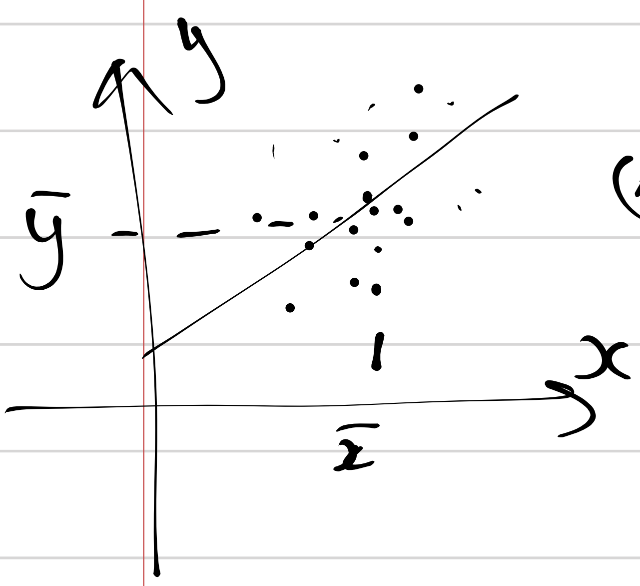
Привет !!

Видно ли / слышно ли ?

При $n >> 0$ научиться строить доверительные интервалы.

$y_i = \beta_1 + \beta_2 \cdot x_i + u_i$ $\qquad$ $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_i$

до вер-х предпосылок:



Ⓐ $\hat{\beta}_2 = \frac{sCov(x,y)}{sVar(x)} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}$

$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \cdot \bar{x}$

Предпосылки

1. Выборка — случайная :

набл №i и набл №j не зависимы $\qquad$ u_7 и x_5 незав

набл №i и набл №j один. распр.

$x_5 \sim x_6$

$x_5 \not\sim u_{10}$ $\qquad$ $u_7 \sim u_{10}$

2. $y_i = \beta_1 + \beta_2 \cdot x_i + u_i$

регр. с МНК $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$ $\Bigg\} \Rightarrow$ $\hat{\beta}_2 = \dots$; $\hat{\beta}_1 = \dots$

3. $E(u_i | x_i) = 0$ экзогенность

[4] гомоскедастичность $Var(u_i | x_i) = \sigma^2$

без 4-ой предп-ки из ① $\Rightarrow$ $Var(u_i | x_i) = h(x_i)$

Ⓐ y_i - расх. i-й семьи на фрукты

x_i - число человек в i-ой семье

гомоск: $\mathrm{Var}(u_i|x_i) = \sigma^2$

гетероск: $\mathrm{Var}(u_i|x_i) = h(x_i)$

Упр

$$E(\underset{\sim}{x_i^2} + w_i | x_i) = x_i^2 + E(w_i|x_i)$$

известная сл.вел.

$$E(x_i \cdot w_i | x_i) = x_i \, E(w|x_i)$$

$$\mathrm{Var}(x_i + w_i | x_i) \overset{?}{=} \mathrm{Var}(w_i|x_i)$$

$$\mathrm{Var}(x_i \cdot w_i | x_i) \overset{?}{=} x_i^2 \cdot \mathrm{Var}(w_i|x_i)$$

$$\hat\beta_2 = \frac{\sum (x_i - \bar x)\cdot(y_i - \cancel{\bar y})}{\sum (x_i - \bar x)^2} =$$

в честь др. М. Ганди

настоящая условная дисперсия

вектор $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$\boxed{\mathrm{Var}(\hat\beta_2 | x_1, \ldots, x_n)} = \mathrm{Var}\left(\frac{\sum (x_i - \bar x)\cdot y_i}{\sum (x_i - \bar x)^2} \,\middle|\, x\right) =$$

$\sigma^2 = \mathrm{Var}(u_i|x_i)$ не знаем!

$$= \left(\frac{1}{\sum (x_i - \bar x)^2}\right)^2 \cdot \mathrm{Var}\left((x_1 - \bar x)\cdot y_1 + (x_2 - \bar x)\cdot y_2 + \ldots \,\middle|\, x\right) =$$

по предп.

