# Modelo de Tobit com erros nas variáveis

João Italo Dias França

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Estatística

Orientador: Prof. Dr. Heleno Bolfarine

Ao longo deste trabalho o autor recebeu auxílio financeiro da CAPES/CNPq.

São Paulo, 27 de maio de 2010

# Modelo de Tobit com erros nas variáveis

Este exemplar corresponde à redação
final da dissertação/tese devidamente corrigida
e defendida por João Italo Dias França
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Heleno Bolfarine (Orientador) - IME-USP
- Profa. Dra. Mônica Carneiro Sandoval - IME-USP
- Prof. Dr. Caio Lucidius Naberezny Azevedo - IME -UNICAMP

# Agradecimentos

A minha mãe Antônia de Fátima Dias pela educação e criação, pois memos sendo apenas três filhos, teve paciência em educar, apoiar, incenticvar nas horas fundamentais.

Ao meu tio Andrade pelo apoio, incentivo, educação, entre outros ensinamentos de vida. Aos meus irmãos Tercina e Manoel, pois mesmo estando longe, foram presentes na minha vida. À minha avó Tercina, pois mesmo que não esteja mais entre nós, foi de fundamental importancia na minha criação, e estará sempre torcendo por mim.

Ao professor César Teixeira por me apresentar a estátistica. Ao professor/amigo João Maurício por me mostrar, e ajudar a enteder os teoremas de probabilidade e inferência. Ao professor Manoel Campêlo por ajudar no caminho da pesquisa acadêmica. A todos os professores e funcionários do Departamento de Estatística e Matemática Aplicada da UFC que me ajudaram na minha formação.

Ao Michel Helcias, Alexandre Patriota, Rafael Bráz, Caio Lucidius, Juvêncio Santos, Lizandra, Amanda e Fran pela companhia, conversas e estudo. A meus amigos Leonardo Mota e Gabriel Marques, pelas conversas descontraidas e brincadeiras. A todos os meus amigos do M.B. e 0.B. pela companhia, conversas, passeios, e brincadeiras. Ao Samuel Guerra e Victor Hugo, pelas conversas, companhia, brincadeiras, e convivência. A todos amigos de Fortaleza, Belém, Independêcia e São Paulo que de alguma forma contribuiram para este projeto.

Ao professor Heleno Bolfarine pela orientação e paciência. Aos professores(as) Mônica Sandoval, Julia Pavan, Gilberto Avarenga, Márcia Branco pelos cursos ministrados. A CNPq pelo auxílio financeiro.

# Resumo

No presente trabalho desenvolvemos estimadores para o modelo de Tobit com erros nas variáveis para vários casos assumindo parâmetros conhecidos, dando assim continuidade ao trabalho de Wang (1998). O método que utilizamos foi de estimação de momentos em dois passos para cada caso, mostrando suas estimativas e a normalidade assintótica para os dois parâmetros do modelo. Avaliamos através de simulação o comportamento de cada caso quanto ao tamanho da amostra, a imprecisão do conhecimento do parâmetro de cada caso, e os erros do tipo I e tipo II.

**Palavras-chave:** Modelo de Tobit, Erros de medida, Estimador de momentos, Normalidade assintótica

# Abstract

This work develops estimators for the Tobit model with errors in variables for several cases of assumption of known parameters, thus giving continuity to work of Wang (1998). The method used was two-step moment estimators for each case, showing their estimates and the asymptotic normality for two parameters of the model. Evaluated by simulating the behavior of each case as the size of the sample, the imprecision of the knowledge of the parameter is in each case, and the errors type I e type II.

**Keywords:** Tobit model, Measurement error, Moment estimation, Asymptotic normal

# Sumário

# Lista de Tabelas

# Capítulo 1

# Introdução

O modelo com variável dependente censurada ou limitada, também conhecido por modelo de Tobit, tem sido usado em várias áreas, como a da economia (Azevedo, 2004; Souza, 2006). Normalmente é assumida neste modelo a premissa que as variáveis dependentes são observadas sem erro de medição, mas tal suposição nem sempre é verdadeira em casos reais, o que pode resultar em estimadores inconsistentes e levará a conclusões erradas. Como exemplo podemos supor um estudo que deseje prever a interleucina 2(IL-2) que está ligada a indução da maturação de linfócitos B a partir da proteína c-reativa(PCR) que está relacionada a infecções ou estimos inflamátorios. Na medida PCR estamo exposto a erros do laborátorio e erros sistemáticos(máquina).

Com a existência de vários estudos sobre modelos com erros nas variáveis, Wang (1998) propôs um modelo de Tobit com erro na(s) variável(is) independente(s). Como é bem conhecido nos modelos com erros nas variáveis, o modelo de Tobit com erro também não é identificável, assim o autor estudou somente o caso da razão de variâncias dos erros sendo conhecida, não estudando outros casos, por exemplo, o caso onde somente a variância do erro é conhecido. Assim temos como objetivo completar (adicionar) o trabalho de Wang estudando outros casos, ou seja, a estimativa dos parâmetros com o conhecimento de outros parâmetros, que não estão estudados na literatura.

## 1.1 Organização do trabalho

O corpo do trabalho apresenta-se organizado em capítulos conforme resumo abaixo.

O Capítulo 2 apresenta um resumo do modelo de Tobit normal sem erros nas variáveis que inclui: definição do modelo, discussão dos procedimentos de estimação dos parâmetros, e propriedades assitóticas como a distribuição assintótica.

Capítulo 3 traz uma breve introdução a modelo com erros, e após apresentamos o modelo de Tobit com erros nas variáveis, falando um pouco do seu problema de identificabilidade, e mostrando a estimação dos parâmetros (para cada caso) e sua normalidade assintótica.

Foram feitas simulações no Capítulo 4 com seus resultados para comparação de cada caso nos seguintes aspectos: tamanho amostrais, imprecisão do parâmetro conhecido, erro do tipo I e erro do tipo II.

As conclusões encontram-se no Capítulo 5, assim como sugestões para estudos futuros.

No Apêndice A o teorema de Fuller (1987) utilizado para a normalidade assintótica e as rotinas computacionais desenvolvidas para a simulação do Capítulo 4.

# Capítulo 2

# Modelo de Tobit

## 2.1 Introdução

Neste capítulo apresentamos o modelo de Tobit, também conhecido como modelo de regressão normal censurado. O nome modelo de Tobit foi dado por sua primeira aparição ser de Tobin (1958), que sugeriu um modelo para um problema de despesas familiares com bens duráveis usando um modelo de regressão que leva em conta o fato de que as despesas não podem ser negativas. Neste trabalho ele chamou de modelo de variável dependente limitada. Também define o estimador de máxima verossimilhança e propôs um procedimento iterativo para encontrá-lo.

Amemiya (1973) em seu artigo prova que o estimador de máxima verossimilhança, dos parâmetros do modelo, é fortemente consistente e segue distribuição assintótica normal. Demonstrou que o estimador inicial que Tobin propôs é inconsistente, e propõem um estimador consistente e de distribuição assintótica normal, mostrando também a eficiência assintótica do estimador no segundo passo do processo iterativo (método de Newton-Raphson).

Olsen (1978) provou que o estimador em questão é único. Posteriormente Fair (1978) propôs um procedimento computacional para o cálculo do estimador, sendo este mais fácil de programar(implementar) do que o método de Newton-Raphson

## 2.2 Definição do modelo

Consideramos o modelo de Tobit definido como:

$$\begin{aligned} y_t &= max\{y_t^*, 0\} \\ y_t^* &= \beta' x_t + u_t, \qquad t = 1, 2, \ldots, N \end{aligned} \tag{2.1}$$

onde $\beta \in \mathbb{R}^k$ é um vetor de parâmetros de tamanho $k$ desconhecido, $x_t$ é um vetor de covariáveis de tamanho $k$ conhecido (constantes), e $u_t$ é o erro com distribuição $N(0, \sigma^2)$ e independentes. As variáveis $y_t$ e $x_t$ são observaveis, e $y_t^*$ não é observado. O interesse é estimar $\theta = (\beta', \sigma^2)'$.

O modelo (2.1) não é um caso particular do modelo mais geral que é definido como:

$$\begin{aligned} y_t &= max\{y_t^*, c_t\} \\ y_t^* &= \beta' x_t + u_t \ , \end{aligned} \tag{2.2}$$

onde $c_t$ é uma constate conhecida. Aplicando a transformação $\eta_t = y_t - c_t$, temos:

$$\begin{aligned} \eta_t &= max\{y_t^* - c_t, 0\} \\ y_t^* - c_t &= (\beta', -1) \begin{pmatrix} x_t \\ c_t \end{pmatrix} + u_t \end{aligned} \tag{2.3}$$

ou seja, temos novamente o modelo (2.1).

Note que o modelo (2.1) pode se reparametrizado multiplicando ambos os lados por $h = 1/\sigma$,ou seja,

$$hy_t = \max(hy_t^*, 0)$$

$$hy_t^* = B_t'x_t + v \tag{2.4}$$

onde $B' = \beta'/\sigma$ e $v_t = u_t/\sigma$ é o erro agora com distribuição $N(0,1)$. O interesse neste modelo é estimar $\theta = (B, h)'$.

Considerando o modelo (2.1), seja $N$ o tamanho amostral. Definimos $H$ como um subconjunto dos inteiros $\{1, 2, ..., N\}$ tal que $y_t = 0$, e $N_0$ como o número de elementos de $H$. O complemento de $H$ no subconjunto dos inteiros $\{1, 2, ..., N\}$ denominamos de $\overline{H}$ e $N_1 = N - N_0$ o número de elementos. Consideramos que $H$ e $\overline{H}$ tenham pelo menos um elemento para não se tornar um caso trivial. E sem perda de generalidade ordenamos os pares amostrais $(y_t, x_t)$ tal que os $y_t > 0$ venham primeiro que os pares em que $y_t = 0$.

