média da variável resposta pertence ao intervalo $(0.05, 0.5)$ as distribuições destes resíduos apresentam leve assimetria, o que novamente é mais evidente no resíduo combinado. Porém, quando a média da variável resposta pertence ao intervalo $(0.5, 0.9)$, para todos os valores do grau de heterogeneidade de dispersão dos dados, o resíduo combinado possui distribuição mais próxima da distribuição normal padrão relativamente aos outros resíduos.

4.2.1 Modelos mal especificados

Um dos objetivos da análise de resíduos é verificar se o modelo postulado está bem especificado. Aqui consideramos um erro de especificação do modelo comumente cometido: usar dispersão constante quando o processo gerador de dados tem dispersão variável. As distribuições empíricas dos resíduos também foram investigadas neste contexto. Aqui, o processo de geração dos dados foi baseado no modelo em (4.2), no entanto no processo de estimação considerou-se o modelo (4.1). Foram considerados três cenários distintos: $\lambda =$

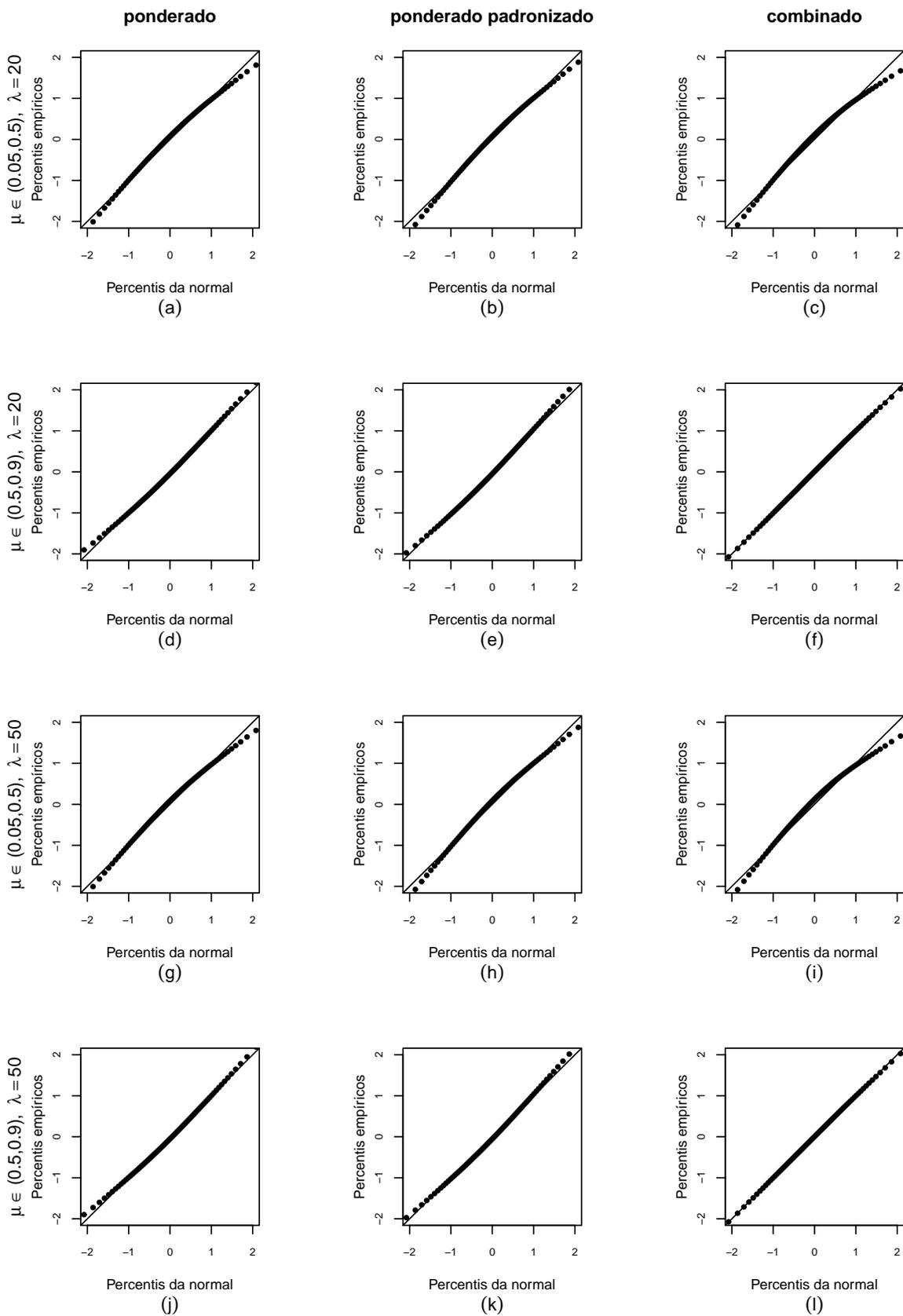


Figura 4.3: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 80$.

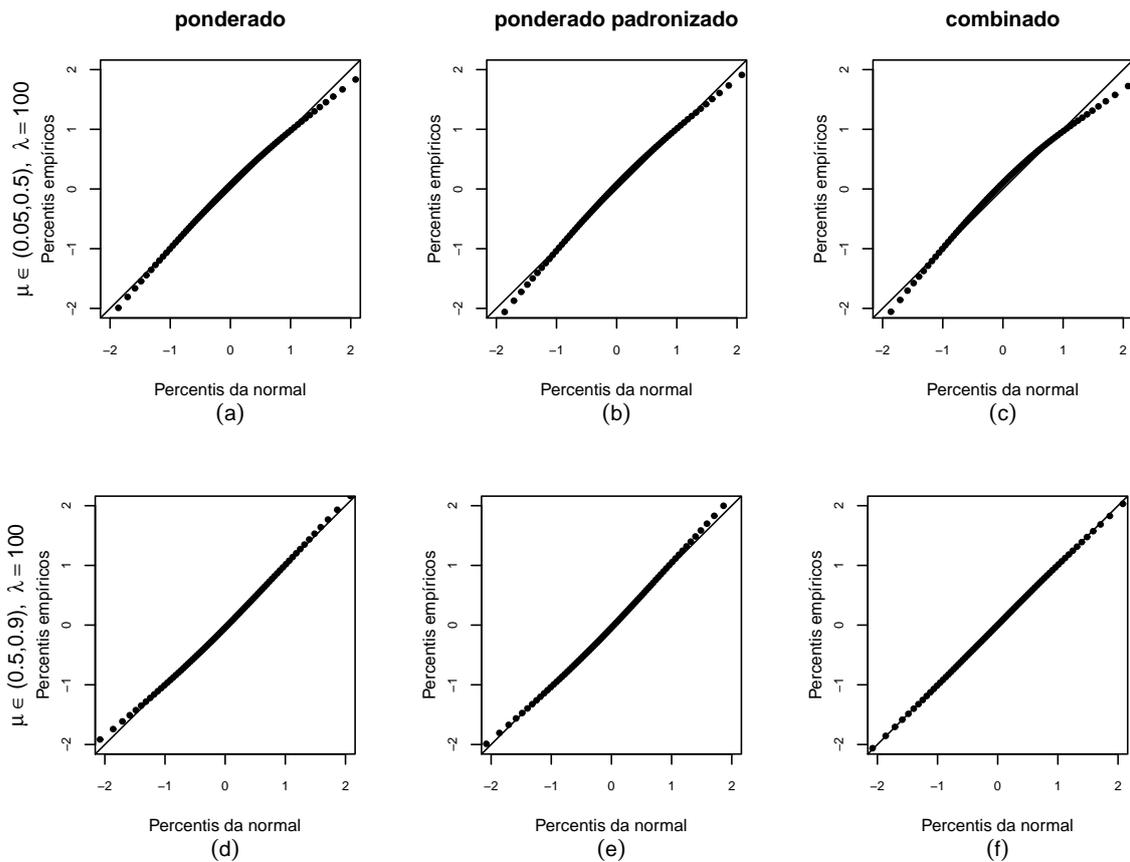


Figura 4.4: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $t = 1, \dots, 80$.

(20, 50, 100). A média da variável resposta assume valores em (0.2, 0.9) e mais uma vez utilizamos um tamanho de amostra consideravelmente pequeno, 20.

Nas Tabelas 4.7, 4.8 e 4.9 é possível observar que todas as medidas sofrem oscilação quando comparadas ao modelo corretamente especificado. O erro-padrão é a medida que mais varia. Por exemplo, na Tabela 4.9, quando $\lambda = 100$, no resíduo ponderado o menor valor observado é 0.486 e o maior é 2.114, no resíduo ponderado padronizado o menor valor observado é 0.59 e o maior 2.322, já no resíduo combinado o menor valor é 0.492 e o maior é 2.779.

Para a análise gráfica foi utilizada uma amostra de tamanho 80 e também foram considerados três cenários com variação no grau de heterogeneidade, $\lambda = (20, 50, 100)$, além de dois cenários diferentes para a variação da média da variável resposta: $\mu \in (0.05, 0.5)$ e $\mu \in (0.5, 0.9)$. Nas Figuras 4.5 e 4.6 nota-se em todos os gráficos que a distribuição empírica dos resíduos não é bem aproximada pela distribuição normal.

Em geral avaliamos a distribuição empírica dos resíduos quando os modelos estão bem especificados. Mas na prática, se existir alguma violação das suposições estabelecidas para o modelo postulado, as distribuições desses resíduos podem mudar consideravelmente. Por

Tabela 4.7: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 20$, quando $\lambda = 20$, $\mu \in (0.2, 0.9)$. O modelo é estimado supondo dispersão constante.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.017	0.019	0.023	0.962	1.053	0.962	-0.048	-0.047	-0.402	2.903	2.903	3.108
2	0.417	0.469	0.386	1.475	1.672	1.489	0.345	0.332	0.185	2.488	2.490	2.544
3	0.170	0.193	0.179	0.863	0.981	0.849	-0.381	-0.377	-0.637	3.153	3.149	3.467
4	-0.029	-0.035	0.011	0.631	0.705	0.631	-0.021	-0.038	-0.260	2.876	2.880	2.930
5	-0.008	-0.012	0.029	0.658	0.750	0.659	-0.106	-0.127	-0.368	3.037	3.051	3.267
6	-0.077	-0.102	-0.032	0.584	0.793	0.589	0.013	0.029	-0.227	2.972	2.957	3.031
7	-0.087	-0.102	-0.069	0.832	0.996	0.837	0.217	0.228	-0.075	2.879	2.892	2.798
8	0.100	0.114	0.121	0.796	0.926	0.786	-0.147	-0.157	-0.432	2.941	2.949	3.169
9	-0.071	-0.075	-0.117	1.304	1.387	1.320	0.065	0.064	-0.334	2.599	2.597	2.733
10	-0.059	-0.067	-0.041	0.867	0.984	0.869	0.098	0.095	-0.241	3.044	3.051	3.082
11	-0.106	-0.142	-0.113	1.041	1.324	1.059	0.141	0.110	-0.156	2.607	2.595	2.618
12	0.112	0.122	0.099	1.164	1.442	1.180	0.316	0.250	0.129	2.599	2.568	2.630
13	-0.086	-0.105	-0.041	0.572	0.696	0.576	0.086	0.086	-0.153	3.041	3.033	3.025
14	-0.133	-0.156	-0.092	0.625	0.729	0.635	0.029	0.018	-0.230	3.079	3.079	3.124
15	0.059	0.063	0.063	0.953	1.025	0.945	-0.266	-0.264	-0.571	3.041	3.041	3.324
16	0.004	0.006	-0.040	1.307	1.470	1.303	-0.046	-0.040	-0.415	2.534	2.526	2.690
17	0.126	0.154	0.135	0.862	1.057	0.849	-0.339	-0.335	-0.565	2.933	2.914	3.184
18	0.156	0.168	0.057	1.708	1.881	1.732	0.291	0.280	-0.036	2.399	2.387	2.409
19	-0.035	-0.038	-0.032	0.987	1.065	0.987	-0.010	-0.009	-0.357	2.891	2.892	3.014
20	-0.165	-0.193	-0.159	0.864	1.020	0.878	0.424	0.425	0.193	3.080	3.076	2.909

Tabela 4.8: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 20$, quando $\lambda = 50$, $\mu \in (0.2, 0.9)$. O modelo é estimado supondo dispersão constante.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.013	0.014	0.018	0.973	1.066	0.969	0.000	0.001	-0.380	2.942	2.942	3.079
2	0.603	0.679	0.563	1.628	1.852	1.653	0.251	0.235	0.095	2.290	2.298	2.407
3	0.234	0.265	0.248	0.787	0.894	0.768	-0.368	-0.371	-0.627	3.155	3.154	3.473
4	-0.059	-0.070	-0.014	0.600	0.672	0.604	-0.069	-0.093	-0.314	2.869	2.876	2.986
5	-0.035	-0.044	0.008	0.620	0.707	0.622	-0.082	-0.110	-0.323	2.802	2.811	2.931
6	-0.115	-0.155	-0.063	0.527	0.716	0.535	0.044	0.063	-0.200	3.074	3.059	3.072
7	-0.144	-0.169	-0.122	0.797	0.953	0.809	0.182	0.193	-0.108	2.764	2.775	2.733
8	0.122	0.140	0.150	0.740	0.860	0.729	-0.178	-0.191	-0.462	2.893	2.895	3.167
9	-0.122	-0.130	-0.157	1.229	1.307	1.247	0.119	0.117	-0.309	2.794	2.793	2.857
10	-0.095	-0.108	-0.064	0.762	0.865	0.770	0.052	0.051	-0.265	3.065	3.066	3.111
11	-0.187	-0.247	-0.212	1.144	1.449	1.176	0.236	0.195	-0.088	2.605	2.572	2.563
12	0.073	0.073	0.057	1.160	1.442	1.180	0.358	0.288	0.157	2.655	2.625	2.677
13	-0.099	-0.121	-0.048	0.536	0.651	0.542	0.070	0.071	-0.173	3.025	3.012	3.045
14	-0.179	-0.210	-0.130	0.549	0.640	0.560	0.144	0.130	-0.110	3.167	3.164	3.135
15	0.151	0.162	0.167	0.816	0.878	0.800	-0.259	-0.257	-0.567	3.165	3.166	3.494
16	-0.014	-0.013	-0.071	1.394	1.567	1.394	-0.080	-0.070	-0.476	2.587	2.574	2.806
17	0.212	0.257	0.228	0.760	0.932	0.741	-0.287	-0.297	-0.516	2.920	2.924	3.142
18	0.358	0.388	0.223	1.947	2.142	1.976	0.267	0.255	-0.060	2.293	2.282	2.398
19	-0.061	-0.066	-0.044	0.893	0.964	0.896	-0.007	-0.006	-0.354	2.997	2.997	3.150
20	-0.222	-0.258	-0.213	0.815	0.962	0.832	0.461	0.475	0.213	3.203	3.224	2.971

exemplo, no modelo de regressão beta sob omissão da modelagem da dispersão, se usarmos os limites convencionais $(-2, 2)$ para detectar pontos aberrantes estaremos cometendo um grave erro, pois como vimos a distribuição dos resíduos é fortemente assimétrica e por isso esses limites não deveriam ser usados. Nossa proposta é que essa questão seja solucionada

Tabela 4.9: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1-\mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 20$, quando $\lambda = 100$, $\mu \in (0.2, 0.9)$. O modelo é estimado supondo dispersão constante.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.012	0.014	0.024	0.936	1.025	0.930	0.044	0.044	-0.337	2.889	2.895	3.004
2	0.676	0.766	0.631	1.667	1.905	1.705	0.240	0.221	0.050	2.344	2.356	2.549
3	0.279	0.315	0.298	0.710	0.807	0.690	-0.455	-0.465	-0.742	3.481	3.493	4.048
4	-0.094	-0.109	-0.041	0.561	0.628	0.566	-0.002	-0.027	-0.242	2.750	2.747	2.821
5	-0.049	-0.060	0.001	0.579	0.661	0.581	-0.024	-0.055	-0.268	2.794	2.795	2.888
6	-0.127	-0.171	-0.068	0.498	0.676	0.507	0.033	0.050	-0.218	3.091	3.064	3.127
7	-0.132	-0.154	-0.103	0.755	0.901	0.766	0.186	0.198	-0.124	2.831	2.856	2.831
8	0.144	0.164	0.177	0.696	0.809	0.684	-0.241	-0.258	-0.536	2.967	2.983	3.305
9	-0.144	-0.154	-0.192	1.283	1.363	1.314	0.092	0.091	-0.383	2.795	2.791	2.907
10	-0.104	-0.118	-0.063	0.699	0.793	0.705	0.137	0.136	-0.196	3.142	3.155	3.124
11	-0.205	-0.269	-0.236	1.152	1.456	1.198	0.189	0.147	-0.151	2.567	2.533	2.551
12	0.087	0.091	0.068	1.182	1.473	1.205	0.417	0.343	0.204	2.753	2.712	2.774
13	-0.123	-0.150	-0.063	0.486	0.590	0.492	0.146	0.141	-0.095	2.983	2.976	2.916
14	-0.206	-0.242	-0.150	0.508	0.592	0.523	0.089	0.070	-0.171	3.198	3.200	3.177
15	0.197	0.213	0.221	0.752	0.809	0.733	-0.297	-0.297	-0.634	3.414	3.417	3.843
16	-0.051	-0.054	-0.123	1.459	1.637	1.454	-0.024	-0.014	-0.442	2.565	2.549	2.720
17	0.249	0.301	0.266	0.738	0.907	0.719	-0.413	-0.433	-0.684	3.264	3.285	3.659
18	0.411	0.445	0.241	2.114	2.322	2.179	0.231	0.219	-0.099	2.186	2.176	2.352
19	-0.071	-0.077	-0.045	0.831	0.897	0.836	-0.017	-0.018	-0.375	3.026	3.026	3.112
20	-0.271	-0.315	-0.259	0.762	0.901	0.778	0.619	0.637	0.362	3.436	3.469	3.093

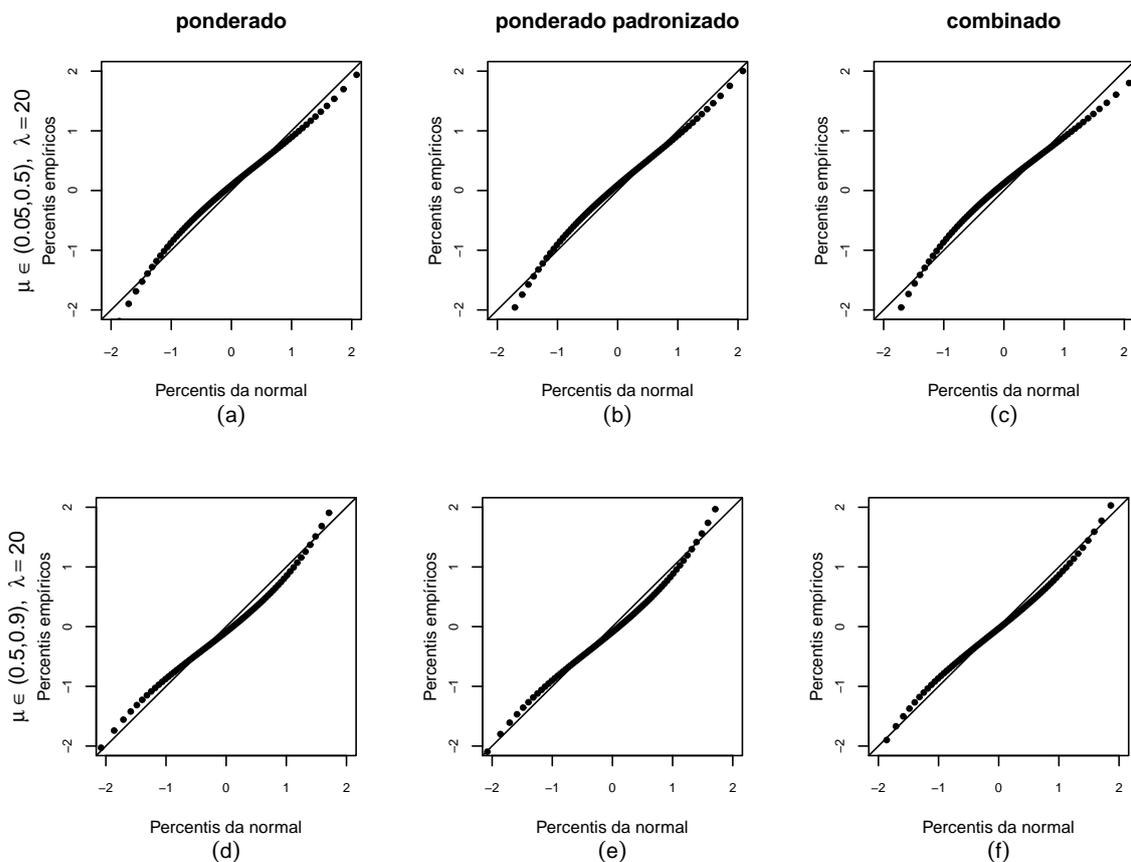


Figura 4.5: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1-\mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 80$. O modelo é estimado supondo dispersão constante.

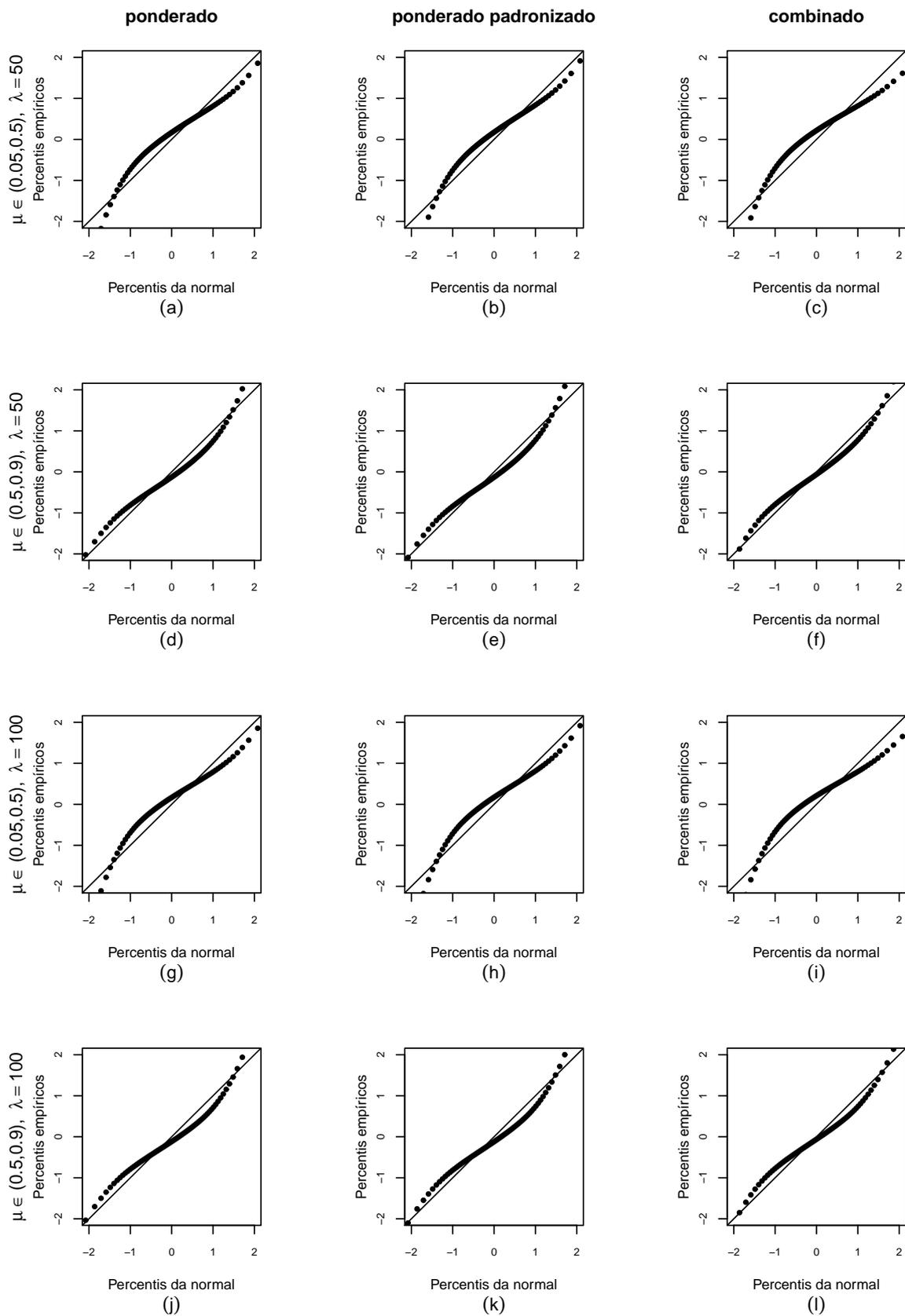


Figura 4.6: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, $x_t \sim U(0, 1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \dots, 80$. O modelo é estimado supondo dispersão constante.

com base na utilização dos quantis empíricos gerados a partir dos envelopes simulados. Mesmo sob a violação de alguma suposição, a distribuição do resíduo estimada a partir dos quantis gerados para a construção do envelope ainda é bem próxima da distribuição verdadeira. Assim, para a construção dos gráficos dos resíduos *versus* índices das observações ou *versus* covariadas, ou ainda *versus* o preditor linear poderíamos usar como ponto de corte $(\omega_{0.025}, \omega_{0.975})$, que são respectivamente os quantis 2.5% e 97.5% da distribuição empírica dos resíduos gerados para a construção do envelope. Para obter estes quantis, diferentemente do que é feito para a construção das bandas do envelope, inicialmente calculamos o valor médio para cada observação dos resíduos simulados e já ordenados. Em seguida, calculamos $\omega_{0.025}$ e $\omega_{0.975}$. O custo computacional é muito baixo, dado que os resíduos já foram simulados para a construção do envelope.

