Свойства оценок

На разных выборках – разные оценки (за счет случайного характера остатков)

1. **Состоятельность** (при росте выборки оценка $\hat{\theta}$ стремится к истинному значению θ) (асимптотическое свойство, часто проявляется при огромных размерах выборок).

   ## Симметрично распределенная случайная величина ξ, истинное среднее $\theta = M\xi$:

   $\hat{\theta}_1 = \bar{x}$, $\hat{\theta}_2 = \left(x_{\min} + x_{\max}\right)/2$ – состоятельные оценки.

   Состоятельная оценка может быть сколь угодно далекой от истинного значения

   ## Средняя зарплата в отрасли, где работает N человек (заинтересованы завысить):

   $\hat{\theta} = \begin{cases} \theta_0, & n < N \\ \bar{x}, & n = N \end{cases}$ – при любом объеме выборки n, кроме сплошного обследования, получаем сколь угодно завышенный результат θ_0.

2. **Несмещенность** ($M\hat{\theta} = \theta$ при любом объеме выборки)

   M – усреднение по всем выборкам данного объема

   (характеристика «хороших свойств» оценки при каждом конечном объеме выборки).

$$M\hat{\Theta} = M\left(\left(X^T X\right)^{-1} X^T Y\right) = M\left(\left(X^T X\right)^{-1} X^T (X\Theta + \varepsilon)\right) = M\left(\Theta + \left(X^T X\right)^{-1} X^T \varepsilon\right) =$$

$$= M\Theta + \left(X^T X\right)^{-1} X^T M\varepsilon = M\Theta = \Theta \text{ – несмещенная.}$$

$$M\hat{\sigma}^2 = M\left(\frac{1}{n}\left(Y - \hat{Y}\right)^T \left(Y - \hat{Y}\right)\right)$$

$$Y - \hat{Y} = X\Theta + \varepsilon - X\hat{\Theta} = X\Theta + \varepsilon - X\left(X^T X\right)^{-1} X^T (X\Theta + \varepsilon) = X\Theta + \varepsilon - X\Theta - X\left(X^T X\right)^{-1} X^T \varepsilon =$$

$$= \left(E_n - X\left(X^T X\right)^{-1} X^T\right)\varepsilon = Z\varepsilon .$$

$Z^T = Z$ – симметричная, $Z^2 = ZZ = Z$ – идемпотентная, $M\left(\varepsilon^T Z\varepsilon\right) = \sigma^2 tr\, Z$

$tr\, Z$ – след матрицы Z (сумма ее диагональных элементов)

$$M\hat{\sigma}^2 = M\left(\frac{1}{n}\left(Y - \hat{Y}\right)^T \left(Y - \hat{Y}\right)\right) = M\left(\frac{1}{n}(Z\varepsilon)^T Z\varepsilon\right) = \frac{1}{n} M\left(\varepsilon^T Z^T Z\varepsilon\right) = \frac{1}{n} M\left(\varepsilon^T Z\varepsilon\right) = \frac{1}{n}\sigma^2 tr\, Z =$$

$$= \frac{\sigma^2}{n} tr\left(E_n - X\left(X^T X\right)^{-1} X^T\right) = \frac{\sigma^2}{n}\left(tr\, E_n - tr\left(\left(X^T X\right)^{-1} X^T X\right)\right) = \frac{\sigma^2}{n}\left(n - (p+1)\right) \text{ – смещенная.}$$

Несмещенная оценка дисперсии остатков:

$$\hat{\sigma}^2 = \frac{n}{n-p-1}\hat{\sigma}^2_{ММП} = \frac{1}{n-p-1}\sum_{i=1}^{n}\left(y_i - \hat{\theta}_0 - \hat{\theta}_1 x_i^{(1)} - \ldots - \hat{\theta}_p x_i^{(p)}\right)^2 = \frac{1}{n-p-1}\left(Y - X\hat{\Theta}\right)^T \left(Y - X\hat{\Theta}\right).$$

$\hat{\sigma}^2 = \dfrac{1}{24-3-1}\left(948^2 + (-1192)^2 + \ldots + 1575^2\right) = 1938734$, $\hat{\sigma} = \sqrt{1938734} = 1392{,}4$.

Ковариационная матрица оценок параметров:

$$\hat{\Sigma}_{\hat{\Theta}} = M\left(\left(\hat{\Theta} - \Theta\right)\left(\hat{\Theta} - \Theta\right)^T\right) = \hat{\sigma}^2\left(X^T X\right)^{-1}.$$

$$\hat{\Sigma}_{\hat{\Theta}} = 1938734\begin{pmatrix} 4{,}83 & -1{,}39 & -0{,}00003 & -0{,}17 \\ -1{,}39 & 0{,}414 & -0{,}000005 & 0{,}0326 \\ -0{,}00003 & -0{,}000005 & 0{,}00000002 & 0{,}00001 \\ -0{,}17 & 0{,}0326 & 0{,}00001 & 0{,}0483 \end{pmatrix} = \begin{pmatrix} 9355484 & -2694758 & -50{,}15 & -328915 \\ -2694758 & 803096 & -9{,}44 & 63228 \\ -50{,}15 & -9{,}44 & 0{,}0396 & 22{,}41 \\ -328915 & 63228 & 22{,}41 & 93598 \end{pmatrix}.$$

Наиболее важны диагональные элементы – квадраты среднеквадратических ошибок s_l оценок коэффициентов θ_l, $l = 0,1,\ldots,p$.

