Анализ временных рядов

[THM] Вольда : $\forall$ слаб. стац. процесс предст. в виде $MA(\infty)$

Метод Юла – Волкера: пусть y_t – сл. стац. сл. процесс

$ACF(y_t)$: $\rho_s = \gamma_s / \gamma_0$

$PACF(y_t)$: φ_{kk} : $y_t^* = \varphi_{k1} y_{t-1}^* + \ldots + \varphi_{kk} y_{t-k}^* + \varepsilon_t$

$R^2 = 1$ – кат. регрессия

$$AIC(p,q) = \log \hat{\sigma}^2 + \alpha \frac{p+q}{T}$$

$$BIC(p,q) = \log \hat{\sigma}^2 + \log T \cdot \frac{p+q}{T}$$ – ас. правд. рез.

Тест Бокса – Пирса

низкая мощность, ≥ 150

Тест Люнга – Бокса

$H_0 : \rho_1 = \ldots = \rho_s = 0$ (предп. о WN)

$H_1 : \sum^s \rho_i^s > 0$

$$Q(s) = \underbrace{n}_{\text{число набл.}} \sum_{k=1}^{s} \underbrace{\hat{\rho}_k^2}_{\text{выб. AC k-пор.}}$$

$$Q(s) = n(n+2) \sum_{k=1}^{s} \frac{\hat{\rho}_k^2}{n-k}$$

$\overset{d}{\sim} \chi^2(s)$ ← число пров. лагов

$\overset{d}{\sim} \chi^2(s-p-q)$ – для ост. мод. $ARIMA(p,d,q)$

Тест Бройша – Годфри

$y_t = \alpha_0 + \alpha_1 x_{1t} + \ldots + \alpha_k x_{kt} + \varepsilon_t$

1) оц. МНК $\hat{\varepsilon}_t$ – ост.

2) $\hat{\varepsilon}_t = \alpha_0' + \alpha_1' x_{1t} + \ldots + \alpha_k' x_{kt} + v_t + \gamma_1 \hat{\varepsilon}_{t-1} + \ldots + \gamma_s \hat{\varepsilon}_{t-s} \Rightarrow R^2_{aux}$

$H_0 : \gamma_1 = \ldots = \gamma_s = 0$

$H_1 : \sum \gamma_k^2 > 0$

$$BG = (n - \underbrace{s}_{\text{пред. пор. AC}}) R^2_{aux} \overset{d}{\sim} F$$

$\overset{d}{\sim} \chi^2(s)$

Тест Дарбина – Уотсона $H_0 : \rho = 0$; $H_1 : \rho > 0$

$$DW = \frac{\sum^T (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum^T \hat{\varepsilon}_t^2} \overset{d}{\sim} 2 - 2\rho \in [0,4]$$

Тест Харке – Бера $H_0 : s = 0, k = 3$

$$JB = (T - p - q - 1)\left(\frac{\hat{S}^2}{6} + \frac{(\hat{k}-3)^2}{24}\right) \sim \chi^2_2$$

Врем. ряды

Одномерные → предск. завтр. знач. → наивные мод.; ARIMA (с автоподб.), ETS (θ - model), TBATS, PROPHET

Одномерные → предск. волатильности

Многомерные → структ.

Ряд → стационарный (стаб. во времени)

Ряд → нестационарный

$$\begin{cases} E(y_t) = \mu_h \\ Cov(y_t, y_{t-k}) = \gamma_k \end{cases} \Leftrightarrow \begin{cases} E(y_t) = const \\ Var(y_t) = const \\ Cov(y_t, y_{t-k}) = const \end{cases}$$

ex ε_i - i.i.d. $E(\varepsilon_i) = 0$ $Var(\varepsilon_i) = 7$ => стац. процесс

ex - // -

(y_t) : $y_0 = 0$

$y_1 = \varepsilon_1$

...

$y_t = \varepsilon_1 + \ldots + \varepsilon_t$

$E(y_t) = 0$

$Var(y_t) = 7t$

$cov(y_t, y_{t-k}) = 7(t-k)$ → $cov(y_7, y_3) = E(y_7, y_3) - \overset{0}{E(y_7)} \cdot E(y_3)$

$= E((\varepsilon_1 + \ldots + \varepsilon_7)(\varepsilon_1 + \ldots + \varepsilon_3)) = k_{min} Var(\varepsilon_i) = 3 \cdot 7 = 21$

=> y_t нестац.

ex - // -

$y_t = \varepsilon_t - 2\varepsilon_{t-1} + 3\varepsilon_{t-2}$, $\forall t$

a) $E(y_t) = 0$

b) $Var(y_t) = 7 + 4 \cdot 7 + 9 \cdot 7 = 98$

c) $Cov(y_t, y_{t-k})$

$k > 2$ $Cov = 0$

$k = 2$ $Cov = 3 \cdot 7 = 21$

$k = 1$ $cov = -2 \cdot 1 \cdot 7 + 3(-2) \cdot 7 = -56$

$k = 0$ $Cov = Var = 98$

=> y_t - стац.

Слабостац. процесс

- $E(y_t) = \mu$
- $Var(y_t) = \sigma_y^2$
- $Cov(y_t, y_{t-k}) = \gamma_k$

Сильно стац. процесс

- $P(y_t \le c) = F(c)$
- $P(y_t \le c_1,\ y_{t-1} \le c_2) = F(c_1, c_2)$
- $P(y_t \le c_1,\ y_{t-1} \le c_2,\ y_{t-2} \le c_3) = F(c_1, c_2, c_3)$

ex ε_i – ind.

$\varepsilon_{2k+1} \sim$

ε	4	6
p	1/2	1/2

$\varepsilon_{2k} \sim N(5;1)$

$$E(\varepsilon_i) = \left\{ \begin{array}{c} 1/2 \cdot 4 + 1/2 \cdot 6 \\ 5 \end{array} \right\} = 5$$

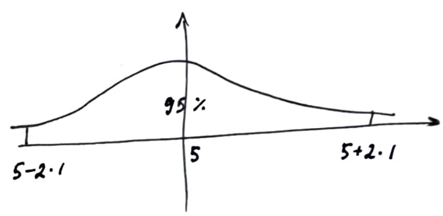
$$Var(\varepsilon_i) = \left\{ \begin{array}{c} 1/2 \cdot 16 + 1/2 \cdot 36 - 25 \\ 1 \end{array} \right\} = 1$$

$Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$

$\Rightarrow \{\varepsilon_i\}$ – слабостац.

$P(\varepsilon_1 > 7) = 0$

$P(\varepsilon_2 > 7) \approx 2{,}5\%$



WN
- $E(y_t) = 0$
- $Var(y_t) = \sigma_y^2$
- $cov(y_t, y_{t-k}) = 0$

ex $\varepsilon_t \sim WN$

$y_t = 2 + 0{,}6\, y_{t-1} + \varepsilon_t$

(a) $y_0 = 0$

$y_1 = 2 + \varepsilon_1$

$y_2 = 2 + 0{,}6\,(2 + \varepsilon_1) + \varepsilon_2 = 3{,}1 + 0{,}6\,\varepsilon_1 + \varepsilon_2$

$y_3 = 2 + 0{,}6\,(3{,}1 + 0{,}6\,\varepsilon_1 + \varepsilon_2) + \varepsilon_3 = 3{,}86 + 0{,}36\,\varepsilon_1 + 0{,}6\,\varepsilon_2 + \varepsilon_3$

$E(y_t) = E(2 + 0{,}6\,y_{t-1} + \varepsilon_t) = 2 + 0{,}6\,E(y_{t-1})$

$Var(y_t) = Var(2 + 0{,}6\,y_{t-1} + \varepsilon_t) = 0{,}36\,Var(y_{t-1}) + \underbrace{Var(\varepsilon_t)}_{=const}$

$\{y_t\}$ – не стац.

b) $y_0 = 5 + \varepsilon_0$

$y_1 = 2 + 0{,}6\,(5 + \varepsilon_0) + \varepsilon_1 = 5 + 0{,}6\,\varepsilon_0 + \varepsilon_1$

$y_2 = 2 + 0{,}6\,(5 + 0{,}6\,\varepsilon_0 + \varepsilon_1) + \varepsilon_2 = 5 + 0{,}6^2\varepsilon_0 + 0{,}6\,\varepsilon_1 + \varepsilon_2$

...

$E(y_t) = 5$

$Var(y_t) = Var(2 + 0{,}6\,y_{t-1} + \varepsilon_t) = 0{,}36\,Var(y_{t-1}) + \delta_y^2$

$Var(y_0) = \delta_y^2$

$Var(y_1) = 0{,}36\,\delta_y^2 + \delta_y^2$

$Var(y_2) = 0{,}36\,(0{,}36\,\delta_y^2) + \delta_y^2$

...

$\{y_t\}$ - не стац

c) $y_0 = 5 + \varepsilon_0 + 0{,}6\,\varepsilon_{-1} + 0{,}6^2\,\varepsilon_{-2} + \ldots$

$y_1 = 5 + \varepsilon_1 + 0{,}6\,\varepsilon_0 + 0{,}6^2\,\varepsilon_{-1} + \ldots + 0{,}6^n\,\varepsilon_{-n+1}$

$y_2 = 5 + \varepsilon_2 + 0{,}6\,\varepsilon_1 + 0{,}6^2\,\varepsilon_0 + \ldots + 0{,}6^n\,\varepsilon_{-n+2}$

...

$E(y_t) = 5$

$Var(y_0) = Var(5 + \varepsilon_0 + 0{,}6\,\varepsilon_{-1} + \ldots) = \delta^2 + 0{,}6^2\delta^2 + 0{,}6^4\delta^2 + \ldots = \dfrac{\delta^2}{1 - 0{,}6^2} = \dfrac{\delta^2}{0{,}64}$

$Var(y_1) = \ldots = \dfrac{\delta^2}{0{,}64}$

$Var(y_t) = \ldots = \dfrac{\delta^2}{0{,}64}$

$Cov(y_1, y_2) = Cov(5 + \varepsilon_1 + \ldots,\ 5 + \varepsilon_2 + \ldots) = 0{,}6\,\delta_y^2 + 0{,}6^3\,\delta_y^2 + \ldots = \dfrac{0{,}6\,\delta^2}{0{,}64}$

$Cov(y_t, y_{t-1}) = Cov(5 + 0{,}6\,y_{t-1} + \varepsilon_t,\ y_{t-1}) = 0{,}6\,Var(y_{t-1}) = 0{,}6\,\dfrac{\delta^2}{0{,}64}$

$Cov(y_t, y_{t-2}) = 0{,}6\,Cov(y_{t-1}, y_{t-2}) = 0{,}6^2\,\dfrac{\delta^2}{0{,}64}$

$Cov(y_t, y_{t-k}) = 0{,}6^k\,\dfrac{\delta^2}{0{,}64}$

$\{y_t\}$ - стац.