$$= \left(\frac{1}{\sum (x_i - \bar x)^2}\right)^2 \cdot \left[\mathrm{Var}((x_1 - \bar x) y_1 | x) + \mathrm{Var}((x_2 - \bar x) y_2 | x) + \ldots\right]$$

$$= \left(\frac{1}{\sum (x_i - \bar x)^2}\right)^2 \cdot \left(\sum (x_i - \bar x)^2 \cdot \mathrm{Var}(y_i | x)\right) =$$

$$= \left(\frac{1}{\sum (x_i - \bar x)^2}\right)^2 \left(\sum (x_i - \bar x)^2 \cdot \mathrm{Var}(\underbrace{\beta_1 + \beta_2 x_i}_{\text{фикс}} + u_i | x)\right) =$$

$$= \left(\frac{1}{\sum(\ldots)^2}\right)^2 \cdot \sum (x_i - \bar x)^2 \cdot \mathrm{Var}(u_i | x) =$$

$$= \frac{\sum (x_i - \bar x)^2 \cdot h(x_i)}{\left(\sum (x_i - \bar x)^2\right)^2}$$

Упр [У: гомоскед]

$$\overset{?}{=} \frac{\sigma^2}{\sum (x_i - \bar x)^2}$$

Итог:

при гетероса — при гомоса

$$Var(\hat\beta_2|x) = \frac{\sum(x_i-\bar x)^2 \cdot h(x_i)}{\left(\sum(x_i-\bar x)^2\right)^2} = \frac{\sigma^2}{\sum(x_i-\bar x)^2}$$

не можем посчитать точно!

Как оценить эти неизв-ые кол-ва?

гомоскед:

$$\boxed{\hat\sigma^2 = \frac{RSS}{n-k}}$$

n - число наблюд-ий

k - число оц-х коэфф-в бета

в парной: β_1, β_2: $\hat\sigma^2 = \frac{RSS}{n-2}$

много способов

гетероскедастичности: (HC0, HC1, HC2 ....)

$h_i^{HC0} = \hat u_i^2$, где $\hat u_i = y_i - \hat y_i$

$h_i = Var(u_i|x)$

Софт:

→ Хочу МНК + гомоск $\hat{Var}(\hat\beta_2|x)$

→ Хочу МНК + HC0 $\hat{Var}_{HC0}(\hat\beta_2|x)$

$$\left\{\begin{array}{cc} \hat\beta_2 & CI \\ \shortparallel & \nparallel \\ \hat\beta_2 & CI \end{array}\right.$$

по дефолту: хочу МНК → хочу МНК + гомоск $\hat{Var}(\hat\beta_2|x)$

Упр $n=3000$ (asy CI)

y	x	число набл
2	1	1000
1	2	1000
3	3	1000

$$y = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} \begin{matrix} \updownarrow \text{ 1000 набл} \\ \updownarrow \text{ 1000 набл} \\ \updownarrow \text{ 1000 набл} \end{matrix}$$

а) $\bar{y}, \bar{x}$, X^TX, X^Ty

$\hat{\beta}_2, \hat{\beta}_1$ → скалярно / → матрично — 5 мин

б) RSS, TSS, ESS, R^2

в) $\widehat{Var}(\hat{\beta}_2|x)$ $\quad \hat{V}ar_{НСО}(\hat{\beta}_2|x)$

г) asy CI → для гомоск / → для гетероск.

д) $Var(\hat{\beta}_1|x)$

е) сделать в-ч 2 для $\hat{\beta}_1$

$\bar{x}=2$

$\bar{y}=2$

$$\hat{\beta}_2 = \frac{\Sigma(x_i-\bar{x})\cdot y_i}{\Sigma(x_i-\bar{x})^2} =$$

← 1000 дублей

← 1000 дублей

$$= \frac{(1-2)\cdot 2 + (2-2)\cdot 1 + (3-2)\cdot 3}{(1-2)^2 + (2-2)^2 + (3-2)^2} = \frac{1}{2} = 0.5$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \cdot \bar{x} = 2 - 0.5\cdot 2 = 1$$

