Aula 2.2: Variáveis Truncadas

Econometria I - IBMEC

Marcus L. Nascimento

29 de outubro de 2025

1. Introdução

2. Estimação e Inferência

3. Interpretação dos Coeficientes

Introdução

Introdução

- Assim como nos casos em que $Y \in \mathrm{R}$ e $Y \in \{0, 1\}$, continuamos interessados em aprender sobre Y a partir de um conjunto de variáveis independentes $X_1, X_2, \ldots, X_p$.

- Por conveniência, introduziremos uma notação vetorial, na qual

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \quad \text{e} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}.$$

Logo, $\boldsymbol{x}'\boldsymbol{\beta} = \beta_0 + x_1\beta_1 + x_2\beta_2 + \ldots + x_p\beta_p$.

Introdução

- Nesta etapa do curso, estudaremos o caso em que Y é essencialmente uma variável contínua não negativa que assume valores em zero com probabilidade positiva.

    - Renda oriunda do trabalho;
    - Número de horas trabalhadas;
    - Despesa com bebida alcoólica.

Introdução

- No caso descrito anteriormente, o **modelo Tobit** é comumente aplicado. O modelo expressa a variável resposta em termos de uma variável latente:

$$
\begin{aligned}
y &= \max\{0, y^*\} \\
y^* &= \boldsymbol{x}'\boldsymbol{\beta} + \varepsilon \\
\varepsilon | \boldsymbol{x} &\sim N(0, \sigma^2).
\end{aligned}
$$

- $y = \max\{0, y^*\}$ implica que

$$
y = \begin{cases} y^*, & \text{se } y^* \geq 0 \\ 0, & \text{se } y^* < 0. \end{cases}
$$

- A variável latente y^* satisfaz as suposições do modelo linear clássico, em particular: distribuição normal e homocedástica com média condicional linear.

Estimação e Inferência

Estimação por Máxima Verossimilhança

- Para descrever a função de verossimilhança, iremos inicialmente descrever a densidade da variável resposta condicional nas variáveis independentes separando em dois casos:

  - $Y > 0$:

$$
\begin{aligned}
f(y|\boldsymbol{x}) &= (2\pi\sigma^2)^{-1/2}\exp\left\{-\frac{(y-\mathbf{x}'\boldsymbol{\beta})^2}{2\sigma^2}\right\} \\
&= \sigma^{-1}\phi\left(\frac{y-\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right).
\end{aligned}
$$

  - $Y = 0$:

$$
\begin{aligned}
\mathrm{P}(Y=0|\boldsymbol{x}) &= \mathrm{P}(Y^* < 0|\boldsymbol{x}) \\
&= \mathrm{P}(\varepsilon < -(\boldsymbol{x}'\boldsymbol{\beta})|\boldsymbol{x}) \\
&= \mathrm{P}(\varepsilon/\sigma < -(\boldsymbol{x}'\boldsymbol{\beta})/\sigma|\boldsymbol{x}) \\
&= \Phi(-(\boldsymbol{x}'\boldsymbol{\beta})/\sigma) \\
&= 1-\Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma).
\end{aligned}
$$

Estimação por Máxima Verossimilhança

Supondo $(y_i, x_{i1}, \ldots, x_{ip})$ uma amostra aleatória da população, temos que a função densidade condicional é dada por:

$$f(y_i|\boldsymbol{x}_i) \quad = \quad [1 - \Phi((\boldsymbol{x}_i'\boldsymbol{\beta})/\sigma)]^{1(y_i=0)} \times \left[\sigma^{-1}\phi\left(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)\right]^{1(y_i>0)}.$$

A função **log-verossimilhança**, por sua vez, é dada por:

$$\ell_i(\boldsymbol{\beta}, \sigma) \quad = \quad 1(y_i = 0) \cdot \ln[1 - \Phi((\boldsymbol{x}_i'\boldsymbol{\beta})/\sigma)] + 1(y_i > 0) \cdot \ln\left[\sigma^{-1}\phi\left(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)\right].$$

Estimação por Máxima Verossimilhança

Considerando uma amostra aleatória de tamanho n, a função **log-verossimilhança** é obtida ao tomarmos a soma em i, $i = 1, 2, \ldots, n$:

$$
\begin{aligned}
\ell(\boldsymbol{\beta}, \sigma) &= \sum_{i=1}^{n} \ell_i(\boldsymbol{\beta}, \sigma) \\
&= \sum_{i=1}^{n} \left\{ 1(y_i = 0) \cdot \ln[1 - \Phi((\boldsymbol{x}_i' \boldsymbol{\beta})/\sigma)] + 1(y_i > 0) \cdot \ln\left[\sigma^{-1} \phi\left(\frac{y_i - \boldsymbol{x}_i' \boldsymbol{\beta}}{\sigma}\right)\right]\right\}.
\end{aligned}
$$

O estimador de máxima verossimilhança de β_j, $j = 0, 1, \ldots, p$, denotado por $\hat{\beta}_j$, é obtido através da maximização da log-verossimilhança.