(ex) X_t — i.i.d.

X_t	0	1
P	1/2	1/2

Y_t — i.i.d. $\quad Y_t \sim N(0;1) \leftarrow WN$

$Z_t = X_t \cdot (1 - X_{t-1}) \cdot Y_t$

$E(Z_t) \overset{i.i.d.}{=} E(X_t) \cdot E(1 - X_{t-1}) \cdot E(Y_t) = 0$

$Var(Z_t) = E(X_t^2 \cdot (1 - X_{t-1})^2 \cdot Y_t^2) - 0^2 = E(X_t^2) \cdot E((1 - X_{t-1})^2) \cdot E(Y_t^2) = \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = \frac{1}{4}$

$Cov(Z_t, Z_{t-1}) = E(Z_t \cdot Z_{t-1}) - 0 \cdot 0 = E(X_t \cdot X_{t-1} (1 - X_{t-1}) \cdot (1 - X_{t-2}) \cdot Y_t \cdot Y_{t-1})$

$= \underbrace{E(X_t)}_{=1/2} \cdot \underbrace{E(X_{t-1}(1 - X_{t-1}))}_{=0} \cdot \underbrace{E(1 - X_{t-2})}_{=1/2} \cdot \underbrace{E(Y_t)}_{=0} \cdot \underbrace{E(Y_{t-1})}_{=0} = 0$

$Cov(Z_t, Z_{t-k}) = 0$

$\Rightarrow Z_t$ — WN
- зависимы $\quad Z_{t-1} = -1{,}6 \Rightarrow Z_t = 0$
- од. распр. $\quad P(Z_t > 7) = P(Z_{t-1} > 7)$

[THM] критерий стационарности

ARMA ур-е имеет стац. реш. с $Cov(y_t, \varepsilon_{t+k}) = 0 \ \forall k > 0$ если и т. е. корни хар. ур-я по мод < 1 (после упрощ.)

(ex) $\varepsilon_t \sim WN$

$y_t = 0{,}3 y_{t-1} - 0{,}02 y_{t-2} + \varepsilon_t + 4 \varepsilon_{t-2}$

$y_t = b_1 \cdot \lambda^{t-1}$

$b_1 \cdot \lambda^{t-1} = 0{,}3 b_1 \cdot \lambda^{t-2} - 0{,}02 b_1 \cdot \lambda^{t-3}$

$\lambda^2 = 0{,}3\lambda - 0{,}02$ — хар. ур-е

$\left.\begin{array}{l}\lambda_1 = 0{,}1 \\ \lambda_2 = 0{,}2\end{array}\right\} \Rightarrow$ есть стац. реш. $\Rightarrow \{y_t\}$ — стац.

(ex)

$y_t = 0{,}6 y_{t-1} + 0{,}4 y_{t-2} + \varepsilon_t + 0{,}5 \varepsilon_{t-1}$

$\lambda^2 - 0{,}6\lambda - 0{,}4 = 0$

$\left.\begin{array}{l}\lambda_1 = 1 \\ \lambda_2 = -0{,}4\end{array}\right\} \Rightarrow \{y_t\}$ — не стац.

$y_t = 0{,}6 L y_t + 0{,}4 L^2 y_t + \varepsilon_t + 0{,}5 L \varepsilon_t$

(ex) $y_t = 0{,}6 y_{t-1} + 0{,}4 y_{t-2} + \varepsilon_t - \varepsilon_{t-1}$

$(1 - 0{,}6L - 0{,}4L^2) y_t = (1 - L) \varepsilon_t$

$+0{,}4(1 - L)(2{,}5 + L) y_t = (1 - L) \varepsilon_t$

$0{,}4 L y_t + y_t = \varepsilon_t$

$0{,}4 + \lambda = 0$

$\lambda = -0{,}4 \Rightarrow \{y_t\}$ — стац.

(ex) $y_t = y_{t-1} - 4 y_{t-2} + \varepsilon_t + 0{,}7 \varepsilon_{t-1}$

$\lambda^2 - \lambda + 4 = 0$

$\lambda_{1,2} = 0{,}5 \pm i\sqrt{3{,}75}$

$|\lambda_1| = \sqrt{\frac{1}{4} + \frac{15}{4}} = 2$

$|\lambda_2| = 2$

$\Rightarrow \{y_t\}$ — не стац.

ex

$y_t = \varepsilon_t - \varepsilon_{t-1}$ $\quad \varepsilon_t \sim WN$

$y_t = 0{,}7 y_{t-1} + \varepsilon_t + 8$

$y_t = 2 + 1{,}3 y_{t-1} - 0{,}4 y_{t-2} + \varepsilon_t$ } – стац

$Cov(y_t, \varepsilon_{t+k}) = 0$

1) $E(y_t) = 0$

$\gamma_k = Cov(y_t, y_{t-k})$ $\quad\quad \rho_k = corr(y_t, y_{t-k})$

$\gamma_0 = 2\sigma^2$ $\quad\quad \gamma_0 = 1$

$\gamma_1 = -\sigma^2$ $\quad\quad \gamma_1 = -1/2$

$\gamma_2 = 0$ $\quad\quad \gamma_2 = 0$

...

$pcorr(y_1, y_3 ; y_2) = ?$

3) $E(y_t) = E(2 + 1{,}3 y_{t-1} - 0{,}4 y_{t-2} + \varepsilon_t) = 2 + 1{,}3 E(y_{t-1}) - 0{,}4 E(y_{t-2}) + 0$

	$y_t = 2 + 1{,}3 y_{t-1} - 0{,}4 y_{t-2} + \varepsilon_0$
y_{t-1}	$\gamma_1 = 0 + 1{,}3\gamma_0 - 0{,}4\gamma_1 + 0$
y_{t-2}	$\gamma_2 = 0 + 1{,}3\gamma_1 - 0{,}4\gamma_0 + 0$
y_{t-3}	$\gamma_3 = 0 + 1{,}3\gamma_2 - 0{,}4\gamma_1 + 0$
y_t	$\gamma_0 = 0 + 1{,}3\gamma_1 - 0{,}4\gamma_2 + \sigma_\varepsilon^2$

$: \gamma_0$, т.к. $\rho_k = corr(y_t, y_{t-k}) = \frac{\gamma_k}{\sqrt{\gamma_0^2}} = \frac{\gamma_k}{\gamma_0}$

$\rho_1 = 0 + 1{,}3 - 0{,}4\rho_1 + 0 \Rightarrow \rho_1 = 13/14$

$\rho_2 = 0 + 1{,}3\rho_1 - 0{,}4 + 0 \Rightarrow \rho_2 = \ldots$

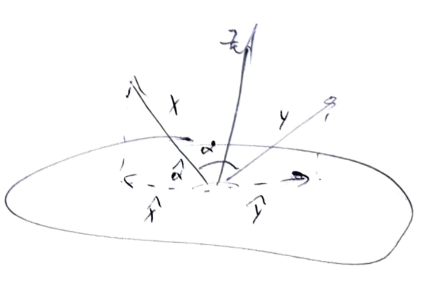
$\rho_3 = 0 + 1{,}3\rho_2 - 0{,}4\rho_1 + 0 \Rightarrow \rho_3 = \ldots$

$\gamma_0 = 1{,}3\rho_1\gamma_0 - 0{,}4\rho_2\gamma_0 + \sigma_\varepsilon^2 \Rightarrow \gamma_0 = \frac{\sigma_\varepsilon^2}{1 - 1{,}3\rho_1 + 0{,}4\rho_2} = \ldots \Rightarrow \ldots$

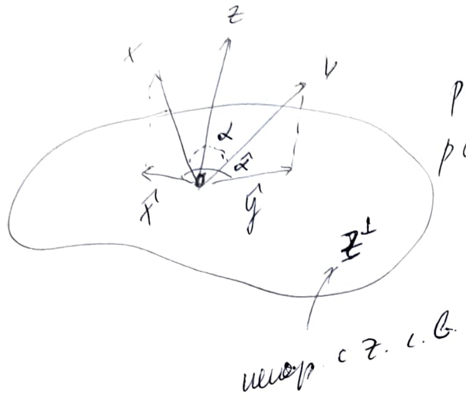
def Частные корреляции

$Corr(X, Y) = \cos(\alpha)$

$pCorr(X, Y, Z) = \cos(\hat{\alpha}) = corr(\hat{X}, \hat{Y})$



Частная корреляция



$Corr(X,Y)=\cos(\alpha)$

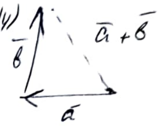
$pCorr(X,Y;Z)=\cos(\hat{\alpha})$

$pCorr(X,Y;Z)=corr(\hat{X},\hat{Y})$

X	$\vec{a}$
$Var(X)$	$\lVert\bar{a}\rVert^2$
$Cov(X,Y)$	$\bar{a}\cdot\bar{b}=\langle\bar{a},\bar{b}\rangle$
$Corr(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$	$\cos(\bar{a},\bar{b})=\frac{\langle\bar{a},\bar{b}\rangle}{\lVert\bar{a}\rVert\cdot\lVert\bar{b}\rVert}$

$Var(X+Y)=Var(X)+Var(Y)$

если $Cov(X,Y)=0$



$\lVert\bar{a}+\bar{b}\rVert^2=\lVert a\rVert^2+\lVert b\rVert^2$,

если $\bar{a}\perp\bar{b}$

$\langle\bar{a},\bar{b}\rangle=0$

ex Монетка подбр. 101 раз

	Р	О
X_i	0	1
p	1/2	1/2

$S=X_1+\ldots+X_{101}$ (← число Орлов)