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ \vdots & 2 \\ \vdots & 3 \\ 1 & 3 \end{bmatrix} \quad 3000\times 2$$

$$X^T = \begin{bmatrix} 1 1 1 1 1 1 1 & 1 1 1 \\ 1 1 1 1 2 2 2 2 3 3 3 3 \end{bmatrix}$$

$$X^TX = \begin{bmatrix} 3000 & 6000 \\ 6000 & 14000 \end{bmatrix} \qquad y = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}$$

$$X^Ty = \begin{bmatrix} 6000 \\ 13000 \end{bmatrix} \qquad \hat{\beta} = \begin{bmatrix} 3000 & 6000 \\ 6000 & 14000 \end{bmatrix}^{-1} \begin{bmatrix} 6000 \\ 13000 \end{bmatrix} =$$

1000^{-1} $\qquad$ 1000

$$= \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 6 \\ 13 \end{bmatrix} = \frac{1}{3 \cdot 14 - 6^2} \cdot \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 13 \end{bmatrix} =$$

$$= \frac{1}{6} \cdot \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 13 \end{bmatrix} = \frac{1}{6} \begin{pmatrix} 14 \cdot 6 - 6 \cdot 13 \\ -6 \cdot 6 + 3 \cdot 13 \end{pmatrix} =$$

$$= \boxed{\begin{matrix} 1 \\ 0.5 \end{matrix}} = \hat\beta \qquad \hat y_i = \hat\beta_1 + \hat\beta_2 x_i = 1 + 0{,}5 x_i$$

каждого i × 1000 дублей:

y	x	$\hat y$	$\hat u$	$(x_i-\bar x)^2 \cdot \hat u_i^2$
2	1	1.5	0.5	0,25
1	2	2	−1	0
3	3	2.5	0,5	0,25

$\hat u_i = y_i - \hat y_i$

$$RSS = \sum \hat u_i^2 = \sum (y_i - \hat y_i)^2 = (0{,}5^2 + (-1)^2 + 0{,}5^2) \times 1000 = 1500$$

$$TSS = \sum (y_i - \bar y)^2 = 1000 \times [(2-2)^2 + (1-2)^2 + (3-2)^2] = 2000$$

$$ESS = TSS - RSS = 500$$

$$R^2 = \frac{ESS}{TSS} = \frac{1}{4}$$

гетероскед:

$$\widehat{Var}_{НСО}(\hat\beta_2 | x) = \frac{\sum (x_i - \bar x)^2 \cdot \hat u_i^2}{(\sum (x_i - \bar x)^2)^2} =$$

$$= \frac{(0.25 + 0 + 0{,}25) \cdot 1000}{(\underbrace{((1-2)^2 + (2-2)^2 + (3-2)^2) \times 1000}_{})^2} =$$

$$= \frac{0{,}5}{2 \cdot 1000} = \frac{1}{4000}$$

гомоскед

$$\widehat{Var}(\hat\beta_2 | x) = \frac{1500/(3000-2)}{2 \cdot 1000} \approx$$

$$= \frac{3}{4 \cdot 2998} \approx \frac{1}{4000}$$

НСО гетер CI

асу

$$\left[0.5 - 1.96 \cdot \sqrt{\frac{1}{4000}} ; 0.5 + 1.96 \sqrt{\frac{1}{4000}}\right]$$

гомоскед.

$$\left[0.5 - 1.96 \sqrt{\frac{3}{4 \cdot 2998}} : 0.5 + 1.96 \sqrt{\frac{3}{4 \cdot 2998}}\right]$$

ⓓ/з

① Выведите удобную ф-лу для $\hat\beta$,

$$\hat\beta_1 = \bar y - \hat\beta_2 \bar x$$ /верная/

зав

$$\hat\beta_1 = \frac{\sum w_i \cdot y_i}{\text{[illegible]}}$$?

только x

по аналогии с

$$\hat\beta_2 = \frac{\sum (x_i - \bar x) \cdot y_i}{\sum (x_i - \bar x)^2}$$

② $Var(\hat\beta_1 | x)$?

③ $\widehat{Var}(\hat\beta_1 | x)$ гомоск

$\widehat{Var}_{HCO}(\hat\beta_1 | x)$ гетероск

④ CI_{asy} → гомоск

→ HCO гетероск.