$s_0 = \sqrt{9335484} = 3059$, $s_1 = \sqrt{803096} = 896$, $s_2 = \sqrt{0{,}0396} = 0{,}1990$, $s_3 = \sqrt{93598} = 305{,}94$.

$$\boxed{\hat{y}_i = \underset{(3059)}{17094} - \underset{(896)}{2938}\,x_i^{(1)} + \underset{(0{,}1990)}{1{,}1327}\,x_i^{(2)} + \underset{(305{,}94)}{600{,}86}\,x_i^{(3)}}.$$

3. **Эффективность** (оценка обладает наименьшим случайным разбросом).

$$M\left(\hat{\theta}_{eff}-\theta\right)^2=\min_{\hat{\theta}\in M}\left(\hat{\theta}-\theta\right)^2$$ – оценка, эффективная в классе M.

Свойства оценок, справедливые для нормальной КЛММР

$t_l=\dfrac{\hat{\theta}_l-\theta_l}{s_l}\sim t(n-p-1)$ – распределена по закону Стьюдента.

Проверка гипотезы H_0: $\theta_l=\theta_l^0$, в т.ч. о значимости регрессора (при $\theta_l^0=0$):

1. Задаем уровень значимости α.
2. Находим эмпирическую точку $t_l=\left(\hat{\theta}_l-\theta_l^0\right)/s_l$.
3. Находим критическую точку $t_{\text{крит}}=t_\alpha(n-p-1)=СТЬЮДРАСПОБР(\alpha;\ n-p-1)$.
4. Если $|t_l|>t_{\text{крит}}$, то H_0 – отвергается, иначе – принимается.

$\theta_0=0$ $\theta_1=0$ $\theta_2=0$ $\theta_3=0$

$$t_0=\frac{17094}{3059}=5{,}59 \quad t_1=\frac{-2938}{896}=-3{,}28 \quad t_2=\frac{1{,}1327}{0{,}1990}=5{,}69 \quad t_3=\frac{600{,}86}{305{,}94}=1{,}96$$

$t_{\text{крит}}=t_{0,05}(20)=2{,}09$

Гипотеза $H_0:\theta_l=0$ отвергается для $\theta_0,\theta_1,\theta_2$ и принимается для θ_3 при $\boxed{\alpha=0{,}05}$, регрессор $x^{(3)}$ – незначим

При $\boxed{\alpha=0{,}1}$ $t_{\text{крит}}=1{,}72$, $H_0:\theta_l=0$ отвергается для всех θ_l, все регрессоры значимы.

При $\boxed{n=1000}$ $t_{\text{крит}}=1{,}96$, $H_0:\theta_l=0$ отвергается для всех θ_l, все регрессоры значимы.

Построение доверительного интервала для θ_l:

1. Задаем доверительную вероятность γ.
2. $\theta_l\in\left[\hat{\theta}_l-t_{1-\gamma}(n-p-1)s_l;\ \hat{\theta}_l+t_{1-\gamma}(n-p-1)s_l\right]$.

$\theta_0\in[10714;23474]$, $\theta_1\in[-4807;-1069]$, $\theta_2\in[0{,}7176;1{,}5478]$, $\theta_3\in[-37;1239]$ при $\gamma=0{,}95$.

Проверка гипотезы H_0: $R^2=0$ о значимости модели:

1. Задаем уровень значимости α.
2. Находим эмпирическую точку $F_{\text{эмп}}=\dfrac{\hat{R}^2_{y.X}}{1-\hat{R}^2_{y.X}}\dfrac{n-p-1}{p}$.
3. Находим критическую точку $F_{\text{крит}}=F_\alpha(p;n-p-1)=FРАСПОБР(\alpha;\ p;\ n-p-1)$.
4. Если $F_{\text{эмп}}>F_{\text{крит}}$, то H_0 – отвергается, иначе – принимается, линейная модель неадекватна.

$\hat{R}^2_{y.X}$ – квадрат множественного коэффициента корреляции, равен коэффициенту детерминации.

Способы расчета:

1. $\hat{R}^2_{y.X}=1-\dfrac{|\hat{R}|}{|\hat{R}|_{00}}$, $\hat{R}=\begin{pmatrix}1 & -0{,}389 & 0{,}613 & 0{,}088\\ -0{,}389 & 1 & 0{,}152 & -0{,}269\\ 0{,}613 & 0{,}152 & 1 & -0{,}391\\ 0{,}088 & -0{,}269 & -0{,}391 & 1\end{pmatrix}$, $\hat{R}^2_{y.X}=1-\dfrac{0{,}2536}{0{,}7832}=0{,}6762$.

2. $$\hat{R}^2_{y.X}=\hat{K}_d(y.X)=1-\frac{S^2_\varepsilon}{S^2_y}=1-\frac{\frac{1}{24-3-1}\left(948^2+(-1192)^2+...+1575^2\right)}{\frac{1}{24-3-1}\left(2817^2+(-2083)^2+...+2117^2\right)}=1-\frac{1938734}{5986667}=0{,}6762.$$

$F_{\text{эмп}}=\dfrac{0{,}6762}{1-0{,}6762}\dfrac{24-3-1}{3}=13{,}92$, $F_{\text{крит}}=F_{0,05}(3;20)=3{,}10$, H_0 отвергается, линейная модель значима при $\alpha=0{,}05$.