ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

$A_n^k = \frac{n!}{(n-k)!}$ $\hat{A}_n^k = n^k$

$P_n = A_n^n = n!$ $\hat{P} = \frac{n!}{n_1!\ldots n_k!}$

$C_n^k = \frac{n!}{(n-k)!\,k!}$ $\hat{C}_n^k = C_{n+k-1}^{k}$

$(a+b)^n = \sum C_n^k a^k b^{n-k}$; $C_n^k = C_{n-1}^{k-1} + C_{n-1}^{k}$

$P(A) \overset{d}{\approx} W(A) = \frac{m}{n}$; $P(A) = \frac{\mu(A)}{\mu(\Omega)}$, $A \subseteq \Omega \subseteq \mathbb{R}^d$

$P(A+B) = P(A) + P(B) - \underbrace{P(AB)}_{=0,\ \text{если } A,B \text{ несовм.}}$

$P(AB) = P(B)\cdot P(A|B) = P(A)\cdot P(B|A)$ $P(AB) = P(A)\cdot P(B)$, A,B - незав.

$P(A_1\cdot\ldots\cdot A_n) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cdot A_2)\cdot\ldots P(A_n|A_1\cdot\ldots\cdot A_{n-1})$

$P(A_1+A_2|B) = P(A_1|B) + P(A_2|B)$, $A_1\cdot A_2 = \varnothing$

$P(A) = P(H_1)\cdot P(A|H_1)+\ldots+P(H_k)\cdot P(A|H_k)$, A происх. с 1 из $H_1, \ldots H_n$

$P(H_i|A) = \frac{P(H_i)\cdot P(A|H_i)}{P(A)}$, $i = \overline{1,k}$

① Пос-ть незав. событий, формула Бернулли:

– исп. независимы, зак. одним из 2 исходов, в-ть наст. A одинакова

$P_n(m) = \frac{n!}{m!\,(n-m)!}\,p^m q^{n-m}$ (m раз из n - A произойдет)

$P(m_1 \le k \le m_2) = \sum_{m_1}^{m_2} P_n(k)$

② Теорема Муавра-Лапласа, $n \to \infty$

$P_n(m) \approx \varphi(x)\frac{1}{\sqrt{npq}}$, где $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $x = \frac{m-np}{\sqrt{npq}}$

$P_n(a \le x \le b) \to \int_a^b \varphi(x)\,dx$ (вер. A наст. m раз)

Ф-я Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$

$P_n(m_1 \le m \le m_2) = \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right)$

при $x \ge 5$ $\Phi(x) = 0,5$; $x < 0$ $\Phi(-x) = -\Phi(x)$

Ф-ла Пуассона $P_n(m) \approx \frac{e^{-\lambda}\lambda^m}{m!}$, где $\lambda = np$ ($p < 0,1$, $n \ge 12$)

$P(|m/n - p| \le \varepsilon) \approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$ $\left(\overset{d}{\approx} 1\right)$

Наивероятн. число наст. A $np - q \le m_0 \le np + p$

② Пуассоновский (простейший) поток – [$P_t(m) \approx \frac{e^{-\lambda t}(\lambda t)^m}{m!}$]

стационарный ($P_t(m)$ зависит только от t, m),

ординарный ($P(m > 1 \mid t = o(t)) \to 0$),

с отсутствием последствия ($P_t(m)$ зав. только от F_t)

★ N_t – пуас. поток:

где N_t ($N(t)$) – число происш. за $[0; t]$

1) $N_0 = 0$

2) $N_t - N_s$ зав только от $(t - s)$; $N_{t+s} - N_t \sim N_s \sim Poiss(\lambda s)$

3) $(s_1; t_1]$, $(s_2; t_2)$ – непересек. промежутки $\Rightarrow (N_{t_1} - N_{s_1}); (N_{t_2} - N_{s_2})$ – независимые

4) $P(N_t - N_0 = 1) = \lambda t + o(t)$

$P(N_t - N_0 > 1) = o(t)$

★ W_t – вин. процесс, $t \in [0; +\infty)$

1) $W_0 = 0$

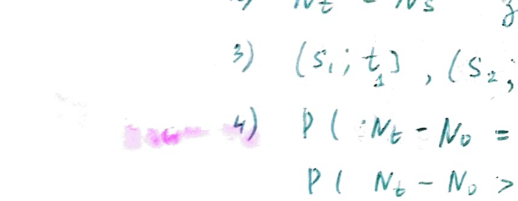
2) $W_{t+s} - W_t$ не зав от F_t

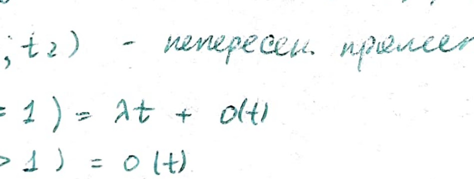
3) $W_{t+s} - W_t \sim W_s \sim N(0; s)$

4) W_t – непрер.

Маркова – послед. случ. соб. : $X_{t+s} - X_t$ не зав от F_t

$= P(X \le x) = \int_{-\infty}^{x} f_X(t)\, dt$, $\forall x \in R$ (непрер. слева)





$\int_{-\infty}^{\infty} f_X(t)\, dt = 1$

$P(X = x) = 0$

$F(x, y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(t_1, t_2)\, dt_1 dt_2$

$P((X, Y) \in \Omega) = \iint_{\Omega} f(x, y)\, dx dy$

$F(x,y) = F_X(x) \cdot F_Y(y)$

$f(x,y) = f_X(x) \cdot f_Y(y)$ $\Leftrightarrow$ X, Y – независимы

$f_Y(y) = f_X(x) \cdot |x'_y|$, где $x = \varphi(y)$

Пусть, $Z = \varphi(x,y) = x+y$

$F_Z(Z) = P(x+y < Z) = \int_{-\infty}^{z-x} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy$

$f_Z(Z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy$

④ $E(X) = \sum x_i p_i = \int_{-\infty}^{\infty} x f_X(x)\,dx$ $\qquad \bar{X} = \frac{1}{n}\sum x_i$

$E(c) = c$

$E(aX+bY) = aE(X) + b \cdot E(Y)$

$E(XY) = E(X) \cdot E(Y)$, X, Y – незав.