$L=X_1+\ldots+X_{51}\sim Bin(n=51,p=1/2)$

$R=X_{51}+\ldots+X_{101}$

a) $Var(L)=51\cdot 1/2\cdot 1/2$

$Var(R)=51\cdot 1/2\cdot 1/2$

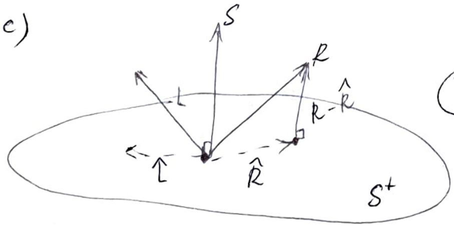
$Var(S)=101\cdot 1/2\cdot 1/2$

$Cov(R,S)=Var(X_i)\cdot 51=Cov(R,R)=\frac{51}{4}$

$Cov(R,L)=Var(X_i)=1/4$

b) $corr(R,L)=\frac{Cov(R,L)}{\sqrt{Var(R)Var(L)}}=\frac{1/4}{\sqrt{51/4\cdot 51/4}}=\frac{1}{51}$ (>0)

c)



$\Rightarrow\ \hat{R}=R-\lambda\cdot S=R-\frac{51}{101}\cdot S$

$Cov(\hat{R},S)=0$

$Cov(R-\lambda S,S)=0$

$Cov(R,S)-\lambda Cov(S,S)=0$

$51/4-\lambda\cdot 101/4=0\Rightarrow\lambda=51/101$

$\hat{L}=L-\gamma\cdot S=L-51/101\cdot S$

$Cov(\hat{L},S)=Cov(L-\gamma S,S)=$

$=Cov(L,S)-\gamma Cov(S,S)=$

$=51/4-\gamma\cdot 101/4=0\Rightarrow\gamma=51/101$

$\alpha\lesssim 90$

$\hat{\alpha}\approx 180$

$pCorr(R,L;S)=Corr(\hat{R},\hat{L})=\frac{Cov(\hat{R},\hat{L})}{\sqrt{Var(\hat{R})\cdot Var(\hat{L})}}=\ldots=\frac{\frac{1}{4}-\frac{51^2}{101\cdot 4}}{\frac{51}{4}\cdot\frac{50}{101}}=-\frac{50}{51}$

(ex) Стац. ли процесс ($cov(y_t, \varepsilon_{t+k}) = 0 \quad \forall k > 0$) ?

a) $y_t = 0{,}6\, y_{t-1} - 1{,}09\, y_{t-2} + 3 + \varepsilon_t - \varepsilon_{t-1}$ → MA: $\mu - 1 = 0$, $\mu = 1$

$\lambda^2 - 0{,}6\lambda + 1{,}09 = 0$

$\lambda_1 = 0{,}3 - i \qquad |\lambda_1| = \sqrt{0{,}09 + 1} = \sqrt{1{,}09} > 1$

$\lambda_2 = 0{,}3 + i \qquad |\lambda_2| = \sqrt{1{,}09}$

Если после удаления общих корней у AR и MA части, все AR-корни $|\lambda| < 1$, то процесс стац. и не завис. от будущ.

b) $y_t = 2{,}5\, y_{t-1} - y_{t-2} + 3 + \varepsilon_t - 2\varepsilon_{t-1}$

$\lambda^2 - 2{,}5\lambda + 1 = 0 \qquad \mu - 2 = 0$

$\lambda_1 = 0{,}5, \quad |\lambda_1| < 1 \qquad \mu = 2$

$\lambda_2 = 2$

=> есть стац. реш.

$y_t = 2{,}5 L y_t - L^2 y_t + 3 + \varepsilon_t - 2L\varepsilon_t$

$(1 - 2{,}5L + L^2)\, y_t = 3 + (1 - 2L)\,\varepsilon_t$

$(1 - 2L)(1 - \frac{1}{2}L)\, y_t = (1 - 2L)\,\varepsilon_t$ → $3 = (1 - 2L)\cdot c$, $c = -3$

$(1 - \frac{1}{2}L)\, y_t = \varepsilon_t$

$$y_t = \frac{\varepsilon_t}{1 - \frac{1}{2}L} = \left(1 + \left(\tfrac{1}{2}L\right) + \left(\tfrac{1}{2}L\right)^2 + \ldots + \ldots\right)\varepsilon_t$$

$$= \varepsilon_t + \tfrac{1}{2}\varepsilon_{t-1} + \tfrac{1}{4}\varepsilon_{t-2} \ldots$$

$(1 - \frac{1}{2}L)\, y_t = -3 + \varepsilon_t$

$$y_t = \frac{1}{1 - \frac{1}{2}\cancel{L}}(-3) + \varepsilon_t + \tfrac{1}{2}\varepsilon_{t-1} + \tfrac{1}{4}\varepsilon_{t-2} + \ldots$$

$$y_t = -6 + \varepsilon_t + \tfrac{1}{2}\varepsilon_{t-1} + \ldots$$

(ex) $pcorr(R, S; Z_1, Z_2) = Corr(\hat{R}, \hat{S})$

$\hat{R} = R - ?Z_1 - ?Z_2$ не кор. с Z_1, Z_2

$\hat{S} = S - ?Z_1 - ?Z_2$ не кор. с Z_1, Z_2

$\varphi_{kk} = pCorr(y_t, y_{t-k}; y_{t-1}, y_{t-2}, \ldots, y_{t-k+1})$

$\varphi_{11} = pCorr(y_1, y_2; \varnothing) = corr(y_1, y_2)$

Thm. (Юл - Волкер)

Если процесс y_t стац. и y_t пр. в виде

$y_t = \varphi_{31} y_{t-1} + \varphi_{32} y_{t-2} + \varphi_{33} y_{t-3} + u$,

где u не кор. с $y_{t-1}, y_{t-2}, y_{t-3}$

то $\varphi_{33} = pCorr(y_t, y_{t-3}; y_{t-1}, y_{t-2})$

(ex) $y_t = 0{,}7\, y_{t-1} + \varepsilon_t + 6$

$\gamma_0 = Cov(0{,}7 y_{t-1} + \varepsilon_t + 6, 0{,}7 y_{t-1} + \varepsilon_t + 6) = 0{,}7^2 \gamma_0 + \sigma_\varepsilon^2 \Rightarrow \gamma_0 = \dfrac{\sigma_\varepsilon^2}{0{,}51}$

$\gamma_1 = Cov(0{,}7 y_{t-1} + \varepsilon_t + 6, y_{t-1}) = 0{,}7\gamma_0 = \dfrac{0{,}7 \cdot \sigma_\varepsilon^2}{0{,}51}$

$\gamma_2 = Cov(0{,}7 y_{t-1} + \varepsilon_t + 6, y_{t-2}) = 0{,}7\gamma_1 = \dfrac{0{,}7^2 \cdot \sigma_\varepsilon^2}{0{,}51}$

...

$\gamma_k = Cov(0{,}7 y_{t-1} + \varepsilon_t + 6; y_{t-k}) = \dfrac{0{,}7^k \sigma_\varepsilon^2}{0{,}51}$

$\rho_k = 0{,}7^k$

$\varphi_{11} = \rho_1 = 0{,}7$

φ_{22}: $y_t = \varphi_{21} \cdot y_{t-1} + \varphi_{22} \cdot y_{t-2} + u$

$y_t = 0{,}7\, y_{t-1} + 0 \cdot y_{t-2} + \varepsilon_t + 6$ ← стац.

все усл. реор. вып. $\Rightarrow \varphi_{22} = 0$

$\varphi_{33} = 0$

$\varphi_{kk} = 0, \; k > 1$

(ex) $y_t = \varepsilon_t + 0{,}7\varepsilon_{t-1} + 9$

$\gamma_0 = Cov(\varepsilon_t + 0{,}7\varepsilon_{t-1} + 9, \varepsilon_t + 0{,}7\varepsilon_{t-1} + 9) = \sigma^2 + 0{,}49\sigma^2 = 1{,}49\sigma^2$

$\gamma_1 = Cov(\varepsilon_t + 0{,}7\varepsilon_{t-1} + 9, \varepsilon_{t-1} + 0{,}7\varepsilon_{t-2} + 9) = 0{,}7\sigma^2$

$\gamma_k = 0, \; k > 1$

$\varphi_{11} = \rho_1 = 0{,}7/1{,}49$

φ_{22}: ...

φ_{22}: $y_t = \varphi_{21} y_{t-1} + \varphi_{22} y_{t-2} + u$

$$\begin{cases} Cov(u, y_{t-1}) = 0 \\ Cov(u, y_{t-2}) = 0 \end{cases} \qquad \begin{cases} Cov(y_t - \varphi_{21} y_{t-1} - \varphi_{22} y_{t-2}, y_{t-1}) = 0 \\ Cov(y_t - \varphi_{21} y_{t-1} - \varphi_{22} y_{t-2}, y_{t-2}) = 0 \end{cases} \qquad \begin{cases} \gamma_1 - \varphi_{21}\gamma_0 - \varphi_{22}\gamma_1 = 0 \\ \gamma_2 - \varphi_{21}\gamma_1 - \varphi_{22}\gamma_0 = 0 \end{cases}$$

$$\begin{cases} 0{,}7\sigma_\varepsilon^2 - \varphi_{21} \cdot 1{,}49\sigma_\varepsilon^2 - \varphi_{22} \cdot 0{,}7\sigma_\varepsilon^2 = 0 \\ 0 - \varphi_{21} \cdot 0{,}7\sigma_\varepsilon^2 - \varphi_{22} \cdot 1{,}49\sigma_\varepsilon^2 = 0 \end{cases}$$

$$0{,}7^2 - \varphi_{22} \cdot 0{,}7^2 + \varphi_{22} \cdot 1{,}49^2 = 0$$

$$\varphi_{22} = \frac{0{,}7^2}{0{,}7^2 - 1{,}49^2}$$

HW $y_t = 0{,}5\, y_{t-1} + \varepsilon_t - \varepsilon_{t-1} + 7$, y_t - стац., $\varepsilon_t \sim WN$