Devemos ressaltar que a investigação da distribuição do resíduo e a proximidade desta distribuição da distribuição normal padrão é bastante útil. Se sob especificação correta do modelo a distribuição do resíduo é bem aproximada pela normal padrão, o gráfico normal de probabilidades com envelopes simulados deve apontar para esta boa aproximação; caso contrário o problema não é do resíduo, mas sim do modelo postulado, que provavelmente está mal especificado.

A seguir apresentamos um exemplo simulado já utilizando como ponto de corte $(\omega_{0.025}, \omega_{0.975})$. O exemplo também tem como objetivo avaliar o comportamento dos resíduos do ponto de vista de diagnóstico quando a modelagem da dispersão é erroneamente desconsiderada.

Consideramos como modelo verdadeiro

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + \beta_2 x_{t2}, \\ \log(\phi_t) &= \gamma_1 + \gamma_2 z_{t2}, \quad t = 1, \dots, 40. \end{aligned}$$

Os valores da covariável x foram gerados através da distribuição uniforme: $x_t \sim U(0, 1)$, $t = 1, \dots, 40$ e definimos $z_{t2} = 0$ para $t = 1, \dots, 20$ e $z_{t2} = 1$ para $t = 21, \dots, 40$, fazendo com que haja dois grupos com precisão diferentes. Os valores dos parâmetros do submodelo da média são $\beta_1 = 0.8$ e $\beta_2 = -2.5$, fazendo com que $\mu \in (0.15, 0.7)$. Para o submodelo da precisão definimos quatro cenários diferentes para obter diferentes graus de heterogeneidade dos dados: $\lambda = 27$, dado que $\gamma = (5.7, -3.3)$, $\lambda = 66$, dado que $\gamma = (5.7, -4.2)$, $\lambda = 109$, dado que $\gamma = (5.7, -4.7)$ e $\lambda = 403$, dado que $\gamma = (5.4, -6.0)$. No entanto, para estimação dos parâmetros desconsideramos a dispersão variável dos dados e modelamos apenas a média da variável resposta.

Quando o ajuste é considerado bom, os gráficos dos resíduos *versus* índices das observações não apresentam padrão na distribuição dos pontos, que são distribuídos de forma aleatória em torno de 0 e dentro dos limites estabelecidos pelo gráfico. No entanto,

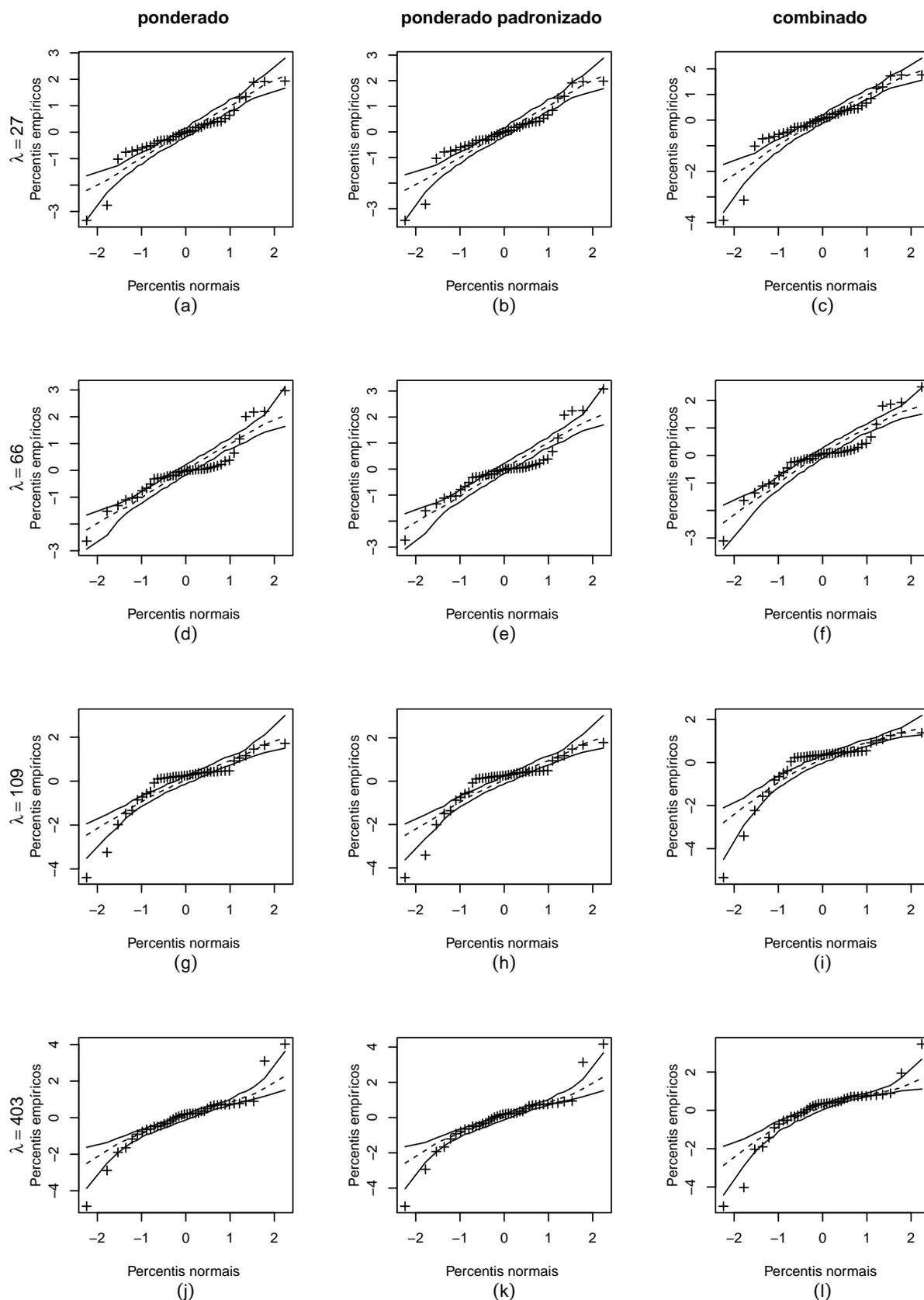


Figura 4.7: Gráficos normais de probabilidades com envelopes simulados. Modelo $\log[\mu_t/(\mu_t)] = \beta_1 + \beta_2 x_{t2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 40$, $z_{t2} = 0$ para $t = 1, \dots, 20$ e $z_{t2} = 1$ para $t = 21, \dots, 40$, $\mu \in (0.15, 0.7)$. O modelo é estimado supondo dispersão constante.

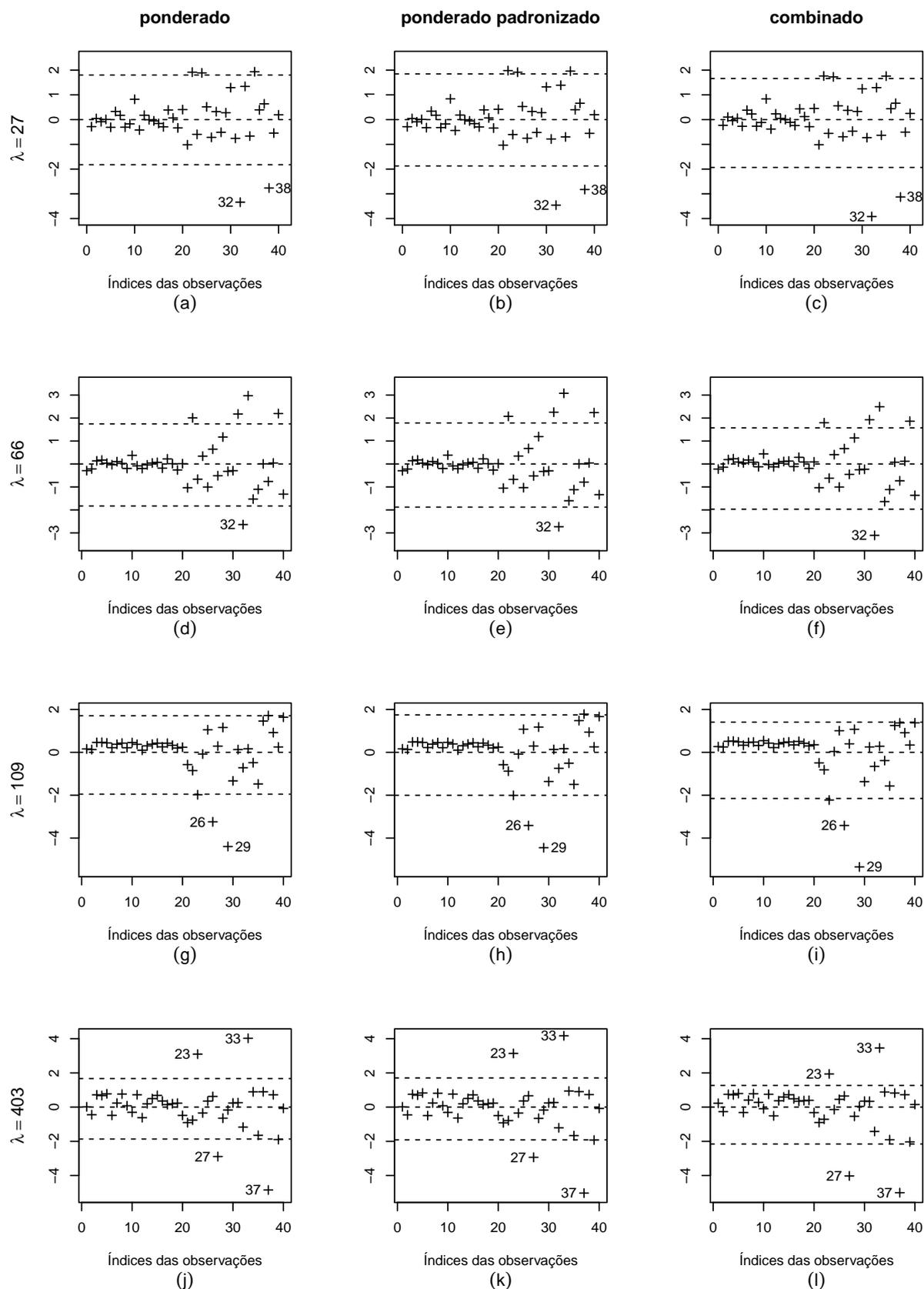


Figura 4.8: Gráficos de resíduos. Modelo $\log[\mu_t/(\mu_t)] = \beta_1 + \beta_2 x_{t2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2}$, $x_t \sim U(0, 1)$, $t = 1, \dots, 40$, $z_{t2} = 0$ para $t = 1, \dots, 20$ e $z_{t2} = 1$ para $t = 21, \dots, 40$, $\mu \in (0.15, 0.7)$. O modelo é estimado supondo dispersão constante.

nas Figuras 4.7 e 4.8 há um certo padrão na dispersão dos pontos nos gráficos. Na Figura 4.8 nas primeiras observações a amplitude dos resíduos é menor e ao longo das observações torna-se maior, o que evidencia a existência dos dois grupos com dispersão diferentes. Esse padrão se torna mais evidente à medida que λ aumenta. Nota-se que os gráficos dos diferentes resíduos apresentam o mesmo padrão, de onde se conclui que há má qualidade no ajuste do modelo. Na Figura 4.7, os gráficos de envelopes confirmam a falta de qualidade do ajuste, que também fica mais visível à medida que λ aumenta e o padrão apresentado indica uma provável dispersão variável nos dados.

Também é possível notar que o resíduo combinado destacou mais enfaticamente a ocorrência de pontos aberrantes na omissão da modelagem da dispersão, em especial quando o grau de heterogeneidade da dispersão se torna maior. Por exemplo, comparando com os outros resíduos, o resíduo combinado destacou mais a observação 32 como ponto aberrante no cenário em que $\lambda = 27$ (Figura 4.8-c); o mesmo ocorre com a observação 32 no cenário em que $\lambda = 66$ (Figura 4.8-f) e a observação 29 quando $\lambda = 109$ (Figura 4.8-i).

4.3 Aplicação

Os dados desta aplicação foram utilizados por Zerbinatti (2008) e estão relacionados a fornecimento de gás natural. O ideal é que o fornecimento de gás não gere custos adicionais às companhias de gás (casos em que há superestimação na demanda de gás) e nem que haja deficiência na qualidade do serviço de fornecimento (quando há subestimação). O fator de simultaneidade (ou seja, a relação percentual entre a potência verificada com que trabalham simultaneamente vários aparelhos servidos por um determinado trecho da tubulação, e a soma da capacidade máxima desses aparelhos) possui grande importância para o dimensionamento das redes de distribuição de gás. É a partir desse fator que se obtém a vazão máxima que provavelmente ocorre em um trecho da tubulação, tal que $Q_p = F \times Q_{max}$, em que Q_p é a vazão máxima provável de um trecho, F é o fator de simultaneidade e Q_{max} é a vazão máxima possível. O fator de simultaneidade é um número adimensional situado no intervalo $(0, 1)$, por isso interpreta-se a vazão máxima provável em um trecho como um percentual da vazão máxima possível. Ou seja, o fator de simultaneidade representa o percentual da vazão máxima possível que provavelmente se pratica em um determinado trecho de tubulação.

A variável resposta é o fator de simultaneidade e a variável explicativa é a potência computada. Zerbinatti (2008) concluiu que o modelo de regressão beta que melhor se ajustou aos dados foi

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 \log(x_t), \quad t = 1, \dots, 42,$$

com $y_t \sim B(\mu_t, \phi)$. As estimativas obtidas no ajuste desse modelo são apresentadas na Tabela 4.10, de onde se conclui que a covariável envolvida é significativa ao nível de significância de 5%.

Tabela 4.10: Estimativas para o modelo logito do gás.

parâmetro	estimativa	erro-padrão	p -valor
β_1	-1.712	0.067	0.000
β_2	-0.793	0.066	0.000
ϕ	4.3738	0.219	

Para a análise dos resíduos foram construídos os gráficos da Figura 4.9, já utilizando os quantis empíricos. Nos gráficos dos resíduos *versus* logaritmo da covariável e resíduos *versus* preditor linear observa-se uma certa concentração de pontos, padrão não esperado quando o ajuste do modelo é bom. Já os gráficos de envelopes simulados não indicam violação severa das suposições do modelo.

Como os gráficos dos resíduos não foram muito satisfatórios sugerimos outros modelos para serem investigados. A fim de selecionar o melhor modelo para os dados, além da análise de resíduos utilizaremos adicionalmente alguns critérios de seleção de modelos: o pseudo- R^2 (FERRARI; CRIBARI-NETO, 2004), denotado por R_p^2 e pertencente ao intervalo $(0, 1)$ sendo que valores mais próximos de 1 indicam melhor ajuste; o pseudo R^2 penalizado, \bar{R}_p^2 ; uma medida baseada em uma razão de verossimilhanças, R_{RV}^2 e essa medida penalizada, \bar{R}_{RV}^2 , que foram propostas por Bayer e Cribari-Neto (2014). Para o modelo utilizado por Zerbini (2008), os valores das medidas R_p^2 , \bar{R}_p^2 , R_{RV}^2 e \bar{R}_{RV}^2 foram 0.7241, 0.7099, 0.6910 e 0.6539, respectivamente.

Entre os modelos sugeridos, primeiro consideramos outra função de ligação para o modelo da média:

$$-\log(-\log(\mu_t)) = \beta_1 + \beta_2 \log(x_t), \quad t = 1, \dots, 42.$$

A Tabela 4.11 apresenta as estimativas para esse modelo. Os valores das medidas R_p^2 , \bar{R}_p^2 , R_{RV}^2 e \bar{R}_{RV}^2 foram 0.6905, 0.6746, 0.7218 e 0.68843, respectivamente. Os gráficos de resíduos e de envelopes simulados para cada um dos resíduos são apresentados na Figura 4.10, onde se observa um padrão similar aos gráficos quando se considerou a função de ligação logito (Figura 4.9), no entanto, agora há mais observações consideradas como pontos aberrantes.

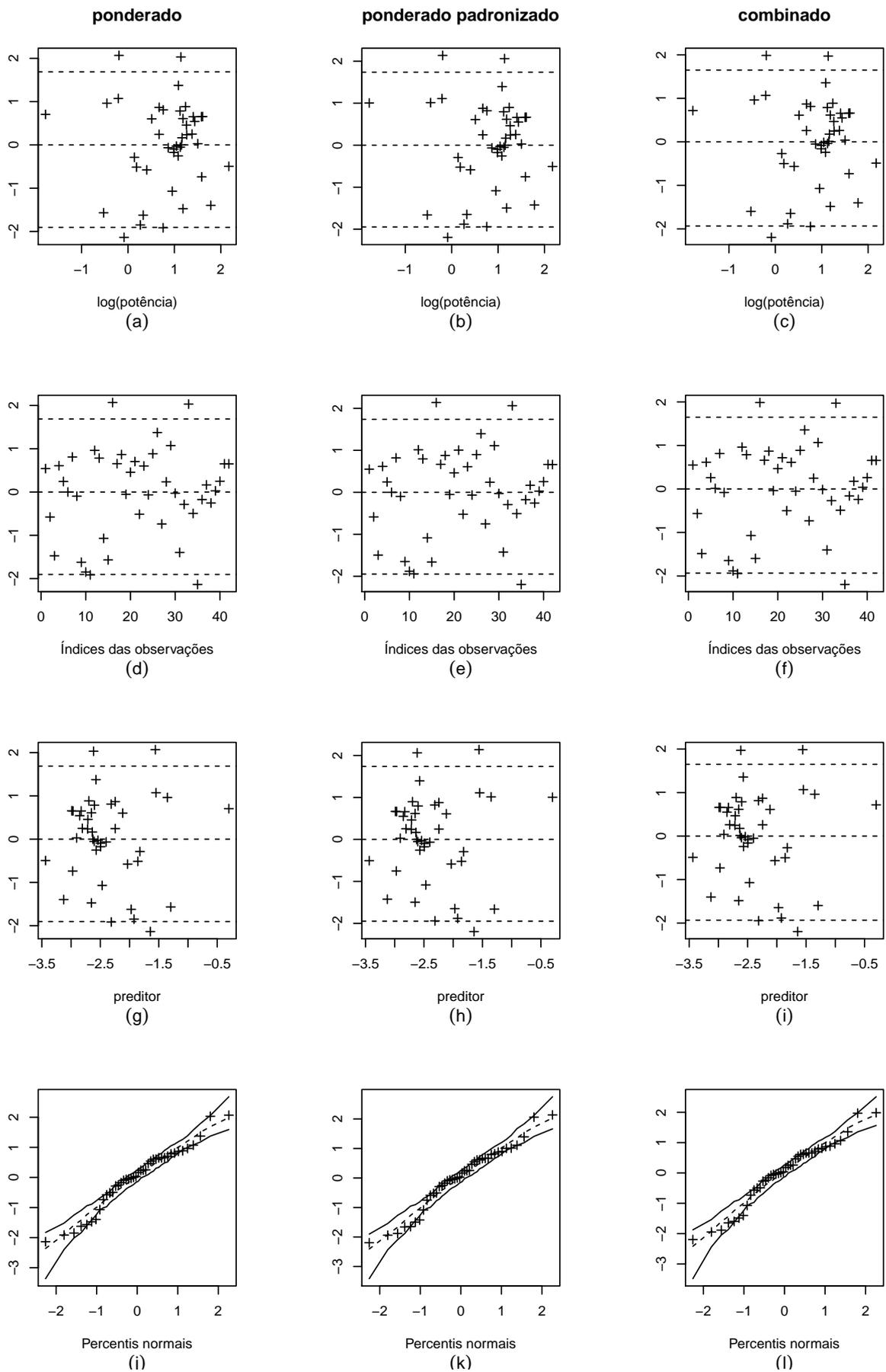


Figura 4.9: Gráficos de resíduos. Dados relacionados ao gás natural. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 \log(x_t)$, $t = 1, \dots, 42$.

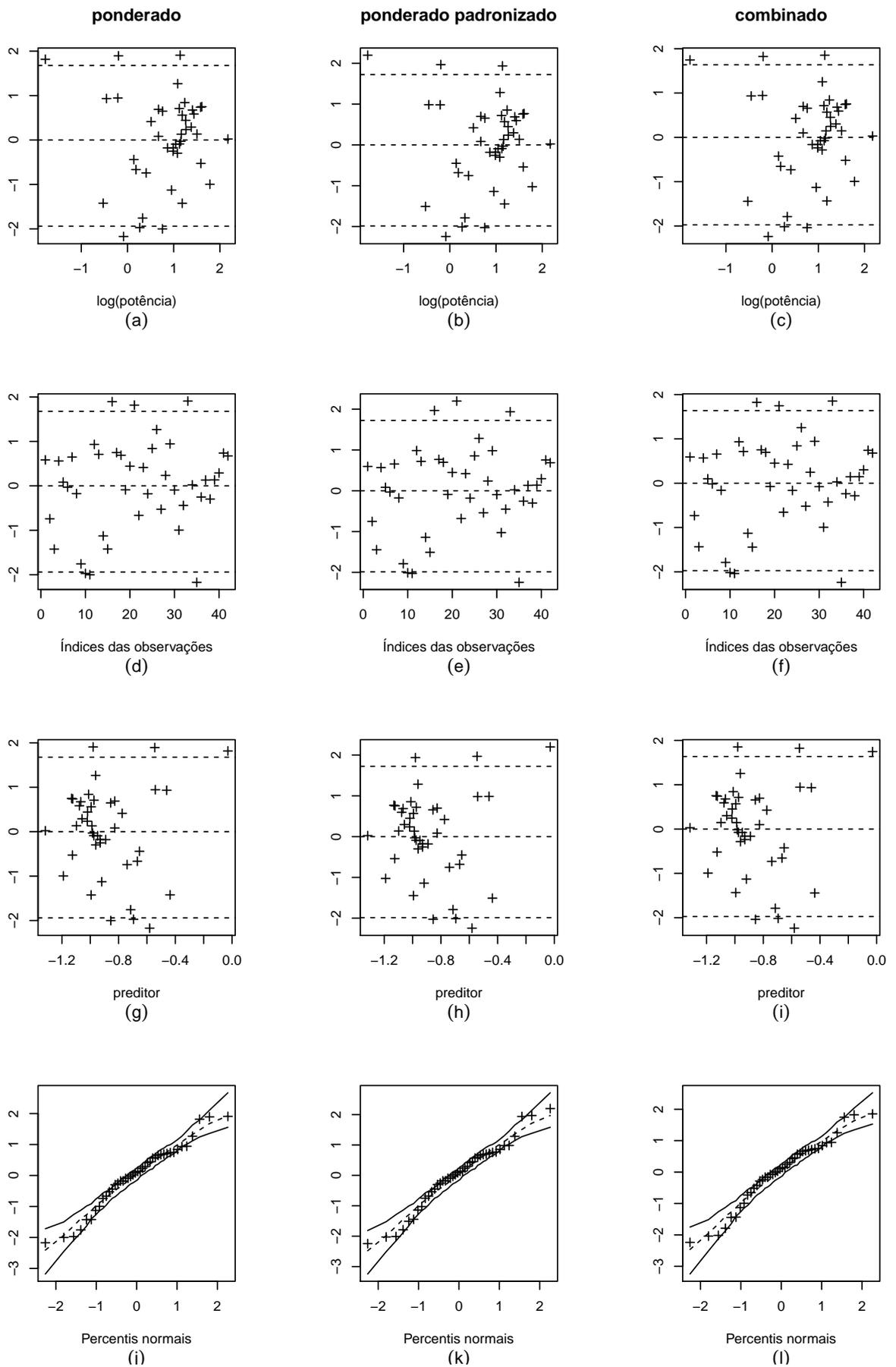


Figura 4.10: Gráficos de resíduos. Dados relacionados ao gás natural. Modelo estimado: $-\log(-\log(\mu_t)) = \beta_1 + \beta_2 \log(x_t)$, $t = 1, \dots, 42$.