$E(X - M(X)) = E(X) - E(E(X)) = 0$

$Var(X) = E[(X - E(X))^2] = E(X^2) - E^2(X)$ $\qquad \hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$

$Var(X) = \sum p_i (x_i - M(X))^2 = \int_{-\infty}^{\infty} f(x)(x - M(X))^2\,dx$

$Var(X) \geq 0$, $Var(c) = 0$

$Var(X \pm Y) = Var(X) + Var(Y) \pm 2cov(X,Y)$

$Var(-aX+b) = a^2 Var(X)$

$\sigma(X) = \sqrt{Var(X)}$

$\sigma(aX) = |a| \cdot \sigma(X)$

$cov_{YX} = \frac{1}{n}\sum\sum(x-\bar{x})(y-\bar{y}) = \overline{XY} - \bar{X}\cdot\bar{Y}$

$Cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X) \cdot E(Y)$

$cov(aX+b, cY+d) = ac \cdot Cov(X,Y)$

$cov(Y_1, Y_2) = \sum_i^n \sum_j^m a_i b_j \, cov(X_i, Y_j)$

$Cov^2(X,Y) \leq Var(X) \cdot Var(Y)$

$Corr(X,Y) = \dfrac{Cov(X,Y)}{\sigma(X)\cdot\sigma(Y)}$

$corr(X,X) = 1$, $|corr(X,Y)| \leq 1$

$Corr(X,Y) = 0$, X, Y – нез.

$corr(aX+b, cY+d) = sign(ac) \cdot corr(X,Y)$

$\alpha_k = E(X^k) = \int_{-\infty}^{\infty} x^k \cdot f(x)\,dx$

$\mu_k = E[(X - E(X))^k] = \int_{-\infty}^{\infty} (x - E(X))^k f(x)\,dx$

$A_X = \mu_3/\sigma^3$ – коэф. асимметрии $\qquad \gamma_X = \mu_4/\sigma^4 - 3$ – коэф. эксцесса

(5) $P(X=k) = C_n^k p^k q^{n-k}$, $X \sim Bin(n,p)$ – биномиальное

$E(X) = np$ $\qquad E(X_i) = p$

$E(X^2) = np(q+np)$

$Var(X) = npq$ $\qquad Var(X_i) = pq$

$F_X(x) = P(X \le x) = \sum_{k=0}^{x} C_n^k p^k q^{n-k}$

$X \sim Bin(n,p),\ Y \sim Bin(n, 1-p) \Rightarrow P_X(k) = P_Y(n-k)$

$X \sim Bin(n_1, p),\ Y \sim Bin(n_2, p) \Rightarrow X+Y \sim Bin(n_1+n_2, p)$

$Bin(n,p) \overset{d}{\approx} N(np, npq)$

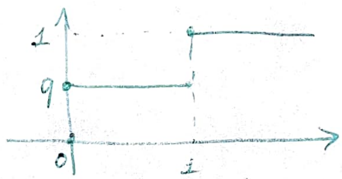
$Bin(n, \lambda/n) \overset{d}{\approx} P(\lambda)$

(6) Распр. Бернулли (геометрическое)

$(X \sim Bin(1,p))$

$P(X=k) = \begin{cases} p, & k=0 \\ q, & k=1 \end{cases}$

$F_X(x) = \begin{cases} 0, & x<0 \\ q, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$



$E(X) = p$

$Var(X) = pq$

$E(X^k) = p$

$P(X=k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$, $X \sim Poiss(\lambda)$, $\lambda > 0$ – Пуассоновское

$F_X(x) = \dfrac{\Gamma(k+1, \lambda)}{k!}$

$E(X) = \lambda$

$Var(X) = \lambda$

$Y_i \sim P(\lambda_i)$ – незав., $i = \overline{1,n}$ $\Rightarrow Y = \sum_{i=1}^{n} Y_i \sim P(\sum_{i=1}^{n} \lambda_i)$

$Y = Y_1 + Y_2$, $Y_i \sim P(\lambda_i) \Rightarrow Y_1 | Y = y \sim Bin\left(y, \dfrac{\lambda_1}{\lambda_1+\lambda_2}\right)$

$P(\lambda) \xrightarrow{\lambda\to\infty} N(\lambda, \lambda)$, где $N(\mu, \sigma^2)$ "гаус." распр.

$f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)$

$E(X) = \mu$; $Var(X) = \sigma^2$

$X \sim Exp(\lambda)$ – показательное

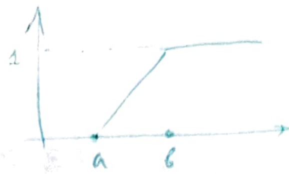
$F_X(x) = 1 - e^{-\lambda x}$; $f(x) = \lambda e^{-\lambda x}$, $x > 0$

$E(X) = 1/\lambda$, $Var(X) = 1/\lambda^2$

$P(X > s+t \mid X \ge s) = P(X > t)$ – отс. памяти

9 $X \sim U[a;b]$ — равномерное

$$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1, & x \geq b \end{cases}$$



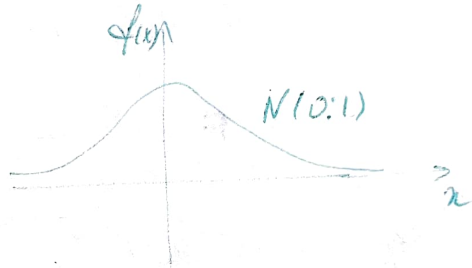
$$f(x) = \begin{cases} 0 & x \notin [a;b] \\ \frac{1}{b-a}, & x \in [a;b] \end{cases}$$

$E(X) = \frac{a+b}{2}$

$Var(X) = \frac{(b-a)^2}{12}$

10 $X \sim N(\mu, \sigma^2)$ — нормальное (гауссовское)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



$E(X) = \mu$

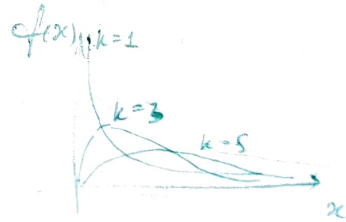
$Var(X) = \sigma^2$

11 χ^2_n — хи-квадрат с n ст. свободы

$X = \sum_{i=1}^{k} Z_i^2 \sim \chi^2_k$, где $Z_i \sim N(0;1)$ — i.i.d.