Эконометрика (6)

ex) $y_t = 0{,}5 y_{t-1} + \varepsilon_t - \varepsilon_{t-1} + 7$

$\varphi_{11} = p\,corr(y_t, y_{t-1};\ ) = corr(y_t, y_{t-1}) = \rho_1$

★ $\left(\varphi_{42} = p\,corr(y_t, y_{t-2}; y_{t-1}, y_{t-3}, y_{t-4})\right)$

$$\begin{cases} \gamma_0 = cov(0{,}5 y_{t-1} + \varepsilon_t - \varepsilon_{t-1} + 7;\ 0{,}5 y_{t-1} + \varepsilon_t - \varepsilon_{t-1} + 7) = 0{,}5^2 \gamma_0 - \cancel{0{,}5\sigma_\varepsilon^2} + \cancel{\sigma_\varepsilon^2} + \sigma_\varepsilon^2 - 0{,}5\sigma_\varepsilon^2 \\ \gamma_1 = cov(0{,}5 y_{t-1} + \varepsilon_t - \varepsilon_{t-1} + 7;\ y_{t-1}) = 0{,}5\gamma_0 - Cov(\varepsilon_{t-1}; y_{t-1}) = 0{,}5\gamma_0 - \sigma_\varepsilon^2 \end{cases}$$

$y_{t-1} = \varepsilon_{t-1} + ?\,\varepsilon_{t-2} + \ldots$

$$\Rightarrow \begin{cases} \gamma_0 = 4/3\,\sigma_\varepsilon^2 \\ \gamma_1 = -1/3\,\sigma_\varepsilon^2 \end{cases} \Rightarrow \rho_1 = \frac{\gamma_1}{\gamma_0} = -\frac{1}{4}$$

$\gamma_2 = cov(0{,}5 y_{t-1} + \varepsilon_t - \varepsilon_{t-1} + 7;\ y_{t-2}) =$

$= 0{,}5\,\gamma_1 = -\frac{1}{6}\gamma_0$

$\rho_2 = -\frac{1}{6} : \frac{4}{3} = -\frac{1}{8}$

φ_{22}: $y_t = \varphi_{21} \cdot y_{t-1} + \varphi_{22} \cdot y_{t-2} + u_t$

$$\begin{cases} Cov(u_t; y_{t-1}) = 0 \\ Cov(u_t; y_{t-2}) = 0 \end{cases} \quad \begin{cases} Cov(y_t - \varphi_{21} \cdot y_{t-1} - \varphi_{22} \cdot y_{t-2};\ y_{t-1}) = 0 \\ Cov(y_t - \varphi_{21} \cdot y_{t-1} - \varphi_{22} \cdot y_{t-2};\ y_{t-2}) = 0 \end{cases}$$

$$\begin{cases} \gamma_1 - \varphi_{21} \cdot \gamma_0 - \varphi_{22} \cdot \gamma_1 = 0 \\ \gamma_2 - \varphi_{21} \cdot \gamma_1 - \varphi_{22} \cdot \gamma_0 = 0 \end{cases}$$

$$\begin{cases} \rho_1 - \varphi_{21} - \varphi_{22} \cdot \rho_1 = 0 \\ \rho_2 - \varphi_{21}\rho_1 - \varphi_{22} = 0 \end{cases} \quad \varphi_{21} = \rho_1(1 - \varphi_{22})$$

$\rho_2 + (\varphi_{22} - 1)\rho_1^2 - \varphi_{22} = 0$

$-\frac{1}{8} + (\varphi_{22} - 1) \cdot \frac{1}{16} - \varphi_{22} = 0 \quad \Rightarrow \quad \varphi_{22} = -0{,}2$

HW $y_t = \varepsilon_t + \alpha\,\varepsilon_{t-1}$

$AR(p) \nearrow$ стац процесс, не смотр. в буд.

[ARMA(p, 0)] $\searrow$ $y_t = ?\,y_{t-1} + \dots + ?\,y_{t-p} + c + \varepsilon_t$

$MA(q) \to y_t = \varepsilon_t + ?\,\varepsilon_{t-1} + \dots + ?\,\varepsilon_{t-p}$

[ARMA(0, q)]

$\nearrow$ стац, и см. в буд.

$ARMA(p, q) \to$ упрощ запись

$\searrow$ $y_t = \underbrace{?\,y_{t-1} + \dots + ?\,y_{t-p}}_{p \text{ прошл. } y} + \varepsilon_t + \underbrace{?\,\varepsilon_{t-1} + \dots + ?\,\varepsilon_{t-q}}_{q \text{ прошл. } \varepsilon} + const$

$\nearrow$ упрощ запись

$ARIMA(p, d, q)$ $\to$ q лаг по ε_t; p число $|\lambda_i| < 1$; число $\lambda_i = 1$

a) $y_t = 0{,}7\,y_{t-1} - 0{,}12\,y_{t-2} + \varepsilon_t - 0{,}4\,\varepsilon_{t-1}$ => ARMA(1,0); AR(1)

$(1 - 0{,}7L + 0{,}12L^2)\,y_t = (1 - 0{,}4L)\,\varepsilon_t$

$0{,}12\,(L - \frac{10}{3})\cancel{(L - \frac{10}{4})} = (1 - \cancel{\frac{4}{10}L})\,\varepsilon_t$

$-\frac{4}{10}$

$-\frac{3}{10}(L - \frac{10}{3}) = \varepsilon_t$

b) $y_t = 0{,}7\,t_{-1} + 0{,}3\,y_{t-2} + \varepsilon_t - 0{,}5\,\varepsilon_{t-1}$ => не ARMA, т.к. не стац.

$L_1 = 1 \quad L_2 = -\frac{10}{3}$ => ARIMA(1, 1, 1)

$\lambda_1 = 1 \quad \lambda_2 = -0{,}3$

лаг L : $L \cdot y_t = y_{t-1}$ $\qquad \Delta y_t = y_t - y_{t-1}$, $\Delta = 1 - L$

сез. лаг L^{12} : $L^{12} \cdot y_t = y_{t-12}$ $\qquad \Delta_{12} y_t = y_t - y_{t-12}$

ARMA(1;2) : $(1 - \gamma L) \cdot y_t = (1 + \alpha_1 \cdot L + \alpha_2 L^2) \cdot \varepsilon_t + const$

ARMA(1;0) – SARMA(2;1) : $(1 - \gamma L)(1 - \delta_1 L^{12} - \delta_2 L^{24}) \cdot y_t = (1 + \alpha_1 \cdot L^{12}) \cdot \varepsilon_t + const$

ex $(1 - \gamma_1 L - \gamma_2 L^2) \cdot (1 - L) \cdot y_t = (1 + \delta_1 L^{12}) \cdot (1 + \alpha_1 \cdot L + \alpha_2 L^2) \cdot \varepsilon_t$

ARIMA(2;1;2) – SMA(1)

ex $(1 - 0{,}7 L^{12}) y_t = (1 + 0{,}5 L) \cdot \varepsilon_t + 4$ $\qquad y_t - 0{,}7 y_{t-12} = \varepsilon_t + 0{,}5 \varepsilon_{t-1} + 4$

forecast
sophisthse
ggplot2

$T = 100$ $\qquad y_{100} = 8$ $\qquad \varepsilon_{100} = -1$

$\varepsilon_t \sim N(0; 9)$ $\qquad y_{89} = 6$

$y_{90} = 7$

a) $E(y_{101} | F_{100}) = E(0{,}7 y_{89} + \varepsilon_{101} + 0{,}5 \varepsilon_{100} + 4 \,|\, F_{100}) = 0{,}7 \cdot \underbrace{y_{89}}_{6} + \underbrace{E(\varepsilon_{101} | F_{100})}_{0} + 0{,}5 \cdot \underbrace{E(\varepsilon_{100} | F_{100})}_{-1} + 4 = 7{,}7$

$E(y_{102} | F_{100}) = E(0{,}7 y_{90} + \varepsilon_{102} + 0{,}5 \varepsilon_{101} + 4 \,|\, F_{100}) =$

$= 0{,}7 \cdot \underbrace{y_{90}}_{7} + \underbrace{E(\varepsilon_{102} | F_{100})}_{0} + 0{,}5\, \underbrace{E(\varepsilon_{101} | F_{100})}_{0} + 4 = 8{,}9$

b) 95% пред. инт. для y_{101} и y_{102}

y_{101} : $\{E(y_{101}) \pm 1{,}96\, \underbrace{se}_{3}\}$ $\qquad y_{102}$: $\{E(y_{102}) \pm 1{,}96 \cdot \underbrace{se}_{\sqrt{11{,}25}}\}$

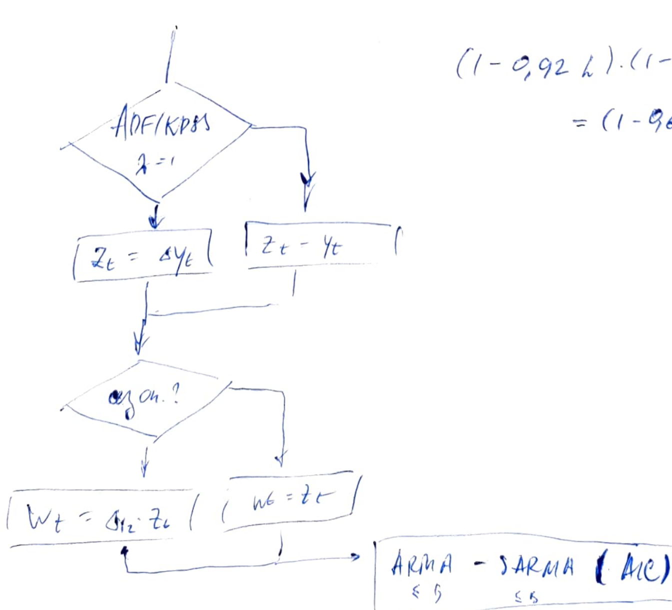
$Var(\underbrace{y_{101} - E(y_{101} | F_{100})}_{\text{ошибка прогноза}}) = Var(\cancel{0{,}7 y_{89}} + \varepsilon_{101} + \cancel{0{,}5 \varepsilon_{100}} + \cancel{4} - \cancel{0{,}7 y_{89}} - \cancel{0{,}5 \varepsilon_{100}} - \cancel{4}) = 9$

$Var(y_{102} - E(y_{102} | F_{100})) = Var(\cancel{0{,}7 y_{90}} + \varepsilon_{102} + 0{,}5 \varepsilon_{101} + \cancel{4} - \cancel{0{,}7 y_{90}} - \cancel{4})$

$= Var(\varepsilon_{102} + 0{,}5 \varepsilon_{101}) = 9 + 0{,}5^2 \cdot 9 = \frac{9 \cdot 5}{4} = 11{,}25$

ex

$(1 - 0{,}92 L) \cdot (1 - 0{,}85 L^{12}) \cdot (1 - L^{12}) \cdot y_t$
$= (1 - 0{,}68 L) \cdot (1 - 0{,}94 L^{12}) \cdot \varepsilon_t$



ex $y_t = \beta \cdot y_{t-1} + u_t \qquad u_t \sim N(0; \sigma_u^2)$

y_t – стац (не зав. от $u_{t+1}, u_{t+2} \dots$)

$y_1 = 0{,}1$
$y_2 = -0{,}2$
$y_3 = 0{,}2$

1) $E(y_t) = \beta E(y_t) + 0$

$E(y_t) = 0$

$\beta = 1$ – не подх., т.к. y_t – стац.