Tabela 4.11: Estimativas para o modelo log-log do gás.

parâmetro	estimativa	erro-padrão	p -valor
β_1	-0.608	0.034	0.000
β_2	-0.325	0.032	0.000
ϕ	4.270	0.219	

Outros modelos a serem investigados são os que supõem dispersão variável. Primeiro consideramos

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 \log(x_t)$$

$$\log(\phi_t) = \gamma_1 + \gamma_2 \log(x_t), \quad t = 1, \dots, 42.$$

A Tabela 4.12 apresenta as estimativas dos parâmetros desse modelo, de onde se conclui que o logaritmo da covariável não é significativo para o modelo da dispersão quando se considera o nível de significância de 5%. Assim, o modelo se reduz a um modelo de dispersão constante.

Tabela 4.12: Estimativas para o modelo logito e log do gás.

parâmetro	estimativa	erro-padrão	p -valor
β_1	-1.717	0.090	0.000
β_2	-0.797	0.084	0.000
γ_1	4.001	0.326	0.000
γ_2	0.541	0.295	0.066
λ	8.509		

Por último, consideramos o modelo

$$-\log(-\log(\mu_t)) = \beta_1 + \beta_2 \log(x_t)$$

$$\log(\mu_t) = \gamma_1 + \gamma_2 \log(x_t), \quad t = 1, \dots, 42.$$

A Tabela 4.13 apresenta as estimativas dos parâmetros desse modelo, de onde se conclui que a covariável é significativa para ambos submodelos ao nível de significância de 5%. Os valores das medidas R_p^2 , \bar{R}_p^2 , R_{RV}^2 e \bar{R}_{RV}^2 são 0.7241, 0.7023, 0.7426 e 0.7069, respectivamente. Os gráficos de resíduos e de envelopes simulados para cada um dos resíduos são apresentados na Figura 4.11, onde se observa um padrão similar aos gráficos dos modelos anteriores, no entanto, agora há menos observações aberrantes.

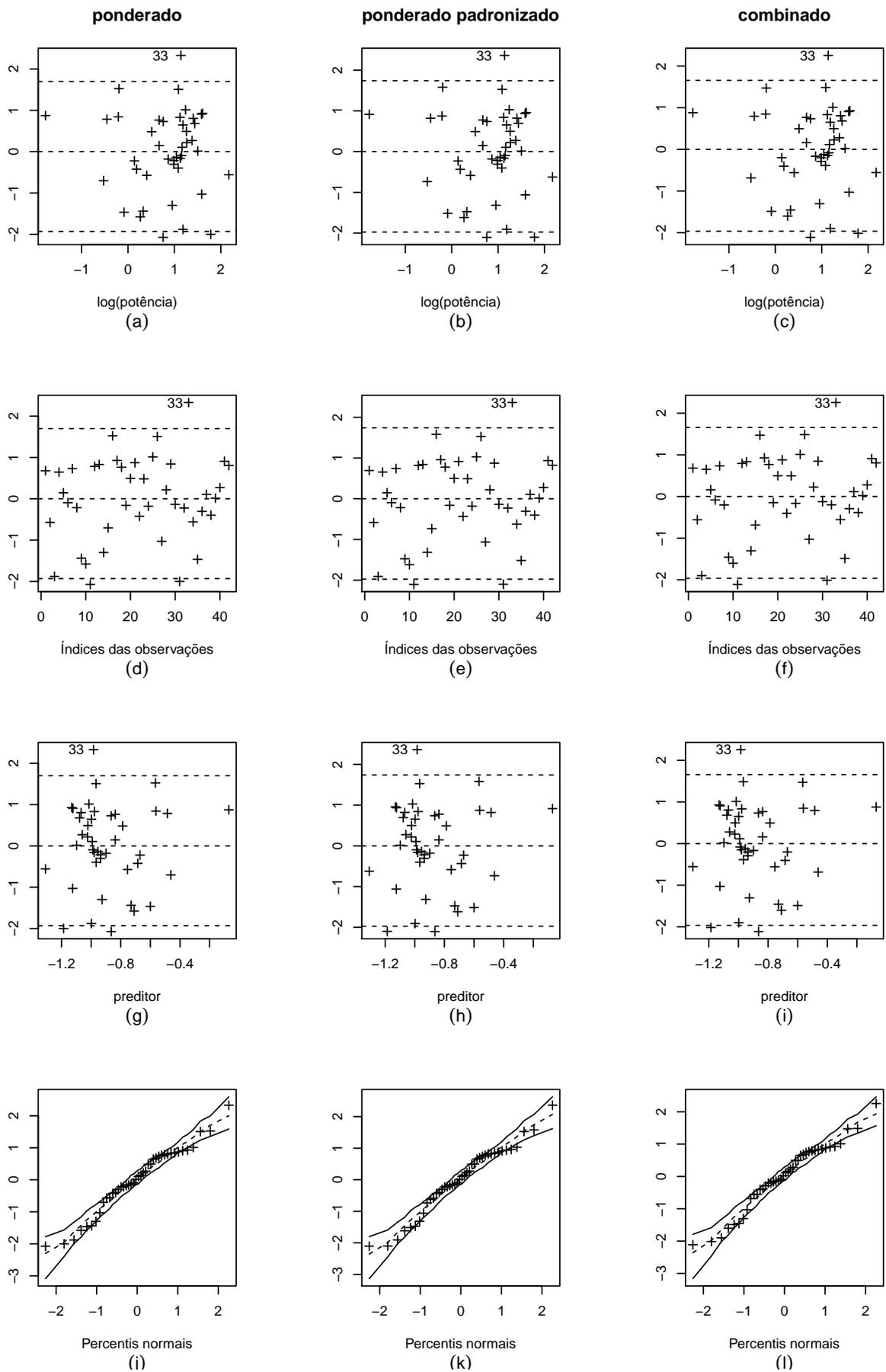


Figura 4.11: Gráficos de resíduos. Dados relacionados ao gás natural. Modelo estimado: $-\log(-\log(\mu_t)) = \beta_1 + \beta_2 \log(x_t)$ e $\log(\mu_t) = \gamma_1 + \gamma_2 \log(x_t)$, $t = 1, \dots, 42$.

Tabela 4.13: Estimativas para o modelo log-log e log do gás.

parâmetro	estimativa	erro-padrão	p -valor
β_1	-0.627	0.045	0.000
β_2	-0.313	0.038	0.000
γ_1	3.811	0.325	0.000
γ_2	0.772	0.294	0.008
λ	21.163		

Como através dos gráficos dos resíduos dos modelos considerados observa-se um padrão parecido, consideraremos adicionalmente as medidas de qualidade de ajuste para complementar a seleção do melhor modelo para os dados. A Tabela 4.14 contém todas essas medidas, de onde se conclui que o melhor modelo avaliado, em quase todos os critérios, é o que considera dispersão variável e função de ligação log-log e log para os modelos da média e dispersão, respectivamente. O modelo considerado possui duas das medidas de ajuste maiores e uma igual ao modelo com dispersão constante e com a função de ligação logito, modelo que foi utilizado por Zerbinatti (2008).

Tabela 4.14: Medidas de ajuste dos modelos competidores. Dados de gás natural.

Medidas	Modelos		
	logit	log-log	log-log e log
R_p^2	0.72410	0.69050	0.72410
\bar{R}_p^2	0.70995	0.67463	0.70232
R_{RV}^2	0.69106	0.72186	0.74266
\bar{R}_{RV}^2	0.65393	0.68843	0.70692

No gráfico de resíduos do modelo selecionado, Figura 4.11, a observação 33 é destacada como atípica. No entanto, ao reestimar o modelo sem esta observação concluiu-se que este ponto não é influente, pois a sua exclusão não alterou as conclusões dos resultados inferenciais.

Modelo de regressão beta não linear

Simas *et al.* (2010) propuseram uma extensão aos modelos citados anteriormente. Nesta extensão a média e o parâmetro de dispersão (ϕ^{-1}) são relacionados a uma estrutura de regressão não linear. Uma das características dos modelos não lineares é que eles são deduzidos a partir de suposições teóricas por pesquisadores de determinada área de estudo, tais como agricultura, biologia, econometria, engenharia, química, etc.

Sejam y_1, \dots, y_n variáveis independentes, em que cada y_n segue a distribuição (2.1). No modelo de regressão beta não linear a média e o parâmetro de precisão satisfazem as seguintes relações funcionais:

$$g(\mu_t) = \eta_{1t} = f_1(x_t^\top; \beta) \quad \text{e} \quad h(\phi_t) = \eta_{2t} = f_2(z_t^\top, \gamma), \quad (5.1)$$

em que $\beta = (\beta_1, \dots, \beta_k)^\top$ e $\gamma = (\gamma_1, \dots, \gamma_q)^\top$ são vetores de parâmetros desconhecidos, com $k + q < n$, η_{1t} e η_{2t} são os preditores, e $x_t^\top = (x_{t1}, \dots, x_{tk_1})$ e $z_t^\top = (z_{t1}, \dots, z_{tq_1})$ são observações em k_1 e q_1 covariáveis fixas e conhecidas, com $t = 1, \dots, n$, $k_1 \leq k$ e $q_1 \leq q$. Geralmente as covariáveis utilizadas para modelar a precisão é um subconjunto das que foram utilizadas na modelagem da média. As funções de ligação, $g(\cdot)$ e $h(\cdot)$, são estritamente monótonas e duas vezes diferenciáveis e $f(\cdot)$ é uma função contínua e diferenciável, tal que as matrizes de derivadas $J_1 = \partial\eta_1/\partial\beta$ e $J_2 = \partial\eta_2/\partial\gamma$ têm postos k e q , respectivamente.

O logaritmo da função de verossimilhança é dado pela Equação (3.1). O vetor escore é denotado por $(U_\beta(\beta, \gamma), U_\gamma(\beta, \gamma))^\top$. Os componentes do vetor escore para β , vetor de

dimensão k , são dados por

$$\begin{aligned} \frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} &= \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial}{\partial \mu_t} \ell_t(\mu_t, \phi_t) \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i}, \quad i = 1, \dots, k \\ &= \sum_{t=1}^n \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i}, \end{aligned} \quad (5.2)$$

em que

$$y_t^* = \log \left(\frac{y_t}{1 - y_t} \right) \quad \text{e} \quad \mu_t^* = \psi(\mu_t \phi_t) - \psi((1 - \mu_t) \phi_t).$$

Em forma matricial

$$U_\beta(\beta, \gamma) = J_1^\top \Phi T (y^* - \mu^*),$$

em que $J_1^\top = \partial \eta_1 / \partial \beta$ é uma matriz de ordem $n \times k$, $\Phi = \text{diag}\{\phi_1, \dots, \phi_n\}$, $T = \text{diag}\{1/g'(\mu_1), \dots, 1/g'(\mu_n)\}$.

Os componentes do vetor escore para γ , vetor de dimensão q , são dados por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j}, \quad j = 1, \dots, q, \quad (5.3)$$

em que

$$\frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} = \mu_t (y_t^* - \mu_t^*) + \log(1 - y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t) = a_t.$$

Além disso, $d\phi_t/d\eta_{2t} = 1/h'(\phi_t)$. Em forma matricial pode ser expressa por

$$U_\gamma(\beta, \gamma) = J_2^\top H a,$$

em que $J_2^\top = \partial \eta_2 / \partial \gamma$ é uma matriz de ordem $n \times q$, $H = \text{diag}\{1/h'(\phi_1), \dots, 1/h'(\phi_n)\}$ e $a = (a_1, \dots, a_n)^\top$.

Para a construção da matriz informação de Fisher é necessário obter as segundas derivadas de $\ell(\beta, \gamma)$ em relação aos parâmetros. A partir de (5.2), as segundas derivadas em relação a β_i e β_j , com $i, j = 1, \dots, k$, são

$$\begin{aligned} \frac{\partial \ell(\mu, \gamma)}{\partial \beta_i \partial \beta_j} &= \frac{\partial}{\partial \beta_j} \left(\sum_{t=1}^n \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} \right) \\ &= \sum_{t=1}^n \phi_t \left[\frac{\partial}{\partial \beta_j} (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} + (y_t^* - \mu_t^*) \frac{\partial}{\partial \beta_j} \left(\frac{1}{g'(\mu_t)} \right) \frac{\partial \eta_{1t}}{\partial \beta_i} \right] \\ &\quad + \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial}{\partial \beta_j} \left(\frac{\partial \eta_{1t}}{\partial \beta_i} \right) \\ &= \sum_{t=1}^n \phi_t \left\{ -q_t \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{\partial \eta_{1t}}{\partial \beta_j} + (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial^2 \eta_{1t}}{\partial \beta_i \partial \beta_j} \right\}, \end{aligned}$$

em que

$$q_t = \left\{ \phi_t [\psi'((1 - \mu_t) \phi_t) + \psi'(\mu_t \phi_t)] + (y_t^* - \mu_t^*) \frac{g''(\mu_t)}{g'(\mu_t)} \right\} \frac{1}{[g'(\mu_t)]^2}.$$

De (3.6), $E(y^* - \mu^*) = 0$, de onde segue que

$$\begin{aligned} E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \beta_j}\right) &= \sum_{t=1}^n E\left(\phi_t \left[-q_t \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{\partial \eta_{1t}}{\partial \beta_j}\right]\right) \\ &= -\sum_{t=1}^n \left(\phi_t \left[\phi_t v_t \frac{1}{\{g'(\mu_t)\}^2}\right] \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{\partial \eta_{1t}}{\partial \beta_j}\right), \end{aligned}$$

em que

$$v_t = \psi'((1 - \mu_t)\phi_t) + \psi'(\mu_t\phi_t).$$

Em notação matricial, temos que

$$E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta \partial \beta^\top}\right) = -J_1^\top \Phi W J_1,$$

em que $W = \text{diag}\{w_1, \dots, w_n\}$, com $w_t = \phi_t v_t \{1/[g'(\mu_t)]^2\}$, $t = 1, \dots, n$.

A partir de (5.3), a segunda derivada de $\ell(\beta, \gamma)$ com relação à γ é dada por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_l \partial \gamma_s} &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_s} \left(\frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_l} \right), \quad l, s = 1, \dots, q \\ &= \sum_{t=1}^n \left[\frac{\partial a_t}{\partial \gamma_l} \frac{1}{h'(\phi_t)} \frac{\partial \eta_{2t}}{\partial \gamma_s} + a_t \frac{\partial}{\partial \gamma_l} \left(\frac{1}{h'(\phi_t)} \right) \frac{\partial \eta_{2t}}{\partial \gamma_s} + a_t \frac{1}{h'(\phi_t)} \frac{\partial}{\partial \gamma_l} \left(\frac{\partial \eta_{2t}}{\partial \gamma_s} \right) \right] \\ &= \sum_{t=1}^n \left[r_t - a_t \frac{h''(\phi_t)}{h'(\phi_t)} \right] \frac{1}{h'(\phi_t)} \frac{\partial \eta_{2t}}{\partial \gamma_l} \frac{\partial \eta_{2t}}{\partial \gamma_s} + a_t \frac{1}{h'(\phi_t)} \frac{\partial^2 \eta_{2t}}{\partial \gamma_s \partial \gamma_l}, \end{aligned}$$

em que

$$r_t = -\mu_t^2 \psi'(\mu_t \phi_t) - (1 - \mu_t) \psi'((1 - \mu_t)\phi_t) + \psi'(\phi_t).$$

Como $E(\log(1 - y_t)) = -\psi(\phi_t) + \psi((1 - \mu_t)\phi_t)$, visto em (3.8), segue que $E(a_t) = 0$.

Assim,

$$E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_l \partial \gamma_s}\right) = -\sum_{t=1}^n \left[\mu_t^2 \psi'(\mu_t \phi_t) + (1 - \mu_t)^2 \psi'((1 - \mu_t)\phi_t) - \psi'(\phi_t) \right] \frac{1}{[h'(\phi_t)]^2} \frac{\partial \eta_{2t}}{\partial \gamma_l} \frac{\partial \eta_{2t}}{\partial \gamma_s},$$

que em notação matricial tem a forma

$$E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma \partial \gamma^\top}\right) = -J_2^\top D^* J_2,$$

em que $D^* = \text{diag}\{d_1^*, \dots, d_n^*\}$, com

$$d_t^* = \left[\mu_t^2 \psi'(\mu_t \phi_t) + (1 - \mu_t)^2 \psi'((1 - \mu_t)\phi_t) - \psi'(\phi_t) \right] \frac{1}{[h'(\phi_t)]^2}.$$

Derivando (5.2) com relação a γ_l , $l = 1, \dots, q$, tem-se que

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_l} &= \frac{\partial}{\partial \gamma_l} \left\{ \sum_{t=1}^n \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} \right\} \\ &= \sum_{t=1}^n \frac{\partial \phi_t}{\partial \gamma_l} (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} + \phi_t \frac{\partial}{\partial \gamma_l} (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} \\ &= -\sum_{t=1}^n [c_t - (y_t^* - \mu_t^*)] \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{1}{h'(\phi_t)} \frac{\partial \eta_{2t}}{\partial \gamma_l}, \end{aligned}$$

em que

$$c_t = \phi_t [\mu_t \psi'(\mu_t \phi_t) - (1 - \mu_t) \psi'((1 - \mu_t) \phi_t)].$$

Como $E(y_t^* - \mu_t^*) = 0$, tem-se que

$$E \left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_l} \right) = - \sum_{t=1}^n c_t \frac{1}{g'(\mu_t)} \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{1}{h'(\phi_t)} \frac{\partial \eta_{2t}}{\partial \gamma_l}.$$

Em forma matricial,

$$E \left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta \partial \gamma^\top} \right) = -J_1^\top C T H J_2,$$

em que $C = \text{diag}\{c_1, \dots, c_n\}$, $T = \text{diag}\{1/g'(\mu_1), \dots, 1/g'(\mu_n)\}$ e $H = \text{diag}\{1/h'(\phi_1), \dots, 1/h'(\phi_n)\}$.

Finalmente, a matriz de informação de Fisher é dada por

$$K = K(\beta, \gamma) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\gamma} \\ K_{\gamma\beta} & K_{\gamma\gamma} \end{pmatrix},$$

em que $K_{\beta\beta} = J_1^\top \Phi W J_1$, $K_{\beta\gamma} = K_{\gamma\beta}^\top = J_1^\top C T H J_2$ e $K_{\gamma\gamma} = J_2^\top D^* J_2$.

Sob certas condições de regularidade e para tamanhos de amostras grandes

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \sim N_{k+q} \left(\begin{pmatrix} \beta \\ \gamma \end{pmatrix}, K^{-1} \right),$$

aproximadamente, $\hat{\beta}$ e $\hat{\gamma}$ sendo os estimadores de máxima verossimilhança de β e γ , respectivamente. Esses estimadores não têm forma fechada, sendo necessário o uso de um algoritmo de otimização não linear, por isso necessário um chute inicial para cada estimador.

Quando $k_1 = k$ e $q = q_1$, uma proposta de chute inicial utiliza a linearização da função com base na expansão em série de Taylor. Seja $f(x_t, \beta)$ a função não linear. Sua expansão até a primeira ordem ao redor do ponto $\beta^{(0)}$ é

$$f(x_t, \beta) \approx f(x_t, \beta^{(0)}) + \sum_{i=1}^k \left[\frac{\partial f(x_t, \beta)}{\partial \beta_i} \right]_{\beta=\beta^{(0)}} (\beta_i - \beta_i^{(0)}),$$

em que $\beta^{(0)} = (\beta_1^{(0)}, \dots, \beta_k^{(0)})$ é um chute inicial. Assim,

$$f(x_t, \beta) = f(x_t, \beta^{(0)}) + \sum_{i=1}^k j_{ti}^{(0)} (\beta_i - \beta_i^{(0)}). \quad (5.4)$$

Analogamente ao que é feito no modelo normal não linear em que $y = f(x_t, \beta)$, aqui consideraremos $g(y_t) = f(x_t, \beta)$. Adicionalmente, supondo $\theta_i^{(0)} = (\beta_i - \beta_i^{(0)})$, segue que

$$g(y_t) - f(x_t, \beta^{(0)}) = \sum_{i=1}^k j_{ti}^{(0)} \theta_i^{(0)},$$

que pode ser visto como um modelo linear e o estimador de mínimos quadrados de $\theta^{(0)}$ será dado por

$$\hat{\theta}^{(0)} = (J^{(0)\top} J^{(0)})^{-1} J^{(0)\top} (g(y) - f(x, \beta^{(0)})), \quad (5.5)$$

em que $J^{(0)} = [\partial\eta_1/\partial\beta]_{\beta=\beta^{(0)}}$ e $\hat{\theta}_i^{(0)} = (\hat{\beta}_i - \beta_i^{(0)})$, então $\hat{\beta}_i^{(1)} = \hat{\theta}_i^{(0)} + \beta_i^{(0)}$. Essa relação pode ser generalizada para construir um processo iterativo tal que $\beta_i^{(m)} = \theta_i^{(m-1)} + \beta_i^{(m-1)}$. Generalizando (5.5) temos que

$$\beta^{(m)} = (J^{(m-1)\top} J^{(m-1)})^{-1} J^{(m-1)\top} (g(y) - f(x, \beta^{(m-1)})) + \beta^{(m-1)}.$$

Nossa proposta é considerar

$$\beta_{NL}^{(0)} = (J_1^{(0)\top} J_1^{(0)})^{-1} J_1^{(0)\top} (g(y) - f(x, \beta_L^{(0)})) \quad \text{e} \quad \beta_L^{(0)} = (X^\top X)^{-1} X^\top g(y).$$

Já para o modelo da precisão o mesmo raciocínio será utilizado, mas agora o parâmetro será γ e na equação equivalente a (5.4) considera-se $h(\phi_t) = f(z_t, \gamma)$, o que implica

$$\gamma_{NL}^{(0)} = \left(J_2^{(0)\top} J_2^{(0)} \right)^{-1} J_2^{(0)\top} \left(h(\phi_{NL}^{(0)}) - f(z, \gamma_L^{(0)}) \right) \quad \text{e} \quad \gamma_L^{(0)} = (Z^\top Z)^{-1} Z^\top h(\phi_L^{(0)}). \quad (5.6)$$

Aqui, $J_2^{(0)} = [\partial\eta_2/\partial\gamma]_{\gamma=\gamma^{(0)}}$,

$$\phi_{Lt}^{(0)} = \frac{\check{\mu}_{Lt}(1 - \check{\mu}_{Lt})}{\check{\sigma}_{Lt}^2} - 1,$$

com $\check{\mu}_{Lt} = g^{-1}(\hat{\eta}_{1L}) = g^{-1}(x_t^\top \beta_L^{(0)})$ e $\check{\sigma}_{Lt}^2 = \check{\epsilon}_L^\top \check{\epsilon}_L / [(n - k)g'(\check{\mu}_{Lt})^2]$ em que $\check{\epsilon}_L = g(y) - \check{\mu}_L$.

Finalmente, em (5.6), segue que

$$\phi_{NL}^{(0)} = \frac{\check{\mu}_{NLt}(1 - \check{\mu}_{NLt})}{\check{\sigma}_{NLt}^2} - 1,$$

em que $\check{\mu}_{NLt} = g^{-1}(f_1(x_t^\top; \beta_{NL}))$, $\check{\sigma}_{NLt}^2 = \check{\epsilon}_{NL}^\top \check{\epsilon}_{NL} / [(n - k)g'(\check{\mu}_{NLt})]$ com $\check{\epsilon}_{NL} = g(y) - \check{\mu}_{NL}$.

No entanto, em situações que $k_1 < k$ e/ou $q_1 < q$, ou seja, o número de covariadas é inferior ao número de parâmetros em pelo menos um dos submodelos, antes do chute inicial acima descrito é necessário fornecer chutes iniciais baseados na teoria (modelo matemático) ou nos dados. Um exemplo desse tipo é apresentado na seção de aplicação a dados reais deste capítulo.

5.1 Resíduos

Os resíduos para o modelo de regressão beta não linear são construídos de forma semelhante ao modelo linear. Nas seções seguintes além do resíduo combinado, serão apresentados os resíduos ponderado e ponderado padronizado, que Rocha e Simas (2011) generalizaram para o modelo de regressão beta não linear.

5.1.1 Resíduo ponderado

O processo iterativo Scoring de Fisher para estimar β é dado por

$$\beta^{(m+1)} = \beta^{(m)} + (J_1^\top \Phi^{(m)} W^{(m)} J_1)^{-1} J_1^\top \Phi^{(m)} T^{(m)} (y^* - \mu^{*(m)}).$$

Na forma de um processo iterativo de mínimos quadrados ponderados o processo é dado por

$$\beta^{(m+1)} = (J_1^\top \Phi^{(m)} W^{(m)} J_1)^{-1} J_1^\top \Phi^{(m)} W^{(m)} u_1^{(m)},$$

em que $u_1^{(m)} = J_1 \beta_1^{(m)} + W^{-1(m)} T^{(m)} (y^* - \mu^{*(m)})$.