Além disso, definimos:

$$\begin{aligned}
F_t &= F(\beta' x_t, \sigma^2) = \int_{-\infty}^{\beta' x_t} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(u/\sigma)^2} du \\
f_t &= f(\beta' x_t, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\beta' x_t/\sigma)^2} \\
\Phi_t &= F_t = \int_{-\infty}^{\frac{\beta' x_t}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
\phi_t &= \sigma f_t = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\beta' x_t/\sigma)^2} \\
\gamma_t &= \frac{\phi_t}{1 - \Phi_t} \\
y' &= (y_1, y_2, \ldots, y_{N_1}) \\
\boldsymbol{X}' &= (x_1, x_2, \ldots, x_{N_1}) \\
\overline{\boldsymbol{X}}' &= (x_{N_1+1}, x_{N_1+2}, \ldots, x_N) \\
\overline{\gamma}' &= (\gamma_{N_1+1}, \gamma_{N_1+2}, \ldots, \gamma_N)
\end{aligned}$$

onde $\phi_t$ e $\Phi_t$ são, respectivamente, a função densidade e distribuição de uma distribuição normal padrão no ponto $\frac{\beta' x_t}{\sigma}$. O vetor $y'$ é de dimensão $1 \times N_1$, a matriz $\boldsymbol{X}'$ é de dimensão

$k \times N_1$, a matriz $\overline{\boldsymbol{X}}'$ é de dimensão $k \times N_0$ e o vetor $\overline{\gamma}'$ é de dimensão $1 \times N_0$.

A função de verossimilhança do modelo é então dada por:

$$
\begin{aligned}
L &= \prod_{t=1}^{N} f(x;\theta,x_t) \\
L &= \prod_{t=1}^{N} [1-F(\beta' x_t,\sigma^2)]^{I(y_t\leq 0)} \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_t-\beta' x_t)^2}\right]^{I(y_t>0)} \\
L &= \prod_{t\in H} [1-F(\beta' x_t,\sigma^2)] \prod_{t\in \overline{H}} \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_t-\beta' x_t)^2}\right]. \qquad (2.5)
\end{aligned}
$$

O logaritmo da função de verossimilhança pode ser escrito como

$$
l = \log L = \sum_{t\in H} \log(1-F_t) - \frac{N_1}{2}\log\sigma^2 2\Pi - \frac{1}{2\sigma^2}\sum_{t\in\overline{H}}(y_t-\beta' x_t)^2. \qquad (2.6)
$$

Calculando a primeira e segunda derivadas de 2.6 temos:

$$
\begin{aligned}
\frac{\partial l}{\partial \beta} &= -\sum_{t\in H}\frac{f_t}{1-F_t}x_t + \frac{1}{\sigma^2}\sum_{t\in\overline{H}}(y_t-\beta' x_t)x_t, \qquad (2.7) \\
\frac{\partial l}{\partial \sigma^2} &= \frac{1}{2\sigma^2}\sum_{t\in H}\frac{\beta' x_t f_t}{1-F_t} - \frac{N_1}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{t\in\overline{H}}(y_t-\beta' x_t)^2 \qquad (2.8) \\
\frac{\partial l}{\partial \beta \partial \beta'} &= -\sum_{t\in H}\frac{f_t}{(1-F_t)^2}\left[f_t - \frac{1}{\sigma^2}(1-F_t)\beta' x_t\right] x_t x_t' - \frac{1}{\sigma^2}\sum_{t\in\overline{H}} x_t x_t' \qquad (2.9) \\
\frac{\partial l}{\partial (\sigma^2)^2} &= \frac{1}{4\sigma^4}\sum_{t\in H}\frac{f_t}{(1-F_t)^2}\left[\frac{1}{\sigma^2}(1-F_t)(\beta' x_t)^3 - 3(1-F_t)\beta' x_t - (\beta' x_t)^2 f_t\right] \\
&\quad + \frac{N_1}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{t\in\overline{H}}(y_t-\beta' x_t)^2 \qquad (2.10)
\end{aligned}
$$

$$\begin{aligned}\frac{\partial l}{\partial\beta\partial\sigma^2} &= -\frac{1}{2\sigma^2}\sum_{t\in H}\frac{f_t}{(1-F_t)^2}\left[\frac{1}{\sigma^2}(1-F_t)(\beta' x_t)^2-(1-F_t)-\beta' x_t f_t\right]x_t \\ &\quad -\frac{1}{\sigma^4}\sum_{t\in\overline{H}}(y_t-\beta' x_t)x_t \end{aligned} \tag{2.11}$$

Igualando as derivadas a 0 (zero), definidas como equações normais, obteremos suas raízes (estimador de máxima verossimilhança). Usando as suposições:

**Suposição 1** *Sejam $\theta = (\beta', \sigma^2)'$ e $\theta_0 = (\beta_0', \sigma_0^2)'$. O espaço de parâmetro $\theta$ é compacto, não contendo a região $\sigma^2 \leq 0$, e contendo uma vizinhança aberta de $\theta_0$;*

**Suposição 2** *$x_t$ é limitado, e a função distribuição empírica $G_N$, definida por $G_N(x) = j/N$, onde $j$ é o numero de pontos $x_1, x_2, ..., x_N$ menores ou iguais a $x$, converge para a função distribuição (designada $G$), e*

**Suposição 3** $\lim_{N\to\infty}(1/N)\sum_{t=1}^{N} x_t x_t'$ *é positiva definida.*

Amemiya (1973) provou os seguintes resultados:

**Teorema 1** *Sob as suposições 1,2, e 3, as equações normais têm raízes fortemente consistentes.*

**Teorema 2** *Sendo $\hat{\theta}_N$ as raízes fortemente consistentes das equações normais, então sob as suposições 1, 2 e 3, temos:*

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \to N\left(\mathbf{0}, \left[\frac{\partial^2 Q(\theta_0)}{\partial\theta\partial\theta'}\right]^{-1}\right),$$

*com*

$$\frac{\partial^2 Q(\theta_0)}{\partial\theta\partial\theta'} = -\lim\frac{1}{N}\begin{bmatrix} \sum_{t=1}^{N} a_t x_t x_t' & \sum_{t=1}^{N} b_t x_t \\ \sum_{t=1}^{N} b_t x_t & \sum_{t=1}^{N} c_t \end{bmatrix},$$

*onde*

$$
\begin{aligned}
a_t &= -\frac{1}{\sigma^2}\left(z_t\phi_t - \frac{\phi_t^2}{1-\Phi_t} - \Phi_t\right) , \\
b_t &= \frac{1}{2\sigma_0^3}\left(z_t^2\phi_t + \phi_t - \frac{z_t\phi_t^2}{1-\Phi_t}\right) , \\
c_t &= -\frac{1}{4\sigma_0^4}\left(z_t^3\phi_t + z_t\phi_t - \frac{z_t^2\phi_t^2}{1-\Phi_t} - 2\Phi_t\right) ,
\end{aligned}
$$

*e* $z_t = \frac{\beta_0' x_t}{\sigma_0}$.

Olsen (1978) mostrou que a matriz $\frac{\partial^2 \log L_r}{\partial\theta\partial\theta'}$($L_r$ é a função de verossimilhança do modelo reparametrizado 2.4) é negativa semi-definida. Usando este fato com os teoremas de 1 e 2 de Amemiya (1973), temos:

**Teorema 3** *Seja* $\hat{\theta}_N$ *a raiz que soluciona* $\frac{\partial \log L_r}{\partial\theta} = 0$ *e suponha de que a matriz* $\frac{\partial^2 \log L_r}{\partial\theta\partial\theta'}$ *seja não singular. Então,* $\hat{\theta}_N$ *corresponde ao único máximo da função de verossiminhança e é um estimador consistente do verdadeiro valor de* $\theta$ *tal que,*

$$\sqrt{N}(\hat{\theta}_N - \theta) \to N\left(\mathbf{0}, C^{-1}\right)$$

*onde*

$$C = \lim_{N\to\infty} -\frac{1}{N}\frac{\partial^2 \log L_r}{\partial\theta\partial\theta'}$$

.

## 2.3 Processo Iterativo

Para o cálculo das raízes das equações $\frac{\partial \log L}{\partial \theta} = 0$, utilizaremos método iterativo de Newton-Raphson para obter estás soluções, visto que estas raízes não são tratáveis analíticamente. O método é descrito a seguir:

1. Escolha um $\theta$ inicial $(\tilde{\theta}_1)$ para inicializar o processo
2. Enquanto $|\tilde{\theta}_k - \tilde{\theta}_{k+1}| \geq \epsilon$ repita

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k - \left[\frac{\partial^2 \log L(\tilde{\theta}_{k+1})}{\partial\theta\partial\theta'}\right]^{-1} \frac{\partial \log L(\tilde{\theta}_{k+1})}{\partial\theta}$$

com $k = \{1, 2, \ldots\}$

Tobin (1958) propôs inicializar o processo de estimação da seguinte forma: resolvendo as equações (2.7) e (2.8) quando igualadas a 0(zero), e pré-multiplicando (2.7) por $\beta'/2\sigma^2$ e adicionando a (2.8) encontramos:

$$\sigma^2 = \frac{\sum_{t\in\overline{H}}(y_t - \beta' x_t)y_t}{N_1} \tag{2.12}$$

de (2.7) temos

$$-\frac{1}{\sigma}\sum_{t\in H}\frac{\phi_t}{1-\Phi_t} + \frac{1}{\sigma}\sum_{t\in\overline{H}}(y_t - \beta' x_t)x_t = 0 \tag{2.13}$$

Aproximando $\phi_t/(1-\Phi_t)$ por uma função linear $a + b\frac{\beta' x_t}{\sigma}$ e substituindo em (2.13) obtemos;

$$-\frac{1}{\sigma}\sum_{t\in H}(a+b\frac{\beta' x_t}{\sigma})+\frac{1}{\sigma}\sum_{t\in \overline{H}}(y_t-\beta' x_t)x_t=0 \tag{2.14}$$

Resolvendo a equação (2.14) para $\beta$ e substituindo em (2.12), obtemos as raízes da equação para $\sigma$. Se as raízes forem imaginárias deve-se procurar alternativas. Caso sejam reais, escolhemos uma das raízes arbitrariamente. Determinado o $\sigma$, encontramos $\beta$ através de (2.14). Como este estimador depende da escolha de $a$ e $b$, Amemiya (1973) mostrou que ele é não consistente dependendo da escolha destes. Para ser consistente deve satisfazer:

$$a\sigma_0 \lim \frac{1}{N}\sum(1-F_{0t})+b\lim\frac{1}{N}\sum(1-F_{0t})x_t x_t'\beta_0=\sigma_0^2\lim\frac{1}{N}\sum f_{0t}x_t, \tag{2.15}$$

onde $F_{0t}=F(\beta_0' x_t,\sigma_0^2)$ e $f_{0t}=f(\beta_0' x_t,\sigma_0^2)$.