$Var(y_t) = \beta^2 \cdot Var(y_t) + \sigma_u^2$

$Var(y_t) = \dfrac{\sigma_u^2}{1-\beta}$

$y_t \sim N\left(0; \dfrac{\sigma_u^2}{1-\beta^2}\right)$

$E(y_2 | y_1) = E(\beta y_1 + u_2 | y_1) = \underset{const}{\beta y_1} + \underset{E(u_1)=0}{E(u_2|y_1)} = \beta y_1$

$Var(y_2|y_1) = Var(\beta y_1 + u_2 | y_1) = Var(u_2|y_1) = \sigma_u^2$

$E(y_3|y_2) = \beta y_2$

$Var(y_3|y_2) = \sigma_u^2$

2) $R \sim N(\mu, \sigma_R^2) \qquad f(z) = (2\pi\sigma_R^2)^{-1/2} \cdot e^{-\frac{1}{2}\frac{(z-\mu)^2}{\sigma_R^2}}$

$$f(y_1) = \frac{1}{\sqrt{2\pi\frac{\sigma_u^2}{1-\beta^2}}} \cdot e^{-\frac{1}{2}\frac{y_1^2}{\frac{\sigma_u^2}{1-\beta^2}}}$$

$$f(y_2|y_1) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \cdot e^{-\frac{1}{2}\frac{(y_2-\beta y_1)^2}{\sigma_u^2}}$$

$$\ell = \ln f(y_1, y_2, y_3 | \beta, \sigma^2) = \ln(f(y_3|y_1,y_2) \cdot f(y_1, y_2))$$
$$= \ln(f(y_3|y_1,y_2) \cdot f(y_2|y_1) \cdot f(y_1))$$
$$= \ln(f(y_3|y_1,y_2)) + \ln(f(y_2|y_1)) + \ln(f(y_1))$$
$$= \ln f(y_1) + \ln f(y_2|y_1) + \ln f(y_3|y_2) \to \max_{\beta, \sigma^2}$$

$$\ell = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\frac{\sigma_u^2}{1-\beta^2}\right) - \frac{1}{2}\cdot\frac{y_1^2}{\frac{\sigma_u^2}{1-\beta^2}}$$
$$-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma_u^2 - \frac{1}{2}\cdot\frac{y_2 - \beta y_1^2}{\sigma_u^2}$$
$$-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma_u^2 - \frac{1}{2}\frac{y_3 - \beta y_2^2}{\sigma_u^2} \to \max_{\beta, \sigma^2}$$

$$\tilde{\ell} = \ell(y_2, y_3 \mid y_1, \beta, \sigma^2) = \ln f(y_2 \mid y_1) + \ln f(y_3 \mid y_2) \to \max_{\beta, \sigma^2}$$

$$\frac{\partial \tilde{\ell}}{\partial \beta} = \frac{y_1 (y_2 - \hat{\beta} y_1)}{\hat{\sigma}_u^2} + y_2 \frac{(y_3 - \hat{\beta} y_2)}{\hat{\sigma}_u} = 0 \Rightarrow \hat{\beta} = \frac{y_1 y_2 + y_2 y_3}{y_1^2 + y_2^2}$$

$$\frac{\partial \tilde{\ell}}{\partial \sigma^2} = -\frac{1}{2} \frac{1}{\hat{\sigma}^2} \cdot 2 - \frac{1}{2} \cdot -1 \frac{(y_2 - \beta y_1)^2}{\hat{\sigma}_u^4} - \frac{1}{2}$$

ex

~~ytllll~~

$$(1-L)(1 - 0{,}23L - 0{,}2L^2 + 0{,}093L^3)(1 - 0{,}33L^{12} - 0{,}19L^{24}) \cdot y_t = u_t$$

~~= ut~~

ex $y_t \sim AR(1)$ стац.

$x_t \sim AR(1)$ стац.

$S_t = y_t + x_t$ – ?

$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t$ $(1-\alpha_1 L)\cdot y_t = \alpha_0 + \varepsilon_t$

$x_t = \beta_0 + \beta_1 x_{t-1} + u_t$ $(1-\beta_1 L)\cdot x_t = \beta_0 + \varepsilon_t$

$$x_t + y_t = \frac{\alpha_0}{1-\alpha_1 L} + \frac{\varepsilon_t}{1-\alpha_1 L} + \frac{\beta_0}{1-\beta_1 L} + \frac{u_t}{1-\beta_1 L}$$

$$= C + \frac{u_t}{1-\beta_1 L} + \frac{\varepsilon_t}{1-\alpha_1 L}$$

1 $\frac{7}{1-0{,}8L} = (1 + 0{,}8L + (0{,}8L)^2 + \ldots)\cdot 7 = \frac{7}{1-0{,}8} = 35$

$$\underbrace{(1-\alpha_1 L)(1-\beta_1 L)}_{AR(2)}\cdot S_t = \tilde{C} + \underbrace{u_t(1-\alpha_1 L) + \varepsilon_t(1-\beta_1 L)}$$

$z_t = \varepsilon_t - \beta_1\varepsilon_{t-1} + u_t - \alpha_1 u_{t-1}$ 2

$E(z_t) = 0$

$Var(z_t) = \sigma_\varepsilon^2 + \beta_1^2\cdot\sigma_\varepsilon^2 + \sigma_u^2 + \alpha_1^2\cdot\sigma_u^2$

$Cov(z_t, z_{t-1}) = -\beta_1\sigma_\varepsilon^2 - \alpha_1\cdot\sigma_u^2$

$Cov(z_t, z_{t-k}) = 0,\ k>1$

$z_t \sim MA(1)$

2 z_t – стац.

$Cov(z_t, z_{t-4}) \neq 0$

$Cov(z_t, z_{t-k}) = 0,\ k>4$

$\Rightarrow z_t \sim MA(4)$

$S_t \sim ARMA(2;1)$ для β_1^*, α_1^* м. б. $AR(2)$

$\beta_1 = \alpha_1$ $AR(1)$

ex $y_t \sim AR(1)$

$x_t \sim MA(1)$

$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t$ $(1-\alpha_1 L)y_t = \alpha_0 + \varepsilon_t$

$x_t = \beta_0 + u_t + \beta_1 u_{t-1}$ $x_t = \beta_0 + (1+\beta_1 L)u_t$

$$S_t = y_t + x_t = \frac{\alpha_0}{1-\alpha_1 L} + \frac{\varepsilon_t}{1-\alpha_1 L} + \beta_0 + (1+\beta_1 L)u_t$$

$(1-\alpha_1 L)S_t = C_0 + \varepsilon_t + (1+\beta_1 L)(1-\alpha_1 L)u_t \Rightarrow S_t \sim ARMA(1,2)$

ex $y_t \sim MA(2)$ $y_t = \alpha_0 + (1+\alpha_1 L + \alpha_2 L^2)\varepsilon_t$

$x_t \sim MA(2)$ $x_t = \beta_0 + (1+\beta_1 L + \beta_2 L^2)u_t$

$S_t = \alpha_0 + \beta_0 + (1+\alpha_1 L + \alpha_2 L^2)\varepsilon_t + (1+\beta_1 L + \beta_2 L^2)\cdot u_t \Rightarrow S_t \sim MA(2)$

$\alpha_2^*\ ?\ \beta_2^* \Rightarrow S_t \sim MA(1)$:

$\hat{\alpha}_1; \hat{\alpha}_2; \hat{\beta}_1, \hat{\beta}_2 \Rightarrow S_t \sim WN$

(ex) $y_t = 0{,}7\, y_{t-1} + 2 + \varepsilon_t + 0{,}5\, \varepsilon_{t-1}$; cm ay. , $Var(\varepsilon_t) = 9$

$y_{100} = 4$

$\varepsilon_{100} = -1$

a) $E(y_{101} \mid F_{100}) = E(0{,}7\, y_{100} + 2 + \varepsilon_{101} + 0{,}5\, \varepsilon_{100} \mid F_{100}) = 0{,}7 \cdot 4 + 2 + 0 + 0{,}5 \cdot (-1) = 4{,}3$

$E(y_{102} \mid F_{100}) = E(0{,}7 \cdot y_{101} + 2 + \varepsilon_{102} + 0{,}5\, \varepsilon_{101} \mid F_{100}) = 0{,}7 \cdot \underbrace{E(y_{101} \mid F_{100})}_{= 4{,}3} + 2 = 5{,}01$

$\lim_{h\to\infty} \underbrace{E(y_{100+h} \mid F_{100})}_{\theta} = \nearrow E(y_{100+h}) = E(y_t)$

$\searrow 0{,}7 \lim_{h\to\infty} \underbrace{E(y_{100+h-1} \mid F_{100})}_{\theta} + 2 + E(\varepsilon_{100+h} \mid F_{100}) + 0{,}5\, E(\varepsilon_{100+h-1} \mid F_{100})$

$\Rightarrow \theta = 0{,}7\theta + 2$

$\theta = \frac{20}{3}$

b) $Var(y_{101} \mid F_{100}) = Var(y_{100} \cdot 0{,}7 + 2 + \varepsilon_{101} + 0{,}5\, \varepsilon_{100} \mid F_{100}) = \sigma_\varepsilon^2 = 9$

$Var(y_{102} \mid F_{100}) = Var(y_{101} \cdot 0{,}7 + 2 + \varepsilon_{102} + 0{,}5\, \varepsilon_{101} \mid F_{100}) = Var(y_{101} \mid F_{100}) \cdot 0{,}49 + \sigma_\varepsilon^2 + 0{,}25 \cdot \sigma_\varepsilon^2 = 15{,}66$

$\lim_{h\to\infty} \underbrace{Var(y_{100+h} \mid F_{100})}_{\theta} = \nearrow Var(y_{100+h}) = Var(y_t)$

$\searrow 0{,}49 \lim_{h\to\infty} \underbrace{Var(y_{100+h-1} \mid F_{100})}_{\theta} + 0 + \lim_{h\to\infty} Var(\varepsilon_{100+h} \mid F_{100}) + 0{,}25 \cdot \lim_{h\to\infty} Var(\varepsilon_{100+h-1} \mid F_{100})$

$\Rightarrow \theta = 0{,}49\theta + 9 + 0{,}25 \cdot 9 \Rightarrow \theta = 22{,}06$

ETS & TBATS

ETS: "forecast. Principles & practice" (R!) App 2 ...