Após convergência,

$$\hat{\beta} = (J_1^\top \hat{\Phi} \hat{W} J_1)^{-1} J_1^\top \hat{\Phi} \hat{W} u_1, \quad (5.7)$$

com

$$u_1 = J_1 \hat{\beta}_1 + \hat{W}^{-1} \hat{T} (y^* - \hat{\mu}^*),$$

que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de $\Phi^{1/2} W^{1/2} u_1$ em $\Phi^{1/2} W^{1/2} J_1$. O resíduo dessa regressão é dado por

$$r^\beta = \hat{\Phi}^{1/2} \hat{W}^{1/2} u_1 - \hat{\Phi}^{1/2} \hat{W}^{1/2} J_1 \hat{\beta} = \hat{\Phi}^{1/2} \hat{W}^{1/2} (u_1 - J_1 \hat{\beta}) = \hat{\Phi}^{1/2} \hat{W}^{-1/2} \hat{T} (y^* - \hat{\mu}^*), \quad (5.8)$$

ou seja,

$$r_t^\beta = \frac{y_t^* - \hat{\mu}_t^*}{\sqrt{\hat{v}_t}} \quad (5.9)$$

é o resíduo ponderado, em que

$$v_t = \psi'((1 - \mu_t) \phi_t) + \psi'(\mu_t \phi_t) = \text{Var}(y_t^*). \quad (5.10)$$

5.1.2 Resíduo ponderado padronizado

Para a padronização do resíduo (5.9) baseada na variância de u_1 , consideramos que $\text{Cov}(\hat{\beta}) \approx (J_1^\top \Phi W J_1)^{-1}$ e reescrevemos a Equação (5.7) como

$$(J_1^\top \Phi W J_1) \hat{\beta} = J_1^\top \Phi W u_1. \quad (5.11)$$

Considerando $\hat{W} \approx W$ e $\hat{\Phi} \approx \Phi$ em (5.11), segue que

$$\begin{aligned} (J_1^\top \Phi W J_1) \text{Cov}(\hat{\beta}) (J_1^\top \Phi W J_1)^\top &\approx (J_1^\top \Phi W) \text{Cov}(u_1) (J_1^\top \Phi W)^\top \\ (J_1^\top \Phi W J_1) (J_1^\top \Phi W J_1)^{-1} (J_1^\top \Phi W J_1) &\approx J_1^\top \Phi W \text{Cov}(u_1) W \Phi J_1 \\ \text{Cov}(u_1) &\approx W^{-1} \Phi^{-1}. \end{aligned}$$

Assim, $\widehat{\text{Cov}}(u_1) \approx \hat{W}^{-1} \hat{\Phi}^{-1}$.

Baseando-se em (5.7) o resíduo em (5.8) também pode ser expresso como

$$\begin{aligned} r^\beta &= \widehat{\Phi}^{1/2} \widehat{W}^{1/2} (u_1 - J_1 (J_1^\top \widehat{\Phi} \widehat{W} J_1)^{-1} J_1^\top \widehat{\Phi} \widehat{W} u_1) \\ &= (\widehat{\Phi}^{1/2} \widehat{W}^{1/2} - \widehat{\Phi}^{1/2} \widehat{W}^{1/2} J_1 (J_1^\top \widehat{\Phi} \widehat{W} J_1)^{-1} J_1^\top \widehat{\Phi} \widehat{W}) u_1. \end{aligned}$$

Então, considerando que $\text{Cov}(u_1) \approx W^{-1} \Phi^{-1}$, $\widehat{W} \approx W$ e $\widehat{\Phi} \approx \Phi$, segue que

$$\begin{aligned} \text{Cov}(r^\beta) &\approx (\Phi^{1/2} W^{1/2} - \Phi^{1/2} W^{1/2} J_1 (J_1^\top \Phi W J_1)^{-1} J_1^\top \Phi W) \text{Cov}(u_1) \\ &\quad \times (\Phi^{1/2} W^{1/2} - \Phi^{1/2} W^{1/2} J_1 (J_1^\top \Phi W J_1)^{-1} J_1^\top \Phi W)^\top \\ &\approx (I - \Phi^{1/2} W^{1/2} J_1 (J_1^\top \Phi W J_1)^{-1} J_1^\top \Phi^{1/2} W^{1/2}) \\ &\approx (I - G). \end{aligned}$$

Logo, $\widehat{\text{Cov}}(r^\beta) \approx (I - \widehat{G})$, com $\widehat{G} = \widehat{\Phi}^{1/2} \widehat{W}^{1/2} J_1 (J_1^\top \widehat{\Phi} \widehat{W} J_1)^{-1} J_1^\top \widehat{\Phi}^{1/2} \widehat{W}^{1/2}$, em que G é a matriz de projeção de $\Phi^{1/2} W^{1/2} u_1$ em $\Phi^{1/2} W^{1/2} J_1$ e é uma matriz simétrica e idempotente.

Assim, o resíduo ponderado padronizado é dado por

$$r_{pt}^\beta = \frac{r^{\beta t}}{\widehat{\text{Cov}}(r_t^\beta)} = \frac{y_t^* - \widehat{\mu}_t^*}{\sqrt{\widehat{v}_t(1 - \widehat{g}_{tt})}}, \quad (5.12)$$

em que g_{tt} é o t -ésimo elemento da matriz diagonal G e v_t está definido em (5.10).

5.1.3 Resíduo combinado

A construção do resíduo combinado é feita de forma similar ao modelo linear. O processo iterativo Scoring de Fisher para estimar γ é dado por

$$\gamma^{(m+1)} = \gamma^{(m)} + (J_2^\top D^{*(m)} J_2)^{-1} J_2^\top H^{(m)} a^{(m)},$$

em que m representa os passos necessários até a convergência do processo e $a = (a_1, \dots, a_n)$ com

$$a_t = \mu_t(y_t^* - \mu_t^*) + \log(1 - y_t) - \psi((1 - \mu_t)\phi_t) + \psi(\phi_t). \quad (5.13)$$

O processo iterativo de mínimos quadrados ponderados é expresso por

$$\gamma^{(m+1)} = (J_2^\top D^{*(m)} J_2)^{-1} J_2^\top D^{*(m)} u_2^{(m)},$$

em que $u_2^{(m)} = J_2 \gamma^{(m)} + D^{*(m)} H^{(m)} a^{(m)}$, com $\eta_2 = (\eta_{21}, \dots, \eta_{2n}) = J_2 \gamma$. Após convergência do processo temos que

$$\widehat{\gamma} = (J_2^\top \widehat{D}^* J_2)^{-1} J_2^\top \widehat{D}^* \widehat{u}_2, \quad (5.14)$$

com $\widehat{u}_2 = J_2 \widehat{\gamma} + \widehat{D}^{*-1} \widehat{H} \widehat{a}$, que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de $D^{*1/2} J_2$ em $D^{*1/2} u_2$, onde o resíduo ordinário é dado por

$$r^\gamma = \widehat{D}^{*1/2} u_2 - \widehat{D}^{*1/2} J_2 \widehat{\gamma} = \widehat{D}^{*1/2} (J_2 \widehat{\gamma} + \widehat{D}^* \widehat{H} \widehat{a} - J_2 \widehat{\gamma}) = (\widehat{D}^*)^{-1/2} \widehat{H} \widehat{a}. \quad (5.15)$$

Assim, o resíduo da t -ésima observação é

$$r_t^\gamma = \frac{a_t}{\sqrt{\hat{\xi}_t}}, \quad (5.16)$$

em que

$$\xi_t = \mu_t^2 \psi'(\mu_t \phi_t) + (1 - \mu_t)^2 \psi'((1 - \mu_t) \phi_t) - \psi'(\phi_t) = \text{Var}(a_t) \quad (5.17)$$

e a_t está definido em (5.13).

O resíduo combinado é construído com componentes dos resíduos obtidos através dos processos iterativos Scoring de Fisher de β e γ , sendo definido por

$$r_t^{\beta\gamma} = (y_t^* - \hat{\mu}_t^*) + \hat{a}_t. \quad (5.18)$$

A variância de (5.18) pode ser obtida através de (3.7), (3.9) e (3.10), ou seja,

$$\text{Var}(y_t^* + \mu_t^* y_t^* + \log(1 - y_t)) = (1 + \mu_t)^2 \psi'(\mu_t \phi_t) + \mu_t^2 \psi'((1 - \mu_t) \phi_t) - \psi'(\phi_t) = \zeta_t.$$

Desta forma, o resíduo combinado padronizado é definido por

$$r_t^{\beta\gamma} = \frac{(y_t^* - \hat{\mu}_t^*) + \hat{a}_t}{\sqrt{\hat{\zeta}_t}}, \quad (5.19)$$

em que $\hat{\zeta}_t$ é dado por ζ_t avaliado em $\hat{\mu}_t$ e $\hat{\phi}_t$.

Uma outra forma de padronizar o resíduo combinado é considerando a variância de u_1 e u_2 . Para encontrar a covariância de u_2 reescrevemos a Equação (5.14) como

$$(J_2^\top \hat{D}^* J_2) \hat{\gamma} = J_2^\top \hat{D}^* u_2. \quad (5.20)$$

Dado que $\text{Cov}(\hat{\gamma}) \approx (J_2^\top \hat{D}^* J_2)^{-1}$ e considerando $\hat{D}^* \approx D^*$ na Equação (5.20), segue que

$$\begin{aligned} (J_2^\top D^* J_2) \text{Cov}(\hat{\gamma}) (J_2^\top D^* J_2)^\top &\approx (J_2^\top \hat{D}^*) \text{Cov}(u_2) (J_2^\top \hat{D}^*)^\top \\ (J_2^\top D^* J_2) (J_2^\top \hat{D}^* J_2)^{-1} (J_2^\top D^* J_2)^\top &\approx (J_2^\top \hat{D}^*) \text{Cov}(u_2) (J_2^\top \hat{D}^*)^\top \\ \text{Cov}(u_2) &\approx D^{*-1}. \end{aligned}$$

Assim, $\widehat{\text{Cov}}(u_2) \approx \hat{D}^{*-1}$.

Baseando-se em (5.14), o resíduo em (5.15) pode ser expresso por

$$r^\gamma = (\hat{D}^{*1/2} - \hat{D}^{*1/2} J_2 (J_2^\top \hat{D}^* J_2)^{-1} J_2^\top \hat{D}^*) u_2.$$

Assim, considerando $\hat{D}^* \approx D^*$, segue que

$$\begin{aligned} \text{Cov}(r^\gamma) &\approx (D^{*1/2} - D^{*1/2} J_2 (J_2^\top D^* J_2)^{-1} J_2^\top D^*) \text{Cov}(u_2) \\ &\quad \times (D^{*1/2} - D^{*1/2} J_2 (J_2^\top D^* J_2)^{-1} J_2^\top D^*)^\top \\ &\approx (I - D^{*1/2} J_2 (J_2^\top D^* J_2)^{-1} J_2^\top D^{*1/2}) \\ &\approx (I - G^*). \end{aligned}$$

Logo, $\widehat{\text{Cov}}(r^\gamma) \approx (I - \widehat{G}^*)$, em que G^* é a matriz de projeção de $D^{*1/2}u_2$ em $D^{*1/2}J_2$ e é uma matriz simétrica e idempotente.

Portanto, a padronização do resíduo em (5.16) que leva em consideração a covariância de u_2 é dada por

$$r_{pt}^\gamma = \frac{r_t^\gamma}{\sqrt{\widehat{\text{Cov}}(r_t^\gamma)}} = \frac{a_t}{\sqrt{\widehat{\xi}_t(1 - \widehat{g}_{tt}^*)}}, \quad (5.21)$$

em que g_{tt}^* é o t -ésimo elemento da diagonal principal de G^* e ξ está definido em (5.17).

Assim, para a padronização do resíduo combinado que se baseia nas covariâncias de u_1 e u_2 temos que

$$\text{Var}(r_{pt}^\beta + r_{pt}^\gamma) = \text{Var}(r_{pt}^\beta) + \text{Var}(r_{pt}^\gamma) + 2\text{Cov}(r_{pt}^\beta, r_{pt}^\gamma),$$

com

$$\begin{aligned} \text{Cov}(\Phi^{1/2}W^{-1/2}T(y^* - \mu^*), D^{*-1/2}Ha) &= \Phi^{1/2}W^{-1/2}\text{Cov}[y^*, \mu^*y^* + \log(1 - y)]HD^{*-1/2} \\ &= \Phi^{1/2}W^{-1/2}[\text{Var}(y^*) + \text{Cov}(y^*, \log(1 - y))]HD^{*-1/2} \\ &= \Phi^{1/2}W^{-1/2}BHD^{*-1/2} = M \end{aligned}$$

em que $\text{Var}(y_t^*)$ e $\text{Cov}(y_t^*, \log(1 - y_t))$ são dados em (3.7) e (3.10), respectivamente, $B = \text{diag}\{b_1, \dots, b_n\}$, em que $b_{tt} = \mu_t^*v_t - \psi'((1 - \mu_t)\phi_t)$ e $M = \text{diag}\{m_1, \dots, m_n\}$, em que $m_{tt} = v_t^{-1/2}[\mu_t^*v_t - \psi'((1 - \mu_t)\phi_t)]\xi_t$, com $t = 1, \dots, n$.

Assim, temos que

$$\text{Var}(r_{pt}^\beta + r_{pt}^\gamma) \approx (1 - g_{tt}) + (1 - g_{tt}^*) + 2m_{tt}.$$

Logo, o resíduo combinado padronizado baseado nas covariâncias de u_1 e u_2 é dado por

$$r_{pt}^{\beta\gamma} = \frac{r_t^\beta + r_t^\gamma}{\sqrt{(1 - g_{tt}) + (1 - g_{tt}^*) + 2m_{tt}}}. \quad (5.22)$$

Neste trabalho não iremos investigar o comportamento deste último resíduo. Nosso objetivo é propor um novo resíduo para a classe de modelos de regressão beta que seja capaz de captar bem o processo gerador dos dados, assim como já faz o resíduo ponderado padronizado proposto por Espinheira *et al.* (2008), mas que não tenha o custo computacional do cálculo de matrizes de projeção.

5.2 Avaliação numérica

Para investigar as distribuições empíricas dos resíduos ponderado (5.9), ponderado padronizado (5.12) e combinado (5.19) nos modelos de regressão beta não lineares, consideramos uma forma de não linearidade e realizamos simulações de Monte Carlo com 5000

réplicas em diferentes cenários. A primeira análise foi baseada nas médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos. Para isso consideramos uma amostra de tamanho 20, em que y_1, \dots, y_{20} são variáveis aleatórias independentes e $y_t \sim B(\mu_t, \phi_t)$, $t = 1, \dots, 20$. A média da variável resposta e o parâmetro de dispersão satisfazem as relações funcionais

$$\begin{aligned} \log \frac{\mu_t}{1 - \mu_t} &= \beta_1 + x_t^{\beta_2}, \\ \log(\phi_t) &= \gamma_1 + z_t^{\gamma_2}. \end{aligned} \quad (5.23)$$

Os valores das covariáveis foram gerados de forma independente através da distribuição uniforme considerando diferentes intervalos: $x_t \sim U(0.1, 1.1)$, $x_t \sim U(0.4, 1.4)$, $z_t \sim U(0.4, 1.4)$ e $z_t \sim U(0.5, 1.5)$, o intervalo utilizado para a distribuição está especificado em cada tabela. Utilizamos três cenários diferentes para a média da variável resposta: $\mu \in (0.1, 0.2)$, dado que $\beta_1 = -2.4$ e $\beta_2 = 0.8$; $\mu \in (0.3, 0.8)$, dado que $\beta = (-1.0, 2.8)$ e $\mu \in (0.6, 0.9)$, dado que $\beta = (-0.5, -0.6)$. Para o modelo da dispersão consideramos dois cenários diferentes: $\lambda = 20$, dado que $\gamma = (3.0, 3.7)$ e $\lambda = 100$, dado que $\gamma = (0.8, 4.15)$.

Tabela 5.1: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(0.1, 1.1)$ e $z_t \sim U(0.4, 1.4)$, $t = 1, \dots, 20$, $\lambda = 20$ e $\mu \in (0.1, 0.2)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.028	0.028	0.031	0.953	0.962	0.951	-0.300	-0.300	-0.436	2.783	2.789	2.892
2	0.018	0.018	0.020	0.950	0.968	0.947	-0.332	-0.323	-0.461	2.722	2.723	2.831
3	0.007	0.007	0.010	0.953	0.979	0.952	-0.163	-0.162	-0.292	2.717	2.729	2.783
4	0.014	0.014	0.014	0.986	1.012	0.986	-0.155	-0.154	-0.260	2.723	2.724	2.775
5	-0.046	-0.049	-0.045	0.989	1.081	0.991	-0.103	-0.090	-0.201	2.779	2.679	2.812
6	-0.016	-0.016	-0.015	0.966	0.987	0.966	-0.247	-0.245	-0.338	2.828	2.834	2.896
7	0.031	0.033	0.029	1.122	1.159	1.121	-0.132	-0.131	-0.183	2.489	2.471	2.509
8	0.000	-0.000	0.005	0.931	0.963	0.930	-0.151	-0.147	-0.304	2.706	2.707	2.795
9	0.055	0.062	0.055	0.983	1.177	0.982	-0.094	-0.101	-0.115	1.962	1.900	1.968
10	-0.031	-0.033	-0.032	1.032	1.063	1.032	-0.124	-0.128	-0.198	2.660	2.653	2.679
11	-0.003	0.000	0.000	0.948	0.978	0.945	-0.192	-0.166	-0.377	2.703	2.692	2.818
12	-0.122	-0.129	-0.123	1.104	1.172	1.105	-0.019	-0.013	-0.057	2.430	2.391	2.429
13	0.034	0.035	0.032	1.061	1.083	1.060	-0.151	-0.152	-0.235	2.801	2.793	2.838
14	0.008	0.008	0.006	1.046	1.064	1.046	-0.130	-0.130	-0.217	2.648	2.642	2.676
15	-0.024	-0.031	-0.021	0.700	1.094	0.700	-0.025	-0.001	-0.067	3.158	2.505	3.168
16	0.046	0.046	0.049	0.948	0.956	0.944	-0.234	-0.236	-0.368	2.744	2.751	2.835
17	0.030	0.033	0.030	1.033	1.181	1.033	-0.071	-0.074	-0.092	1.895	1.885	1.897
18	0.037	0.036	0.039	0.966	0.980	0.962	-0.283	-0.283	-0.433	2.668	2.677	2.788
19	0.030	0.031	0.027	1.112	1.138	1.112	-0.166	-0.164	-0.237	2.757	2.732	2.792
20	-0.051	-0.053	-0.053	1.072	1.113	1.073	-0.12	-0.118	-0.185	2.795	2.758	2.824

A partir das Tabelas 5.1 a 5.6 observa-se que as médias dos resíduos são bem semelhantes e próximas de zero, a similaridade dos valores também ocorre entre os erros-padrão, que são bem próximos de um. O resíduo combinado apresenta assimetria negativa, o que também ocorre em alguns casos para os outros resíduos, com exceção dos cenários em que a média da variável resposta pertence ao intervalo (0.6, 0.9) (Tabelas 5.5 e 5.6). O resíduo combinado apresentou-se mais assimétrico quando comparado aos outros resíduos, mas em

geral possui valores de assimetria próximos de zero. Com relação à curtose, os resíduos apresentaram valores semelhantes e, em geral, próximos de três. Portanto, os valores das medidas utilizadas são de maneira geral próximos ao que caracteriza a distribuição normal. Vale ressaltar que o tamanho da amostra utilizado é consideravelmente pequeno.

Tabela 5.2: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(0.1, 1.1)$ e $z_t \sim U(0.5, 1.5)$, $t = 1, \dots, 20$, $\lambda = 100$ e $\mu \in (0.1, 0.2)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.046	0.046	0.052	0.870	0.874	0.861	-0.739	-0.739	-0.987	3.618	3.618	4.005
2	0.161	0.161	0.166	0.754	0.760	0.740	-0.813	-0.816	-1.025	3.624	3.639	3.969
3	0.010	0.010	0.012	0.964	0.977	0.959	-0.427	-0.427	-0.698	3.014	3.017	3.352
4	-0.014	-0.014	-0.016	1.026	1.042	1.025	-0.309	-0.310	-0.528	2.922	2.922	3.116
5	-0.099	-0.110	-0.100	1.021	1.096	1.025	-0.158	-0.150	-0.345	2.919	2.804	3.055
6	-0.028	-0.029	-0.029	0.999	1.013	0.997	-0.497	-0.497	-0.675	3.136	3.141	3.355
7	0.019	0.020	0.015	1.129	1.164	1.127	-0.154	-0.155	-0.238	2.459	2.452	2.500
8	0.011	0.011	0.019	0.904	0.920	0.898	-0.469	-0.463	-0.793	3.101	3.104	3.519
9	0.069	0.082	0.070	0.944	1.169	0.943	-0.112	-0.123	-0.139	1.994	1.952	2.000
10	-0.045	-0.046	-0.048	1.067	1.089	1.068	-0.361	-0.361	-0.509	3.044	3.037	3.195
11	0.019	0.019	0.027	0.885	0.892	0.873	-0.638	-0.636	-0.983	3.365	3.362	3.865
12	-0.172	-0.183	-0.175	1.096	1.181	1.099	-0.072	-0.050	-0.131	2.362	2.324	2.371
13	-0.003	-0.003	-0.009	1.112	1.129	1.112	-0.319	-0.320	-0.471	2.946	2.938	3.084
14	-0.003	-0.003	-0.009	1.078	1.091	1.076	-0.252	-0.252	-0.432	2.895	2.889	3.030
15	-0.006	-0.001	-0.001	0.607	1.081	0.607	-0.019	0.007	-0.069	3.268	2.476	3.273
16	0.022	0.022	0.027	0.913	0.917	0.905	-0.600	-0.600	-0.868	3.346	3.349	3.722
17	0.073	0.083	0.073	1.010	1.180	1.009	-0.129	-0.132	-0.159	1.955	1.943	1.965
18	0.041	0.041	0.046	0.874	0.877	0.864	-0.806	-0.804	-1.052	3.716	3.718	4.116
19	-0.002	-0.002	-0.007	1.123	1.146	1.122	-0.205	-0.205	-0.321	2.669	2.660	2.735
20	-0.093	-0.097	-0.097	1.105	1.142	1.108	-0.265	-0.265	-0.373	2.807	2.774	2.881

Tabela 5.3: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(0.4, 1.4)$ e $z_t \sim U(0.4, 1.4)$, $t = 1, \dots, 20$, $\lambda = 20$ e $\mu \in (0.3, 0.8)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	-0.010	-0.010	-0.002	0.949	0.958	0.949	-0.046	-0.046	-0.251	2.655	2.662	2.721
2	0.010	0.011	0.015	0.964	0.978	0.961	-0.106	-0.101	-0.310	2.520	2.529	2.610
3	-0.016	-0.016	-0.012	0.942	0.970	0.943	0.112	0.112	-0.029	2.678	2.694	2.671
4	0.006	0.007	0.007	0.985	1.012	0.985	0.067	0.067	-0.055	2.652	2.652	2.647
5	-0.015	-0.017	-0.015	0.986	1.081	0.987	0.087	0.079	-0.001	2.728	2.626	2.720
6	0.012	0.012	0.013	0.974	0.989	0.973	-0.078	-0.080	-0.213	2.700	2.710	2.758
7	-0.012	-0.013	-0.016	1.145	1.186	1.145	0.015	0.016	-0.056	2.491	2.474	2.495
8	-0.010	-0.009	-0.004	0.915	0.950	0.917	0.163	0.160	0.014	2.700	2.716	2.670
9	0.014	0.018	0.015	0.990	1.205	0.990	-0.028	-0.026	-0.057	1.912	1.841	1.914
10	-0.031	-0.031	-0.032	1.033	1.055	1.033	-0.017	-0.015	-0.130	2.684	2.677	2.687
11	-0.004	-0.003	0.000	0.945	0.970	0.946	0.187	0.176	-0.005	2.615	2.597	2.579
12	0.002	0.002	-0.000	1.160	1.204	1.160	-0.047	-0.049	-0.104	2.348	2.337	2.357
13	-0.006	-0.007	-0.009	1.062	1.085	1.062	-0.035	-0.033	-0.146	2.681	2.664	2.706
14	0.014	0.014	0.012	1.039	1.056	1.039	0.011	0.012	-0.109	2.707	2.702	2.722
15	0.004	0.012	0.008	0.653	1.072	0.653	-0.006	-0.000	-0.058	3.481	2.511	3.491
16	0.013	0.013	0.020	0.942	0.950	0.941	-0.004	-0.004	-0.194	2.658	2.665	2.703
17	0.003	0.003	0.003	1.032	1.192	1.032	-0.001	-0.001	-0.031	1.889	1.876	1.891
18	-0.006	-0.006	-0.000	0.970	0.983	0.970	-0.026	-0.026	-0.256	2.568	2.580	2.653
19	-0.011	-0.011	-0.014	1.095	1.122	1.095	0.003	0.002	-0.088	2.636	2.613	2.649
20	-0.007	-0.007	-0.010	1.087	1.114	1.087	-0.054	-0.054	-0.151	2.740	2.717	2.765