· если $X_i \sim N(\mu, \sigma^2)$, i.i.d., то $Y = \sum_{i=1}^{k} \left(\frac{X_i - \mu}{\sigma}\right) \sim \chi^2_k$

$$f(x) = \Gamma(2, k/2) = \frac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}$$



$$F(x) = \frac{\gamma(k/2, x/2)}{\Gamma(k/2)}$$

$X \sim \chi^2_{k_1}$, $Y \sim \chi^2_{k_2}$ => $X + Y \sim \chi^2_{k_1+k_2}$

$E(X) = k$

$Var(X) = 2k$

$Y \sim \chi^2_k \overset{d}{\approx} N(k, 2k)$, т.е. $\frac{Y-k}{\sqrt{2k}} \xrightarrow{d} N(0;1)$

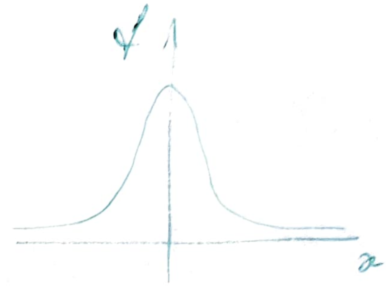
· $X \sim \chi^2_2 \equiv Exp(1/2)$

· $X \sim \chi^2_{k_1}$, $Y \sim \chi^2_{k_2}$ => $F = \frac{X/k_1}{Y/k_2} \sim F(k_1, k_2)$ — Фишера

(12) $t = \dfrac{Y_0}{\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n} Y_i^2}} \sim t_n$ — стьюдента.

где $Y_i \sim N(0,1)$, i.i.d

$f_t(y) = \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\dfrac{y^2}{n}\right)^{-\frac{n+1}{2}}$



$E(t) = 0, \quad n > 1$

$var(t) = \dfrac{n}{n-2}, \quad n > 2$

$\{t_n\} \sim t_n \Rightarrow t_n \xrightarrow{d} N(0,1)$

$t \sim t_n \Rightarrow t^2 \sim F(1, n)$

(13) $F = \dfrac{Y_1/d_1}{Y_2/d_2} \sim F(d_1, d_2)$ — фишера

где $Y_i \sim \chi^2_{d_i}$, $i = 1, 2$

$f(x) = \dfrac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\cdot\Gamma\left(\frac{m}{2}\right)} \cdot \dfrac{n^{n/2}\, m^{m/2}\, x^{n/2-1}}{(nx+m)^{n+m/2}}, \quad x > 0$

$E(F(n,m)) = \dfrac{m}{m-2}, \quad m > 2$

$Var(F(n,m)) = \dfrac{2m^2(n+m-2)}{n\cdot(m-2)^2(m-4)}, \quad m > 4$

$F \sim F(d_1, d_2) \Rightarrow 1/F \sim F(d_2, d_1)$

$F_{d_1,d_2} \xrightarrow{d} \delta(x-1)$, где $\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0, & x \neq 0 \end{cases}$; $\int_{-\infty}^{+\infty} \delta(x)\,dx = 1$

(14) $X \sim N(\mu, \delta^2) \Rightarrow P(X \in (a, b)) = \Phi\left(\frac{b-\mu}{\delta}\right) - \Phi\left(\frac{a-\mu}{\delta}\right)$

$X \sim N(\mu, \delta^2) \Rightarrow P(|X-\mu| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\delta}\right)$ $\left(= 2\Phi\left(\frac{3\delta}{\delta}\right) = 0{,}9973\right)$

$X \sim N(\mu, \delta^2) \Rightarrow (\mu - 3\delta; \mu + 3\delta)$ наир. 99,7 % значений

(15) Н-во Маркова $P(X \geq \varepsilon) \leq \dfrac{E(X)}{\varepsilon}$, $\forall \varepsilon > 0$

Н-во Чебышева $P(|X - E(X)| < \varepsilon) \geq 1 - \dfrac{Var(X)}{\varepsilon^2}$ $\forall \varepsilon > 0$

WLLN $\lim\limits_{n\to\infty} P(|\overline{X}_n - E(X_i)| < \varepsilon) = 1$ $\forall \varepsilon > 0$ $\Rightarrow \overline{X}_n \xrightarrow{P} \mu$

где $\overline{X}_n = \frac{1}{n}\sum X_i$ — выб. среднее

$E(X_i) = \mu$, X_i — i.i.d.

SLLN $P(\lim\limits_{n\to\infty} \overline{X}_n = \mu) = 1$

Теор. Бернулли $\lim\limits_{n\to\infty} P(|\frac{m}{n} - p| < \varepsilon) = 1$ $\quad \frac{m}{n} \xrightarrow{P} p$

Теор. Пуассона p_i ; m_n/n – част. появл. A в n исп.

$$\lim_{n\to\infty} P(|\frac{m_n}{n} - \overline{p}| < \varepsilon) = 1$$

Центр. пред. теор:

Классическая ЦПТ

X_i ~ i.i.d. : $E(X_i) = \mu$; $Var(X_i) = \sigma^2$

$X = \sum^n X_i$

$\frac{X - \mu n}{\sigma\sqrt{n}} \xrightarrow{d} N(0;1) \Rightarrow \sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{d} N(0;1)$

Локальная ЦПТ

$\{X_i\}$ – абс. непр. $\Rightarrow$ сх. плотность

$\frac{X - \mu n}{\sigma\sqrt{n}} \xrightarrow{d} N(0;1)$ – абс. интегрир.

16 · $\lim\limits_{n\to\infty} P(|F_n(x) - F(x)| < \varepsilon) = 1 \Rightarrow F_n(x) \xrightarrow{P} F(x)$

где $F_n(x) = \frac{n_x}{n}$, n_x – число знач. выб. $(< x)$

· $\bar{x} = \frac{1}{n}\sum x_i$ – выб. среднее

$\hookrightarrow \frac{1}{n}\sum(x_i - \bar{x}) = 0$

$\bar{x} = \frac{1}{n}(\bar{x}_1 \cdot n_1 + \ldots + \bar{x}_k \cdot n_k)$

· $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$ – выб. дисп.