Hyndman R.

↓

error + trend + seasonality → A – Additive, M – Multiplicative, N – не входит

ETS(ANN)

$$\begin{cases} y_t = l_t + \varepsilon_t, & \varepsilon_t \sim N(0; \sigma_\varepsilon^2) \\ l_t = l_{t-1} + \alpha \varepsilon_t \end{cases} \Rightarrow \hat\alpha, \hat\sigma^2, \hat l_0$$

↑ level

ETS(A,N,N)

ex $y_1 = 5;\ y_2 = 6;\ y_3 = 7;\ \hat\alpha = 0{,}5 \Rightarrow \hat l_0 - ?$

RSS → min

$$l_t = l_{t-1} + \alpha (y_t - l_t)$$

$$l_t = \frac{l_{t-1} + \alpha y_t}{1+\alpha} = \frac{1}{1+\alpha} l_{t-1} + \frac{\alpha}{1+\alpha} \cdot y_t$$

$$\hat y_{t+1|t} = l_t \qquad \hat y_{1+1|t} = l_0$$

$$\hat y_{2+1|t} = l_1 = \frac{2}{3} l_0 + \frac{1}{3} \underset{5}{\underset{\parallel}{y_1}}$$

$$\hat y_{3+1|t} = l_2 = \frac{2}{3} \cdot l_1 + \frac{1}{3} \cdot \underset{6}{\underset{\parallel}{y_2}}$$

$$\min_{l_0} RSS = (\underset{5}{\underset{\parallel}{y_1}} - l_0)^2 + \left(\underset{6}{\underset{\parallel}{y_2}} - \frac{2}{3} l_0 - \frac{5}{3}\right)^2 + \left(7 - \frac{2}{3} \underset{\frac{2}{3} l_0 + \frac{1}{3}\cdot 5}{\underset{\parallel}{l_1}} - \frac{6}{3}\right)^2$$

~~ИСКЛ~~

$$\frac{\partial RSS}{\partial l_0} = 45 - 9\hat l_0 + 36 - 4\hat l_0 - 10 + 28 - \frac{16}{3}\hat l_0 - \frac{40}{3} - 8 =$$

$$\Rightarrow \hat l_0 =$$

ex AAN, $\varepsilon_t \sim N(0; \sigma_\varepsilon^2)$

$$\begin{cases} y_t = l_{t-1} + b_{t-1} + \varepsilon_t \\ l_t = l_{t-1} + b_{t-1} + \alpha \varepsilon_t & \text{— } \hat{?} \text{ ср. знач. } y_t \\ b_t = b_{t-1} + \beta \varepsilon_t & \text{— } \hat{?} \text{ скорость роста} \end{cases}$$

$$\hat{y}_{t+h|t} = l_t + b_t \cdot h$$

y_i		
6	$l_0 + b_0$	$\hat{\alpha} = \frac{1}{2}$
7	$l_1 + b_1$	$\beta = \frac{1}{2}$
8	$l_2 + b_2$	

ex $\varepsilon_t = y_t - l_{t-1} - b_{t-1}$

$$\begin{cases} l_t = l_{t-1} + b_{t-1} + \alpha(y_t - l_{t-1} - b_{t-1}) \\ b_t = b_{t-1} + \beta(y_t - l_{t-1} - b_{t-1}) \end{cases}$$

$$\begin{cases} l_t = \ldots \\ b_t = \ldots \end{cases}$$

$$\min_{l_0} (6 - (l_0 + b_0))^2 + \ldots \Rightarrow \hat{l}_0$$

Эконометрика (10)

$ETS(A, A, A)$ $m=2$, полугод. данные

y_t – набл. показатель

l_t – сглаж. y_t, очищ. от сезонности

b_t – темп роста

s_t – сезонный эффект

ε_t – случ. составляющая

★ $ETS(AAA)$ $m=12$ $\Rightarrow$ 17 пар.

$ETS(AAN)$ $\Rightarrow$ 5 пар.

ex ETS: $y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$, $\varepsilon_t \sim N(0; \sigma^2)$

$$l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t$$
$$b_t = b_{t-1} + \beta\varepsilon_t$$
$$s_t = s_{t-m} + \gamma\cdot\varepsilon_t$$

$\boxed{s_0 + s_{-1} = 0}$ усл. нормировки

1) $\theta = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \sigma^2 \end{pmatrix} \begin{matrix} b_0 \\ l_0 \\ s_0 \\ (s_{-1}) \end{matrix}$ – вектор параметров модели $\Rightarrow \theta = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \sigma^2 \\ b_0 \\ l_0 \\ s_0 \end{pmatrix}$

2) Василий: θ, y_1 $\Rightarrow$? $\varepsilon_1, l_1, b_1, s_1$

3) ? $l_t = \alpha\cdot(\ldots) + (1-\alpha)(\ldots)$ $\quad l_t = l_{t-1} + b_{t-1} + \alpha(y_t - l_{t-1} - b_{t-1} + s_{t-m})$

4) ? $b_t = \ldots \longrightarrow$

? $s_t = \ldots \longrightarrow$

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1})$$
$$b_t = \beta(y_t - s_{t-m} - l_{t-1}) + (1-\beta)\cdot b_{t-1}$$
$$s_t = \gamma(y_t - b_{t-1} - l_{t-1}) + (1-\gamma)\cdot s_{t-m}$$

5) $\ln f(y_1|\theta) = \ln\left[\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\exp\left(-\frac{1}{2\sigma^2}(y_1 - \underbrace{\mu}_{l_0+b_0+s_{-1}})^2\right)\right]$

$\ln f(y_2|y_1, \theta) = \ln\left[\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\exp\left(-\frac{1}{2\sigma^2}(y_2 - \underbrace{\mu}_{l_1+b_1+s_0})^2\right)\right]$

$y_1|\theta \sim N(l_0 + b_0 + s_{-1}; \sigma^2)$

6) $\ln l(\theta) = \ln f(y_1|\theta) + \ln f(y_2|y_1, \theta) + \ln f(y_3|y_2, y_1, \theta) \ldots + \ln f(y_n|y_{n-1}, \ldots, y_1, \theta) \to \max\limits_{\theta}$

ex $ETS(AAA)$, $m=2$

$\hat\alpha = 0.1$, $\hat\beta = 0.2$, $\hat\gamma = 0.3$, $\hat l_0 = 5$, $\hat b_0 = 1$, $\hat s_0 = -2$, $\hat s_{-1} = 2$

$y_1 = 6$; $y_2 = 8$ $\quad \hat\varepsilon_1 = 6 - 5 - 1 - 2 = -2 \longrightarrow \hat\varepsilon_1 = -2$

a) $\hat l_1 = 5 + 1 + 0.1\cdot\underbrace{\hat\varepsilon_1}_{=-2} = 5.8$ $\quad \hat b_1 = 0.6$ $\quad \hat s_1 = 1.4$

$\hat l_2 = \ldots$ $\quad \hat b_2 = \ldots$ $\quad \hat s_2 = \ldots$ $\quad \hat\varepsilon_2 = \ldots$

д) $y_1, \dots, y_{100}$

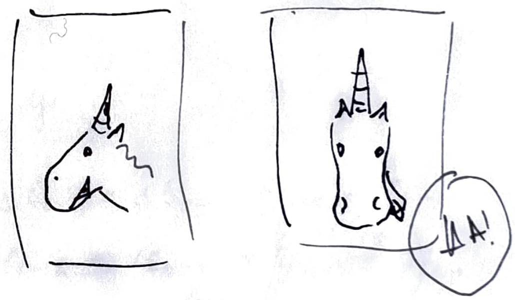
$y_{100} = 70, \quad b_{100} = 2, \quad \hat{s}_{95} = 3, \quad \hat{l}_{100} = 72$

y_{100+1} – 95%

$E(y_{101} \mid y_{100}, \dots, y_1, \theta) = l_{100} + b_{100} + s_{95} \Rightarrow \hat{y}_{101|100} = 77$

$Var(y_{101} \mid y_{100}, \dots, y_1, \theta) = 16$

$[77 - 1{,}96 \cdot \sqrt{16}\ ;\ 77 + 1{,}96 \cdot \sqrt{16}] \approx \{77 \pm 8\}$



Эконометрика (13)

ETS (ANN)

$$\begin{cases} y_t = l_{t-1} + \varepsilon_t \\ l_t = l_{t-1} + \alpha \varepsilon_t \end{cases}$$

ETS (AAN)

$$\begin{cases} y_t = l_{t-1} + b_{t-1} + \varepsilon_t \\ l_t = l_{t-1} + b_{t-1} + \alpha \varepsilon_t \\ b_t = b_{t-1} + \beta \varepsilon_t \end{cases}$$

$y_t \xrightarrow{?} \Delta y_t \Rightarrow$ стац. $\Rightarrow ARIMA(p, d, q)$

a) $l_{t-1} = y_t - \varepsilon_t$

$y_t - \varepsilon_t = y_{t-1} - \varepsilon_{t-1} + \alpha \varepsilon_t$

$y_t = y_{t-1} + (1+\alpha)\varepsilon_t - \varepsilon_{t-1} \Rightarrow ARIMA(0;1;1)$