Tabela 5.4: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(0.4, 1.4)$ e $z_t \sim U(0.5, 1.5)$, $t = 1, \dots, 20$, $\lambda = 100$ e $\mu \in (0.3, 0.8)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	0.022	0.022	0.040	0.910	0.913	0.904	-0.195	-0.194	-0.746	2.877	2.878	3.470
2	0.047	0.047	0.064	0.875	0.879	0.858	-0.360	-0.358	-0.883	3.019	3.027	3.650
3	-0.010	-0.010	-0.003	0.937	0.951	0.941	0.259	0.257	-0.073	2.803	2.805	2.737
4	-0.004	-0.004	-0.007	1.026	1.042	1.028	0.154	0.153	-0.126	2.796	2.792	2.800
5	0.017	0.017	0.015	1.029	1.096	1.027	0.263	0.246	0.083	2.913	2.827	2.843
6	-0.018	-0.018	-0.018	1.001	1.009	1.001	-0.183	-0.184	-0.486	2.834	2.836	3.113
7	0.019	0.019	0.011	1.143	1.181	1.143	-0.043	-0.043	-0.159	2.381	2.376	2.408
8	-0.065	-0.066	-0.054	0.865	0.881	0.877	0.514	0.513	0.158	3.100	3.116	2.909
9	-0.002	-0.004	-0.001	0.941	1.197	0.941	0.005	0.004	-0.031	1.982	1.921	1.984
10	-0.025	-0.025	-0.031	1.073	1.088	1.074	-0.108	-0.106	-0.333	2.774	2.765	2.875
11	-0.050	-0.050	-0.036	0.859	0.867	0.873	0.588	0.587	0.082	3.274	3.271	3.050
12	-0.023	-0.024	-0.029	1.152	1.194	1.152	-0.042	-0.043	-0.134	2.298	2.294	2.313
13	0.016	0.016	0.007	1.112	1.129	1.111	-0.039	-0.037	-0.252	2.686	2.678	2.755
14	0.004	0.004	-0.003	1.088	1.099	1.087	0.051	0.050	-0.191	2.742	2.737	2.775
15	0.009	0.019	0.016	0.514	1.072	0.514	-0.025	-0.063	-0.094	4.056	2.624	4.079
16	0.012	0.012	0.027	0.923	0.926	0.918	0.031	0.030	-0.464	2.735	2.737	2.991
17	0.010	0.012	0.009	1.017	1.210	1.017	-0.020	-0.020	-0.059	1.895	1.885	1.899
18	0.001	0.001	0.023	0.903	0.906	0.897	-0.118	-0.117	-0.781	2.938	2.940	3.543
19	0.005	0.005	-0.003	1.142	1.165	1.142	-0.022	-0.023	-0.182	2.566	2.557	2.589
20	-0.022	-0.023	-0.031	1.138	1.159	1.137	-0.076	-0.076	-0.248	2.685	2.672	2.738

Tabela 5.5: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta,\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(0.1, 1.1)$ e $z_t \sim U(0.4, 1.4)$, $t = 1, \dots, 20$, $\lambda = 20$ e $\mu \in (0.6, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$	r^β	r_p^β	$r^{\beta,\gamma}$
1	-0.015	-0.015	-0.016	1.030	1.036	1.032	0.147	0.146	-0.052	2.869	2.862	2.825
2	-0.052	-0.051	-0.050	0.948	1.030	0.954	0.226	0.219	0.046	2.730	2.870	2.681
3	-0.010	-0.010	-0.008	0.975	0.990	0.976	0.062	0.061	-0.102	2.682	2.683	2.688
4	0.002	0.002	0.002	0.994	1.015	0.994	0.053	0.056	-0.077	2.655	2.659	2.657
5	-0.011	-0.011	-0.012	1.029	1.061	1.029	0.033	0.033	-0.083	2.675	2.669	2.678
6	-0.023	-0.024	-0.023	0.973	1.014	0.974	0.195	0.195	0.087	2.692	2.697	2.645
7	0.006	0.006	0.005	1.083	1.117	1.083	0.071	0.070	-0.004	2.566	2.560	2.560
8	0.013	0.013	0.012	1.019	1.031	1.019	0.032	0.032	-0.175	2.730	2.721	2.794
9	-0.000	-0.000	0.002	1.000	1.120	1.000	0.026	0.022	-0.008	2.015	1.997	2.015
10	-0.001	-0.001	-0.001	1.020	1.059	1.020	0.134	0.136	0.039	2.707	2.707	2.687
11	0.042	0.049	0.044	0.961	1.060	0.959	0.038	0.044	-0.197	2.686	2.830	2.738
12	0.017	0.022	0.019	0.794	1.114	0.790	0.132	0.077	0.072	2.630	2.532	2.564
13	-0.017	-0.017	-0.017	1.031	1.050	1.031	0.062	0.063	-0.045	2.763	2.758	2.763
14	-0.010	-0.010	-0.010	1.017	1.038	1.017	0.067	0.066	-0.050	2.733	2.728	2.734
15	-0.016	-0.017	-0.013	0.892	1.105	0.892	-0.001	0.000	-0.040	2.195	2.138	2.197
16	-0.003	-0.002	-0.002	1.003	1.011	1.003	0.088	0.088	-0.099	2.710	2.709	2.721
17	-0.011	-0.011	-0.009	1.026	1.120	1.026	0.022	0.023	-0.015	2.050	2.049	2.051
18	0.000	0.000	-0.001	1.070	1.084	1.072	0.083	0.081	-0.087	2.302	2.313	2.306
19	-0.017	-0.017	-0.018	1.071	1.098	1.072	0.046	0.046	-0.048	2.722	2.711	2.723
20	0.011	0.012	0.011	1.001	1.101	1.001	0.112	0.109	0.040	2.716	2.686	2.697

A próxima análise será gráfica e o objetivo é comparar os quantis empíricos dos resíduos com os quantis teóricos da distribuição normal padrão. Para isso consideramos uma amostra de tamanho 80. Para a construção dos gráficos nas Figuras 5.1 e 5.2 consideramos o modelo em (5.23) e os mesmos cenários anteriores. Para compor o novo tamanho amostral foi considerada a mesma amostra das covariadas anteriores, no entanto elas foram replicadas quatro vezes a fim de manter o mesmo grau de heterogeneidade dos dados. Nas

Tabela 5.6: Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^β), ponderado padronizado (r_p^β) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(0.1, 1.1)$ e $z_t \sim U(0.5, 1.5)$, $t = 1, \dots, 20$, $\lambda = 100$ e $\mu \in (0.6, 0.9)$.

t	Média			Erro-padrão			Assimetria			Curtose		
	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$	r^β	r_p^β	$r^{\beta\gamma}$
1	-0.064	-0.064	-0.062	0.971	0.973	0.984	0.461	0.460	-0.083	3.367	3.366	3.090
2	-0.376	-0.379	-0.375	0.714	0.751	0.775	0.618	0.654	0.124	3.464	3.873	3.061
3	0.107	0.108	0.104	1.006	1.013	0.992	0.126	0.125	-0.269	2.753	2.753	2.837
4	0.044	0.045	0.039	1.043	1.055	1.040	0.078	0.078	-0.195	2.663	2.663	2.708
5	0.048	0.048	0.041	1.080	1.104	1.076	0.064	0.061	-0.169	2.732	2.723	2.761
6	-0.096	-0.099	-0.097	0.961	0.988	0.970	0.376	0.374	0.138	2.941	2.944	2.804
7	0.016	0.016	0.012	1.105	1.136	1.104	0.094	0.094	-0.022	2.466	2.463	2.464
8	0.178	0.178	0.182	0.966	0.972	0.936	0.114	0.112	-0.403	2.851	2.852	3.093
9	0.002	0.002	0.004	0.979	1.129	0.979	0.005	0.005	-0.036	2.010	1.998	2.011
10	-0.035	-0.036	-0.039	1.054	1.085	1.058	0.251	0.250	0.056	2.883	2.876	2.795
11	0.255	0.257	0.264	0.912	0.951	0.855	0.177	0.108	-0.475	2.851	2.952	3.225
12	0.096	0.142	0.102	0.731	1.098	0.727	0.049	0.006	-0.019	2.566	2.512	2.526
13	-0.007	-0.008	-0.013	1.085	1.099	1.086	0.111	0.110	-0.090	2.759	2.754	2.750
14	0.012	0.012	0.008	1.057	1.072	1.057	0.107	0.107	-0.129	2.817	2.811	2.833
15	-0.095	-0.122	-0.091	0.850	1.129	0.852	0.056	0.067	0.012	2.145	2.086	2.141
16	0.016	0.016	0.016	0.991	0.995	0.991	0.240	0.238	-0.228	2.832	2.832	2.868
17	0.009	0.011	0.010	1.026	1.142	1.026	0.020	0.021	-0.026	2.011	2.008	2.011
18	-0.034	-0.034	-0.030	0.970	0.978	0.986	0.382	0.375	-0.200	3.001	3.012	2.931
19	-0.003	-0.003	-0.008	1.101	1.122	1.101	0.100	0.099	-0.054	2.623	2.617	2.621
20	0.035	0.039	0.035	1.019	1.111	1.018	0.183	0.185	0.057	2.745	2.728	2.722

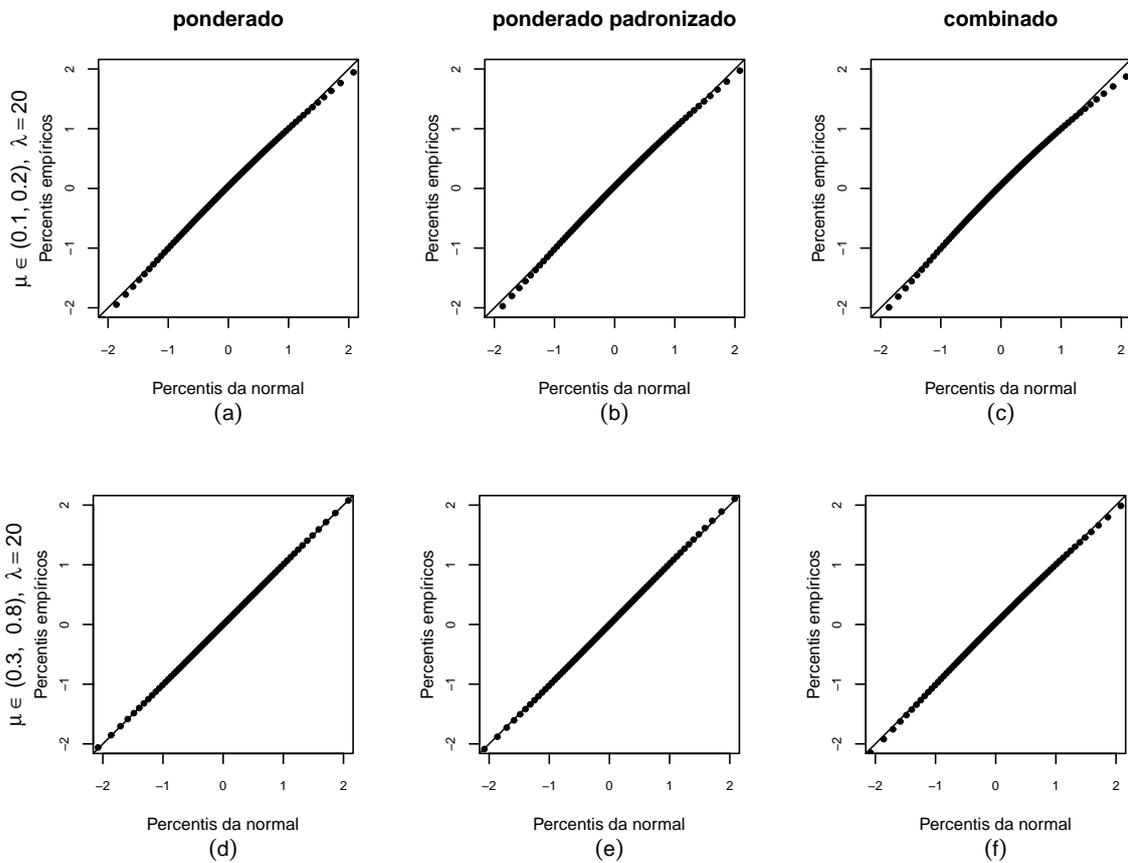


Figura 5.1: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $t = 1, \dots, 80$.

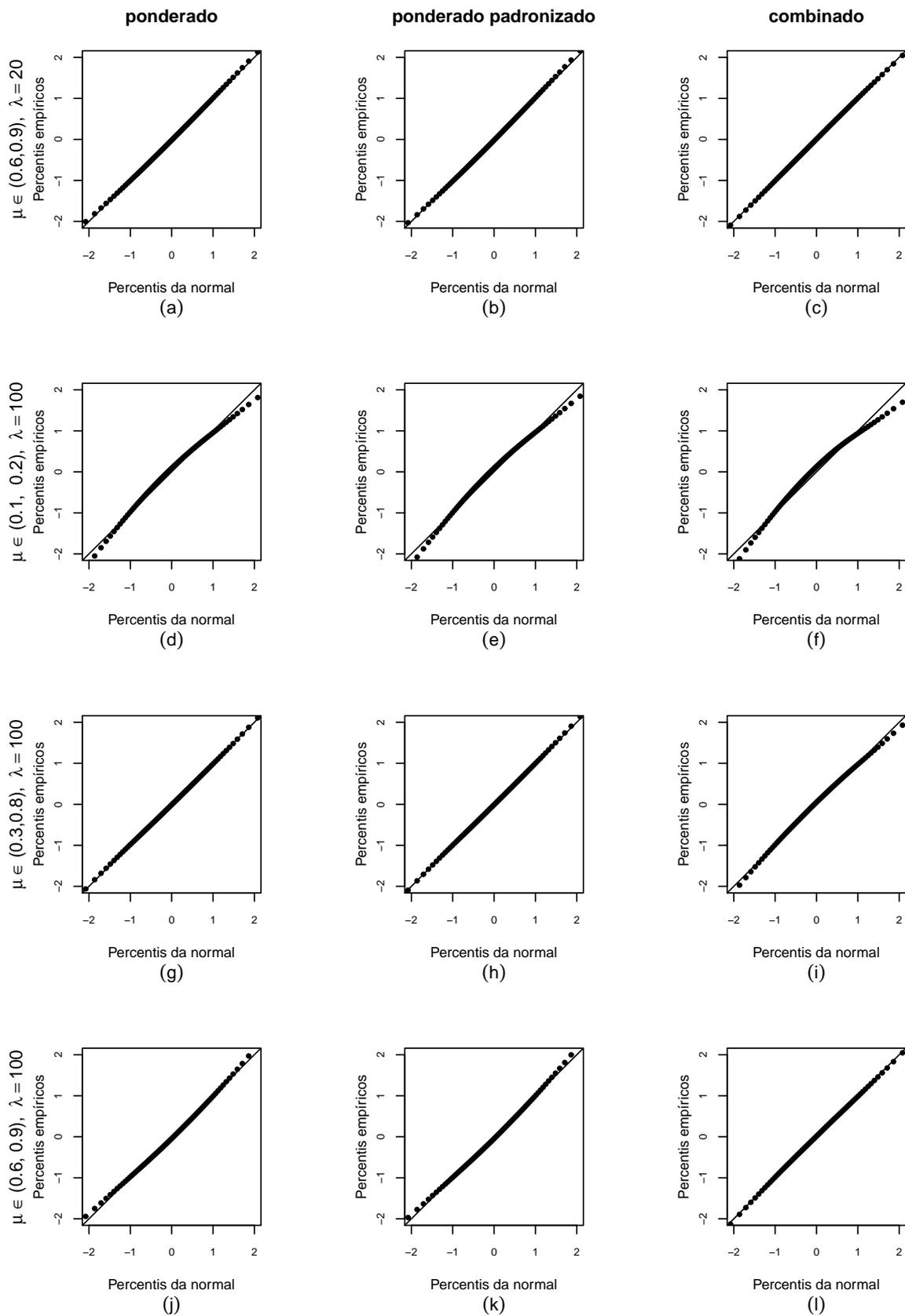


Figura 5.2: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $t = 1, \dots, 80$.

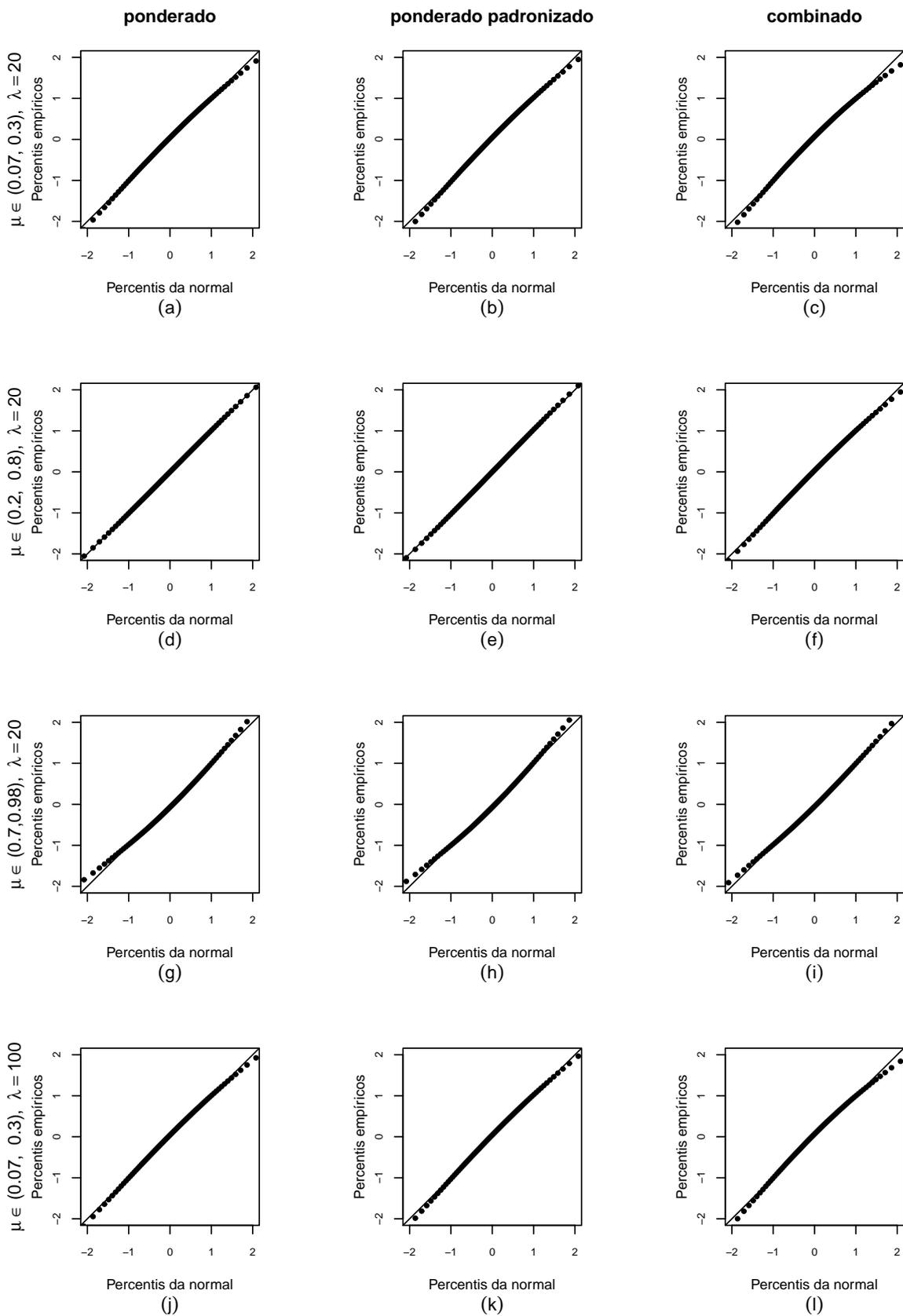


Figura 5.3: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_{2t}^{\beta_2} + \beta_3 x_{3t}$ e $\log(\phi_t) = \gamma_1 + z_{2t}^{\gamma_2} + \gamma_3 z_{3t}$, $t = 1, \dots, 80$.

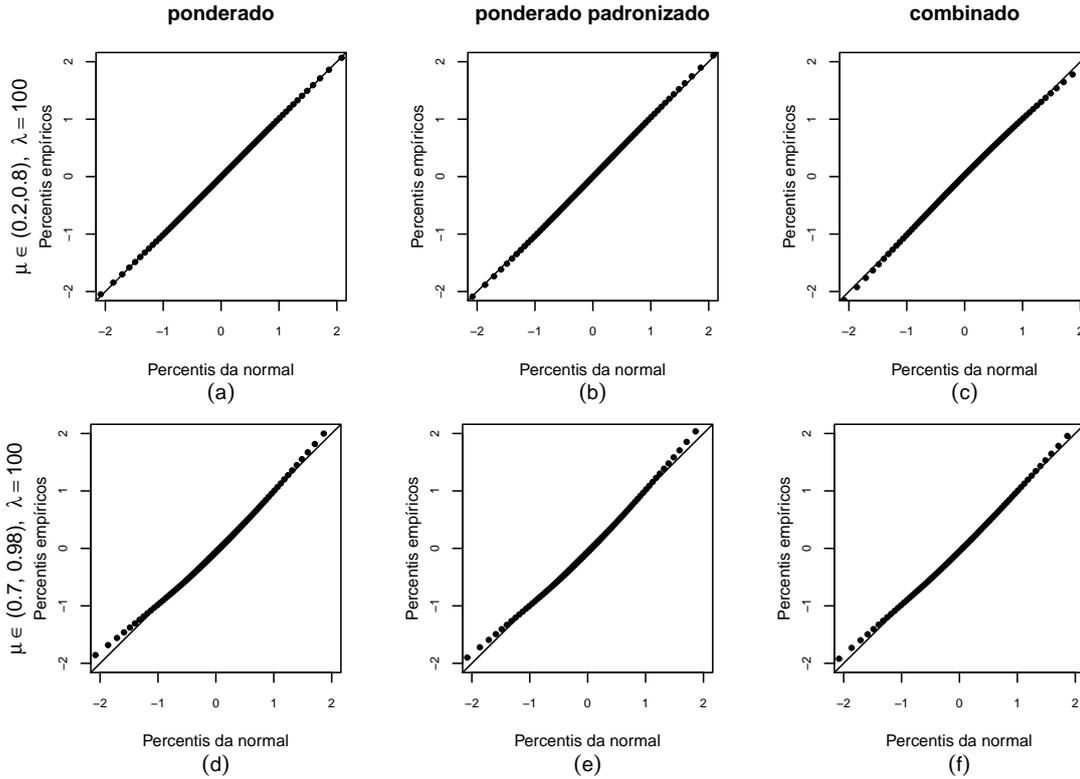


Figura 5.4: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_{2t}^{\beta_2} + \beta_3 x_{3t}$ e $\log(\phi_t) = \gamma_1 + z_{2t}^{\gamma_2} + \gamma_3 z_{3t}$, $t = 1, \dots, 80$.

Figuras 5.1 e 5.2 nota-se que os resíduos possuem distribuição consideravelmente próxima da distribuição normal padrão. O resíduo combinado, como constatado anteriormente, possui assimetria maior do que os demais na maioria dos cenários (Figuras 5.1-c,f e 5.2-f,i).

Para as Figuras 5.3 e 5.4, consideramos as seguintes relações funcionais:

$$\log\left(\frac{\mu_t}{1 - \mu_t}\right) = \beta_1 + x_{2t}^{\beta_2} + \beta_3 x_{3t},$$

$$\log(\phi_t) = \gamma_1 + z_{2t}^{\gamma_2} + \gamma_3 z_{3t}.$$

Os valores das covariáveis foram gerados a partir da distribuição uniforme: $x_{2t} \sim U(1, 2)$, $x_{3t} \sim U(0, 1)$, $z_{2t} \sim U(1, 2)$ e $z_{3t} \sim U(-0.5, 0.5)$. Os cenários para a média da variável resposta foram: $\mu \in (0.07, 0.3)$, dado que $\beta = (-1.6, -1.6, -1.5)$, $\mu \in (0.2, 0.8)$, dado que $\beta = (-1.1, 1.6, -1.7)$ e $\mu \in (0.7, 0.98)$, dado que $\beta = (1.1, 1.6, -1.7)$. Para o submodelo da dispersão consideramos: $\lambda = 20$, dado que $\gamma = (2.1, 1.5, -1.6)$ e $\lambda = 100$, dado que $\gamma = (2.1, 1.78, -2.8)$. Nesses cenários os resíduos também apresentam distribuição empírica próxima da distribuição normal padrão, principalmente quando $\mu \in (0.2, 0.8)$. Em geral, os resíduos apresentam leve assimetria, o que é mais evidente no resíduo combinado.

