Семинар 8.04.25.

20 000

y - число проданных билетов

y^* - реальный спрос $\quad y^* \sim N(\mu, \sigma^2)$

$E(y) = 18000$

$P(y^* \geq 20.000) = \frac{1}{4}$

Нас интересует: $E(y^*) = \mu$ - ?

$P(y^* \geq 20\,000) = \frac{1}{4}$ - согл. аут произошёл

$P(y^* < 20000) = \Phi\left(\frac{20.000 - \mu}{\sigma}\right) = \frac{3}{4}$ → квантиль ст. норм. распр.

$$\frac{20\,000 - \mu}{\sigma} = z_{0.75} = 0.675$$

$$y = \begin{cases} y^*, & y^* < 20.000 \\ 20.000, & y^* \geq 20\,000 \end{cases}$$

$$E[y] = E[y^* | y^* < 20\,000] \cdot P(y^* < 20\,000) + E[y | y^* \geq 20.000] \cdot P(y^* \geq 20.000)$$

$$18.000 = E[y^* | y^* < 20\,000] \cdot \frac{3}{4} + 20\,000 \cdot \frac{1}{4}$$

$$E[y^* | y^* < 20.000] = 17\,333\frac{1}{3}$$

Вспоминаем мат.стат:

$$X \sim N(\mu, \sigma^2)$$

$$E[X | X \geq a] = \mu + \sigma \frac{\varphi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

$$E[X | X < a] = \mu - \sigma \frac{\varphi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

$$\varphi(-x) = \varphi(x), \quad \Phi(-x) = 1 - \Phi(x)$$

$$E[y^* | y^* < 20\,000] = \mu - \sigma \cdot \frac{\varphi\left(\frac{20000-\mu}{\sigma}\right)}{\ldots} = 17333\frac{1}{3}$$

$$\varphi\left(\frac{20000-\mu}{\delta}\right)$$

$$\begin{cases} \mu - \dfrac{\delta \cdot \varphi(0.675)}{0.75} = 17333\frac{1}{3} \\ \dfrac{20\,000-\mu}{\delta} = 0.675 \end{cases}$$

$$\begin{cases} \mu = \mathbb{E}(y^*) = 18360 \\ \delta = 2428. \end{cases}$$

Модель Тобита

$y_i^* = x_i^T\beta + \varepsilon_i$, $\varepsilon_i \sim \mathcal{N}(0, \delta_\varepsilon^2)$ Надо писать $(\varepsilon_i | x_i) \sim \mathcal{N}(0, \delta_\varepsilon^2)$

$$y_i = \begin{cases} y_i^*, & y_i^* > 0 \\ 0, & y_i^* \leq 0 \end{cases}$$

$$\mathbb{E}(y_i^* | x_i) = x_i^T\beta$$

$$\frac{\partial \mathbb{E}(y_i^* | x_i)}{\delta x_{ij}} = \beta_j$$

$$\mathbb{E}(y_i | x_i) = \mathbb{E}(y_i | y_i^* > 0) \cdot \mathbb{P}(y_i^* > 0) + \cancel{\mathbb{E}[y_i | y_i^* \leq 0]}^{\to 0} \cdot \mathbb{P}(y_i^* \leq 0) =$$

$$= \underset{②}{\mathbb{E}(y_i | y_i^* > 0)} \cdot \mathbb{P}(y_i^* > 0) =$$

$$\mathbb{P}(y_i^* > 0 | x_i) = \mathbb{P}(x_i^T\beta + \varepsilon_i > 0 | x_i) = \mathbb{P}(\varepsilon_i > -x_i^T\beta) = \mathbb{P}(\varepsilon_i < x_i^T\beta) =$$

$$= \Phi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right)$$

$$\mathbb{E}[x_i^T\beta + \varepsilon_i | \varepsilon_i > -x_i^T\beta, x_i] = x_i^T\beta + \mathbb{E}(\varepsilon_i | \varepsilon_i > -x_i^T\beta, x_i) =$$

$$= x_i^T\beta + \delta_\varepsilon \frac{\varphi\left(\frac{-x_i^T\beta}{\delta_\varepsilon}\right)}{1 - \Phi\left(\frac{-x_i^T\beta}{\delta_\varepsilon}\right)} = x_i^T\beta + \delta_\varepsilon \frac{\varphi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right)}{\Phi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right)}$$

$$\ominus \left(x_i^T\beta + \delta_\varepsilon \frac{\varphi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right)}{\Phi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right)}\right) \cdot \Phi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right) = x_i^T\beta \cdot \Phi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right) + \delta_\varepsilon \varphi\left(\frac{x_i^T\beta}{\delta_\varepsilon}\right)$$

M.E: $\dfrac{\partial \mathbb{P}(y_i^* > 0 \mid x_i)}{\delta x_{ij}} = \varphi\left(\dfrac{x_i^T\beta}{\delta_\varepsilon}\right) \cdot \dfrac{\beta_j}{\delta_\varepsilon}$