· $\hat{\alpha}_k = \frac{1}{n}\sum x_i^k$; $\hat{\mu}_k = \frac{1}{n}\sum(x_i - \bar{x})^k$

17 $\hat{\theta}_n = \hat{\theta}(X_1, \ldots, X_n)$

– несмещ. $E(\hat{\theta}_n) = \theta$

– сост. $\lim\limits_{n\to\infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1 \Rightarrow \hat{\theta}_n \xrightarrow{P} \theta$

– эффект. $Var[\hat{\theta}_n]$ min по ср. с др.

(Т) $E(\hat{\theta}_n) = \theta$

$Var(\hat{\theta}_n) \xrightarrow{d} 0 \Rightarrow \hat{\theta}_n \xrightarrow{P} \theta$

(18) ГС: $E(X_i) = \mu$, $Var(X_i) = \delta^2$

1) $E(\bar{X}) = \frac{1}{n} \sum E(X_i) = \mu$

$\lim\limits_{n\to\infty} Var(\bar{X}) = \lim\limits_{n\to\infty} \left(\frac{\delta^2}{n}\right) = 0$

2) $E(\hat{\delta}^2) = \frac{n-1}{n} \delta^2 \Rightarrow \hat{\delta}^2$ - смещ.

$S^2 = \frac{n}{n-1} \hat{\delta}^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$ - испр. выб. д.

$\hat{\delta}^2 \xrightarrow{P} \delta^2$

(19) Теор. Слуцкого $x_n \xrightarrow{P} a$, $y_n \xrightarrow{P} b \Rightarrow f(x_n, y_n) \xrightarrow{P} f(a, b)$

Нер. по РФК (если $e = 1$) $e(\hat{\Theta}_n) = \frac{\Delta_n}{Var(\hat{\Theta}_n)}$, где $\Delta_n = \frac{1}{n \cdot J(\theta)}$, $J(\theta) = E\left[\left(\frac{\partial \ln f(x;\theta)}{\partial\theta}\right)^2\right]$

Мет. моментов: $\hat{\alpha}_1 = \bar{x} = E(x)$

$\hat{\mu}_2 = S^2 = Var(x)$

Мет. макс. правд.: $L(x_1, \dots, x_n, \theta) = \prod^n P(x_i, \theta) = \prod^n f(x_i, \theta)$

1. $L(x_1, \dots, x_n, \theta) \to \max\limits_{\theta}$

2. $\frac{\partial \ln L}{\partial \theta} = 0$

МНК $T(a, b) = \sum (ax_i + b - y_i)^2 \to \min\limits_{a, b}$

(20) $P(|\hat{\Theta}_n - \Theta| < \varepsilon) = \gamma$, где γ - надежность , ε - точность

$(\hat{\Theta}_n - \varepsilon ; \hat{\Theta}_n + \varepsilon)$ - д.и.

$P(\hat{\Theta}_1 < \Theta < \hat{\Theta}_2) \xrightarrow{d} \gamma$

↳ 1. $\hat{\Theta}_n - \frac{\delta(\hat{\Theta}_n)}{\sqrt{n}} u_\gamma < \Theta < \hat{\Theta}_n + \frac{\delta(\hat{\Theta}_n)}{\sqrt{n}} \cdot u_\gamma$, где $\Phi(u_\gamma) = \frac{\gamma}{2}$

2. $g(\theta) = \int \frac{d\theta}{\delta(\theta)}$

$g(\hat{\Theta}_n) - \frac{1}{\sqrt{n}} u_\gamma < g(\Theta) < g(\hat{\Theta}_n) + \frac{1}{\sqrt{n}} \cdot u_\gamma$

$\Downarrow$

$< \Theta <$

α - H_0 - верна, но отв-ся

β - H_0 - неверна, но прин-ся

$E(Y) = E[E(Y|X)]$

$E(Y|X) = \mu_Y \Rightarrow cov(Y,X) = 0 \quad cor(Y,X) = 0$

$(=const)$

$E(Y^2) = \sigma_Y^2 + \mu_Y^2 \quad (E(Y) = \mu_Y,\ Var(Y) = \sigma_Y^2)$

$E(XY) = \sigma_{XY} + \mu_X \mu_Y \quad (cov(X,Y) = \sigma_{XY} \quad E(X)\cdot E(Y) = \mu_X \mu_Y)$

Корреляционное н-во $|cor(X,Y)| \leq 1$

$|\sigma_{XY}| \leq \sqrt{\sigma_X^2 \sigma_Y^2}$

$Var(S) = Var(E(S|Z)) + E(Var(S|Z))$

$Var(AY) = A\, Var(Y) \cdot A'$

$(AB)' = B'A'$

$(A^{-1})' = (A')^{-1}$

$(X'X)' = X'X'' = X'X$

$((X'X)^{-1})' = (X'X)^{-1}$

Эконометрика

① Предположения МНК

a) ПЛР: 1. $E(u_i | X) = 0$

2. (X_i, y_i) – i.i.d.

3. $0 < E(X_i^4) < \infty$

$0 < E(y_i^4) < \infty$ $\Rightarrow$ $\gamma_4 = \dfrac{E(y_i - \mu_Y)^4}{\sigma_Y^2} \in [-2; \infty)$

b) МЛР: 1. $E(u_i | X) = 0$

2. $(X_{1i}, \dots, X_{ki}, y_i)$ – i.i.d.

3. $\gamma_4 \in [-2; \infty)$

4. Отсутствует сов. мультиколлинеарность

c) с контр. переменной: 1'. $E(u_i | X_i, W_i) = E(u_i | W_i)$

d) $Y_i = X_i' \beta + u_i$:

1. $E(U | X) = 0$

2. (X_i, y_i) – i.i.d.