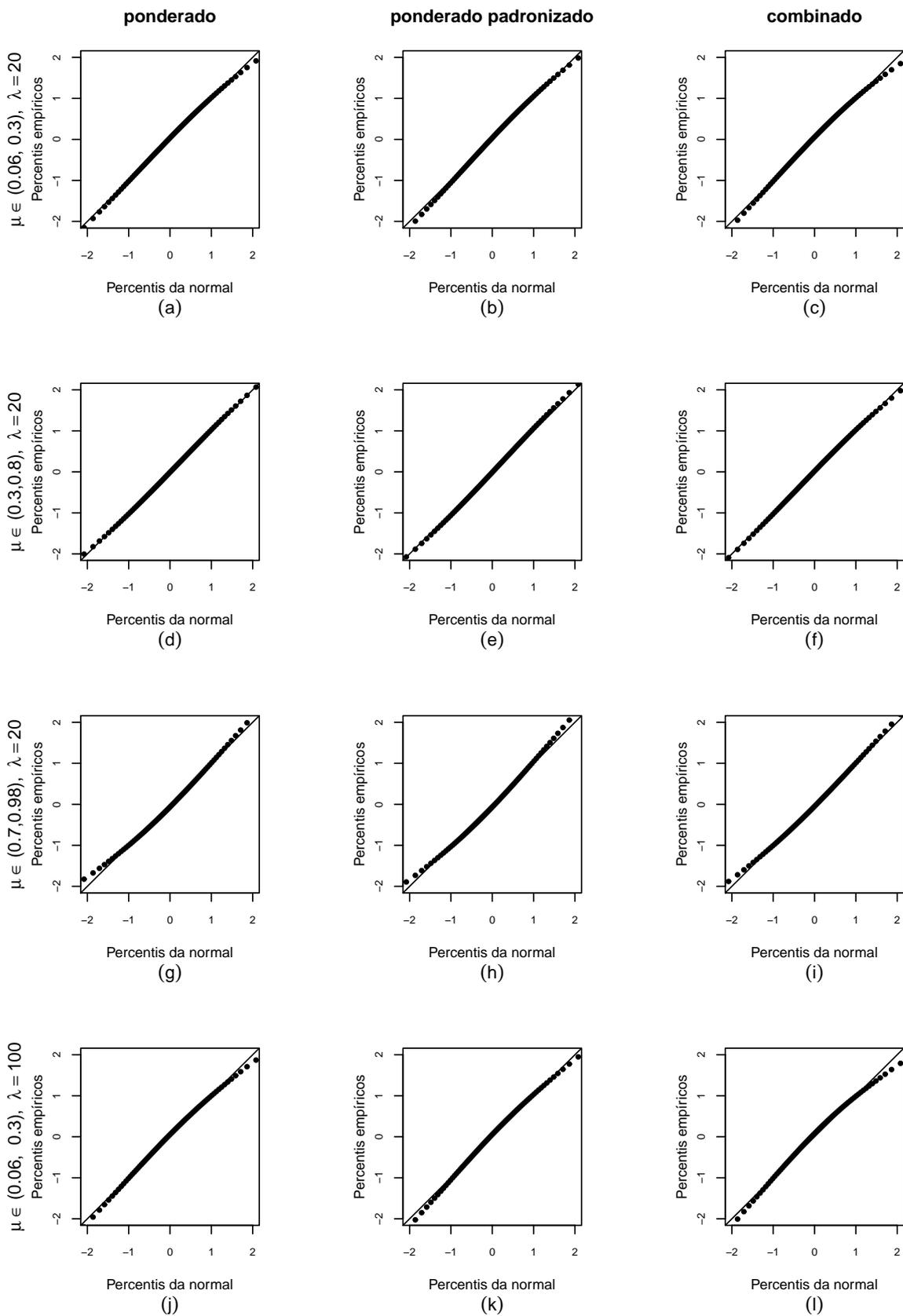


Figura 5.5: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_{2t}^{\beta_2} + \beta_3 x_{3t} + \beta_4 x_{4t} + \beta_5 x_{5t}$ e $\log(\phi_t) = \gamma_1 + z_{2t}^{\gamma_2} + \gamma_3 z_{3t} + \gamma_4 z_{4t} + \gamma_5 z_{5t}$, $t = 1, \dots, 80$.

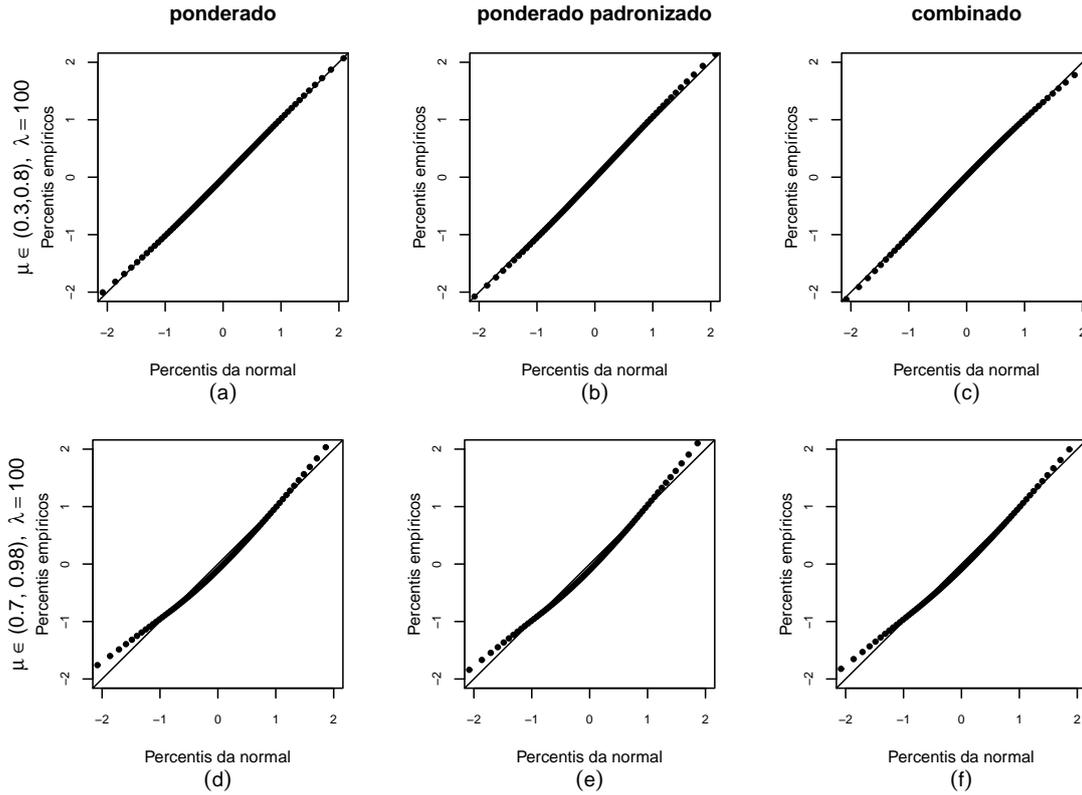


Figura 5.6: Gráficos normais de probabilidades dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_{2t}^{\beta_2} + \beta_3 x_{3t} + \beta_4 x_{4t} + \beta_5 x_{5t}$ e $\log(\phi_t) = \gamma_1 + z_{2t}^{\gamma_2} + \gamma_3 z_{3t} + \gamma_4 z_{4t} + \gamma_5 z_{5t}$, $t = 1, \dots, 80$.

Nos gráficos das Figuras 5.5 e 5.6, foram consideradas as seguintes relações funcionais:

$$\log\left(\frac{\mu_t}{1 - \mu_t}\right) = \beta_1 + x_{2t}^{\beta_2} + \beta_3 x_{3t} + \beta_4 x_{4t} + \beta_5 x_{5t},$$

$$\log(\phi_t) = \gamma_1 + z_{2t}^{\gamma_2} + \gamma_3 z_{3t} + \gamma_4 z_{4t} + \gamma_5 z_{5t}.$$

Os valores das covariadas foram gerados a partir da distribuição uniforme: $x_{2t} \sim U(1, 2)$, $x_{it} \sim U(0, 1)$, $z_{2t} \sim U(1, 2)$ e $z_{it} \sim U(-0.5, 0.5)$, com $i = 3, 4, 5$. Foram considerados os seguintes cenários: $\mu \in (0.06, 0.3)$, dado que $\beta = (-1.21, -1.31, -0.85, -0.82, -0.78)$, $\mu \in (0.3, 0.8)$, dado que $\beta = (-1.3, 0.7, -1.3, 0.9, 1.1)$ e $\mu \in (0.7, 0.98)$, quando definimos $\beta = (1.6, 1.5, 0.9, -1.2, -1.3)$. Para o submodelo da precisão consideramos $\lambda = 20$, dado que $\gamma = (2.6, 1.1, -1.3, -1.2, 1.9)$ e $\lambda = 100$, quando definimos $\gamma = (1.8, 2.1, 2.3, 2.2, 2.2)$. Nestes cenários a distribuição empírica de cada resíduo também é bem próxima da normal e novamente o resíduo combinado foi o que apresentou maior assimetria quando comparado aos outros resíduos, principalmente na Figura 5.5-c,1.

5.2.1 Modelos mal especificados

Para verificar se os resíduos conseguem detectar a má especificação do modelo postulado consideramos alguns tipos possíveis de má especificação. Um possível tipo de erro ocorre

quando se considera um modelo linear, porém há um modelo não linear que se adequaria melhor aos dados. Nesse sentido, realizamos simulações onde os dados foram gerados considerando o modelo

$$\begin{aligned}\log \frac{\mu_t}{1 - \mu_t} &= \beta_1 + x_t^{\beta_2} \\ \log(\phi_t) &= \gamma_1 + \gamma_2 z_t.\end{aligned}$$

Os valores das covariáveis foram gerados de forma independente a partir da distribuição uniforme: $x_t \sim U(0.1, 1.1)$ e $z_t \sim U(0, 1)$, $t = 1, \dots, 40$. Consideramos dois cenários diferentes para a média da variável resposta: $\mu \in (0.05, 0.1)$, dado que $\beta = (-3.0, 1.0)$ e $\mu \in (0.4, 0.7)$, dado que $\beta = (-0.3, 1.7)$. Para o submodelo da dispersão consideramos três cenários: $\lambda = 20$, dado que $\gamma = (3.0, 3.0)$; $\lambda = 50$, dado que $\gamma = (2.5, 4.0)$ e $\lambda = 100$, dado que $\gamma = (2.0, 4.65)$.

Para a estimação dos parâmetros consideramos

$$\begin{aligned}\log \frac{\mu_t}{1 - \mu_t} &= \beta_1 + \beta_2 x_t \\ \log(\phi_t) &= \gamma_1 + \gamma_2 z_t.\end{aligned}$$

Nas Figuras 5.7 a 5.12 apresentamos gráficos de resíduos e gráficos normais de probabilidade com envelopes simulados para cada um dos resíduos. Nota-se que os gráficos dos diferentes resíduos utilizados apresentam o mesmo padrão na dispersão dos pontos. Os gráficos de resíduos *versus* covariável e resíduos *versus* preditor mostram a não aleatoriedade dos pontos, o que indica a má qualidade do ajuste, a mesma conclusão sendo obtida a partir dos gráficos de envelopes simulados.

Outro possível erro de especificação é cometido ao considerar linearidade nos parâmetros do submodelo da precisão quando na verdade há não linearidade. Nas simulações realizadas os dados foram gerados considerando as relações funcionais em (5.23). Os valores das covariadas foram gerados através da distribuição uniforme, $x_t \sim U(1, 2)$ e $z_t \sim U(0.4, 1.4)$, $t = 1, \dots, 40$. Foram considerados dois cenários diferentes para a variação da média da variável resposta: $\mu \in (0.36, 0.84)$, quando foi definido que $\beta = (-1.6, 1.7)$ e $\mu \in (0.88, 0.96)$, dado que $\beta = (1.0, 1.2)$. Para o submodelo da precisão definimos que $\gamma = (2.6, 4.6)$, o que conduziu a $\lambda = 100$.

Na estimação dos parâmetros consideramos

$$\begin{aligned}\log \frac{\mu_t}{1 - \mu_t} &= \beta_1 + x_t^{\beta_2} \\ \log(\phi_t) &= \gamma_1 + \gamma_2 z_t.\end{aligned}\tag{5.24}$$

Nas Figuras 5.13 e 5.14 apresentamos gráficos de resíduos e gráficos normais de probabilidade com envelopes simulados para cada um dos resíduos. Aqui os gráficos dos diferentes resíduos também apresentam o mesmo padrão e novamente a falta de qualidade do ajuste

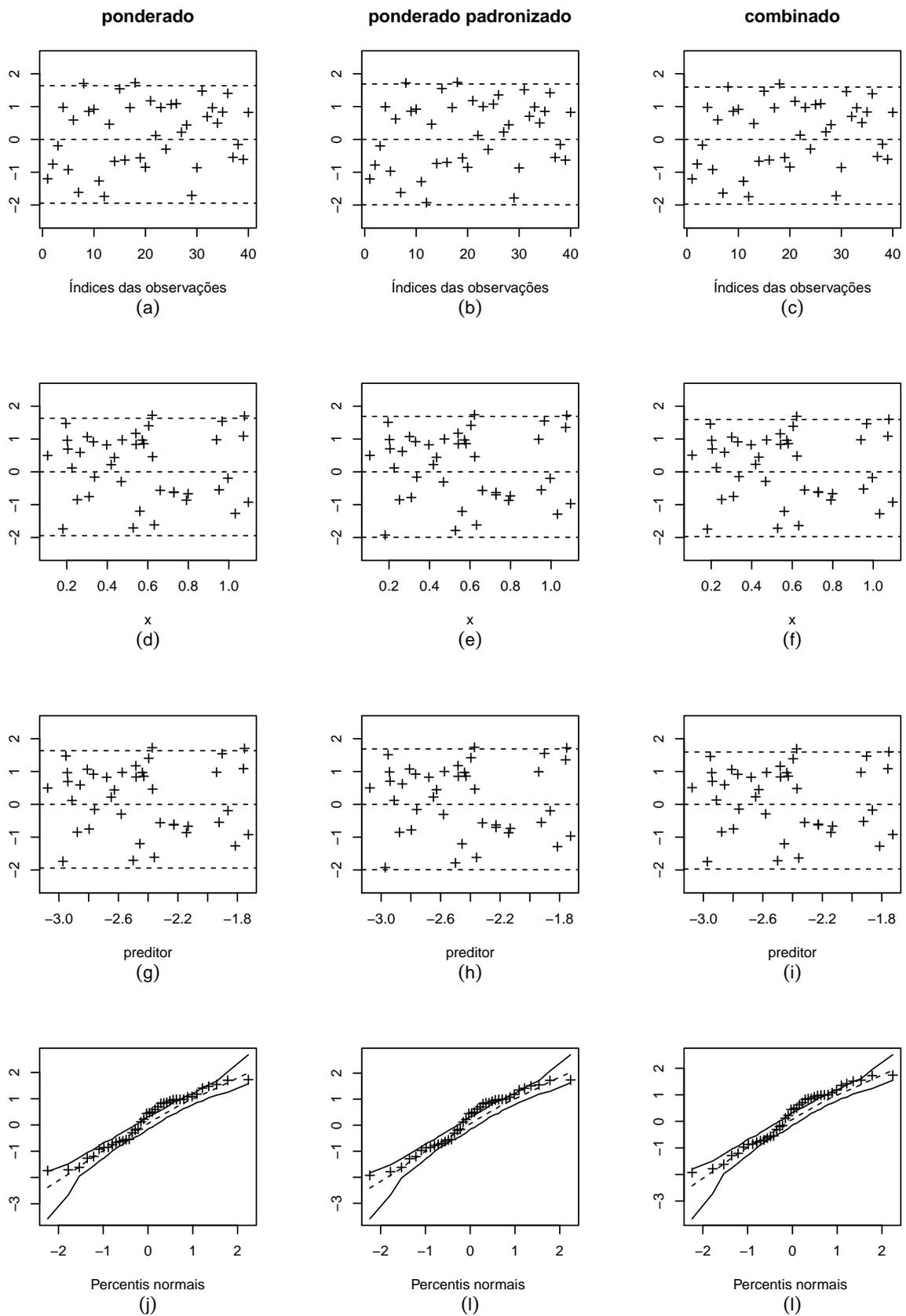


Figura 5.7: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.1, 1.1)$, $z_t \sim U(0, 1)$, $t = 1, \dots, 40$, $\lambda = 20$ e $\mu \in (0.05, 0.1)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_t$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

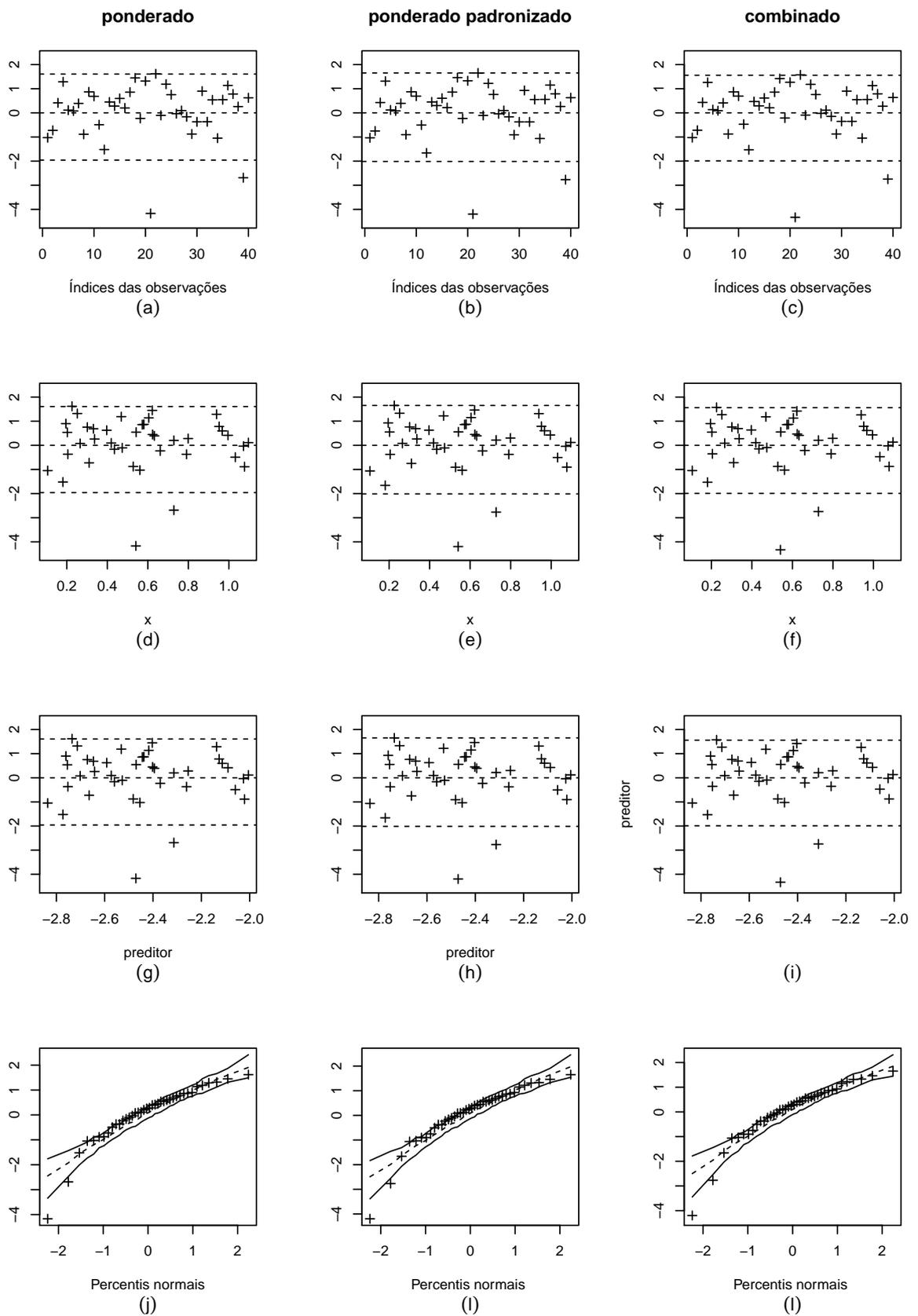


Figura 5.8: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.1, 1.1)$, $z_t \sim U(0, 1)$, $t = 1, \dots, 40$, $\lambda = 50$ e $\mu \in (0.05, 0.1)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_t$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

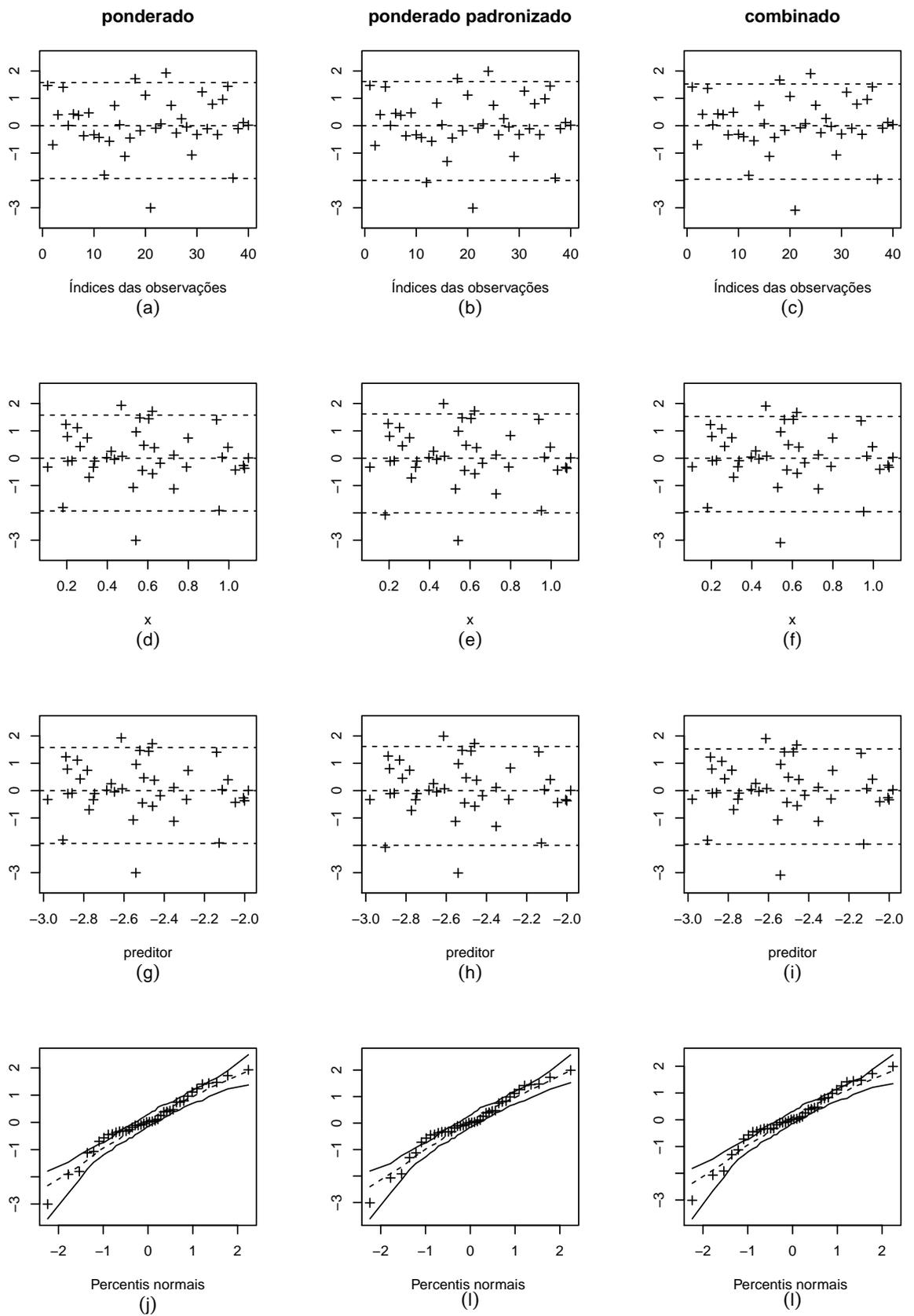


Figura 5.9: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.1, 1.1)$, $z_t \sim U(0, 1)$, $t = 1, \dots, 40$, $\lambda = 100$ e $\mu \in (0.05, 0.1)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_t$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