Для норм. станд. распр: $\varphi'(x) = -x\varphi(x)$

M.E: $\dfrac{\delta \mathbb{E}(y_i \mid x_i)}{\delta x_{ij}} = \Phi\left(\dfrac{x_i^T\beta}{\delta_\varepsilon}\right) \cdot \dfrac{\beta_j}{\delta_\varepsilon}$

Доп. занятие. $n = n_1 + n_2 + n_3$

$(x=1,\ y=1) - n_1$

$(x=1,\ y=0) - n_2$

$(x=0,\ y=0) - n_3$

$$y_i = \begin{cases} 1, & y_i^* > 0 \\ 0, & y_i^* \le 0 \end{cases}$$

$y_i^* = x_i^T\beta + \varepsilon_i, \quad \varepsilon_i \sim iid(0, \delta_\varepsilon^2)$

$y_i^* = \beta_1 + \beta_2 x_i + \varepsilon_i$

$\mathbb{E}(y_i / x_i) = \mathbb{P}(y_i = 1 / x_i) = F_\varepsilon(\beta_1 + \beta_2 x_i)$

Составим функцию правдоподобие:

$$L = \left(F_\xi(\beta_1 + \beta_2)\right)^{n_1} \cdot \left(1 - F_\xi(\beta_1 + \beta_2)\right)^{n_2} \cdot \left(1 - F_\varepsilon(\beta_1)\right)^{n_3}$$

$y=1, x=1$: $\mathbb{P}(y_i = 1 / x_i = 1) = F_\xi(\beta_1 + \beta_2)$

$y=0, x=1$: $\mathbb{P}(y_i = 0 / x_i = 1) = 1 - F_\xi(\beta_1 + \beta_2)$

$y=0, x=0$: $\mathbb{P}(y_i = 0 / x_i = 0) = 1 - F_\xi(\beta_1)$

Перейдём к логарифму:

$$l = n_1 \ln F(\gamma) + n_2 \ln(1-F(\gamma)) + n_3 \ln(1-F(\beta_1)) \to \max_{\beta_1,\beta_2}$$

$$\begin{cases} \dfrac{\delta l}{\delta \beta_1} = -n_3 \cdot \dfrac{1}{1-F(\beta_1)} \cdot \varphi(\beta_1) = 0 \; ! \\[2ex] \dfrac{\delta l}{\delta \gamma} = \dfrac{n_1 f(\gamma)}{F(\gamma)} - \dfrac{n_2 f(\gamma)}{1-F(\gamma)} = 0 \end{cases}$$

проблема: $\varphi(\beta_1)=0$ невозм., т.е нужно $\beta_1 \to 0$.

Вернёмся к tobit модели:

$$y_i^* = x_i^T\beta + \varepsilon_i$$

$$y = \begin{cases} y_i^*, & y_i^* > c \\ a, & y_i^* \le c \end{cases}$$

пример с з.п:
Пусть $c = 70.000$
$a = 15.000$

$$E(y) = E(y^* | y^* > 70) \cdot P(y^* > 70) + 15 \cdot P(y^* \le 70)$$

↑ дают вклад.

Функция правдоподобия:

$$L = \prod_{i: y_i = a} \overbrace{P(y_i^* \le c)}^{= P(\varepsilon_i \le c - x_i^T\beta) = \Phi\left(\frac{c - x_i^T\beta}{\sigma_\varepsilon}\right)} \prod_{i: y_i > c \text{ (или } a)} f_{y_i^*} \to \max_\beta$$

$$l = \sum_{i: y_i = a} \ln \Phi\left(\frac{c - x_i^T\beta}{\sigma_\varepsilon}\right)$$

$$\int_{-\infty}^{x''} \frac{1}{\sqrt{2\pi}} e^{-\frac{c - x_i^T\beta}{\sigma_\varepsilon}} dx$$

не роляет вообще (поэтому можем с положить нулем)

Тесты LM и Вальда

Задача 1.

$$L = \prod_{i=1}^{n} \mathbb{P}(y_i = 1)^{y_i} \cdot \mathbb{P}(y_i = 0)^{1-y_i}$$

$$l = \sum_{i=1}^{n} \left[y_i \ln \mathbb{P}(y_i = 1) + (1-y_i) \ln (\mathbb{P}(y_i = 0)) \right]$$

$\tilde{l}$ - для одного наблюдения:

$$\tilde{l} = y \ln F(\beta_1 + \beta_2 d) + (1-y) \ln(1 - F(\beta_1 + \beta_2 d))$$