3. $\gamma_4 \in [-2; \infty)$

4. X – полного ранга

5. $E(U \cdot U' | X) = \sigma_u^2 \cdot I_n$

(т.к. гомоск. $Var(u_i | X_i) = \sigma_u^2$)

6. $(U | X) \sim N(0_n; \sigma_u^2 \cdot I_n)$

(т.к. $(u_i | X_i) \sim N(0, \sigma_u^2)$)

e) с инстр. переменной

1. $E(u_i | W_{1i}, \dots, W_{ri}) = 0$, где W_i – экзогенные регр., кот. не корр. с u_i или явл. контр. переменными

2. $(X_{1i}, \dots, X_{ki}, W_{1i}, \dots, W_{ri}, Z_{1i}, \dots Z_{ni}, Y_i)$ – i.i.d.

3. $\gamma_4 \in [-2; \infty)$

допущ. Z_k { 4. релевантность Z_k + отс. сов. МК + $corr(X_i, Z_i) \neq 0$

5. экзогенность : $corr(Z_i, u_i) = 0$

② Св-ва МНК-оценок

1) $(\hat\beta_0, \dots, \hat\beta_k) \overset{d}{\sim} N(E, \Sigma)$

для ПЛР: $\hat\beta_0 \overset{d}{\sim} N(\beta_0, \sigma^2_{\hat\beta_0})$

$\hat\beta_1 \overset{d}{\sim} N(\beta_1, \sigma^2_{\hat\beta_1})$

$$Cov(\hat\beta_0, \hat\beta_1 | X) = \frac{-\bar{x} \cdot \sigma^2}{n \sum (x_i - \bar{x})^2}$$

$$Var(\hat\beta_0 | X) = \frac{\sigma^2 \cdot \sum x_i^2}{n \cdot \sum (x_i - \bar{x})^2}$$

$$Var(\hat\beta_1 | X) = \frac{\sigma^2}{n \cdot \sum (x_i - \bar{x})^2}$$

где $\sigma^2_{\hat\beta_0} = \frac{1}{n} \frac{Var[H_i u_i]}{[E(H_i^2)]^2}$, $H_i = 1 - \frac{\mu_x}{E(X_i^2)} \cdot X_i$

$\sigma^2_{\hat\beta_1} = \frac{1}{n} \cdot \frac{Var[(X_i - \mu_x) u_i]}{(Var(X_i))^2}$

2) $\hat\beta_0, \dots, \hat\beta_k$ – несмещенные (усл.), состоятельные

3) $\frac{1}{n} \sum \hat{u}_i = 0$

$\frac{1}{n} \sum \hat{Y}_i = \bar{Y}$

$\sum \hat{u}_i \cdot x_i = 0$

$S_{\hat{u}x} = \overline{x \cdot \hat{u}} - \bar{x} \cdot \bar{\hat{u}} = 0$

4) $TSS = RSS + ESS$

$$\sum (Y_i - \bar{Y})^2 = \underbrace{\sum (Y_i - \hat{Y}_i)^2}_{\sum \hat{u}_i^2} + \underbrace{\sum (\hat{Y}_i - \bar{Y})^2}_{\hat\beta_1^2 \cdot \sum (X_i - \bar{X})^2}$$

5) С контр. переменной $y_i = \beta_0 + \beta_1 x_i + \beta_2 w_i + u_i$

$\hat\beta_1$ – несмещ. с прич.-след. интерпретацией

$\hat\beta_2$ – смещ. в общ. случае

$E(\hat\beta_2) = \beta_2 + \gamma_2$,

т.к. $E(u_i | w_i) = \gamma_1 w_i + \gamma_0$ $\beta \neq 0$

$\hat\beta_2 = \beta_2 + \beta_3 \cdot \frac{Cov(x_2, x_3)}{Var(x_2)} \neq 0$, т.к. $+ \frac{Cov(x_2, u)}{Var(x_2)}$

$\hat\beta_1 \xrightarrow{p} \beta_1 + \underbrace{\cancel{\rho_{xu}}^{0} \frac{\sigma_u}{\sigma_x}}_{\text{вел. смещ.}}$

6) $Var(\hat\beta_j | X) = \frac{\sigma^2}{RSS_j}$, где RSS_j из регр. x_j (x_{-j})

$\widehat{Var}(\hat\beta_j | X) = \frac{\hat\sigma^2}{RSS_j} \Rightarrow se(\hat\beta_j)$

МЛР:

1) $X_{1\times m}$ – i.i.d., $E[(X_i - \mu_x)\cdot(X_i - \mu_x)^T] = \Sigma_x$

то ЦПТ: $\sqrt{n}(\bar{X} - \mu_x) \xrightarrow{d} N(0_m, \Sigma_x)$

2) $\sqrt{n}(\hat{\beta} - \beta) \overset{a}{\sim} N(0_{k+1}, \Sigma_{\sqrt{n}(\hat{\beta}-\beta)})$

$\hat{\beta} \overset{a}{\sim} N(\beta, \Sigma_{\hat{\beta}})$

где $\Sigma_{\sqrt{n}(\hat{\beta}-\beta)} = \left(E(X_i X_i')\right)^{-1} \cdot E((X_i u_i)\cdot(X_i u_i)') \cdot (E(X_i X_i'))^{-1}$

$\Sigma_{\hat{\beta}} = \frac{1}{n} \cdot \Sigma_{\sqrt{n}(\hat{\beta}-\beta)}$

3) $\hat{\beta} \sim N(\beta, \Sigma_{\hat{\beta}|X})$

где $\Sigma_{\hat{\beta}|X} = \sigma_u^2 (X'X)^{-1}$

4) $\mathrm{Var}(\hat{\beta}|X) = \sigma^2 (X' \cdot X)^{-1}$

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

$\frac{RSS}{\sigma^2} | X \sim \chi^2_{n-k}$

$\frac{RSS}{\sigma^2} \sim \chi^2_{n-k}$

③ МНК – оценки:

1) $y_i = \beta_1 \cdot x_i + u_i$

$$\hat\beta_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

2) $y_i = \beta_0 + \beta_1 \cdot x_i + u_i$

$$\partial Q / \partial\beta \Rightarrow \begin{cases} \sum \hat u_i = 0, & \hat u \perp 1 \\ \sum x_i \hat u_i = 0, & \hat u \perp x \end{cases}$$

$$\hat\beta_0 = \bar y - \hat b_1 \cdot \bar x$$

$$\hat\beta_1 = \frac{\sum (x_i - \bar x) \cdot (y_i - \bar y)}{\sum (x_i - \bar x)^2} = \frac{\sum (y_i - \bar y) x_i}{\sum (x_i - \bar x) \cdot x_i} =$$

$$= \frac{S_{xy}}{S_x^2} = \frac{r_{xy} S_x S_y}{S_x^2} = \beta_1 + \frac{\sum (x_i - \bar x) u_i}{\sum (x_i - \bar x)^2}$$

3) $Y_i = X_i' \beta + u_i$

$Y = X' \cdot \beta + U$

$\partial Q / \partial \beta \Rightarrow X' \cdot (Y - X \cdot \beta) = 0$

$X' \cdot Y = X' \cdot X \cdot \beta$

$\hat\beta = (X'X)^{-1} X' \cdot Y = \beta + (X'X)^{-1} \cdot X'U$

4) $P_x = X (X'X)^{-1} X'$

$M_x = I_n - P_x$

$\hat Y = P_x \cdot Y = X \hat\beta$

$\hat U = M_x Y = M_x U = Y - \hat Y$

$$\hat Y' \hat U = Y' \underbrace{P_x' M_x}_{=0} Y = 0$$

Св-ва: $P_x = P_x P_x$ $\quad M_x = M_x \cdot M_x$

$P_x X = X$ $\quad M_x X = 0$

④ Качество приближения

a) $SER = S_{\hat{u}}$

$$S^2_{\hat{u}} = \frac{1}{n-2} \sum \hat{u}_i^2 = \frac{RSS}{n-2} \qquad \left(\bigstar \quad S^2_y = \frac{1}{n-1} \sum (y_i - \bar{y})^2 \right)$$

$$S^2_{\hat{u}} = \frac{1}{n-k} \sum u_i^2 = \frac{RSS}{n-k}$$

b) $$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = \left(\operatorname{corr}(x,y)\right)^2 = r^2_{xy}$$

$\in [0;1]$

$$= 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}^2_y}$$

а) $R^2_{x(y)} = R^2_{y(x)} \iff r_{xy} = r_{yx}$

б) $R^2 = 0,\ ESS = 0 \iff \hat{\beta}_1 = 0$

с) Нарушение св-в при $\beta_0 = 0$

$$\hat{\beta}_1 \neq \frac{S_{xy}}{S^2_x} \Rightarrow \sum \hat{u}_i \neq 0 \Rightarrow TSS \neq RSS + ESS$$

$$R^2_{uc} = 1 - \frac{RSS}{\sum y_i^2} = \frac{\sum \hat{y}_i^2}{\sum y_i^2}$$

c) $\bar{R}^2$ – скорректированный, т.к. R^2 м.б. завышен

$$\bar{R}^2 = 1 - \frac{n-1}{n-k} \cdot \frac{RSS}{TSS} = 1 - \frac{S^2_{\hat{u}}}{S^2_y}$$

$k\uparrow$ $RSS\downarrow$ $\quad$ $k\uparrow$ $\frac{n-1}{n-k}\uparrow$

$R^2\uparrow$ $\quad$ $R^2\downarrow$

d) Информационные критерии

↳ компромисс: точность vs сложность сравн.

$$AIC = \ln \frac{RSS}{n} + \frac{2(k+1)}{n} + 1 + \ln 2\pi$$ (для выб. разн. объем.)

$$BIC = \ln \frac{RSS}{n} + \frac{\ln n \cdot (k+1)}{n} + 1 + \ln 2\pi$$ (больший штраф за $k\uparrow$)

чем меньше, тем лучше модель

$BIC = \ln n \cdot k - 2l \,/\, AIC = 2k - 2l;\quad AIC^* = \frac{2k}{n} - \frac{2L}{n} = \frac{2k}{n} + \frac{1}{n}\cdot\left[n\left(\ln \frac{2\pi RSS}{n} + 1\right)\right]$

5 Теорема Гаусса – Маркова

a) Если 1) $E(u_i | X) = 0$, то $\hat{\beta}$ – BLUE

2) $Var(u_i | X) = \sigma_u^2$

3) $E(u_i u_j | X) = 0$

b) Если 1) $E(U | X) = 0_n$, то $\hat{\beta}$ – BLUE

2) $E(UU' | X) = \sigma_u^2 \cdot I_n$

3) X – полного ранга

6 Теорема Фриша – Во

1) Вычислить остатки $x_k (X_{-k}) \Rightarrow \tilde{x}_k$

2) Вычислить остатки $y (X_{-k}) \Rightarrow \tilde{y}$

3) Оценить $\tilde{y}(\tilde{x}_k)$

<u>теор.</u> МНК-оценка коэф. в регр. $\tilde{y}(\tilde{x}_k)$ равна МНК-оц. β_k при x_k в $y(X)$

интерпретация: $\hat{\beta}_k$ – оценка влияния x_k на y при фикс. x_{-k}

7 Мультиколлинеарность

a) совершенная ($\exists$ регрессор – точная лин. комб. других)

↳ невозможно получить оценки из-за деления на 0

$$\hat{\beta} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(\lambda - \lambda)^2}$$

$$x_{1i} = \lambda \cdot (\underbrace{x_{0i}}_{=1})$$

причины: ошибки специф. ловушка перем.

b) несовершенная

↳ оценки несмещ., неэфф.

↳ $SE(\hat{\beta})$ завышены, t занижены

признаки:

1) $\exists\ |\rho_{x_i x_j}| > 0,7$

2) $Det(\Sigma_{corr}) \approx 0$

3) $\underbrace{SE(\hat{\beta})\ \text{высокие}}_{\text{незнач. коэф.}}$ при $\underbrace{R^2\ \text{высоком}}_{\text{знач. модели}}$

4) введ. новой переменной

$\Rightarrow$ существ. изм. $\hat{\beta}$ при малом. изм. R^2

борьба:

I Использование внешней информации

$$y_i = \beta_0 + \beta_1 \cdot x_i + \beta_2 \cdot z_i + u_i$$

$$\hat{y}' = \beta_0' + \beta_1' \cdot x' \Rightarrow \hat{\beta}_1'$$

$$\underbrace{y - \beta_1 x}_{} = \beta_0 + \beta_2 \cdot z_i + u_i$$

II Ограничения: $\beta_3 = \beta_4 \Rightarrow \beta_3 (x_i + z_i)$

III Др. приемы и нов. методы

⑧ Гетероскедастичность

$Var(u_i|X) \neq const$

Следствия: 1) $\hat{\beta}$ – ас. несм., ас. норм., сост.