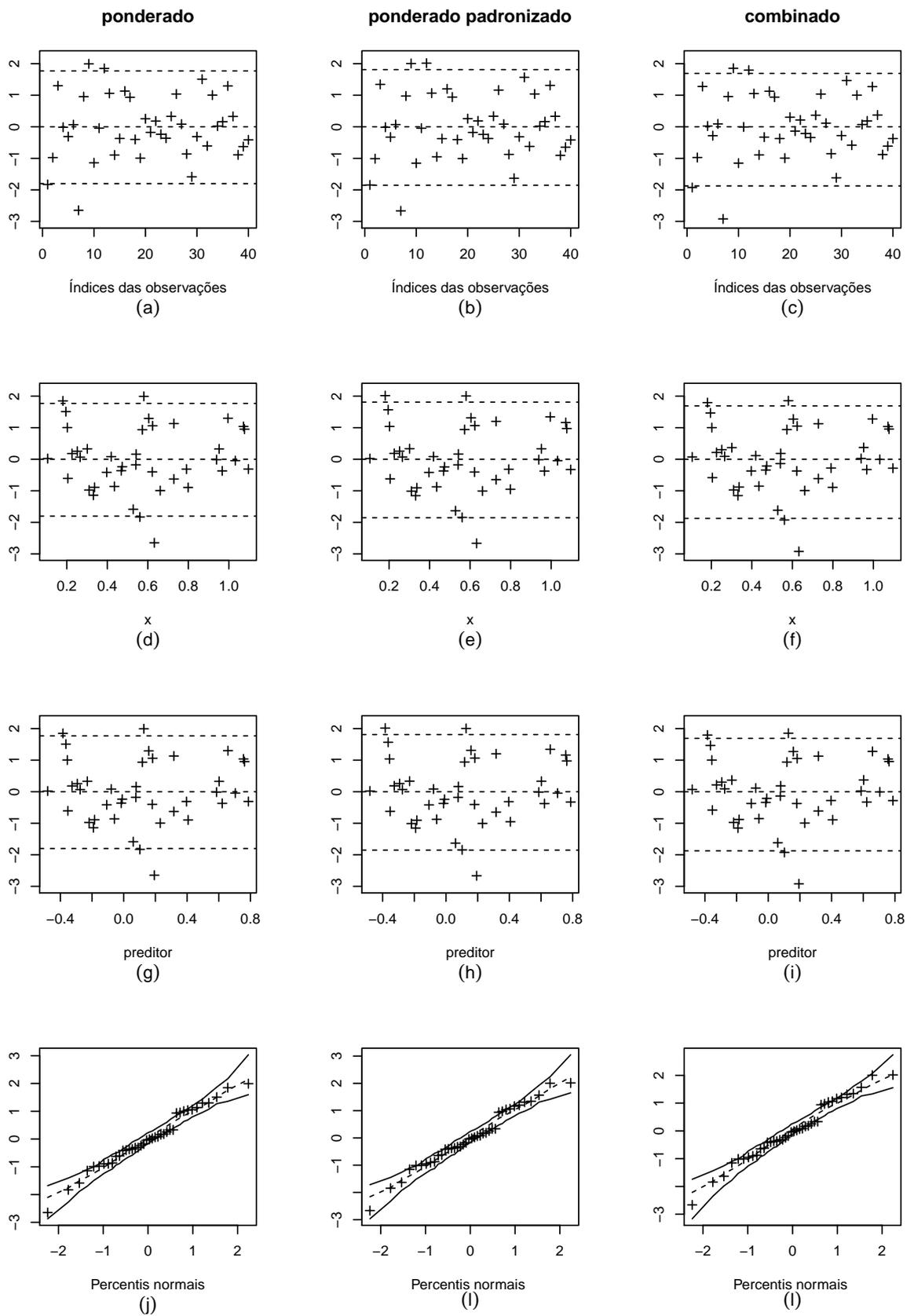


Figura 5.10: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.1, 1.1)$, $z_t \sim U(0, 1)$, $t = 1, \dots, 40$, $\lambda = 20$ e $\mu \in (0.4, 0.7)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_t$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

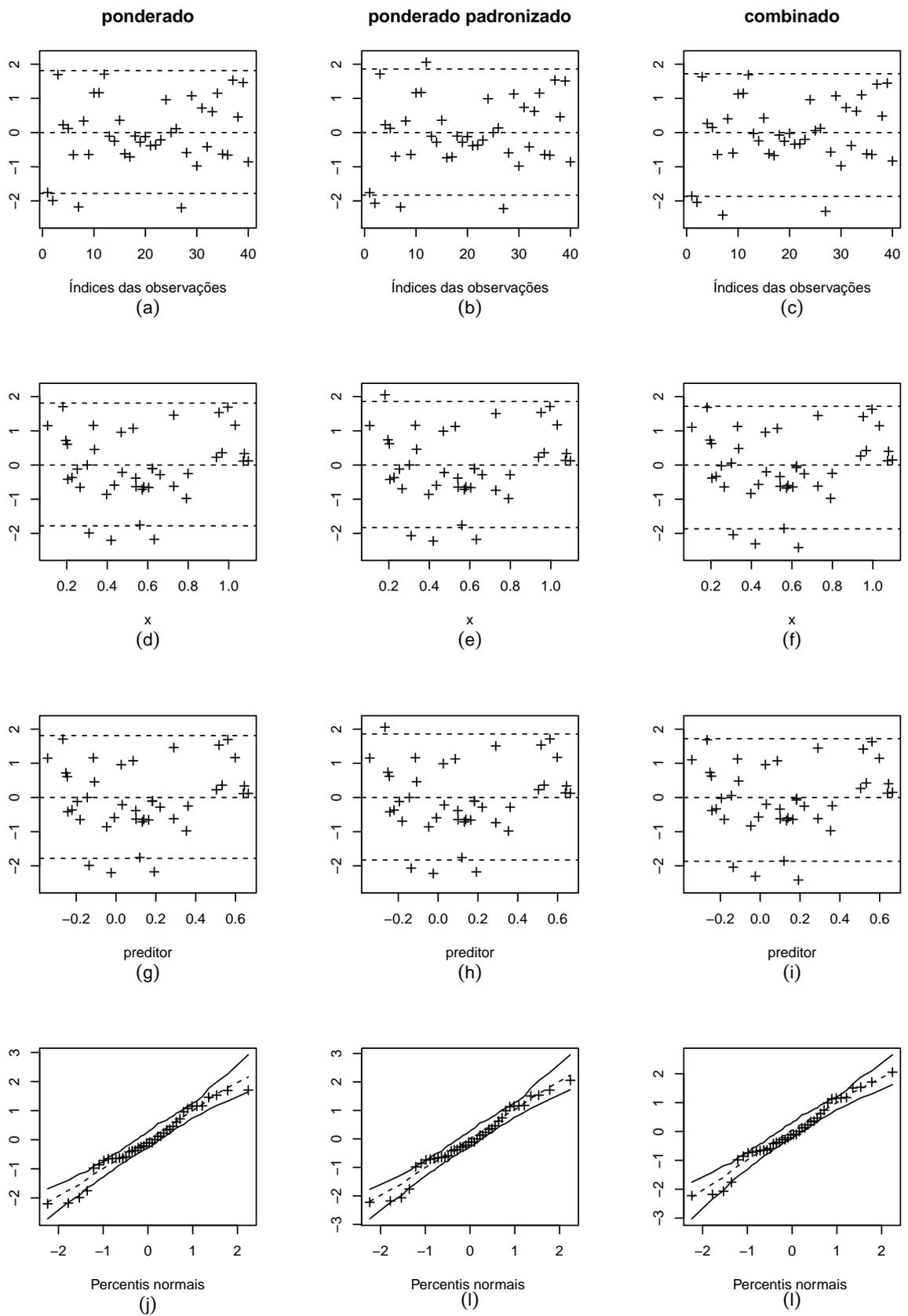


Figura 5.11: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.1, 1.1)$, $z_t \sim U(0, 1)$, $t = 1, \dots, 40$, $\lambda = 50$ e $\mu \in (0.4, 0.7)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_t$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

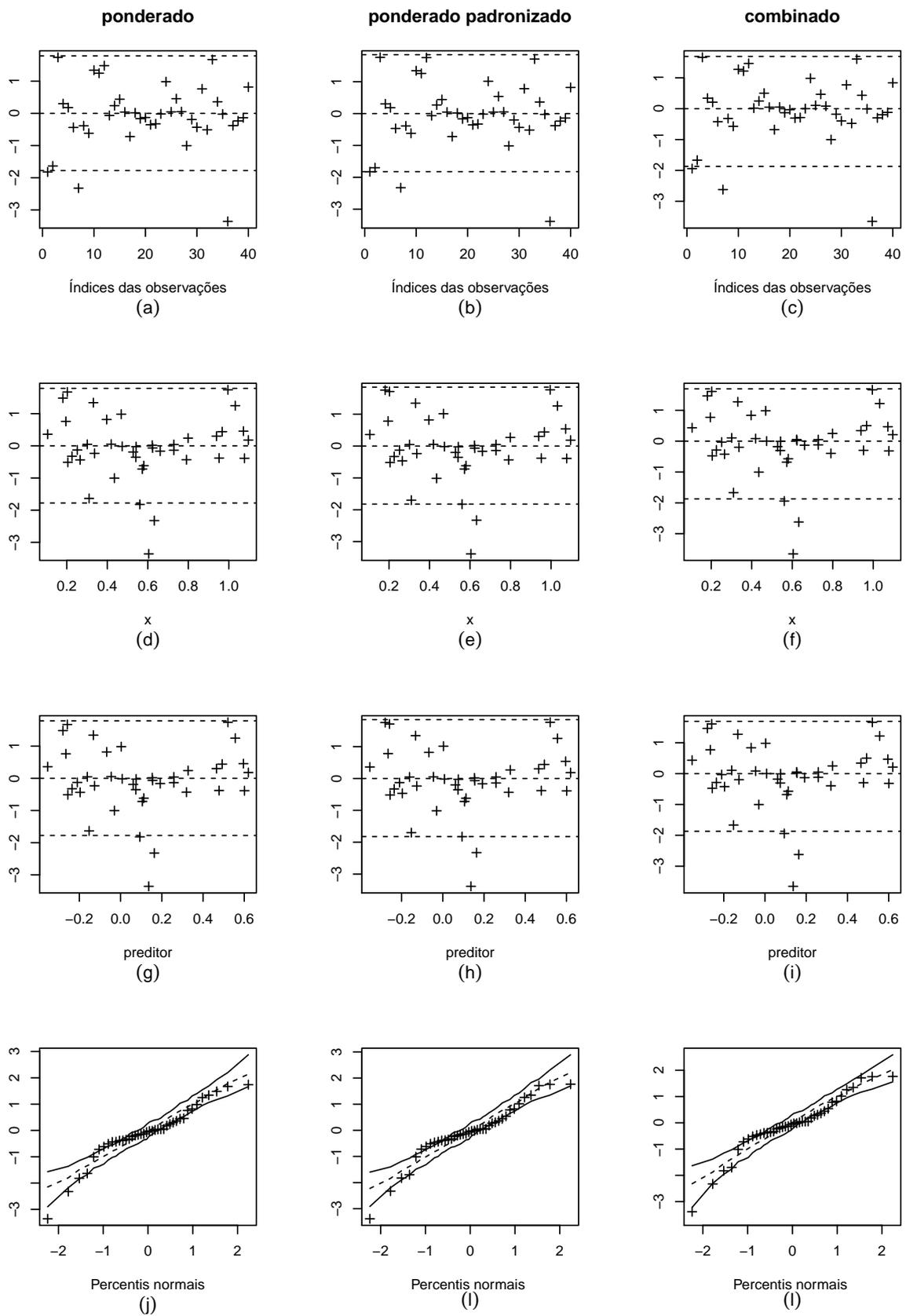


Figura 5.12: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.1, 1.1)$, $z_t \sim U(0, 1)$, $t = 1, \dots, 40$, $\lambda = 100$ e $\mu \in (0.4, 0.7)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + \beta_2 x_t$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

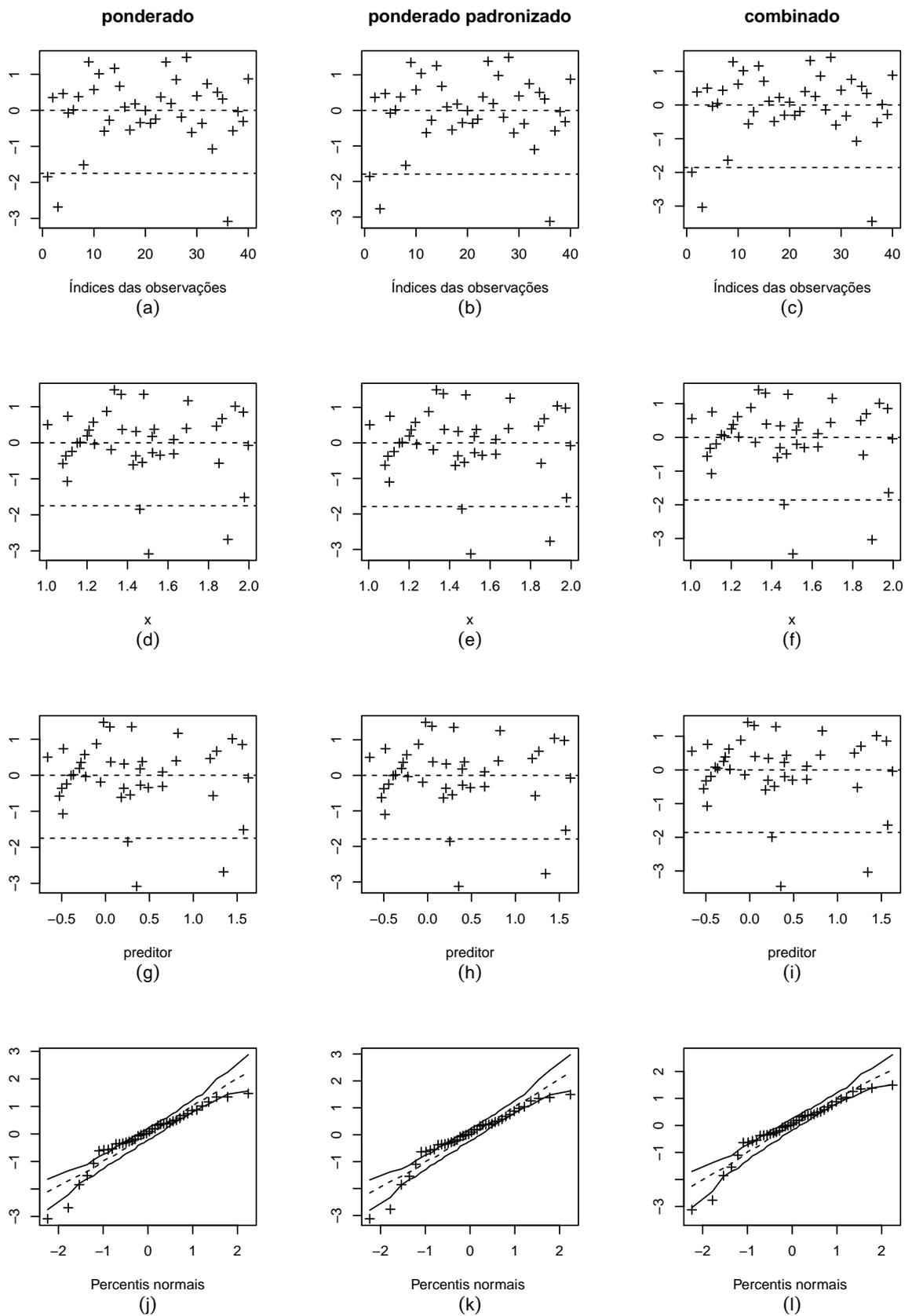


Figura 5.13: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(1, 2)$, $z_t \sim U(0.4, 1.4)$, $t = 1, \dots, 40$, $\lambda = 100$ e $\mu \in (0.36, 0.84)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

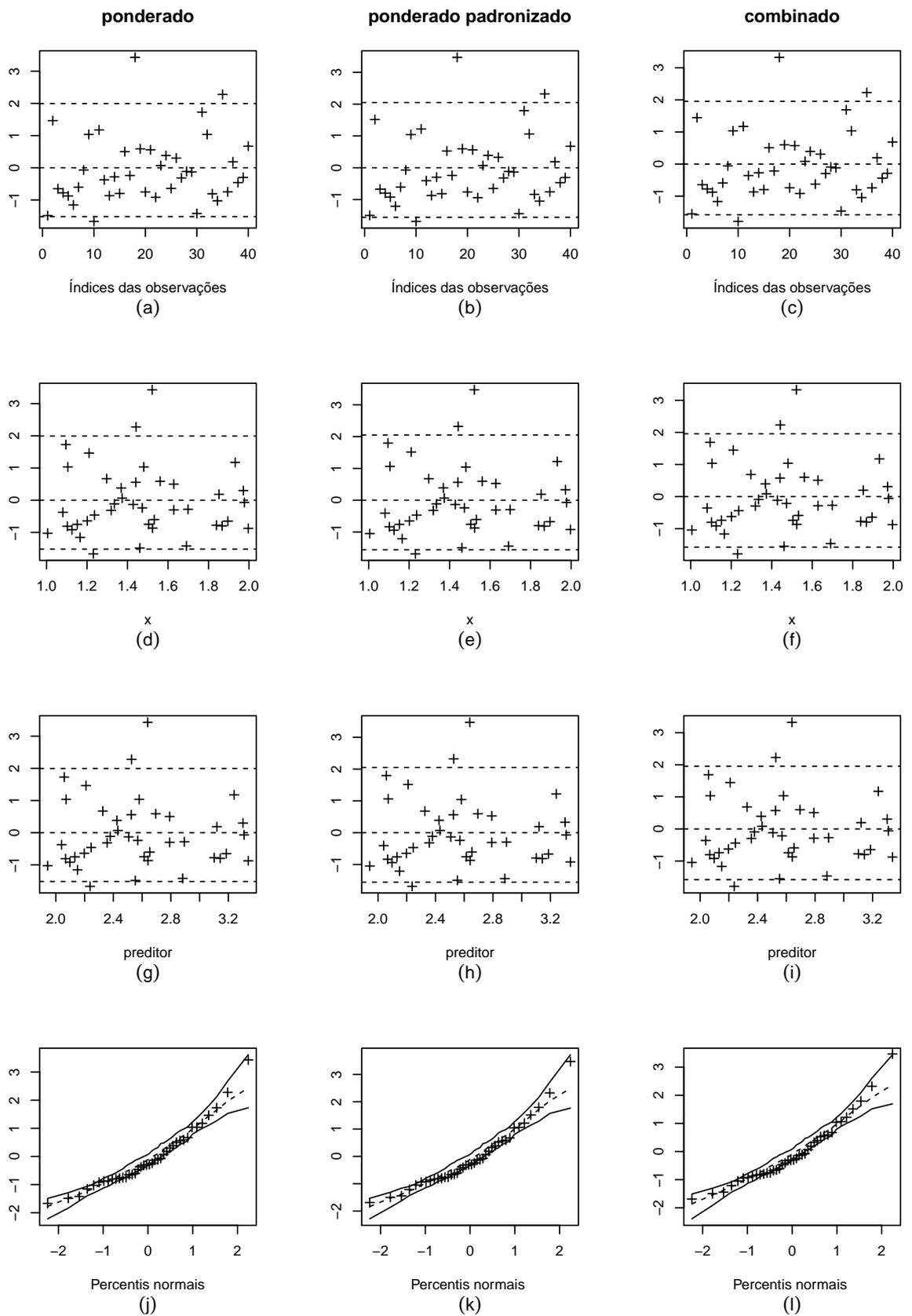


Figura 5.14: Gráficos de resíduos. Modelo verdadeiro: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + z_t^{\gamma_2}$, $x_t \sim U(1, 2)$, $z_t \sim U(0.4, 1.4)$, $t = 1, \dots, 40$, $\lambda = 100$ e $\mu \in (0.88, 0.96)$. Modelo estimado: $\log(\mu_t/[1 - \mu_t]) = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$.

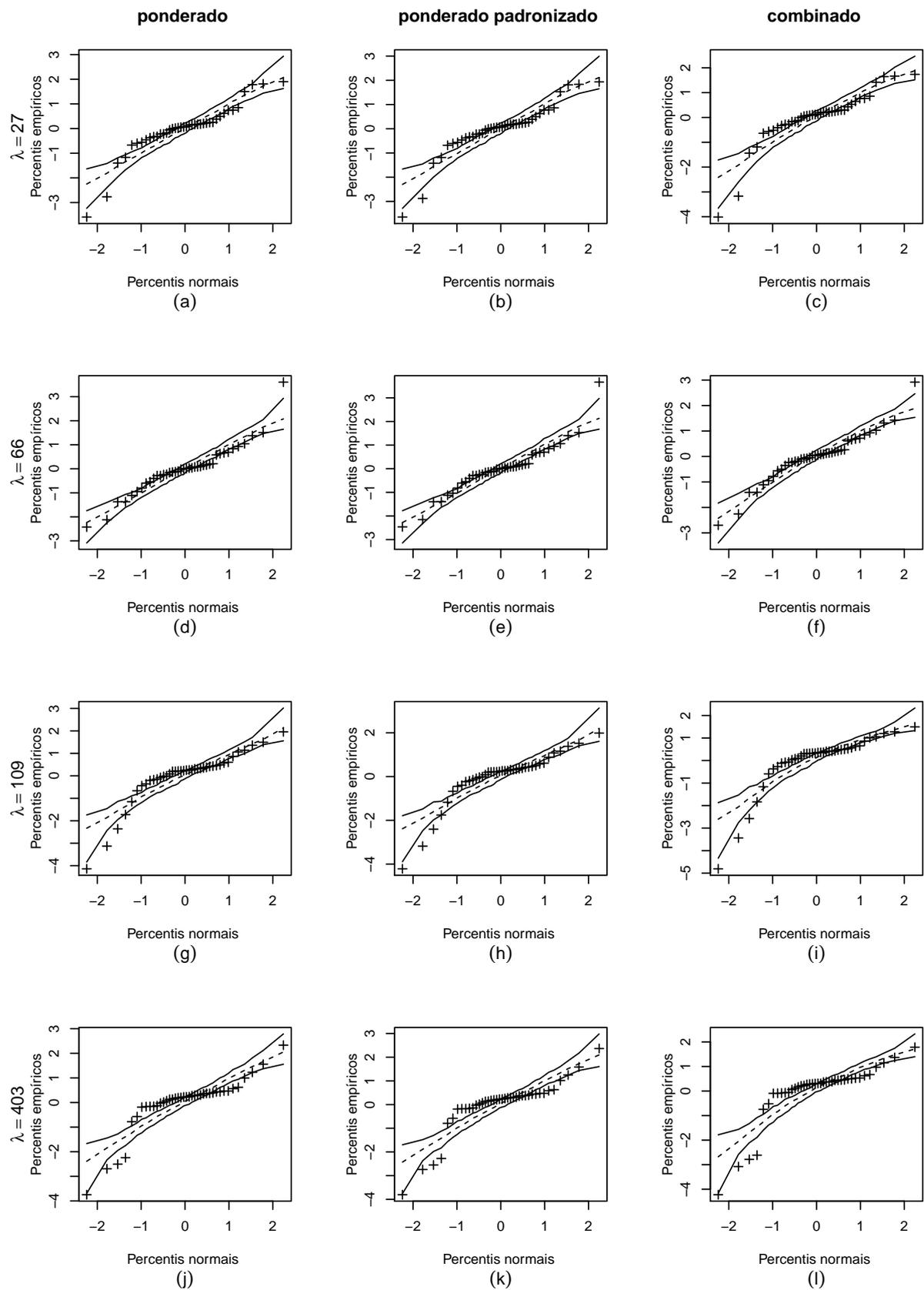


Figura 5.15: Gráficos normais de probabilidades com envelopes simulados. Modelo $\log[\mu_t/(\mu_t)] = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.4, 1.4)$, $t = 1, \dots, 40$, $z_t = 0$ para $t = 1, \dots, 20$ e $z_t = 1$ para $t = 21, \dots, 40$, $\mu \in (0.3, 0.8)$. O modelo é estimado supondo dispersão constante.