Estimação por Máxima Verossimilhança

OBSERVAÇÕES:

- Assim como nos casos logit e probit, não é possível obter fórmulas fechadas para os estimadores do modelo tobit.
  - Métodos numéricos são aplicados para encontrar as soluções do problema.
- Vale recordar também que sob condições bastante gerais, o estimador de máxima verossimilhança é
  - Consistente;
  - Assintoticamente normal;
  - Assintoticamente eficiente.
- A partir dos resultados de normalidade assintótica, é possível testar a hipótese $\mathrm{H}_0 : \beta_j = 0$ de maneira usual utilizando uma estatística t $\hat{\beta}_j/\mathrm{se}(\hat{\beta}_j)$.

Teste de Razão de Verossimilhança

- Nos casos em que desejamos testar hipóteses acerca de um grupo de variáveis:

$$\begin{aligned} H_0 &: \beta_{p-q+1} = \beta_{p-q+2} = \ldots = \beta_p = 0 \\ H_1 &: H_0 \text{ não é verdadeira,} \end{aligned}$$

  onde $q < p$ é o tamanho do grupo a ser testado, podemos aplicar o teste da razão de verossimilhanças.

- Como vimos, a estatística de razão de verossimilhança é dada por:

$$LR = 2[\ell_{ur} - \ell_r]$$

  onde ℓ_{ur} é o valor da log-verossimilhança para o modelo irrestrito e ℓ_r é o valor da log-verossimilhança para o modelo restrito.

- LR possui aproximadamente uma distribuição χ^2_q.

Interpretação dos Coeficientes

Valor Esperado (Condicional e Incondicional)

- É importante observar que β_j mede o efeito parcial de X_j em $\mathrm{E}(Y^*|\boldsymbol{X} = \boldsymbol{x})$, onde Y^* é a variável latente.

- Estamos interessados, no entanto, em explicar a variável Y e precisamos caracterizar o valor esperado de Y como função de $X_1, X_2, \ldots, X_p$.

$$\begin{aligned}
\mathrm{E}(Y|\boldsymbol{X} = \boldsymbol{x}) &= \mathrm{P}(Y = 0|\boldsymbol{X} = \boldsymbol{x})\mathrm{E}(Y|Y = 0, \boldsymbol{X} = \boldsymbol{x}) \\
&+ \mathrm{P}(Y > 0|\boldsymbol{X} = \boldsymbol{x})\mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x}) \\
&= \mathrm{P}(Y > 0|\boldsymbol{X} = \boldsymbol{x})\mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x}) \\
&= \Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)\mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x}).
\end{aligned}$$

- A partir da equação acima, notamos que para caracterizar $\mathrm{E}(Y|\boldsymbol{X} = \boldsymbol{x})$ é necessário descrever $\mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x})$.

Valor Esperado (Condicional e Incondicional)

Para obter $\mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x})$, aplicamos o seguinte resultado para variáveis aleatórias com distribuição normal: Se $Z \sim N(0, 1)$, então $\mathrm{E}(Z|Z > c) = \phi(c)/[1 - \Phi(c)]$ para qualquer constante c.

$$
\begin{aligned}
\mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x}) &= \boldsymbol{x}'\boldsymbol{\beta} + \mathrm{E}(\varepsilon|\varepsilon > -(\boldsymbol{x}'\boldsymbol{\beta})) \\
&= \boldsymbol{x}'\boldsymbol{\beta} + \sigma\mathrm{E}(\varepsilon/\sigma|\varepsilon/\sigma > -(\boldsymbol{x}'\boldsymbol{\beta})/\sigma) \\
&= \boldsymbol{x}'\boldsymbol{\beta} + \sigma\phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)/\Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma),
\end{aligned}
$$

já que $\phi(-c) = \phi(c)$, $1 - \Phi(-c) = \Phi(c)$ e ε/σ possui uma distribuição normal padrão independente de $X_1, \ldots, X_p$.

Valor Esperado (Condicional e Incondicional)

Em resumo, temos que

$$\mathrm{E}(Y|Y>0, \boldsymbol{X}=\boldsymbol{x}) \quad = \quad \boldsymbol{x}'\boldsymbol{\beta} + \sigma\lambda((\boldsymbol{x}'\boldsymbol{\beta})/\sigma),$$

onde $\lambda(c) = \phi(c)/\Phi(c)$ é denominada razão de Mills inversa.