$\Delta y_{t+1} = (1+\alpha)\cdot \varepsilon_t + \varepsilon_{t+1} \Rightarrow$ стац. $MA(1)$

b) $b_{t-1} = y_t - l_{t-1} - \varepsilon_t$

$y_t - l_{t-1} - \varepsilon_t = y_{t-1} - l_{t-2} - \varepsilon_{t-1} + \beta \varepsilon_t$

$y_t = y_{t-1} + l_{t-1} - l_{t-2} + (1+\beta)\varepsilon_t - \varepsilon_{t-1}$

$l_{t-1} - l_{t-2} = y_t - y_{t-1} -$

$y_t = l_{t-1} + b_{t-1} + \varepsilon_t \Rightarrow \Delta y_t = \Delta l_{t-1} + \Delta b_{t-1} + \varepsilon_t - \varepsilon_{t-1}$

$\Delta l_t = b_{t-1} + \alpha \varepsilon_t \qquad = b_{t-2} + \alpha\varepsilon_{t-1} + \beta \varepsilon_{t-1}$

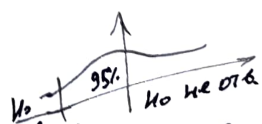
$\Delta b_t = \beta \varepsilon_t \qquad \varepsilon_t + (1+\alpha+\beta)\varepsilon_{t-1}$

$\Delta(\Delta y_t) = \Delta b_{t-2} + \dots$

$= \beta \varepsilon_{t-2}$

$\Rightarrow MA(2) \Rightarrow ARIMA(0;2;2)$

$\Rightarrow$ процесс ↗ стац. ↘ нестац. ↗ $\Delta^d y_t$ – стац. ↘ ?



Тесты на стац.:

a) ADF $\left(t = \dfrac{\hat{\gamma} - 0}{se(\hat{\gamma})} \right.$ по спец. табл.$\left.\right)$ $H_0: \gamma = 1$, $H_A: \gamma < 1$ ← левостор.

b) KPSS – гром. формула по спец. табл. H_0: стац. H_1: есть ед. кор

c) PP

...

ADF: $\Delta y_t = [c] + [\beta t] + \gamma y_{t-1} + [\alpha_1 \Delta y_{t-1} + \alpha_2 \Delta y_{t-2} + \dots + \alpha_p \Delta y_{t-p}]$

Эндогенный регрессор $\quad Cov(x_i, \varepsilon_i) \neq 0$

Экзогенный регрессор $\quad Cov(x_i, \varepsilon_i) = 0$

$$\hat{\beta}_1 = \frac{\widehat{Cov}(x,y)}{\widehat{Var}(x)} \xrightarrow{p} \frac{Cov(x_i, y_i)}{Var(x_i)} = \frac{Cov(x_i, \beta_0 + \beta_1 x_i + \varepsilon_i)}{Var(x_i)}$$

$$= \frac{\beta_1 Cov(x_i, x_i) + Cov(x_i, \varepsilon_i)}{Var(x_i)}$$

$$= \beta_1 + \underbrace{\frac{Cov(x_i, \varepsilon_i)}{Var(x_i)}}_{0}$$

=> при нар. экзогенности МНК ок смещ.

Инстр. пер. 1) экзогенность $\quad Cov(z_i, \varepsilon_i) = 0$

2) релевантность $\quad Cov(z_i, x_i) \neq 0$

TSLS (2SLS)

$\hat{x}_i = \hat{\theta}_1 + \hat{\theta}_2 z_i \quad$ => реш. проблема эндогенности

$$\hat{\beta}_1^{TSLS} = \frac{\widehat{Cov}(\hat{x}, y)}{\widehat{Var}(\hat{x})} = \frac{\widehat{Cov}(\hat{\theta}_0 + \hat{\theta}_1 z, y)}{\widehat{Var}(\cdot)} = \frac{\hat{\theta}_1 \widehat{Cov}(\cdot)}{(\hat{\theta}_1)^2 \cdot Var(z)} =$$

$$\ldots = \frac{\widehat{Cov}(y, z)}{\widehat{Cov}(x, z)}$$

$$\hat{\beta}_1^{2SLS} \xrightarrow{p} \hat{\beta}_1 + \frac{Cov(z, \varepsilon_i)}{Cov(z, x_i)}$$

2SLS

1) $\hat{x} = \theta^T_{\text{экз}} W + \theta^T_{\text{ин}} Z + u_i$

2) $y_i = \beta^T \hat{X} + \beta^T W + \beta^T Z + u_i$

Эндог.: $\text{cov}(z_i^{(k)}, \varepsilon_i) = 0$, $k = \overline{1, m}$

Реш.: на 2 шаге нет гипот. ~~медианы~~

шаге отсут. → $\hat{\beta} = (X^T X)^{-1} X^T y$

ex

$m < k$

$y_i = \beta_0 + \beta_1 x_i^{(1)} + \beta_2 x_i^{(2)} + \varepsilon_i$ $\quad k = 2$

z_i $\quad m = 1$

$\hat{x}_1 = \hat{\theta}_0 + \hat{\theta}_1 z_1$ ⇒ мульт. МК на 2 шаге

$\hat{x}_2 = \hat{\alpha}_0 + \hat{\alpha}_1 z_1$

Исп. спец. процедуры se, s^2

$$se(\hat{\beta}_1) = \sqrt{\frac{s^2}{\sum (x_i - \bar{x})^2} \cdot \frac{1}{\widehat{\text{corr}}(x, z)^2}} \qquad s^2 = \frac{\sum e_i^2}{n-2}$$

Расш. пробит модель

(порядковая пробит модель)

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

$$y_i = \begin{cases} 1, & \text{если } y_i^* \le \gamma_{i,1} \\ \dots \\ k, & \text{если } \gamma_{i,k-1} < y_i^* \le \gamma_{i,k} \\ \dots \\ K, & \text{если } y_i^* > \gamma_{i,k-1} \end{cases}$$, γ_i – пор. значения

Пусть $K = 3$

$$y_i = \begin{cases} 1, & y_i^* \le \gamma_1 \\ 2, & \gamma_1 < y_i^* \le \gamma_2 \\ 3, & y_i^* > \gamma_2 \end{cases}$$

$$P(y_i = 1 | x_i) = \Phi\left(\frac{\gamma_1 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right)$$

$$P(y_i = 2 | x_i) = P(\gamma_1 < y_i^* \le \gamma_2 | x_i)$$

$$= P(\varepsilon_i \le \gamma_2 - \beta_1 - \beta_2 x_{i2} \dots | x_i) - P(\varepsilon_i < \gamma_1 - \beta_1 - \dots | x_i)$$

$$= P\left(\frac{\varepsilon_i}{\sigma} \le \frac{\gamma_2 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} \dots + \dots}{\sigma}\right) - P\left(\frac{\varepsilon_i}{\sigma} < \frac{\gamma_1 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots}{\sigma}\right)$$

$$= \Phi\left(\frac{\gamma_2 - \beta_1}{\sigma} - \frac{(\dots)}{\sigma}\right) - \Phi\left(\frac{\gamma_1 - \beta_1}{\sigma} - \frac{(\dots)}{\sigma}\right)$$

Нормировка $\gamma_1 = 0$, $\sigma = 1$, $\gamma_2 = \gamma$

$$\Rightarrow \quad P(y_i = 1 | x_i) = P(y_i^* < 0 | x_i) = \Phi(-x_i^T \beta)$$

$$P(y_i = 2 | x_i) = P(0 < y_i^* \le \gamma | x_i) = \Phi(\gamma - x_i^T \beta) - \Phi(-x_i^T \beta)$$

$$P(y_i = 3 | x_i) = P(y_i^* > \gamma | x_i) = 1 - \Phi(\gamma - x_i^T \beta)$$

Пусть x_{2i} $\beta_2 > 0$ $\Rightarrow x_2 \uparrow \Rightarrow P(y_i = 1 | x_i) \downarrow$

$$L = \prod \left(\Phi(-x_i^T \beta)\right)^{I(y_i=1)} \left(\Phi(\gamma - x_i^T \beta) - \Phi(-x_i^T \beta)\right)^{I(y_i=2)} \left(1 - \Phi(x_i^T \beta)\right)^{I(y_i=3)}$$

, где $I(y_i = k) = \begin{cases} 1, & y_i = k \\ 0, & y_i \ne k \end{cases}$

Прогнозирование: $\hat{y}_i = k_0$, если $\hat{P}(y_i = k_0 | x_i) = \max\limits_{k = \overline{1,K}} \hat{P}(y_i = k | x_i)$

<. . . .>

$y_i^* = \cancel{\ldots}\ \beta^T X_i + \varepsilon$, $\varepsilon_i \sim N(0, \sigma^2)$ - i.i.d.

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

$\theta_1 = \dfrac{\gamma_i - \beta_i}{\sigma}$ $\quad \hat{\theta}_1 = \dfrac{\hat{\beta}_i - \hat{\gamma}_i}{\hat{\sigma}}$

$\theta_j = \dfrac{\beta_i}{\sigma}$ $\quad \hat{\theta}_j = \dfrac{\hat{\beta}_i}{\hat{\sigma}}$

$\gamma_1 = \ldots = \gamma_n = 0, \ \sigma = 1 \Rightarrow \hat{\theta}_j = \hat{\beta}_j$

$H_0: \varepsilon \sim$ i.i.d. $N(0, \sigma^2)$

(2) $P(y_i = 1 \mid x) = \phi(x_i^T\theta + w_1 (x_i^T\theta)^2 + w_2 (x_i^T\theta)^3)$

$F(z) = \phi(z + w_1 z^2 + w_2 z^3)$

$H_0^*: \ w_1 = w_2 = 0$ $\qquad (H_0^* = H_0)$

оцениваем обе модели $\left.\begin{matrix}1\\2\end{matrix}\right\} \Rightarrow ML \begin{matrix}\nearrow \ln L_1 \\ \searrow \ln L_2\end{matrix}$

Крит. отн. правдоподобия (Likelihood Ratio)

$LR_{стат} = 2(\ln L_2 - \ln L_1) \overset{H_0}{\underset{d}{\sim}} \chi^2(2)$

$LR_{crit} = \chi^2_{1-\alpha}(2)$, α - ур. зн.