A equação (2.15) nem sempre está satisfeita. Por este motivo, Amemiya (1973) propôs o estimador:

$$\hat{\theta}_1=\begin{bmatrix}\hat{\beta}_1\\ \hat{\sigma}_1^2\end{bmatrix}=\left[\sum_{t\in\overline{H}}\begin{bmatrix}x_t\hat{y}_t\\ 1\end{bmatrix}(x_t' y_t,1)\right]^{-1}\sum_{t\in\overline{H}}\begin{bmatrix}x_t\hat{y}_t\\ 1\end{bmatrix}y_t^2, \tag{2.16}$$

onde

$$\hat{y}_t=x_t'(\sum_{t\in\overline{H}}x_t x_t')^{-1}\sum_{t\in\overline{H}}x_t y_t \tag{2.17}$$

que é consistente e assintoticamente normal.

Fair (1978) propôs um procedimento iterativo para calcular os estimadores de máxima verossimilhaça de Tobit que é descrito a seguir:

1. Calcule $\beta_{emq} = (X'X)^{-1}X'y$ e $(X'X)^{-1}\hat{X}'$
2. Esconha um valor inicial para $\beta$. Denotemos de $\beta^{(1)}$.
3. Calcule $\sigma^2$ através de 2.12. Denotemos a raiz quadrada do valor de $\sigma^2$ por $\sigma^{(1)}$.
4. Calcule o vetor $\overline{\gamma}$ usando $\beta^{(1)}$ e $\sigma^{(1)}$.Denotamos este vetor como $\overline{\gamma}^{(1)}$.
5. Calcule $\beta = \beta_{emq} - \sigma(X'X)^{-1}\overline{X}'\overline{\gamma}$ usando $\sigma^{(1)}$ e $\overline{\gamma}^{(1)}$. Denotemos este valor de $\tilde{\beta}^{(1)}$. Calcule $\beta^{(2)} = \beta^{(1)} + \lambda(\tilde{\beta}^{(1)} - \beta^{(1)})$, onde $0 < \lambda \leq 1$.
6. Usando $\beta^{(2)}$ como novo valor de $\beta$. Volte ao passo 3 e repita o processo. Paramos quando $|\beta^{(k+1)} - \beta^{(k)}| < \epsilon$ sendo que $\epsilon$ é o nível de tolerância.

onde $\beta_{emq}$ é a estimativa de mínimos quadrados ordinário, calculado com os dados onde a variável dependente é diferente de zero.

A convergência dependerá do $\lambda$ escolhido. A grande contribuição deste procedimento é sua fácil implementação em comparação ao método de Newton-Raphson.

# Capítulo 3

# Modelo Tobit com erro

Neste capítulo, discutimos estimação no modelo de Tobit com erros nas variáveis. Inicialmente, discutimos o problema de falta de identificabilidade no modelo usual sob a normalidade. Apresentamos a seguir condições para tornar o modelo identificável. Como mostra Wang (1998), o modelo de Tobit com erros nas variáveis também não é identificável. Wang (1998) considerou a razão de variâncias conhecida para tornar o modelo identificável. Discutimos o problema de estimação para a situação da razão de variâncias conhecidas e também para o modelo de regressão simples com erros.

## 3.1 O Modelo de regressão simples com erros nas variáveis

O modelo de regressão simples com erros nas variáveis considera que os pares de $(W_i, Y_i)$, com $i = 1, 2, \ldots, N$, estão relacionados através das equações:

$$\begin{aligned} Y_i &= \beta_1 + \beta_2 x_i + e_i, \\ W_i &= x_i + u_i, \end{aligned} \tag{3.1}$$

com as suposições,

$$\begin{pmatrix} e_i \\ u_i \\ x_i \end{pmatrix} \sim N_3 \left( \begin{pmatrix} 0 \\ 0 \\ \mu_x \end{pmatrix} ; \begin{pmatrix} \sigma_e & 0 & 0 \\ . & \sigma_u & 0 \\ . & . & \sigma_x \end{pmatrix} \right).$$

Sob estas suposições a distribuição conjunta do observável $(W_i, Y_i)$ é tal que

$$\begin{pmatrix} W_i \\ Y_i \end{pmatrix} \sim N_2 \left( \begin{pmatrix} \mu_x \\ \beta_1 + \beta_2 \mu_x \end{pmatrix} ; \begin{pmatrix} \sigma_x + \sigma_u & \beta_2 \sigma_x \\ . & \beta_2^2 \sigma_x + \sigma_e \end{pmatrix} \right).$$

Note que o modelo acima depende de 6 parâmetros, ou seja, $\underset{\sim}{\theta} = (\beta_1, \beta_2, \mu_x, \sigma_x, \sigma_u, \sigma_e)$. Então tomando $\underset{\sim}{\theta}_1 = (1, 1, 1, 1, 1, 1)$ e $\underset{\sim}{\theta}_2 = (-1, 3, 1, \frac{1}{3}, \frac{5}{3}, 1)$ temos que apesar de $\underset{\sim}{\theta}_1 \neq \underset{\sim}{\theta}_2$ eles levam a mesma função de verossimilhança.

Para tornar o modelo identificável é necessário que suposições adicionais sejam incorporadas no modelo. Algumas suposições adicionais que tornam o modelo identificável são:

1. $\lambda = \dfrac{\sigma_u}{\sigma_e}$ conhecido,
2. $\sigma_u$ conhecido (ou $\sigma_e$ conhecido),
3. $\sigma_x$ conhecido,

4. $k_x = \dfrac{\sigma_x}{\sigma_u + \sigma_x}$ conhecido,

5. $\beta_1$ conhecido.

No caso de $\sigma_u$ conhecido, pode-se mostrar que o estimador

$$\widehat{\beta}_{2c} = \frac{S_{WY}}{S_W^2 - \sigma_u}$$

com $S_{WY} = \dfrac{1}{N}\sum_{i=1}^{N}(W_i - \bar{W})Y_i$ e $S_W^2 = \dfrac{1}{N}\sum_{i=1}^{N}(W_i - \bar{W})^2$ é consistente, pois

$$S_{WY} \xrightarrow{P} \beta_2\sigma_x \text{ e } S_W^2 \xrightarrow{P} \sigma_x + \sigma_u$$

No caso de $\lambda = \dfrac{\sigma_e}{\sigma_u}$ ser conhecido, um estimador consistente para $\beta_2$ é dado por

$$\widehat{\beta}_{2m2g} = \frac{S_W^2 - \lambda S_Y^2 + \sqrt{(S_W^2 - \lambda S_Y^2)^2 + 4S_{WY}}}{2S_{WY}}$$

## 3.2 Modelo de Tobit com erros nas variáveis

O modelo de Tobit com erros nas variáveis pode ser escrito da seguinte forma:

$$\begin{aligned} \eta_i &= max(y_i, 0), \\ y_i &= \beta_1 + \beta_2' x_i + e_i, \\ w_i &= x_i + u_i, \end{aligned} \tag{3.2}$$

onde $\eta_i \in \mathbb{R}_+$, $w_i \in \mathbb{R}^k$ são variáveis observadas, $y_i \in \mathbb{R}$, $x_i \in \mathbb{R}^k$ são variáveis não observadas, $\beta_2$, $\beta_1$ são os parâmetros do modelo, e $e_i$, $u_i$ os erros. Também assumimos que:

$$\begin{pmatrix} e_i \\ u_i \\ x_i \end{pmatrix} \sim N_{2k+1}\left( \begin{pmatrix} 0 \\ 0 \\ \mu_x \end{pmatrix} ; \begin{pmatrix} \sigma_e & 0 & 0 \\ . & \Sigma_u & 0 \\ . & . & \Sigma_x \end{pmatrix} \right)$$

No modelo (3.2) a distribuição amostral (ou verossimilhança) é o produto de $(\eta_i, w_i)$ sendo a amostra i.i.d. . Podemos determinar os momentos do modelo (3.2) a partir dos momentos de $(y_i, w_i)$, que podem ser determinados pelo primeiro e segundo momento de $(y_i, x_i)$ sob normalidade. Assim temos, como no caso do modelo:

$$\begin{aligned} &\mu_y = \beta_1 + \beta_2' \mu_x, \quad \mu_w = \mu_x, \quad \sigma_y = \beta_2' \Sigma_x \beta_2 + \sigma_e, \\ &\sigma_{wy} = \Sigma_x \beta_2, \quad \Sigma_w = \Sigma_x + \Sigma_u. \end{aligned} \tag{3.3}$$

De (3.3) observamos que apresenta mais do que um conjunto de $(\beta_1, \beta_2, \sigma_e, \Sigma_x, \Sigma_u)$ que são compatíveis para uma mesma amostra $(\eta_i, w_i)$. Assim o modelo 3.2 é não identificável em geral.

Salvo os parâmetros $(\beta_1, \beta_2, \sigma_e, \Sigma_x, \Sigma_u)$, todos os outros parâmetros são identificados ($\mu_x$, $\mu_y$, $\sigma_y$, $\sigma_{wy}$). O $\mu_x$ é facilmente identificável. O $\mu_y, \sigma_y$ e $\sigma_{wy}$ são unicamente determinados pelos seguintes momentos:

$$\begin{aligned} E(\eta_i) &= \Phi(\delta)E(\eta_i|\eta_i > 0), \\ E(\eta_i|\eta_i > 0) &= \mu_y + \sqrt{\sigma_y}\phi(\delta)/\Phi(\delta) \\ E(w_i\eta_i|\eta_i > 0) &= \sigma_{wy} + \mu_x E(\eta_i|\eta_i > 0) \end{aligned} \tag{3.4}$$

onde $\delta = \mu_y/\sqrt{\sigma_y}$, $\phi(.)$ é a função de densidade da normal padrão e $\Phi(.)$ é a função distribuição da normal padrão.