OBSERVAÇÕES:

- O valor esperado condicional em $Y > 0$ é $\boldsymbol{x}'\boldsymbol{\beta}$ mais um termo estritamente positivo;
- É possível visualizar por que utilizar o OLS para as observações $y_i > 0$, $i = 1, \ldots, n$, nem sempre estimará $\beta_1, \beta_2, \ldots, \beta_p$.
  - A razão de Mills inversa é uma variável omitida, geralmente correlacionada com elementos de $X_1, \ldots, X_p$.

Valor Esperado (Condicional e Incondicional)

Por fim, a partir de $\mathrm{E}(Y|Y>0,\boldsymbol{X}=\boldsymbol{x})$, conseguimos caracterizar $\mathrm{E}(Y|\boldsymbol{X}=\boldsymbol{x})$:

$$\begin{aligned}
\mathrm{E}(Y|\boldsymbol{X}=\boldsymbol{x}) &= \Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)\times[\boldsymbol{x}'\boldsymbol{\beta}+\sigma\lambda((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)] \\
&= \Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)(\boldsymbol{x}'\boldsymbol{\beta})+\sigma\phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma).
\end{aligned}$$

Observações:

- $\mathrm{E}(Y|\boldsymbol{X}=\boldsymbol{x})$ é uma função não linear de $x_1,\ldots,x_p$ e $\beta_0,\beta_1,\ldots,\beta_p$;
- Pode ser demonstrado que a parcela direita da igualdade é positiva para quaisquer valores de $x_1,\ldots,x_p$ e $\beta_0,\beta_1,\ldots,\beta_p$.
  - Para quaisquer valores de $x_1,\ldots,x_p$ e $\beta_0,\beta_1,\ldots,\beta_p$, $\mathrm{E}(Y|\boldsymbol{X}=\boldsymbol{x})$ é positivo.

Efeitos Parciais

Suponha X_j uma variável contínua e suponha que X_j não está funcionalmente relacionada às demais variáveis independentes. Os efeitos parciais de X_j sobre Y são calculados da seguinte forma:

$$\begin{aligned}\frac{\partial \mathrm{E}(Y|\boldsymbol{X}=\boldsymbol{x})}{\partial x_j} &= \frac{\partial \mathrm{P}(Y>0|\boldsymbol{X}=\boldsymbol{x})}{\partial x_j} \cdot \mathrm{E}(Y|Y>0,\boldsymbol{X}=\boldsymbol{x}) \\ &+ \mathrm{P}(Y>0|\boldsymbol{X}=\boldsymbol{x}) \cdot \frac{\partial \mathrm{E}(Y|Y>0,\boldsymbol{X}=\boldsymbol{x})}{\partial x_j}.\end{aligned}$$

Como $\mathrm{P}(Y>0|\boldsymbol{X}=\boldsymbol{x}) = \Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)$,

$$\frac{\partial \mathrm{P}(Y>0|\boldsymbol{X}=\boldsymbol{x})}{\partial x_j} = (\beta_j/\sigma)\phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma).$$

Efeitos Parciais

Utilizando $d\Phi/dc = \phi(c)$ e $d\phi/dc = -c\phi(c)$, é possível mostrar que $d\lambda/dc = -\lambda(c)[c + \lambda(c)]$. Desta forma,

$$
\begin{aligned}
\frac{\partial \mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x})}{\partial x_j} &= \beta_j - \sigma \frac{\beta_j}{\sigma} \times \frac{\lambda((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)}{(\boldsymbol{x}'\boldsymbol{\beta})/\sigma) + \lambda((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)}. \\
&= \beta_j \times \left[1 - \frac{\lambda((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)}{(\boldsymbol{x}'\boldsymbol{\beta})/\sigma) + \lambda((\boldsymbol{x}'\boldsymbol{\beta})/\sigma)}\right]
\end{aligned}
$$

Efeitos Parciais

Caracterizando $\partial \mathrm{P}(Y|\boldsymbol{X} = \boldsymbol{x})/\partial x_j$ e $\partial \mathrm{E}(Y|Y > 0, \boldsymbol{X} = \boldsymbol{x})/\partial x_j$, temos

$$\frac{\partial \mathrm{E}(Y|\boldsymbol{X} = \boldsymbol{x})}{\partial x_j} = \beta_j \Phi((\boldsymbol{x}'\boldsymbol{\beta})/\sigma).$$

Na prática, podemos calcular os efeitos parciais considerando duas abordagens:

- Efeito parcial na média (PEA):

$$\hat{\beta}_j \Phi((\hat{\beta}_0 + \hat{\beta}_1 \bar{x}_1 + \ldots + \hat{\beta}_p \bar{x}_p)/\hat{\sigma}).$$

- Efeito parcial médio (APE):

$$n^{-1} \hat{\beta}_j \sum_{i=1}^{n} \Phi((\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \ldots + \hat{\beta}_p x_{ip})/\hat{\sigma}).$$