Проверка гип. ~~Н0~~ о НС

$H_0: D(\varepsilon_i \mid x_i) = const$

$y_i = \alpha + \beta x_i + \varepsilon_i$

$D(\varepsilon_i \mid x_i) = e^{k x_i}$, $k > 0$

(3) $P(y_i = 1 \mid x_i) = \Phi\left(\dfrac{\alpha + \beta x_i}{\sqrt{e^{k x_i}}}\right)$

$H_0^*: k = 0$

(1) $\to ML \to \ln L_1$

(3) $\to \ln L_3$ $\quad \Rightarrow LR$

Расширения: бинарная → порядковая → мультиномиальная

$u_{ik} = \cancel{...} \ \beta^T x_i + \varepsilon_{ik}$, $k = \overline{1,K}$, $i = \overline{1,n}$

$$p(y_i = k) = P(u_{ik} = \max_{j=\overline{1,k}} u_{ij}) = P(x_{ik}^T \beta + \varepsilon_{ik} > \max_{j=\overline{1,k}/j=k} (x_{ij}^T \beta + \varepsilon_{ij}))$$

$\varepsilon_{ik} \sim F(z) = e^{-e^{-z}}$ — распр. Гумбеля (т.к. норм. не подх.)

$$p(y_i = k) = \frac{\exp(x_{ik}^T \beta)}{\exp(x_{i1}^T \beta) + \ldots + \exp(x_{ik}^T \beta)} = \Big/ : \exp(x_{i1}^T \beta)$$

$$= \frac{\exp(x_{ik}^T \beta - x_{i1}^T \beta)}{1 + \ldots + \exp(x_{ik}^T \beta - x_{i1}^T \beta)} = \left[x_{i1}^T \beta = 0 \right]$$

ест. нормализация

$$= \frac{\exp(x_{ik}^T \beta)}{1 + \ldots + \exp(x_{ik}^T \beta)}$$

x по-разн. влияет на u => умножаем

Мультиномиальная модель

$u_{ik} = (\beta^k)^T x_i + \varepsilon_{ik}$, $i = \overline{1,n}$, $k = \overline{1,k}$

$$p(y_i = k) = \frac{\exp(x_i^T \beta^k)}{\exp(x_i^T \beta^1) + \ldots \exp(x_i^T \beta^k)} = \left[\beta^1 = 0 \right]$$

$$= \frac{\exp(x_i^T \beta^k)}{1 + \exp(x_i^T \beta^2) + \ldots + \exp(\cdot)}$$

$$L(\beta | x_i) = \prod_{i=1}^n \prod_{k=1}^k \left(P(y_i = k)\right)^{d_{ik}} = \prod_{i=1}^n \prod_{k=1}^k \left(\frac{\exp(x_i^T \beta^k)}{1 + \exp(x_i^T \beta^2) + \exp(\cdot)} \right)^{d_{ik}}, \quad d_{ik} = \begin{cases} 1, & y_i = k \\ 0, & y_i \neq k \end{cases}$$

Эконометрика (1)

$p(y_i=1 \mid x_i) = x_i^T\theta$

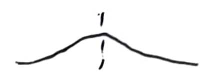
$$\boxed{p(y_i=1 \mid x_i) = G(x_i^T\theta)}$$

$y_i = G(x_i^T\theta) + \varepsilon_i$

$E(\varepsilon_i \mid x_i) = 0$

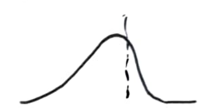
1) $G(x_i^T\theta) = \Phi(x_i^T\theta)$ пробит-модель

2) $G(x_i^T\theta) = \Lambda(x_i^T\theta)$ логит-модель



$\Lambda(z) = \frac{e^z}{1+e^z}$ $\quad \Lambda'(z) = \frac{e^z}{(1+e^z)^2} = \Lambda(z)(1-\Lambda(z))$

3) $G(x_i^T\theta) = E(x_i^T\theta)$ гомпит-модель



$E(z) = 1 - e^{-e^z}$

экстремальная ф-я

$p(y_i=1 \mid x_i) = G(x_i^T\theta)$

$$\frac{\partial p(y_i=1 \mid x_i)}{\partial x_{ik}} = \frac{\partial G(x_i^T\theta)}{\partial x_{ik}}$$

$$\Delta P(y_i=1 \mid x_i) = \frac{\partial G(\cdot)}{\partial x_{ik}} \cdot \Delta x_{ik}$$

для пробит-модели

ф-я плотности

$$\frac{\partial P(y_i=1 \mid x_i)}{\partial x_{ik}} = \frac{\partial \Phi(x_i^T\theta)}{\partial x_{ik}} = \varphi(x_i^T\theta)\cdot\theta_k$$ → интерпретация знака θ_k

---

для логит-модели

$$\frac{\partial P(y_i=1 \mid x_i)}{\partial x_{ik}} = \frac{\partial \Lambda(x_i^T\theta)}{\partial x_{ik}} = \Lambda(x_i^T\theta)\left(1-\Lambda(x_i\theta)\right)\cdot\theta_k$$

ML

$A \quad P(A) = p \qquad \frac{p}{1-p} \leftarrow \frac{\text{произ.}}{\text{не произ.}}$

$A = \{y_i = 1\}$

$$\frac{p(y_i=1 \mid x_i)}{1-p(y_i=1 \mid x_i)} = \frac{p(y_i=1 \mid x_i)}{p(y_i=0 \mid x_i)} = \frac{\Lambda(x_i^T\theta)}{1-\Lambda(x_i^T\theta)} = e^{x_i^T\theta}$$

$\ln(\ \cdot\ ) = x_i^T\theta \quad \leftarrow$ логит $= 0$ шансы равны

логит $> 0 \qquad p > 1-p$

при $\uparrow \Delta x_{ik}$ шансы события $y_i = 1$ против $y_i = 0$ изм. $\approx$ на $\theta_k \cdot \Delta x_{ik} \cdot 100\%$

$$y_i = G(\theta_1 + \theta_2 x_{i2} + \ldots + \theta_k x_{ik}) + \varepsilon_i \qquad i = \overline{1, n}$$

а) $E(\varepsilon_i \mid x_i) = 0$

б) при фикс. x_i $\quad \varepsilon_i$ – i.i.d.

$\Rightarrow$ $y_1, \ldots, y_n$ – i.i.d.

$P(y_i = 1 \mid x_i) = E(y_i \mid x_i) = G(x_i^T\theta)$

$$L = \prod_{i=1}^{n} P(y_i=1 \mid x_i)^{y_i} \cdot P(y_i=0 \mid x_i)^{1-y_i} = \prod_{i=1}^{n} \left(G(x_i^T\theta)\right)^{y_i} \cdot \left(1 - G(x_i^T\theta)\right)^{1-y_i} = L(\theta)$$

$\uparrow$ совм. p получения набора $y_1 \ldots y_n$ ф-я правдоподоб.

$$\ln L = \sum_{i=1}^{n} y_i \ln G(x_i^T\theta) + \sum_{i=1}^{n} (1-y_i) \ln(1 - G(x_i^T\theta))$$

Тривиальная модель

$y_i = G(\theta_1) + \varepsilon_i \qquad E(\varepsilon_i) = 0$

$p(y_i = 1) = G(\theta_1)$

$\hat{p}(y_i = 1) = \frac{\sum y_i}{n} = \bar{y}$ $\leftarrow$ число единиц

Показатели качества моделей бинарного выбора

$y_i = G(x_i^T\theta) + \varepsilon_i$

$$R^2_{predict} = 1 - \frac{\nu_{wrong,1}}{\nu_{wrong,0}}$$

← доля неправ. пр. по выбр. модели

← ~ по трив. модели

коэф. дет. нельзя исп., т.к. модель нелин. по коэф.

$\hat{p}(y_i = 1 | x_i) = G(x_i^T\hat{\theta})$

если $\hat{p}(y_i = 1 | x_i)$ $> 1/2$, $\hat{y}_i = 1$; $\leq 1/2$, $\hat{y}_i = 0$

$G(x_i^T\theta) > 1/2$ | $x_i^T\theta > 0$

$\leq 1/2$ | ≤ 0

$$\hat{y}_i = \begin{cases} 1, & \text{если } x_i^T\theta > 0 \\ 0, & \text{если } x_i^T\theta \leq 0 \end{cases}$$

$$n_{wrong,1} = \sum^n |y_i - \hat{y}_i| = \sum (y_i - \hat{y}_i)^2$$

$$\nu_{wrong,1} = \frac{1}{n}\sum (y_i - \hat{y}_i)^2$$

для тривиальной $p(y_i = 1) = \bar{y}$

если $\bar{y} > 1/2$ => все $\hat{y}_i = 1$

если $\bar{y} \leq 1/2$ => все $\hat{y}_i = 0$

$$\nu_{wrong,0} = \begin{cases} 1 - \bar{y}, & \text{если } \bar{y} > 1/2 \\ \bar{y}, & \text{если } \bar{y} \leq 1/2 \end{cases}$$

$\nexists$

$$R^2_{predict} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{n_{wrong,0}}$$, м.б. < 0

$L_0 \leq L_1 \leq 1$

$\ln L_0 \leq \ln L_1 \leq 0$

$|\ln L_1| \leq |\ln L_0|$

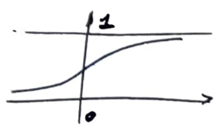
$$\text{pseudo } R^2 = 1 - \frac{1}{1 + 2\frac{\ln L_1 - \ln L_0}{n}}$$

LRI — likelihood ratio index

$$\text{McFadden } R^2 = LRI = 1 - \frac{\ln L_1}{\ln L_0} \in [0; 1)$$

L_1 не прим. 1

из-за асимпт. $L \neq 1$



$$AIC = -\frac{2 \ln L}{n} + \frac{2k}{n}$$

$$BIC(SC) = -\frac{2 \ln L}{n} + \frac{k \cdot \ln k}{n}$$

$$HQ = -\frac{2 \ln L}{n} + \frac{k \cdot \ln(\ln n)}{n}$$

Hannan-
Quinn
IC