Temos então que a distribuição amostral de $(\eta_t, w_t)$ e os primeiros e segundos momentos de $(y_t, w_t)$ são unicamente determinados. Assim para identificar o modelo (3.2), isto é, obter seus parâmetros, é equivalente a determinação única dos parâmetros do modelo pelas equações (3.3) (parâmetros). Fazendo um processo de contagem temos que as equações (3.3) contêm $(k+1)(k+2)$ parâmetros livres e apenas $(k+1)(k+4)/2$ equações. Logo precisamos restringir os parâmetros do modelo para torná-lo identificável. Para se tornar identificável precisamos informar a priori $k(k+1)/2$ parâmetros (a diferença entre o número de parâmetros livres e o número de equações). Isto ocorre quando é dado (conhecido) um dos casos (para qualquer $k$):

1. $\Sigma_u$,
2. $\Sigma_x$,
3. $\Delta = \Sigma_x^{-1}\Sigma_u$, e
4. $\kappa = \Sigma_w^{-1}\Sigma_u$

No caso de $k = 1$ temos alguns casos particulares:

5. $\beta_1$ e,

6. $\sigma_e$.

Os casos 3 e 4 foram estudados por Wang (1998). Vamos estudar os casos demais.

### 3.2.1 Estimação de momentos em dois passos

Vamos considerar um procedimento de estimação em dois passos(EDP), visto que para o cálculo do EMV não sai de forma analítica tendo que aplicar algum procedimento iterativo. Voltando ao procedimento EDP, suponha o conjunto de dados $(\eta_i, w_i)$ com $i = 1, 2, \ldots, N$ em que $N_0$ dos $y_t$'s são iguais a zero e $N_1$ são positivos $(N = N_0 + N_1)$. Vamos assumir que $0 < N_1 < N$ pois $N_1 = 0$ e $N_1 = N$ são casos triviais.

Tendo os momentos amostrais:

$$\begin{aligned} \hat{\mu}_\eta &= \frac{1}{N}\sum_{i=1}^{N} \eta_i \\ \hat{\mu}_w &= \frac{1}{N}\sum_{i=1}^{N} w_i \\ \hat{\Sigma}_w &= \frac{1}{N}\sum_{i=1}^{N} (w_i - \bar{w})(w_i - \bar{w})' \end{aligned} \tag{3.5}$$

calculamos a seguir os estimadores dos momentos condicionais de (3.4) utilizando $\mu_{\eta+}$ (média dos $\eta+$), e $\mu_{wy+}$(média conjunta de $wy+$) que são estimados consistentemente pelos seus estimadores de momentos utilizando os $y_t$'s positivos com seus $x_t$'s correspondentes. Denotamos estes de $\hat{\mu}_{\eta+}$, e $\hat{\mu}_{wy+}$. Pela primeira equação de 3.4 temos que $\hat{\delta} = \Phi^{-1}(\frac{\hat{\mu}_\eta}{\hat{\mu}_{\eta+}}) = \Phi^{-1}(\frac{N_1}{N})$ e das demais equações temos:

$$\hat{\mu}_y = \frac{\hat{\delta}\hat{\mu}_{\eta+}}{\hat{\delta} + \frac{\phi(\hat{\delta})}{\Phi(\hat{\delta})}}, \ \hat{\sigma}_y = \left(\frac{\hat{\mu}_y}{\hat{\delta}}\right)^2, \hat{\sigma}_{wy} = \hat{\mu}_{w\eta+} - \hat{\mu}_w\hat{\mu}_{\eta+}. \tag{3.6}$$

Substituindo $\hat{\mu}_x$, $\hat{\Sigma}_w$ em 3.3 e resolvendo as equações obtemos as estimativas dos parâmetros. Abaixo colocamos as estimativas para cada caso.

1. $\Sigma_u$ conhecido:

$$\begin{aligned} &\hat{\beta}_2 = \hat{\Sigma}_x^{-1}\hat{\sigma}_{wy}, \quad \hat{\beta}_1 = \hat{\mu}_y - \hat{\beta}_2'\hat{\mu}_x, \quad \hat{\mu}_x = \hat{\mu}_w, \\ &\hat{\sigma}_e = \hat{\sigma}_y - \hat{\beta}_2'\hat{\sigma}_{wy}, \quad \hat{\Sigma}_x = \hat{\Sigma}_w - \Sigma_u \end{aligned} \tag{3.7}$$

2. $\Sigma_x$ conhecido:

$$\begin{aligned} &\hat{\beta}_2 = \Sigma_x^{-1}\hat{\sigma}_{wy}, \quad \hat{\beta}_1 = \hat{\mu}_y - \hat{\beta}_2'\hat{\mu}_x, \quad \hat{\mu}_x = \hat{\mu}_w, \\ &\hat{\sigma}_e = \hat{\sigma}_y - \hat{\beta}_2'\hat{\sigma}_{wy}, \quad \hat{\Sigma}_u = \hat{\Sigma}_w - \Sigma_x \end{aligned} \tag{3.8}$$

3. $\beta_1$ conhecido (modelo passando pela origem $\beta_1 = 0$):

$$\begin{aligned} &\hat{\beta}_2 = \frac{\hat{\mu}_y - \beta_1}{\hat{\mu}_x}, \quad \hat{\Sigma}_x = \frac{\hat{\sigma}_{wy}}{\hat{\beta}_2}, \quad \hat{\mu}_x = \hat{\mu}_w, \\ &\hat{\sigma}_e = \hat{\sigma}_y - \hat{\beta}_2'\hat{\sigma}_{wy}, \quad \hat{\Sigma}_u = \hat{\Sigma}_w - \hat{\Sigma}_x \end{aligned} \tag{3.9}$$

4. $\sigma_e$ conhecido:

$$\begin{aligned} &\hat{\beta}_2 = \frac{\hat{\sigma}_y - \sigma_e}{\hat{\sigma}_{wy}}, \quad \hat{\beta}_1 = \hat{\mu}_y - \hat{\beta}_2'\hat{\mu}_x, \quad \hat{\mu}_x = \hat{\mu}_w, \\ &\hat{\Sigma}_x = \frac{\hat{\sigma}_{wy}}{\hat{\beta}_2'}, \quad \hat{\Sigma}_u = \hat{\Sigma}_w - \hat{\Sigma}_x \end{aligned} \tag{3.10}$$

Os estimadores em (3.6) a (3.10) são consistentes, visto que eles são funções contínuas de momentos amostrais. Por exemplo, no caso de $\Sigma_u$ conhecido temos que:

$$\begin{aligned}\hat{\beta}_2 &= \hat{\Sigma}_x^{-1}\hat{\sigma}_{wy} \\ \hat{\beta}_2 &= (\hat{\Sigma}_w - \Sigma_u)^{-1}\hat{\sigma}_{wy}\end{aligned}$$

Sendo que:

$$\hat{\Sigma}_w \xrightarrow{P} \Sigma_w \text{ e } \hat{\sigma}_{wy} \xrightarrow{P} \Sigma_x\beta_2$$

Obtemos que:

$$\begin{aligned}&(\Sigma_w - \Sigma_u)^{-1}\Sigma_x\beta_2 \\ &\Sigma_x^{-1}\Sigma_x\beta_2 = \beta_2\end{aligned}$$

Assim provamos que $\hat{\beta}_2 \xrightarrow{P} \beta_2$ no caso de $\Sigma_u$ conhecido. Os demais casos podem ser provados de maneira similar.

### 3.2.2 Normalidade Assintótica

Seja $\Psi = (\mu_y, \sigma_y, \sigma'_{wy}, \mu_w, \sigma'_w)$ onde:

$$\sigma'_w = \text{vec}(\Sigma_w) = \mu_{w\otimes w} - \mu_w \otimes \mu_w,$$

onde vec é uma operação usual de vetorização (Fütkepohl, 2005) e $\otimes$ o produto de Kronecker. Denotamos a seguir os estimadores de momentos do primeiro passo dado por $\hat{\Psi} = (\hat{\mu}_y, \hat{\sigma}_y, \hat{\sigma}'_{wy}, \hat{\mu}_w, \hat{\sigma}'_w)$. Sendo:

$$\hat{\Psi} - \Psi_0 = \hat{A} \sum_{t=1}^{N} \frac{z_t}{N}, \tag{3.11}$$

onde $\Psi_0$ corresponde aos verdadeiros valores dos parâmetros do modelo, $\hat{A}$ converge em probabilidade para

$$A = \begin{bmatrix} \frac{\mu_{\eta^2+}}{\mu_\eta} & \frac{\mu_y}{\mu_{\eta+}\Phi\delta} & 0 & 0 & 0 \\ 2\sigma_y\left(\frac{\mu_{\eta^2+}}{\mu_y\mu_\eta} - \frac{1}{\delta\phi(\delta)}\right) & \frac{2\sigma_y}{\mu_{\eta+}\Phi(\delta)} & 0 & 0 & 0 \\ 0 & -\frac{\mu_w}{\Phi(\delta)} & \frac{1}{\Phi(\delta)}I_k & -\mu_{\eta+}I_k & 0 \\ 0 & 0 & 0 & I_k & 0 \\ 0 & 0 & 0 & -I_k \otimes \mu_w - \mu_w \otimes I_k & I_k \end{bmatrix}, \tag{3.12}$$

com

$$z_t = (s_t - \Phi(\delta), s_t(\eta_t - \mu_{\eta+}), s_t(w_t\eta_t - \mu_{w\eta+})', (w_t - \mu_w)', (w_t \otimes w_t - \mu_{w\otimes w})')', \tag{3.13}$$

e

$$s_t = \begin{cases} 1, & \eta_t > 0, \\ 0, & \eta_t = 0. \end{cases} \tag{3.14}$$

Desde que $z_t$ seja independentes e identicamente distribuídos, pelo teorema do limite central temos

$$\frac{1}{\sqrt{N}} \sum_{t=1}^{N} z_t \xrightarrow{L} \mathrm{N}(0, \Sigma_z), \tag{3.15}$$

com

$$\Sigma_z = E(z_t z_t') = \Phi(\delta)(\Sigma_1, \Sigma_2), \tag{3.16}$$

onde

$$\Sigma_1 = \begin{bmatrix} 1 - \Phi(\delta) & . & . \\ 0 & \sigma_{\eta+} & . \\ 0 & \mu_{w\eta^2+} - \mu_{\eta+}\mu_{w\eta+} & \Sigma_{w\eta+} \\ \mu_{w+} - \mu_w & \mu_{w\eta+} - \mu_{w+}\mu_{\eta+} & \mu_{ww'\eta+} - \mu_{w+}\mu'_{w\eta+} \\ \mu_{w\otimes w+} - \mu_{w\otimes w} & \mu_{(w\otimes w)\eta+} - \mu_{w\otimes w+}\mu_{\eta+} & \mu_{(w\otimes w)w'\eta+} - \mu_{w\otimes w+}\mu'_{w\eta+} \end{bmatrix} \tag{3.17}$$

e

$$\Sigma_2 = \begin{bmatrix} . & . \\ . & . \\ . & . \\ \Sigma_w/\Phi(\delta) & . \\ (\mu_{(w\otimes w)w'} - \mu_{w\otimes w}\mu'_w)/\Phi(\delta) & \Sigma_{w\otimes w}/\Phi(\delta) \end{bmatrix} \tag{3.18}$$

Wang (1998) provou que:

$$\sqrt{N}(\hat{\Psi} - \Psi_0) \xrightarrow{L} \mathrm{N}(0, A\Sigma_z A') \tag{3.19}$$

onde $A$ é dado por (3.12).