$\hat{Var}$ – неэф., несост.

2) $\frac{RSS}{n-k} \xrightarrow{d} \sigma^2$

3) $\frac{RSS}{\sigma^2} | X \nsim \chi^2_{n-k}$

$t = \frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \nsim t_{n-k}$

$\overset{d}{\nsim} N(0;1)$

$F = \frac{RSS_R - RSS_{UR}/q}{RSS_{UR}/n-k} \nsim F_{q,n-k}$

$\overset{d}{\nsim} \chi^2_2$

ВМНК

а) недоступный

пусть $Var(u_i|X_i) = \lambda \cdot h(X_i)$

$Y_i = \beta_0 + \beta_1 \cdot x_i + u_i \quad | : \sqrt{h(x_i)}$

$\tilde{Y}_i = \beta_0 \cdot \tilde{x}_{0i} + \beta_1 \cdot \tilde{x}_{1i} + \tilde{u}_i \Rightarrow Var(\tilde{u}_i|X_i) = \lambda$

↳ по $Y_i(\tilde{x}_{0i}, \tilde{x}_{1i})$ оцениваем $\hat{\beta}_1$

b) доступный

· $Y_i(X_i) \Rightarrow$ оценить остатки $\hat{u}_i$

· оценить модель $Var(u_i|X_i)$ $[= \theta_0 + \theta_1 \cdot X_i^2]$

· $\hat{u}_i^2(X_i^2) \Rightarrow$ оценки $\hat{\theta}_0 \xrightarrow{p} \theta_0$, $\hat{\theta}_1 \xrightarrow{p} \theta_1 \Rightarrow \hat{Var}(u_i|X_i)$

$\Rightarrow$ доступ. ВМНК

⑨ Автокорреляция

$E(u_i u_j | X) \neq 0$ $\quad$ $Cov(u_i u_j | X) \neq 0$

Следствия: $\hat\beta$ – несмещ., неэф., сост.

$\hat{Var}(\hat u_i)$, $\hat{Var}(\hat\beta)$ – несост., смещ.

SE занижены, t завышены

ОМНК

$\min\limits_{\beta} e^T W e$, где W – м-ца весов (для МНК: $W = I$)

$W = P^T P$:

$$[P(y - X\beta)]^T [P(y - X\beta)] = (y^* - X^*\beta)^T (y^* - X^*\beta)$$

$$\hat\beta_{GLS} = (X^T \Sigma^{-1} \cdot X)^{-1} \cdot X^T \cdot \Sigma^{-1} \cdot Y$$, где $W = \Sigma^{-1}$ – ков. м-ца сл. ош. (исп. в кач-ве вес.)

$$V(\hat\beta_{GLS}) = (X^T \cdot \Sigma^{-1} \cdot X)^{-1}$$

Процедура Дарбина:

в случае AR(1)

$$\boxed{\hat\rho = 1 - \frac{DW}{2}}$$

$Y_t = \beta_0 + \beta_1 \cdot x_t + u_t$,

где $u_t = \rho \cdot u_{t-1} + \varepsilon_t$

то авторегр. преобразование: $\boxed{\tilde Y_t = Y_t - \rho \cdot Y_{t-1}}$

$$\underbrace{Y_t - \rho \cdot Y_{t-1}}_{\tilde Y_t} = \underbrace{\beta_0 (1 - \rho)}_{\beta_0'} + \beta_1 \underbrace{(X_t - \rho \cdot X_{t-1})}_{\tilde X_t} + \underbrace{u_t - \rho \cdot u_{t-1}}_{\varepsilon_t}$$

(без учета первого наблюдения)

Поправка Прайса – Уинстена:

для учета первого набл. x_i, y_i умн. на $k = \sqrt{1 - \rho^2}$

где k – поправочный коэф. для соизмеримости данных

$\Rightarrow \tilde X_1 = \sqrt{1 - \hat\rho^2} \cdot X_1$; $\tilde Y_1 = \sqrt{1 - \hat\rho^2} \cdot Y_1$

Процедура Кохрейна – Оркатта:

$\hat\rho_{i+1}$ – последнее знач. итеративной процедуры (если не сход. $\Rightarrow$ AR более высокого порядка)

1) $Y_t(X_t)$ => оц. остатки $\hat{u}_t$

2) $\hat{u}_t = \rho \cdot \hat{u}_{t-1} + \varepsilon_t$ => оц. пар. $\hat{\rho}$

3) по проц. Дарбина => $\tilde{Y}_t = \beta_0' + \beta_1 \cdot \tilde{X}_t + \varepsilon_t$ => $\hat{\beta}_0, \hat{\beta}_1$

(+ если необх. восст. 1 набл. поправк. Пр.-Уин.)

4) повторять процедуру пока для $\alpha > 0$

$|\rho_i - \rho_{i+1}| < \alpha$

=> $\hat{\rho}_{i+1}$

⑩ Ошибки спецификации: пропущенная переменная

$\Downarrow$ w_i — контрольная переменная

$$Y_i = \beta_0 + \beta_1 \cdot x_i + \beta_2 \cdot w_i + u_i$$

смещ. из-за проп. пер.

$$\boxed{\hat{\beta}_1 \xrightarrow{p} \beta_1 + \rho_{xu} \cdot \frac{\sigma_u}{\sigma_x}}$$

Смещение $\Leftarrow$ $E(u_i | X) \neq 0$, $cov(x_i, u_i) = \rho_{xu} \neq 0$

+ оц. несост.

эффективная w_i: $\rho_{xu} = 0$ при включении в регр.