$^* \mathbb{P}(y_i = 1) = F(\beta_1 + \beta_2 d_i)$ $\qquad \theta = [\beta_1 \ \beta_2]^T$

$$I_1(\theta; d) = \mathbb{E}\left(\frac{\delta \tilde{l}}{\delta \theta} \cdot \frac{\delta \tilde{l}}{\delta \theta^T} \right)$$

$$\begin{cases} \dfrac{\delta \tilde{l}}{\delta \beta_1} = y \cdot \dfrac{1}{F(\beta_1 + \beta_2 d)} \cdot f(\beta_1 + \beta_2 d) - \dfrac{(1-y) f(\beta_1 + \beta_2 d)}{1 - F(\beta_1 + \beta_2 d)} \\ \dfrac{\delta \tilde{l}}{\delta \beta_2} = y \cdot \dfrac{1}{F(\beta_1 + \beta_2 d)} \cdot f(\beta_1 + \beta_2 d) \cdot d - \end{cases}$$

$$- \frac{(1-y) f(\beta_1 + \beta_2 d) \cdot d}{1 - F(\beta_1 + \beta_2 d)}$$

$$\frac{y \cdot f(x) \cdot (1 - F(x)) - (1-y) f(x) F(x)}{F(x) \cdot (1 - F(x))} =$$

$$= \frac{f(x)}{F(x)\cdot(1-F(x))}\cdot(y-F(x))$$

$$\mathbb{E}\left(\frac{f^2(x)}{F^2(x)(1-F(x))^2}\cdot(y-F(x))^2\right) =$$

$$= \frac{f^2(x)}{F^2(x)\cdot(1-F(x))^2}\cdot\underbrace{\mathbb{E}(y^2-2yF(x)+F^2(x))}=$$

$$\textcircled{=}$$

$$E[y] = \mathbb{P}(y=1) = F(x)$$

$$\mathbb{E}[y^2] = E[y] = F(x)$$

$$\textcircled{=}\ \cancel{F(x)}(1-\cancel{F(x)^2})\cdot\frac{f^2(x)}{\cancel{F^2(x)}\cdot(1-F(x))^2}$$

$$= \frac{f^2(x)}{F(x)(1-F(x))}, \quad x = \beta_1 + \beta_2 d.$$

Umo2 $I_1(\theta)$:

$$I_1(\theta;d) = \frac{f^2(x)}{F(x)\cdot(1-F(x))}\cdot\begin{bmatrix}1 & d\\ d & d^2\end{bmatrix}$$

$$I_n = \sum_{i=1}^{n} I_1(\theta;d_i) = \sum_{i=1}^{n}\frac{f^2(\beta_1+\beta_2 d_i)}{F(\beta_1+\beta_2 d_i)\cdot(1-F(\beta_1+\beta_2 d_i))}\cdot\begin{bmatrix}1 & d_i\\ d_i & d_i^2\end{bmatrix}$$

$$= \#\{d_i=0\} \cdot \frac{f^2(\beta_1)}{F(\beta_1)\cdot(1-F(\beta_1))} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} +$$

$$+ \#\{d_i=1\} \frac{f^2(\beta_1+\beta_2)}{F(\beta_1+\beta_2)\cdot(1-F(\beta_1+\beta_2))} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Тест Вальда:

Гипотеза:

Свойства логит модели:

$$\Lambda'(x)=\Lambda(x)\cdot(1-\Lambda(x))$$

H_0: $\beta_2=0$

$$g(\beta_1,\beta_2)=0$$

$$g(\beta_1,\beta_2)=\beta_2=0$$

$$\frac{\delta g}{\delta \theta^T}=[0 \;\; 1]$$

из ММП: $\hat\beta_1=0.47$; $\hat\beta_2=1.57$.

$$I_n=52\cdot \underbrace{\Lambda(0.47)}_{\frac{8}{13}}\cdot \underbrace{(1-\Lambda(0.47))}_{\frac{5}{13}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} +$$

$$+48\cdot\frac{1}{4}\cdot\frac{3}{4}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 21.3 & 9 \\ 9 & 9 \end{bmatrix}$$

$$W:=g^T(\hat\theta_{uk})\cdot\left[\frac{\delta g}{\delta\theta^T}(\hat\theta_{uk})\cdot I^{-1}(\hat\theta_{uk})\cdot\frac{\partial g}{\delta\theta}^T(\hat\theta_{uk})\right]^{-1}$$

$$\hat\beta_2^{uk}\left[[0|1]\begin{bmatrix} 21.3 & 9 \\ 9 & 9 \end{bmatrix}[0,1]^T\right]^{-1}\cdot\hat\beta_2^{uk}$$

$$= \frac{(-1{,}57)^2}{\left[\quad\right]^{-1}} = 11.92.$$

$(3.84; +\infty)$

LM-тест:

$$\frac{\delta l}{\delta \theta}(\hat{\theta}_R) \qquad \frac{\delta l}{\delta \beta_1}(\hat{\beta}_1^R, 0)$$

$$\frac{\delta l}{\delta \beta_2}(\hat{\beta}_1^R, 0)$$