Denotemos uma função $g(H; J|O)$ onde está função calcula a derivada parcial do estimador $H$, dado $O$, em relação a $J$.Para encontrar a distribuição assintótica de $\beta_1$ e $\beta_2$ em cada caso precisamos de $g(\beta_1; \Psi|.)$ e $g(\beta_2; \Psi|.)$. Para auxiliar nos calculos das derivadas parciais usamos Fütkepohl (2005).

- $\Sigma_u$: Temos para $\beta_1$:

$$\begin{aligned} \beta_1 &= \mu_y - \beta_2'\mu_x \\ &= \mu_y - (\Sigma_x^{-1}\sigma_{wy})'\mu_x \\ &= \mu_y - \sigma'_{wy}(\Sigma_w - \Sigma_u)^{-1}\mu_w \end{aligned} \tag{3.20}$$

As derivadas parciais são:

$$\frac{\partial \beta_1}{\partial \mu_y} = 1$$

$$
\begin{aligned}
\frac{\partial \beta_1}{\partial \sigma_y} &= 0 \\
\frac{\partial \beta_1}{\partial \sigma'_{wy}} &= (\Sigma_w - \Sigma_u)^{-1}\mu_w \\
\frac{\partial \beta_1}{\partial \mu_w} &= -\sigma'_{wy}(\Sigma_w - \Sigma_u)^{-1} \\
\frac{\partial \beta_1}{\partial \Sigma_w} &= \sigma'_{wy}(\Sigma_w - \Sigma_u)^{-2}\mu_w
\end{aligned}
\tag{3.21}
$$

Assim:

$$
g(\beta_1; \Psi|\Sigma_u) = \left[\ 1,\ \ 0,\ \ -(\Sigma_w - \Sigma_u)^{-1}\mu_w,\ \ -\sigma'_{wy}(\Sigma_w - \Sigma_u)^{-1},\ \ \sigma'_{wy}(\Sigma_w - \Sigma_u)^{-2}\mu_w\ \right].
$$

Para $\beta_2$:

$$
\begin{aligned}
\beta_2 &= \Sigma_x^{-1}\sigma_{wy} \\
&= (\Sigma_w - \Sigma_u)^{-1}\sigma_{wy}
\end{aligned}
\tag{3.22}
$$

As derivadas parciais são:

$$
\begin{aligned}
\frac{\partial \beta_2}{\partial \mu_y} &= 0 \\
\frac{\partial \beta_2}{\partial \sigma_y} &= 0 \\
\frac{\partial \beta_2}{\partial \sigma_{wy}} &= (\Sigma_w - \Sigma_u)^{-1} \\
\frac{\partial \beta_2}{\partial \mu_w} &= 0 \\
\frac{\partial \beta_2}{\partial \Sigma_w} &= -(\Sigma_w - \Sigma_u)^{-2}\sigma_{wy}
\end{aligned}
\tag{3.23}
$$

Assim:

$$g(\beta_2; \Psi|\Sigma_u) = \left[\ 0,\ \ 0,\ \ (\Sigma_w - \Sigma_u)^{-1},\ \ 0,\ \ -(\Sigma_w - \Sigma_u)^{-2}\sigma_{wy}\ \right].$$

- $\Sigma_x$ conhecido:

  Temos para $\beta_1$:

$$\begin{aligned} \beta_1 &= \mu_y - \beta_2'\mu_x \\ &= \mu_y - (\Sigma_x^{-1}\sigma_{wy})'\mu_x \\ &= \mu_y - \sigma_{wy}'\Sigma_x^{-1}\mu_w \end{aligned} \tag{3.24}$$

As derivadas parciais são:

$$\begin{aligned} \frac{\partial\beta_1}{\partial\mu_y} &= 1 \\ \frac{\partial\beta_1}{\partial\sigma_y} &= 0 \\ \frac{\partial\beta_1}{\partial\sigma_{wy}'} &= -\Sigma_x^{-1}\mu_w \\ \frac{\partial\beta_1}{\partial\mu_w} &= -\sigma_{wy}'\Sigma_x^{-1} \\ \frac{\partial\beta_1}{\partial\Sigma_w} &= 0 \end{aligned} \tag{3.25}$$

Assim:

$$g(\beta_1; \Psi|\Sigma_x) = \left[\ 1,\ \ 0,\ \ -\Sigma_x^{-1}\mu_w,\ \ -\sigma_{wy}'\Sigma_x^{-1},\ \ 0\ \right].$$

Temos para $\beta_2$:

$$\beta_2 = \Sigma_x^{-1}\sigma_{wy} \tag{3.26}$$

As derivadas parciais são:

$$\begin{aligned}
\frac{\partial\beta_2}{\partial\mu_y} &= 0 \\
\frac{\partial\beta_2}{\partial\sigma_y} &= 0 \\
\frac{\partial\beta_2}{\partial\sigma_{wy}} &= \Sigma_x^{-1} \\
\frac{\partial\beta_2}{\partial\mu_w} &= 0 \\
\frac{\partial\beta_2}{\partial\Sigma_w} &= 0
\end{aligned} \tag{3.27}$$

Assim:

$$g(\beta_2;\Psi|\Sigma_x) = \left[\ 0,\ \ 0,\ \ \Sigma_x^{-1},\ \ 0,\ \ 0\ \right].$$

- $\beta_1$ conhecido:

  Temos para $\beta_2$:

$$\begin{aligned}
\beta_2 &= \frac{\mu_y - \beta_1}{\mu_x} \\
&= \frac{\mu_y - \beta_1}{\mu_w}
\end{aligned} \tag{3.28}$$

As derivadas parciais são:

$$\begin{aligned}
\frac{\partial \beta_2}{\partial \mu_y} &= \frac{1}{\mu_w} \\
\frac{\partial \beta_2}{\partial \sigma_y} &= 0 \\
\frac{\partial \beta_2}{\partial \sigma_{wy}} &= 0 \\
\frac{\partial \beta_2}{\partial \mu_w} &= -\frac{\mu_y - \beta_1}{\mu_w^2} \\
\frac{\partial \beta_2}{\partial \Sigma_w} &= 0
\end{aligned} \tag{3.29}$$

Assim:

$$g(\beta_2; \Psi|\Sigma_x) = \left[ \tfrac{1}{\mu_w}, \ 0, \ 0, \ -\tfrac{\mu_y - \beta_1}{\mu_w^2}, \ 0 \right].$$

- $\sigma_e$ conhecido:

  Temos para $\beta_1$:

$$\begin{aligned}
\beta_1 &= \mu_y - \beta_2 \mu_x \\
&= \mu_y - \frac{\sigma_y - \sigma_e}{\sigma_{wy}} \mu_x \\
&= \mu_y - \frac{\sigma_y - \sigma_e}{\sigma_{wy}} \mu_w
\end{aligned} \tag{3.30}$$

As derivadas parciais são:

$$\begin{aligned}
\frac{\partial \beta_1}{\partial \mu_y} &= 1 \\
\frac{\partial \beta_1}{\partial \sigma_y} &= -\frac{\mu_w}{\sigma_{wy}}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \beta_1}{\partial \sigma'_{wy}} &= \frac{\sigma_y - \sigma_e}{\sigma^2_{wy}} \mu_w \\
\frac{\partial \beta_1}{\partial \mu_w} &= -\frac{\sigma_y - \sigma_e}{\sigma_{wy}} \\
\frac{\partial \beta_1}{\partial \Sigma_w} &= 0
\end{aligned} \tag{3.31}$$

Assim:

$$g(\beta_1; \Psi|\sigma_e) = \left[ 1, \ -\tfrac{\mu_w}{\sigma_{wy}}, \ \tfrac{\sigma_y - \sigma_e}{\sigma^2_{wy}}, \ -\tfrac{\sigma_y - \sigma_e}{\sigma_{wy}}, \ 0 \right].$$

Temos para $\beta_2$:

$$\beta_2 = \frac{\sigma_y - \sigma_e}{\sigma_{wy}} \tag{3.32}$$

As derivadas parciais são:

$$\begin{aligned}
\frac{\partial \beta_2}{\partial \mu_y} &= 0 \\
\frac{\partial \beta_2}{\partial \sigma_y} &= \frac{1}{\sigma_{wy}} \\
\frac{\partial \beta_2}{\partial \sigma_{wy}} &= -\frac{\sigma_y - \sigma_e}{\sigma^2_{wy}} \\
\frac{\partial \beta_2}{\partial \mu_w} &= 0 \\
\frac{\partial \beta_2}{\partial \Sigma_w} &= 0
\end{aligned} \tag{3.33}$$

Assim:

$$g(\beta_2; \Psi|\sigma_e) \;=\; \left[\; 0, \;\; \frac{1}{\sigma_{wy}}, \;\; -\frac{\sigma_y - \sigma_e}{\sigma_{wy}^2}, \;\; 0, \;\; 0 \;\right].$$

e aplicando para o caso de interesse o teorema de Fuller (1987) segue para $\beta_i$

$$\sqrt{N}(\beta_i - \beta_{i0}) \xrightarrow{L} \mathrm{N}(0, g(\beta_i; \Psi|C)A\Sigma_z A' g(\beta_i; \Psi|C)'), \tag{3.34}$$

onde $i = (1, 2)$, $C = (\Sigma_u, \Sigma_x, \beta_1, \sigma_e)$, e $\beta_{i0}$ corresponde a verdadeiro valor do parâmetro. A.1

# Capítulo 4

# Simulação

Neste capítulo iremos estudar através de simulações o comportamento dos estimadores de $\tau = (\beta_1, \beta_2)$, para cada caso nos seguintes aspectos:

1. **Tamanho amostral:** Consideraremos vários tamanhos amostrais, desde pequenos (N = 20) a grandes (N = 400), passando por amostras moderadas (N=40, 50, 100, ou 200);

2. **Imprecisão do parâmetro conhecido:** Consideraremos erros de 10 e 20 por cento do parâmetro, sendo este para mais ou para menos. Trabalharemos neste caso com amostras de $N = 200$.