$\Leftrightarrow$ $E(u_i | x_i, w_i) = E(u_i | w_i)$

$= \gamma_0 + \gamma_2 w_i$

$$v_i = u_i - E(u_i | x_i, w_i)$$

$$Y_i = \beta_0 + \beta_1 \cdot x_i + \beta_2 \cdot w_i + \gamma_0 + \gamma_2 w_i + v_i$$

$$= \underbrace{\beta_0 + \gamma_0}_{\delta_0} + \beta_1 \cdot x_i + \underbrace{(\beta_2 + \gamma_2)}_{\delta_2} \cdot w_i + v_i$$

$\Rightarrow$ $\hat{\delta}_0, \hat{\beta}_1, \hat{\delta}_2$ — несмещ. оценки

$$E(\hat{\beta}_1) = \beta_1$$

$$E(\hat{\beta}_2) = \delta_2 = \beta_2 + \underbrace{\gamma_2}$$

смещ. из-за

$\rho_{xw} \neq 0$

кор-ии проп. и конт. пер.

$$\hat{\beta}_2 = \beta_2 + \gamma_2 \underbrace{\rho_{x \cdot w} \cdot \frac{\sigma_w}{\sigma_x}}_{\frac{Cov(x_i, w_i)}{Var(x_i)}} + \underbrace{\frac{Cov(x_i, u_i)}{Var(x_i)}}_{=0}$$

Алгоритм: 1) опр. целевую перем.; пот. смещ. из-за проп. пер.

2) вкл. проп. пер. / вкл. корр. с ней контр. пер w_i

3) оценка "баз." и альт. моделей на чувствительност

(11) Классическая модель с ошибками измерения

$Y_i = \beta_0 + \beta_1 X_i + u_i$; $E(u_i / X_i) = 0$

↑истинное

$Y_i = \beta_0 + \beta_1 \tilde{X}_i + \underbrace{\beta_1 (X_i - \tilde{X}_i) + u_i}_{v_i}$; Пусть $\tilde{X}_i = X_i + w_i$

↑набл.

$\rho_{wx} = \rho_{wu} = 0$

$\Rightarrow \hat{\beta}_1 \xrightarrow{p} \beta_1 - \beta_1 \cdot \frac{\sigma_w^2}{\sigma_x^2 + \sigma_w^2} = \beta_1 \cdot \frac{\sigma_x^2}{\sigma_x^2 + \sigma_w^2}$

$\hat{\beta}_1$ – несмещ., неэф., т.к. $\sigma^2_{v_i} = \sigma^2_{u_i} + \sigma^2_{w_i} > \sigma^2_{u_i}$

(12) Смещение из-за одновременной причинности

$Y_i = \beta_0 + \beta_1 X_i + u_i$

$X_i = \gamma_0 + \gamma_1 Y_i + v_i$

$\Rightarrow cov(X_i, u_i) = \frac{\gamma_1}{1 - \gamma_1 \cdot \beta_1} \cdot \sigma_u^2$

$\hat{\beta}_1 = \beta_1 + \frac{Cov(X_i, u_i)}{\sigma_x^2}$

если $\beta_1 < 1/\gamma_1$, то $\hat{\beta}$ смещ. вниз

(13) Метод инструментальных переменных

при нарушении усл. экзогенности (пер. кор-ют со случ. ош. ⇒ опр. внутри)

Z_i – инструмент: релевантность $corr(Z_i; X_i) \neq 0$; экзогенность $corr(Z_i; u_i) = 0$

$Y_i = \beta_0 + \beta_1 \cdot X_{1i} + \ldots + \beta_k \cdot X_{ki} + \beta_{k+1} \cdot W_{1i} + \ldots \beta_{k+r} W_{ri} + u_i$

$X_{1i}, \ldots, X_{ki}$ – эндогенные регр., потенц. корр. с u_i

$W_{1i}, \ldots, X_{ri}$ – экзогенные регр., не корр. с u_i или контр. пер.

доб. чтобы удост. что Z_i не корр. с u_i

$Z_{1i}, \ldots, Z_{mi}$ – инстр. переменные ($m = k$ коэф. опр. точно; $m > k$ коэф. переопр.; $m < k$ коэф недоопр.)

TSLS (2МНК): $Y_i = \beta_0 + \beta_1 \cdot X_i + \beta_2 W_{1i} + \ldots + \beta_{1+r} W_{ri} + u_i$

$\hat{\beta}_1$ – смещ.

1 шаг: $X_i = \pi_0 + \pi_1 Z_{1i} + \ldots + \pi_m Z_{mi} + \pi_{m+1} \cdot W_{1i} + \ldots + \pi_{m+r} W_{ri} + v_i$

$\Rightarrow \hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_{1i} + \ldots + \hat{\pi}_{m+r} \cdot W_{ri}$

$[\rho_{uz} = 0 \Rightarrow \rho_{\hat{X}u} = 0]$

2 шаг: $Y_i = \beta_0 + \beta_1 \cdot \hat{X}_i + \beta_2 \cdot W_{1i} + \ldots + \beta_{1+r} W_{ri} + u_i \Rightarrow \hat{\beta}_0^{TSLS}, \ldots, \hat{\beta}_{1+r}^{TSLS}$

Эмп. правило тестирования слабых инстр.:

(объясн. малую часть изм. x_i)

$H_0: \pi_1 = 0, \dots \pi_m = 0$

$$F = \frac{RSS^{R} - RSS^{UR} / m}{RSS^{UR} / n-(m+2+1)} \sim F(m, n-(m+2+1))$$

набор инстр слабый, если $F < 10$

Тест на сверхидентифицирующие ограничения (J – стат.)

$$\hat{u}_i^{TSLS} = \delta_0 + \delta_1 Z_{1i} + \dots + \delta_m \cdot Z_{mi} + \delta_{m+1} \cdot W_{1i} + \dots + \delta_{m+r} W_{ri} + e_i$$

$H_0: \delta_1 = \dots = \delta_m = 0$

$$F = \frac{RSS^{R} - RSS^{UR} / m}{RSS^{UR} / n-(m+2+1)} \sim F(m, n-(m+2+1))$$

$J = m \cdot F \overset{d}{\sim} \chi^2_{m-k}$, при гомоск. e_i

где $m-k$ – "степень сверхид",

k – число энд. регр.

если $m = k$, то $J = 0$