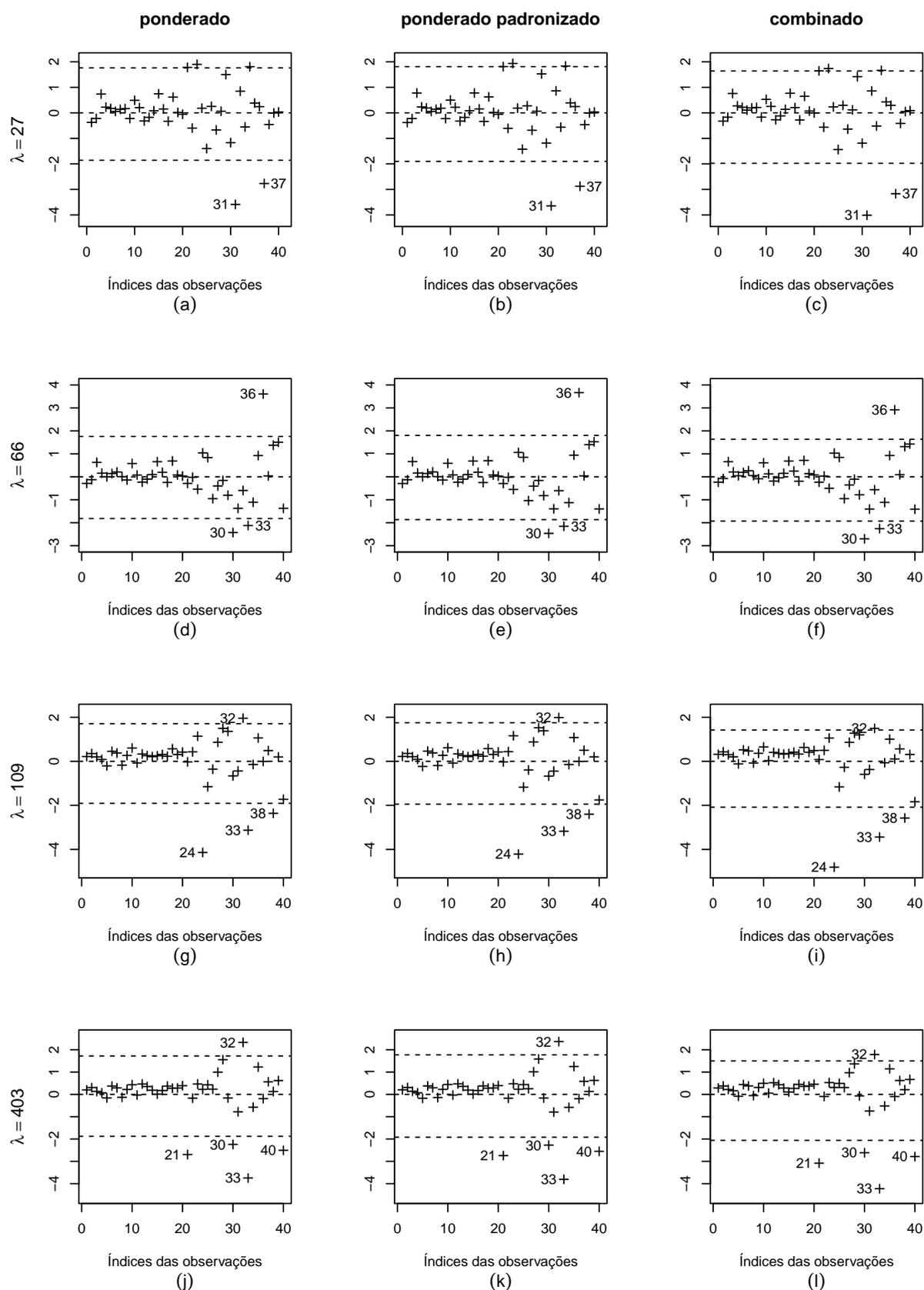


Figura 5.16: Gráficos de resíduos. Modelo $\log[\mu_t/(\mu_t)] = \beta_1 + x_t^{\beta_2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_t$, $x_t \sim U(0.4, 1.4)$, $t = 1, \dots, 40$, $z_t = 0$ para $t = 1, \dots, 20$ e $z_t = 1$ para $t = 21, \dots, 40$, $\mu \in (0.3, 0.8)$. O modelo é estimado supondo dispersão constante.

é evidente já que há a não aleatoriedade na dispersão dos pontos nos gráficos.

Outro tipo de erro comumente cometido é usar dispersão constante quando o processo gerador de dados tem dispersão variável. Neste sentido, realizamos simulações onde os dados foram gerados considerando o modelo em (5.24), em que $x_t \sim U(0.4, 1.4)$, $t = 1, \dots, 40$, e definimos $z_t = 0$ para $t = 1, \dots, 20$ e $z_t = 1$ para $t = 21, \dots, 40$, dessa forma são definidos dois sub-conjuntos de dados, cada um com uma precisão própria. Os parâmetros do submodelo da média foram definidos por $\beta_1 = 1.0$ e $\beta_2 = 2.8$, fazendo com que $\mu \in (0.3, 0.8)$. Para o submodelo da precisão definimos quatro cenários: $\lambda = 27$, dado que $\gamma = (5.7, -3.3)$, $\lambda = 66$, dado que $\gamma = (5.7, -4.2)$, $\lambda = 109$, dado que $\gamma = (5.7, -4, 7)$ e $\lambda = 403$, dado que $\gamma = (7.6, -6.0)$. No entanto, para a estimação dos parâmetros desconsideramos a dispersão variável e modelamos apenas a média da variável resposta.

Os gráficos dos resíduos *versus* índices das observações, Figura 5.16, apresentam um certo padrão na dispersão dos pontos e isso indica que a qualidade do ajuste do modelo não é boa. Pelo padrão é possível observar os dois grupos com dispersões diferentes que foram definidos. O padrão apresentado indica uma provável dispersão variável nos dados. Os gráficos dos diferentes resíduos apresentam a mesma tendência na dispersão dos pontos, variam apenas ao destacar com mais ênfase ou não a ocorrência de pontos aberrantes. Em geral, o resíduo combinado destaca mais enfaticamente a ocorrência de pontos aberrantes. Também é possível notar que quanto maior o grau de heterogeneidade da dispersão dos dados, mais evidente fica que o modelo está mal especificado.

Os gráficos da Figura 5.16 apresentam vários pontos aberrantes, outro indício de que o modelo pode estar mal especificado. Ao retirar separadamente cada um desses pontos e reestimar os parâmetros do modelo, em geral não houve grande variação nas estimativas dos parâmetros. Apenas um caso apresentou maior variação: quando $\lambda = 109$ ao retirar o ponto 24, a estimativa de ϕ sofre um aumento de 40%, mas o gráfico de resíduos *versus* índices das observações ainda se mantém do mesmo padrão, indicando mais uma vez a má qualidade do ajuste. O ponto 24 é mais destacado no gráfico do resíduo combinado.

Os gráficos normais de probabilidades com envelopes simulados, apresentados na Figura 5.15, também indicam a falta de qualidade do ajuste do modelo, que também é mais evidente a medida que o grau de heterogeneidade dos dados aumenta.

5.3 Aplicação

Os dados desta aplicação estão disponíveis em McCullagh e Nelder (1989) e relacionam-se à estimação do menor gasto da mistura de inseticidas e sinérgicos. A análise é baseada em gafanhotos da espécie *Melanopus sanguinipes* com o inseticida carbofuran e o sinérgico piperonyl butoxide (PB), que torna o inseticida mais tóxico. A variável resposta é a

proporção de gafanhotos mortos (y) e as variáveis explicativas são as doses de inseticida (x_1) e sinérgico (x_2). O modelo sugerido por McCullagh e Nelder (1989) é da forma

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 \log(x_{t1} - \beta_3) + \beta_4 \frac{x_{t2}}{x_{t2} + \beta_5}, \quad t = 1, \dots, 15, \quad (5.25)$$

e a distribuição para a variável resposta é a binomial. Aqui utilizaremos o mesmo modelo, mas supomos que $y_t \sim B(\mu_t, \phi)$.

Para a estimação dos parâmetros desse modelo, caso em que há mais parâmetros do que covariadas, antes de começar o processo que foi descrito neste capítulo será necessário fornecer chutes iniciais numéricos para β_3 e β_5 . Para isso é necessário levar em consideração a relação $\log(x_{t1} - \beta_3)$ e neste caso $x_{t1} - \beta_3 > 0$ e $(x_{t2})/(x_{t2} + \beta_5)$, então $x_{t2} + \beta_5 \neq 0$. Note que após fornecer esses valores teremos um modelo com $q = q_1$ e $k = k_1$. Para verificar se esses chutes são adequados observa-se se houve convergência do processo iterativo e se houve poucas iterações até a sua convergência. Por isso os valores iniciais de β_3 e β_5 podem ser diferentes a cada modelo investigado.

Para o modelo (5.25) os chutes iniciais para β_3 e β_5 foram 1.9 e 1.0, respectivamente. Os resultados da estimação são apresentados na Tabela 5.7, de onde se conclui que a variável x_2 , ao nível de significância de 5%, não é significativa para o modelo. No entanto, se retirarmos esta variável e ajustarmos um novo modelo, a variável x_1 torna-se não significativa também.

Tabela 5.7: Estimativas

parâmetro	estimativa	erro-padrão	p -valor
β_1	-3.229	0.952	0.0007
β_2	1.475	0.409	0.0003
β_3	1.564	0.401	0.0001
β_4	1.794	0.510	0.0004
β_5	1.909	2.669	0.4744
ϕ	11.201	0.334	

Como a proposta do exemplo é verificar o menor gasto da mistura de inseticidas e sinérgicos, consideramos o primeiro modelo ajustado e construímos os gráficos dos resíduos e o gráfico de probabilidades com envelopes simulados, que são apresentados na Figura 5.17. Como era esperado, os gráficos mostram a falta de qualidade no ajuste do modelo. O gráfico dos resíduos contra cada uma das covariáveis não apresenta aleatoriedade e nota-se que há diferença na dispersão dos resíduos para as diferentes doses de sinérgico: quando a dose é menor, a dispersão dos resíduos é maior, quando a dose é maior a dispersão é menor. No gráfico de resíduos *versus* inseticida há uma tendencia decrescente no comportamento da dispersão dos resíduos a medida que a dose de inseticida aumenta. Assim, existem

evidências de dispersão variável nos dados, o que é comprovado pelos gráficos normais de probabilidades com envelopes simulados (Figuras 5.17-j,k,l), já que os resíduos cortam as bandas dos envelopes.

Considerando a dispersão dos dados variável, Simas *et al.* (2010) propuseram o seguinte modelo:

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 \log(x_{t1} - \beta_3) + \beta_4 \frac{x_{t2}}{x_{t2} + \beta_5}, \quad t = 1, \dots, 15$$

$$\sqrt{\phi_t} = \gamma_1 + \gamma_2 x_1 + \gamma_3 x_2.$$

Para a estimação dos parâmetros utilizamos como chutes iniciais para β_3 e β_5 os valores 1.5 e 1.8, respectivamente. As estimativas dos parâmetros seguem na Tabela 5.8, de onde é possível concluir que a variável x_2 , ao nível de significância de 5%, não é significativa para o submodelo da média. Como o objetivo do exemplo é verificar o menor gasto da mistura de sinérgico e inseticida, propomos o modelo em (5.26).

Tabela 5.8: Estimativas

parâmetro	estimativa	erro-padrão	p-valor
β_1	-3.282	0.323	0.000
β_2	1.489	0.121	0.000
β_3	1.471	0.194	0.000
β_4	1.714	0.263	0.000
β_5	1.619	1.054	0.124
γ_1	5.354	0.782	0.000
γ_2	0.995	0.074	0.000
γ_3	-0.479	0.024	0.000

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 \log(x_{t1} - \beta_3) + \beta_4 \frac{x_{t2}}{x_{t2} + \beta_5}, \quad t = 1, \dots, 15$$

$$\sqrt{\phi_t} = \gamma_1 + \gamma_2 x_1 + \gamma_3 x_2 + \gamma_4(x_1 x_2). \quad (5.26)$$

Os chutes iniciais para β_3 e β_5 no processo de estimação foram 1.5 e 2.0, respectivamente. As estimativas dos parâmetros desse modelo são apresentadas na Tabela 5.9, de onde se conclui que todas as covariadas são significativas, inclusive a interação entre inseticida e sinérgico no modelo da dispersão. Na Figura 5.18 são apresentados os gráficos de resíduos e o gráfico de probabilidades com envelopes simulados para cada resíduo, de onde se conclui que há boa qualidade no ajuste. Em todos os gráficos há um padrão aleatório na dispersão dos pontos, o que não era observado anteriormente.

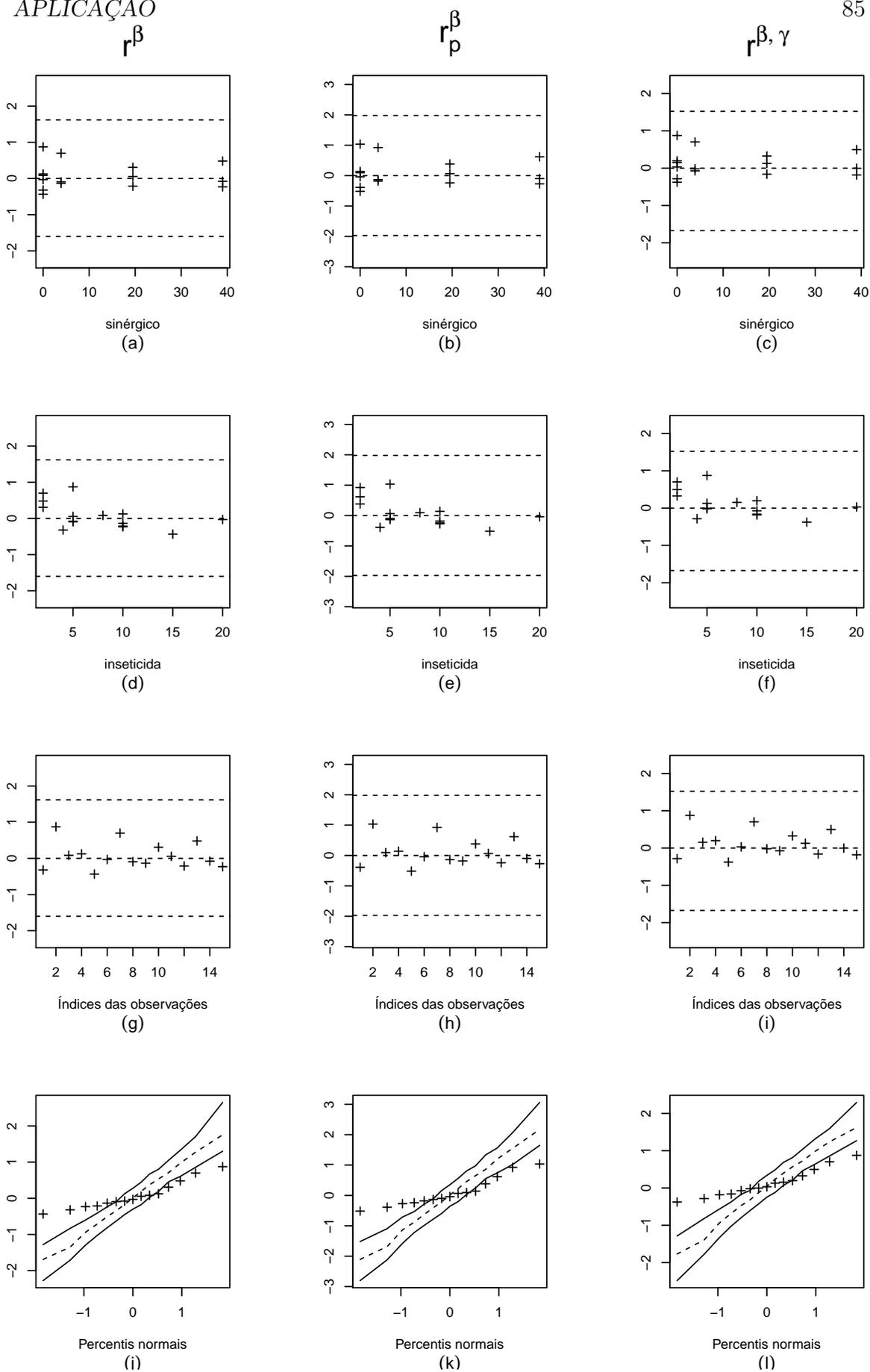


Figura 5.17: Gráficos de resíduos para os dados de inseticida e sinérgico. Modelo estimado supondo dispersão constante

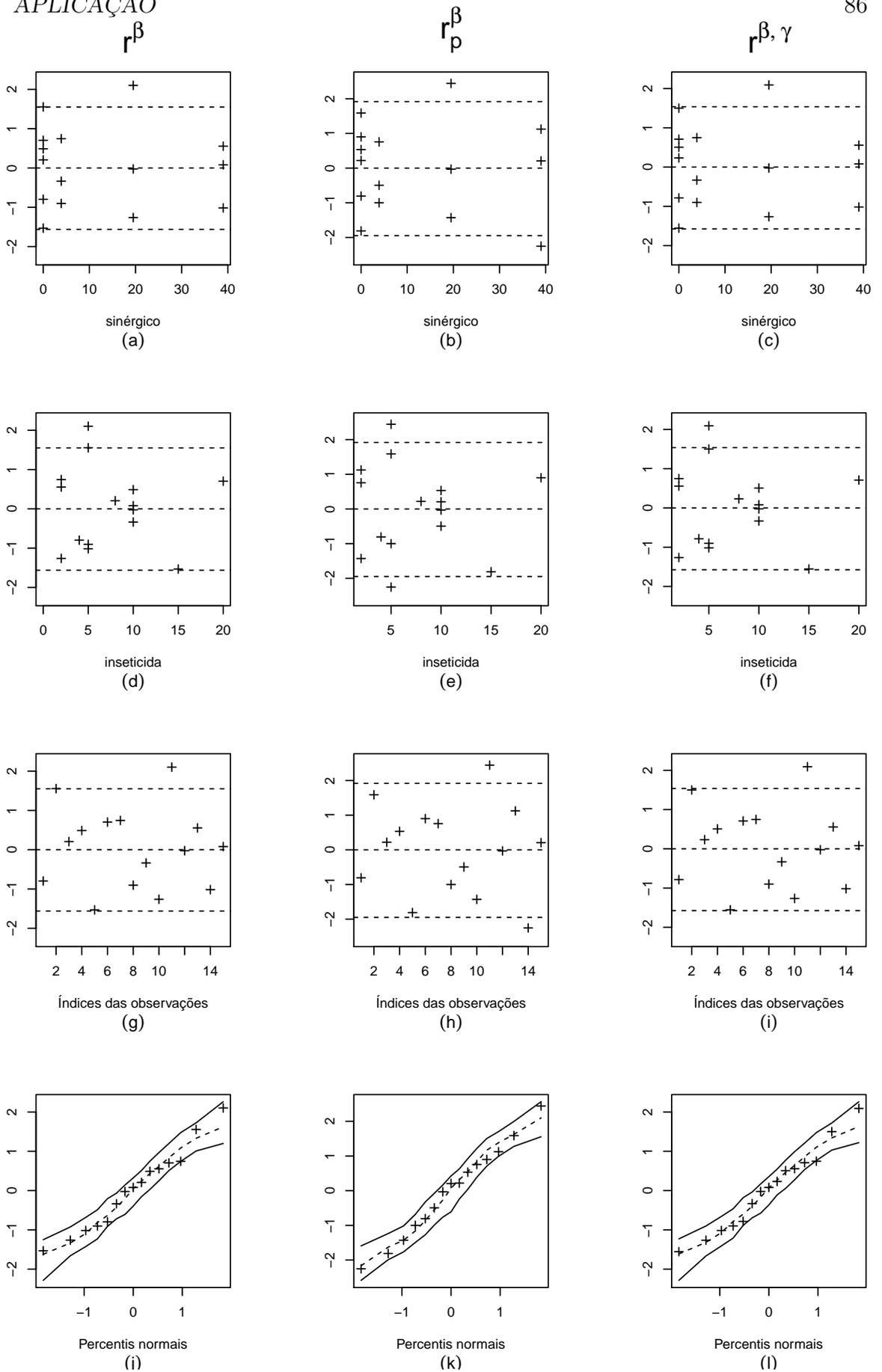


Figura 5.18: Gráficos de resíduos para os dados de inseticida e sinérgico. Modelo estimado supondo dispersão variável

Tabela 5.9: Estimativas

parâmetro	estimativa	erro-padrão	p -valor
β_1	-3.346	0.102	0.0000
β_2	1.518	0.024	0.0000
β_3	1.554	0.046	0.0000
β_4	1.842	0.078	0.0000
β_5	1.721	0.243	0.0000
γ_1	2.298	0.875	0.0000
γ_2	0.802	0.085	0.0000
γ_3	-0.888	0.046	0.0000
γ_4	0.873	0.006	0.0000

Considerações finais

Neste trabalho apresentamos o modelo de regressão beta proposto por Ferrari e Cribari-Neto (2004) e suas extensões: o modelo que considera a dispersão variável, proposto por Smithson e Verkulien (2006), e o modelo não linear proposto por Simas *et al.* (2010). O objetivo foi propor e avaliar um novo resíduo para esses modelos, o resíduo combinado. A construção deste novo resíduo, além de considerar o processo iterativo Scoring de Fisher para estimar β , como foi proposto por Espinheira *et al.* (2008), também leva em consideração o processo iterativo Scoring de Fisher para estimar γ .

Através de simulações de Monte Carlo investigamos a distribuição empírica do resíduo combinado em diferentes cenários, e esta se apresentou consideravelmente próxima da normal, inclusive em situações extremas: em que o tamanho amostral era pequeno, o parâmetro de precisão ou o grau de heterogeneidade dos dados pequenos e a média da variável resposta se encontrava em um dos extremos do intervalo $(0, 1)$. Porém, em alguns cenários considerados a sua distribuição foi assimétrica, cenários que os outros resíduos já existentes na literatura também se apresentaram assimétricos. Para solucionar este problema propusemos que os quantis 2.5% e 97.5% da distribuição empírica dos resíduos gerados a partir dos envelopes simulados sejam utilizados como limites dos gráficos de resíduos.

O novo resíduo foi capaz de dar maior destaque às observações aberrantes quando ocorre a omissão da modelagem da dispersão e quando o grau de heterogeneidade é alto. De forma geral, o novo resíduo tem um comportamento igual ou superior ao usual resíduo ponderado padronizado, porém sem carregar o custo computacional com cálculo da matriz de projeção.

Referências

- ALLEN, D. M. The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, v. 16, p. 125–127, 1974.
- BAYER, F. M.; CRIBARI-NETO, F. Model selection criteria in beta regression with varying dispersion. <http://arxiv.org/abs/1405.3718>. 2014.
- CORDEIRO, G. M. Modelos lineares generalizados. VII Simpósio Nacional de Probabilidade e Estatística, Campinas, SP. 1986.
- COX, D.; SNELL, E. A general definition of residuals. *Journal of the Royal Statistical Society*, v. 30, p. 248–275, 1968.
- DOORNIK, J. A. *An Object-Oriented Matrix Programming Language Ox*. 6. ed. London: Timberlake Consultants Ltd, 2009.
- ESPINHEIRA, P. L.; FERRARI, S. L. P.; CRIBARI-NETO, F. On beta regression residuals. *Journal of Applied Statistics*, v. 35, n. 4, p. 407–419, 2008.
- FERRARI, S. L. P.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, v. 31, n. 7, p. 799–815, 2004.
- FERRARI, S. L. P.; ESPINHEIRA, P. L.; CRIBARI-NETO, F. Diagnostic tools in beta regression with varying dispersion. *Statistica Neerlandica*, v. 65, n. 3, p. 337–351, 2011.
- LEHMANN, E. L.; CASELLA, G. *Theory of Point Estimation*. 2. ed. New York: Springer, 1998.
- LIU, R. Y. Bootstrap procedures under some non-i.i.d. models. *The Annals of Statistics*, v. 16, n. 4, p. 1696–1708, 1988.
- MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. 2. ed. London: Chapman and Hall, 1989.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2014. Disponível em: <<http://www.R-project.org/>>.
- ROCHA, A. V.; SIMAS, A. B. Influence diagnostics in a general class of beta regression models. *Test*, v. 20, n. 1, p. 95–119, 2011.

- SEN, P.; SINGER, J. M. *Large Sample Methods in Statistics: An Introduction With Applications*. 6. ed. New York: Chapman and Hall, 1993.
- SIMAS, A. B.; BARRETO-SOUZA, W.; ROCHA, A. V. Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis*, v. 54, n. 2, p. 348–366, 2010.
- SMITHSON, M.; VERKULIEN, J. A better lemon squeezer? Maximum-likelihood regression with beta-disbruted dependent variables. *Psychological Methods*, v. 11, n. 1, p. 54–71, 2006.
- ZERBINATTI, L. F. M. *Predição de fator de simultaneidade através de modelos de regressão para proporções contínuas*. Dissertação (Mestrado) — Universidade de São Paulo, 2008.