3. **Teste de Hipóteses:** Estimaremos a probabilidade de erro do tipo I e erro do tipo II para $\beta_2$. Em ambos os casos geraremos 100 vezes amostras de tamanho variados (50, 100 e 200) calculando as estimativas e contabilizando quantos destes rejeitamos. Repetimos este processo 1000 vezes e tiramos sua média e desvio padrão da probabilidade do erro. No caso do erro do tipo II trabalhamos com $\beta_2$ iguais a $0.66, 0.72, 0.78, 0.84$, e $0.90$.

Trabalharemos com o modelo (3.2)com $k = 1$ sendo os verdadeiros valores dos parâmetros são $\beta_1 = -6, \beta_2 = 0.6, \sigma_e = \Sigma_u = 18, \mu_x = 20$ e $\Sigma_x = 180$. Estes valores são os mesmo utilizados por Wang (1998). Para os estimadores $\tau$, calculamos as médias e o MAD (*Mean absolute deviations*), dado pela fórmula:

$$MAD(\delta) = E[|\delta - \delta_0|]$$

onde $\delta_0$ é o verdadeiro valor do parâmetro.

Os estimadores foram calculados pelas fórmulas (3.7)-(3.10).

## 4.1 Tamanho amostral

Na Tabela 4.1 observamos que em média as estimativas do parâmetro $\beta_1$ são subestimadas nos casos $\Sigma_u, \sigma_e$, e $\Delta$, enquanto no caso $\Sigma_x$ ela é sobreestimada para tamanhos amostrais pequenos/moderados. Os dois casos de menores MAD foram em $\Delta$ e $\Sigma_u$, respectivamente, para todos os tamanhos amostrais. Os casos $\sigma_e$ e $\Sigma_x$ se alternaram entre o terceiro e quarto no tamanho amostral 100. Porém $\Sigma_x$ teve erros menores do verdadeiro valor do parâmetro na maioria dos tamanhos amostrais. No caso de aplicar somente o modelo de Tobit sem erros(M.T.) nas variáveis observamos viés, sobreestimando o valor do parâmetro, não melhorando com o aumento do tamanho amostral (inconsistência).

Tabela 4.1: Simulação para diferentes tamanhos amostrais

| | Conhecido | | N | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 20 | 40 | 50 | 100 | 200 | 400 |
| $\beta_1$ | $\Sigma_u$ | média | -6,5577 | -6,5926 | -6,2687 | -6,1256 | -6,0923 | -6,0397 |
| | | MAD | 2,5582 | 1,9208 | 1,7689 | 1,1501 | 0,8389 | 0,5625 |
| | $\Sigma_x$ | média | -5,6847 | -5,9863 | -5,9453 | -5,9982 | -6,0084 | -6,0445 |
| | | MAD | 3,9114 | 2,8281 | 2,6143 | 1,8830 | 1,2693 | 0,8908 |
| | $\sigma_e$ | média | -6,5127 | -6,5104 | -6,3886 | -6,1915 | -6,1417 | -6,0553 |
| | | MAD | 4,2358 | 3,0185 | 2,6920 | 1,8229 | 1,3021 | 0,9270 |
| | $\Delta$ | média | -6,2964 | -6,4654 | -6,1768 | -6,0865 | -6,0717 | -6,0342 |
| | | MAD | 2,4480 | 1,8801 | 1,7405 | 1,1460 | 0,8285 | 0,5643 |
| | M.T. | média | -4,9892 | -5,2086 | -4,9702 | -4,9366 | -4,9385 | -4,9223 |
| | | MAD | 2,2184 | 1,6296 | 1,6451 | 1,3192 | 1,1376 | 1,0908 |
| $\beta_2$ | $\Sigma_u$ | média | 0,6219 | 0,6201 | 0,6106 | 0,6030 | 0,6027 | 0,6012 |
| | | MAD | 0,1042 | 0,0779 | 0,0697 | 0,0458 | 0,0334 | 0,0223 |
| | $\Sigma_x$ | média | 0,5796 | 0,5907 | 0,5950 | 0,5961 | 0,5986 | 0,6015 |
| | | MAD | 0,1808 | 0,1301 | 0,1174 | 0,0855 | 0,0587 | 0,0405 |
| | $\beta_1$ | média | 0,5959 | 0,5911 | 0,5979 | 0,5971 | 0,5982 | 0,5993 |
| | | MAD | 0,0517 | 0,0380 | 0,0349 | 0,0239 | 0,0170 | 0,0117 |
| | $\sigma_e$ | média | 0,6191 | 0,6161 | 0,6174 | 0,6060 | 0,6051 | 0,6019 |
| | | MAD | 0,1882 | 0,1330 | 0,1192 | 0,0812 | 0,0577 | 0,0413 |
| | $\Delta$ | média | 0,6090 | 0,6139 | 0,6061 | 0,6010 | 0,6017 | 0,6009 |
| | | MAD | 0,0981 | 0,0757 | 0,0676 | 0,0455 | 0,0330 | 0,0223 |
| | M.T. | média | 0,5484 | 0,5538 | 0,5480 | 0,5447 | 0,5458 | 0,5457 |
| | | MAD | 0,0912 | 0,0699 | 0,0681 | 0,0599 | 0,0551 | 0,0545 |

Nas médias (estimativas) para cada caso de $\beta_2$ temos que $\Sigma_u, \sigma_e$ e $\Delta$ sobreestima, enquanto $\Sigma_x$ e $\beta_1$ sobreestima. O caso onde $\beta_1$ é conhecido observamos menor viés na maioria dos tamanhos amostrais, como também menor MAD entre os outros casos. Já $\Sigma_u$ e $\Delta$ apresentam comportamentos parecidos em relação ao MAD, sendo que no caso $\Delta$ conhecido temos um menor viés. O $\Sigma_x$ e o $\sigma_e$ apresentam MAD parecidos, sendo que $\Sigma_x$ na maioria dos tamanhos

amostrais teve menor viés que $\sigma_e$. No caso M.T. observamos viés próximo de 0.05 em todos os casos e este viés não diminui com o aumento da amostra.

## 4.2 Imprecisão do parâmetro conhecido

Tabela 4.2: Simulação para imprecisão do parâmetro conhecido

| | | | Imprecisão | | | | |
|---|---|---|---|---|---|---|---|
| | | | $-20\%$ | $-10\%$ | 0% | $+10\%$ | $+20\%$ |
| $\beta_1$ | $\Sigma_u$ | média | -5,8755 | -5,8739 | -6,0820 | -6,1610 | -6,2941 |
| | | MAD | 0,8103 | 0,8097 | 0,8061 | 0,8226 | 0,9220 |
| | $\Sigma_x$ | média | -9,1847 | -7,3061 | -5,9597 | -4,8369 | -3,9618 |
| | | MAD | 3,2340 | 1,6872 | 1,2294 | 1,5448 | 2,1554 |
| | $\sigma_e$ | média | -6,7745 | -6,3547 | -6,0788 | -5,7186 | -5,3889 |
| | | MAD | 1,4297 | 1,3676 | 1,2356 | 1,3241 | 1,4131 |
| | $\Delta$ | média | -5,8866 | -5,8719 | -6,0587 | -6,1222 | -6,2403 |
| | | MAD | 0,8122 | 0,8051 | 0,7983 | 0,8150 | 0,9002 |
| $\beta_2$ | $\Sigma_u$ | média | 0,5913 | 0,5941 | 0,6026 | 0,6071 | 0,6137 |
| | | MAD | 0,0326 | 0,0318 | 0,0327 | 0,0326 | 0,0375 |
| | $\Sigma_x$ | média | 0,7567 | 0,6657 | 0,5964 | 0,5409 | 0,4972 |
| | | MAD | 0,1588 | 0,0815 | 0,0564 | 0,0735 | 0,1066 |
| | $\beta_1$ | média | 0,5376 | 0,5704 | 0,5985 | 0,6293 | 0,6593 |
| | | MAD | 0,0624 | 0,0309 | 0,0155 | 0,0308 | 0,0594 |
| | $\sigma_e$ | média | 0,6363 | 0,6181 | 0,6025 | 0,5848 | 0,5684 |
| | | MAD | 0,0643 | 0,0606 | 0,0553 | 0,0587 | 0,0649 |
| | $\Delta$ | média | 0,5919 | 0,5940 | 0,6014 | 0,6052 | 0,6110 |
| | | MAD | 0,0324 | 0,0316 | 0,0321 | 0,0320 | 0,0363 |

Na Tabela 4.2 observamos as estimativas de $\tau$ quando temos uma imprecisão do parâmetro conhecido em cada caso de 20% e 10% para mais ou para menos. Para comparação usamos o mesmo tamanho amostral ($N = 200$), e colocamos também o verdadeiro valor de cada caso (0% de imprecisão). Todos os casos foram sensíveis a mudança de seu verdadeiro valor. A menor amplitude de mudança foi em $\Delta, \Sigma_u$ e $\beta_1$, conhecidos. No caso de $\Sigma_u$ conhecido não houve grandes mudanças nos MAD´s das estimativas de $\beta_1$ e $\beta_2$.

## 4.3 Teste de hipóteses

Testaremos através do teste de Wald, a 5% de significância, as hipóteses:

- **Hipótese Nula** : $\beta_2 = 0,6$
- **Hipótese Alternativa** : $\beta_2 \neq 0,6$

Trabalharemos testando o parâmetro $\beta_2$ para assim comparar todos os casos.

### 4.3.1 Erro Tipo I

A probabilidade do Erro do Tipo I (nível de significância) será estimada pela média de 1000 repetições de 100 simulações do modelo sob a hipótese nula, e calculando nestas 100 simulações a proporção de que o teste rejeitou a hipótese nula. Calculamos o desvio padrão destas 1000 repetições para obter uma noção da variabilidade do erro. Como estamos gerando as observações sob a hipótese nula esperamos que, em torno de 5% das simulações, a hipótese nula seja rejeitada. A tabela 4.3 mostra os resultados.

Tabela 4.3: Erro Tipo I

| Coeficiente | Conhecido | Tamanho | | | | | |
|---|---|---|---|---|---|---|---|
| | | 50 | | 100 | | 200 | |
| | | Média | DP | Média | DP | Média | DP |
| $\beta_2$ | $\Sigma_u$ | 0,0656 | 0,0241 | 0,0587 | 0,0233 | 0,0537 | 0,0222 |
| | $\Sigma_x$ | 0,0935 | 0,0288 | 0,0726 | 0,0270 | 0,0616 | 0,0244 |
| | $\beta_1$ | 0,0459 | 0,0210 | 0,0481 | 0,0219 | 0,0483 | 0,0213 |
| | $\sigma_e$ | 0,0500 | 0,0223 | 0,0491 | 0,0214 | 0,0489 | 0,0219 |
| | $\Delta$ | 0,0647 | 0,0245 | 0,0590 | 0,0241 | 0,0534 | 0,0221 |

Observamos que, na maioria dos diferentes tamanhos amostrais, os casos $\beta_1$ e $\sigma_e$ conhecido se comportaram mais próximos do teórico. Já os casos de conhecido $\Sigma_u$ e $\Delta$ comportam-se de maneiras semelhantes, e ambos aproximam do teórico a partir de $N = 100$. O caso de $\Sigma_x$ conhecido é o caso que mais distânte do teórico em todos os tamanhos amostrais.

### 4.3.2 Erro Tipo II

A probabilidade do Erro do Tipo II será estimada pela média de 1000 repetições de 100 simulações do modelo sob a hipótese alternativa, e calculando nestas 100 simulações a proporção de que o teste aceitou a hipótese nula. Calculamos o desvio padrão destas 1000 repetições para obter uma noção da variabilidade do erro. Variamos o $\beta_2$ de 0.66 a 0.90, de 10 em 10 porcento do valor verdadeiro. A tabela 4.4 mostra os resultados.

Tabela 4.4: Erro Tipo II

| $\beta_2$ | Conhecido | Tamanho | | | | | |
|---|---|---|---|---|---|---|---|
| | | 50 | | 100 | | 200 | |
| | | Média | DP | Média | DP | Média | DP |
| 0.66 | $\Sigma_u$ | 0,8803 | 0,0324 | 0,8170 | 0,0389 | 0,6899 | 0,0459 |
| | $\Sigma_x$ | 0,9428 | 0,0245 | 0,9427 | 0,0246 | 0,9193 | 0,0266 |
| | $\beta_1$ | 0,6926 | 0,0465 | 0,4760 | 0,0479 | 0,1980 | 0,0398 |
| | $\sigma_e$ | 0,9641 | 0,0199 | 0,9414 | 0,0232 | 0,9001 | 0,0300 |
| | $\Delta$ | 0,8894 | 0,0309 | 0,8247 | 0,0383 | 0,6931 | 0,0444 |
| 0.72 | $\Sigma_u$ | 0,6717 | 0,0455 | 0,4462 | 0,0495 | 0,1744 | 0,0376 |
| | $\Sigma_x$ | 0,9405 | 0,0241 | 0,8919 | 0,0303 | 0,7526 | 0,0429 |
| | $\beta_1$ | 0,2311 | 0,0414 | 0,0429 | 0,0203 | 0,0007 | 0,0026 |
| | $\sigma_e$ | 0,9315 | 0,0252 | 0,8436 | 0,0368 | 0,6800 | 0,0466 |
| | $\Delta$ | 0,6863 | 0,0450 | 0,4562 | 0,0495 | 0,1686 | 0,0366 |
| 0.78 | $\Sigma_u$ | 0,3936 | 0,0482 | 0,1346 | 0,0348 | 0,0118 | 0,0112 |
| | $\Sigma_x$ | 0,9103 | 0,0283 | 0,7746 | 0,0416 | 0,4927 | 0,0482 |
| | $\beta_1$ | 0,0470 | 0,0211 | 0,0014 | 0,0038 | 0,0000 | 0,0000 |
| | $\sigma_e$ | 0,8571 | 0,0336 | 0,6794 | 0,0463 | 0,3966 | 0,0505 |
| | $\Delta$ | 0,3975 | 0,0482 | 0,1319 | 0,0328 | 0,0098 | 0,0102 |
| 0.84 | $\Sigma_u$ | 0,1765 | 0,0376 | 0,0260 | 0,0160 | 0,0004 | 0,0019 |
| | $\Sigma_x$ | 0,8546 | 0,0351 | 0,6157 | 0,0518 | 0,2492 | 0,0455 |
| | $\beta_1$ | 0,0088 | 0,0095 | 0,0000 | 0,0006 | 0,0000 | 0,0000 |
| | $\sigma_e$ | 0,7541 | 0,0419 | 0,4958 | 0,0503 | 0,1777 | 0,0382 |
| | $\Delta$ | 0,1755 | 0,0370 | 0,0218 | 0,0145 | 0,0002 | 0,0015 |
| 0.90 | $\Sigma_u$ | 0,0652 | 0,0246 | 0,0043 | 0,0065 | 0,0000 | 0,0003 |
| | $\Sigma_x$ | 0,7756 | 0,0420 | 0,4456 | 0,0499 | 0,1008 | 0,0290 |
| | $\beta_1$ | 0,0020 | 0,0044 | 0,0000 | 0,0003 | 0,0000 | 0,0000 |
| | $\sigma_e$ | 0,6439 | 0,0485 | 0,3321 | 0,0477 | 0,0655 | 0,0255 |
| | $\Delta$ | 0,0608 | 0,0228 | 0,0032 | 0,0057 | 0,0000 | 0,0003 |

Observamos que no caso de $\beta_1$ conhecido temos menor Erro do tipo II, isto é, tem um poder do teste maior que os outros casos, em todos os tamanhos amostrais apresentados. Os casos $\Sigma_u$ e $\Delta$ têm aproximadamente o mesmo erro, mas o erro se tornar proximos ou menores de 5% em $\beta_2 = 0.78$ com $N = 200$, $\beta_2 = 0.84$ com $N = (100, 200)$ e para $\beta_2 = 0.78$ com todos os $N's$. Já os casos $\Sigma_x$ e $\sigma_e$ só tem um bom poder em $\beta_2 = 0.9$ com $N = 200$, sendo que o poder do caso $\sigma_e$ é maior que do $\Sigma_x$.

# Capítulo 5

# Conclusões e estudos futuros

Nesse trabalho estendemos os casos para o modelo de Tobit com erros nas variáveis, trabalhando com casos de conhecimento de outros parâmetros,como $\Sigma_u, \Sigma_x, \sigma_e$ e $\beta_1$ além do $\Delta$ já estudado por Wang (1998). Apresentamos para cada caso os seus estimadores e a normalidade assintótica. Desenvolvemos na linguagem R rotinas para o cálculo das estimativas do modelo de Tobit para cada caso, com $k = 1$ podendo ser facilmente generalizado para outros k's.

Com relação as estimativas observamos no capítulo de simulação que os caso de menor viés e melhor poder(menor erro do tipo II), quando $k = 1$, $\beta_1$. Para $k$ maior que um, os dois melhores casos de conhecimento de parâmetros seriam o $\Delta$, e $\Sigma_u$, sendo estes de comportamentos muito próximos. O caso de $\Sigma_x$ e $\sigma_e$ podem ser usados, para obter bons resultados, em grandes amostras ($N > 200$).

Como estudos futuros, uma das propostas é aplicar o algoritmo EM para calcular as estimativas, e também fazer a parte de análise de resíduos deste modelo.

# Apêndice A

# Apêndice

## A.1 Teorema de Fuller

**Teorema 4** *Seja $g(\boldsymbol{a})$ uma função continua com valor real do vetor $\boldsymbol{a}$ de dimensão $k$ sendo $\boldsymbol{a}$ um elemento de dimensão $k$ no espaço Euclidiano. Assumindo que $g(a)$ tenha as primeiras derivadas continuas no ponto $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_k)'$. Seja o vetor de variáveis aleatórias $\boldsymbol{X_t} = (X_{t1}, X_{t2}, \ldots, X_{tk})'$, $t = 1, 2, \ldots$ independentes e identicamente distribuídas com média $\boldsymbol{\mu}$ e matriz de covariância $\boldsymbol{\Sigma_{XX}}$ . Seja:*

$$\overline{\boldsymbol{X}} = n^{-1} \sum_{t=1}^{n} \boldsymbol{X}_t$$

*Então*

$$n^{1/2}[g(\overline{\boldsymbol{X}}) - g(\boldsymbol{\mu})]$$

*converge em distribuição para uma variável aleatória normal com média zero e matriz de covariância*

$$[g^{(1)}(\boldsymbol{\mu}), g^{(2)}(\boldsymbol{\mu}), \ldots, g^{(k)}(\boldsymbol{\mu})]\boldsymbol{\Sigma_{XX}}[g^{(1)}(\boldsymbol{\mu}), g^{(2)}(\boldsymbol{\mu}), \ldots, g^{(k)}(\boldsymbol{\mu})]',$$

*onde* $g^{(i)}(\boldsymbol{\mu})$ *é a derivada parcial de* $g(\boldsymbol{a})$ *em relação a* $a_i$ *calculada em* $\boldsymbol{a} = \boldsymbol{\mu}$.

## A.2 Rotinas no R

Para simular uma amostra para o modelo de Tobit com erros na variáveis:

```
#### Função para simular o modelo
# n  = tamanho da amostra
# mu = a média de X
# sig = variancias dos erros e de X
# bet = os parametros do modelo
####

amostra<-function(n,mu,sig,bet)
{
e <- rnorm(n,0,sqrt(sig[1]))
u <- rnorm(n,0,sqrt(sig[2]))
x <- rnorm(n,mu,sqrt(sig[3]))
w <- x+u
y <- bet[1]+x*bet[2]+e
ym <- y
ym[ym<=0] <- 0
return(list(ym=ym,y=y,w=w,x=x,e=e,u=u))
}

####
```

Para os estimadores de momentos na primeira fase do processo e calculo da matriz **A** e de variância e covarincia de $z$:

```
##### Função para estimadores de momentos

# Nesta função ym é o eta, e w é ele mesmo do modelo discutido no dissertação
```

```
# Quando a variável conter o número 1 (um) no final significa que estamos trabalhando
# ou w´s positivos.

est_t<-function(ym,w)
{

nt1 <- sum(ym>0)

nt <- length(ym)

correcao <- (nt-1)/nt

correcao1 <- (nt1-1)/nt1

mu_eta <- mean(ym)

mu_w <- mean(w)

sigma_w <- var(w)*correcao

mu_eta1 <- mean(ym[ym>0]) #y_eta+

mu2_eta1 <- mean((ym[ym>0])^2) #mu^2_eta+

sigma_eta1 <- var(ym[ym>0])*correcao1

mu_w1 <- mean(w[ym>0])

mu2_w1 <- mean((w[ym>0])^2)
```

```
mu2_w <- mean(w^2)

mu_w_eta21 <- mean((w*ym^2)[ym>0]) #mu_w_eta2+

mu_w_eta1 <- mean((w*ym)[ym>0]) #mu_w_eta+

mu_w2_eta1 <- mean((w^2*ym)[ym>0]) #mu_w2_eta+

sigma_w_eta1 <- var(ym[ym>0]*w[ym>0])*correcao1 #sigma_w_eta+

mu_w3_eta1 <- mean((w[ym>0])^3*ym[ym>0]) #mu_w3_eta+

mu_w3 <- mean(w^3)

sigma_w2 <- var(w^2)*correcao


sigma_eta_w <- var(cbind(ym,w))[1,2]*correcao



delta <- qnorm(nt1/nt)
mu_y <- (delta*mu_eta1)/(delta+dnorm(delta)/pnorm(delta))
sigma_y <- (mu_y/delta)^2
sigma_w_y <- mu_w_eta1 - mu_w*mu_eta1

return(data.frame(mu_eta, mu_w, sigma_w, mu_eta1, mu2_eta1, sigma_eta1,
mu_w1, mu2_w1, mu2_w, mu_w_eta21, mu_w_eta1, mu_w2_eta1, sigma_w_eta1,
mu_w3_eta1, mu_w3, sigma_w2, sigma_eta_w, nt1, nt, delta, mu_y,
```

```
sigma_y, sigma_w_y))
}

#######################

#aux_est <- est_t
#attach(aux_est)


#### Função para calcular as Matriz A e a matriz de covariancia-variancia de Z

calcular_matrix <- function(est)
{
attach(est)

Sz <- A <- matrix(0,5,5)

A[4,4]<-A[5,5]<-1

A[1,1] <- mu2_eta1/mu_eta

A[1,2] <- mu_y/(mu_eta1*pnorm(delta))

A[2,1] <- 2*sigma_y*(mu2_eta1/(mu_y*mu_eta)-1/(delta*dnorm(delta)))

A[2,2] <- 2*sigma_y/(mu_eta1*pnorm(delta))

A[3,2] <- -mu_w/pnorm(delta)

A[3,3] <- 1/pnorm(delta)
```

```
A[3,4] <- -mu_eta1

A[5,4] <- -2*mu_w


Sz[1,1] <- 1-pnorm(delta)

Sz[2,2] <- sigma_eta1

Sz[3,3] <- sigma_w_eta1

Sz[4,4] <- sigma_w/pnorm(delta)

Sz[5,5] <- sigma_w2/pnorm(delta)


Sz[1,2] <- Sz[2,1] <- 0

Sz[1,3] <- Sz[3,1] <- 0

Sz[1,4] <- Sz[4,1] <- mu_w1-mu_w

Sz[1,5] <- Sz[5,1] <- mu2_w1-mu2_w


Sz[2,3] <- Sz[3,2] <- mu_w_eta21-mu_eta1*mu_w_eta1

Sz[2,4] <- Sz[4,2] <- mu_w_eta1-mu_w1*mu_eta1
```

```
Sz[2,5] <- Sz[5,2] <- mu_w2_eta1-mu2_w1*mu_eta1


Sz[3,4] <- Sz[4,3] <- mu_w2_eta1-mu_w1*mu_w_eta1

Sz[3,5] <- Sz[5,3] <- mu_w3_eta1-mu2_w1*mu_w_eta1


Sz[4,5] <- Sz[5,4] <- (mu_w3-mu2_w*mu_w)/pnorm(delta)

Sz<-pnorm(delta)*Sz

detach(est)

S_phi <- A%*%Sz%*%t(A)

return(list(A,Sz,S_phi))
}

####


#### Função para fazer o teste de Wald a partir de
# n     =  tamanho da amostra
# mu    =  o estimador da amostra
# mu_0  =  o valor do paramentro na hipotese nula
# dif   =  Derivado do caso
# var_s =  matrix de variancia
####
```

```
twald <- function(n,mu,mu_0,dif,var_s)
{
aux <- (mu-mu_0)^2/((dif%*%var_s%*%dif)/n)
p.valor<-pchisq(aux,1,lower.tail=F)
return(c(aux,p.valor))
}

####
```

Para o calculo de $\beta_1$ e $\beta_2$ com diferentes tamanhos amostrais:

```
n<-c(20,40,50,100,200,400)

hh<-list()

for(j in 1:6)
{

beta_su <- 0
alfa_su <- 0

beta_sx <- 0
alfa_sx <- 0

beta_alf <- 0

beta_se <- 0
alfa_se <- 0

beta_delta <- 0
alfa_delta <- 0
```

```
est1 <- matrix(0,nrow=1000,ncol=8)

i <- 1
repeat
{


m <- 20
sig <- c(18,18,180)
beta <- c(-6,0.6)

mod1 <- amostra(n[j],m,sig,beta)

est1[i,] <- est(mod1$ym,mod1$w)


beta_su[i] <- (solve(est1[i,3]-sig[2])*est1[i,23])[1,1]
alfa_su[i] <- est1[i,21]-beta_su[i]*est1[i,2]

beta_sx[i] <- (solve(sig[3])*est1[i,23])[1,1]
alfa_sx[i] <- est1[i,21]-beta_sx[i]*est1[i,2]

beta_alf[i] <- (est1[i,21]-beta[1])/est1[i,2]

beta_se[i] <- (est1[i,22]-sig[1])/est1[i,23]
alfa_se[i] <- est1[i,21]-beta_se[i]*est1[i,2]

beta_delta[i] <- solve(est1[i,3]*solve(diag(1)+0.1))*est1[i,23]
alfa_delta[i] <- est1[i,21]-beta_delta[i]*est1[i,2]
```

```
fit <- vglm(mod1$ym ~ mod1$w, tobit(Lower=0, Upper=Inf))
beta_sem[i] <- coef(fit,matrix=T)[2,1]
alfa_sem[i] <- coef(fit,matrix=T)[1,1]

if(est1[i,7]/est1[i,8]<1&est1[i,7]/est1[i,8]>0&est1[i,6]!=0) i <- i+1


if(i>1000) break()
}
names(est1)<-names(est_t(1,1))
hh[[j]]<-data.frame(alfa_su,alfa_sx,alfa_se,alfa_delta,alfa_sem,beta_su,beta_sx,beta_
}
```

Para o caso com imprecisão nos casos:

```
n<-200

er<-c(0.8,0.9,1,1.1,1.2)

hh1<-list()

for(j in 1:1)
{
for(k in 1:5)
{

beta_su <- 0
alfa_su <- 0

beta_sx <- 0
```

```
alfa_sx <- 0

beta_alf <- 0

beta_se <- 0
alfa_se <- 0

beta_delta <- 0
alfa_delta <- 0

est1 <- matrix(0,nrow=1000,ncol=8)

i <- 1
repeat
{


m <- 20
sig <- c(18,18,180)
beta <- c(-6,0.6)

mod1 <- amostra(n[j],m,sig,beta)

est1[i,] <- est(mod1$ym,mod1$w)


beta_su[i] <- (solve(est1[i,3]-er[k]*sig[2])*est1[i,23])[1,1]
alfa_su[i] <- est1[i,21]-beta_su[i]*est1[i,2]

beta_sx[i] <- (solve(er[k]*sig[3])*est1[i,23])[1,1]
```

```
alfa_sx[i] <- est1[i,21]-beta_sx[i]*est1[i,2]

beta_alf[i] <- (est1[i,21]-er[k]*beta[1])/est1[i,2]

beta_se[i] <- (est1[i,22]-er[k]*sig[1])/est1[i,23]
alfa_se[i] <- est1[i,21]-beta_se[i]*est1[i,2]

beta_delta[i] <- solve(est1[i,3]*solve(diag(1)+er[k]*0.1))*est1[i,23]
alfa_delta[i] <- est1[i,21]-beta_delta[i]*est1[i,2]


if(est1[i,7]/est1[i,8]<1&est1[i,7]/est1[i,8]>0&est1[i,6]!=0) i <- i+1


if(i>1000) break()
}

hh1[[k]]<-data.frame(alfa_su,alfa_sx,alfa_se,alfa_delta,beta_su,beta_sx,beta_alf,beta
}
}
```

# Referências Bibliográficas


Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal, *Econometrica* **41**(6): 997–1016.

Azevedo, A. F. Z. d. (2004). O efeito do mercosul sobre o comércio: uma análise com o modelo gravitacional, *Pesquisa e Planejamento Econômico* **34**: 307–339.

Fair, R. C. (1978). A note on the computation of the tobit estimator, *Econometrica* **45**(7): 1723–1727.

Fütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer.

Fuller, W. A. (1987). *Measurement Error Models*, Wiley.

Olsen, R. J. (1978). Note on the uniqueness of the maximum likelihood estimator for the tobit model, *Econometrica* **46**(5): 1211–1215.

Souza, G. d. S. e. S. (2006). Significância de efeitos técnicos na eficiência de produção da pesquisa agropecuária, *Revista Brasileira de Economia* **60**: 69–86.

Tobin, J. (1958). Estimation of relationships for limited dependent variables, *Econometrica* **26**(1): 24–36.

Wang, L. (1998). Estimation of censored linear errors-in-variables models, *Journal of Econometrics* **84**: 383–400.