Привет !!

Идея – вопрос

Как сильно зависимая выборка „хуже" независимой?

→ KPSS

→ MCMC

$y_1, \ldots, y_n$	$y_1, \ldots, y_n$
эталонная нез-ая один. распр	зависимая y_i ~ один. распр.
	↓ как-то по другому

95% CI μ

$$\left[\bar{y} - 1.96\sqrt{\frac{\sigma^2}{n}};\ \bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}}\right]$$

$$\operatorname{Var}(\bar{y}) = \frac{\sigma^2}{n}$$

$$\sigma^2 = \operatorname{Var}(y_i)$$

$$\operatorname{Var}(\bar{y}) \neq \frac{\sigma^2}{n}$$

(y_t) – стац. процесс

Опр 1 Долгосрочная дисперсия – (идея) значение дисперсии независимых (y_t), к-рое давало бы такой же дов-ый интервал для μ.

Long run variance

$$\operatorname{Var}(\bar{y}) = \frac{\lambda^2}{n} + o\left(\frac{1}{n}\right)$$

λ^2 – долг. дисперсия.

$$\lambda^2 = \lim_{n\to\infty}\left[n \cdot \operatorname{Var}\left(\frac{y_1 + \ldots + y_n}{n}\right)\right]$$

effective sample size

Опр 2 эфф-ое кол-во наблюдений
— (идея) сколько нужно взять н.у.р.в.-х (y_t) с такой же дисперсией, как есть, чтобы дов. инт для μ имел ту же ширину.

$$n_{eff} = n \cdot \frac{Var(y_t)}{\lambda^2}, \quad \lambda^2 = \lim_{n\to\infty} n \cdot Var(\bar{y})$$

/ESS/

Упр.

MA(1) $y_t = u_t + 2u_{t-1}$ $\quad Var(u_t) = 16$

u_t - б. шум.

а) $Var(y_t)$, долг. дисп. y_t λ^2

б) $n = 100$ $\quad n_{eff}$?

$$Var(y_t) = Var(u_t + 2u_{t-1}) = \sigma^2 + 4\sigma^2 = 5\sigma^2 = 5 \cdot 16 = 80$$

$$Var(\bar{y}) = \frac{\lambda^2}{n} + o\left(\frac{1}{n}\right)$$

$$n Var(\bar{y}) = \lambda^2 + o(1) \to \lambda^2$$

$$n \cdot Var\left(\frac{y_1 + y_2 + \ldots + y_n}{n}\right) \to ?$$

$$\frac{1}{n} \cdot \left[n \cdot Var(y_1) + 2(n-1) \cdot Cov(y_1, y_2)\right] \to ?$$

$$\lambda^2 = Var(y_1) + 2 \cdot Cov(y_1, y_2) =$$

$$= 5\sigma^2 + 2 \cdot Cov(u_1 + 2u_0; u_2 + 2u_1) =$$

$$= 5\sigma^2 + 4\sigma^2 = 9\sigma^2 = 9 \cdot 16 = 144.$$

б) $n = 100$ $\quad n_{eff} = 100 \cdot \frac{80}{144} \approx 56$

част. исключение

y_1, y_1, y_1, y_1, y_1, ⋮

λ^2 не сх-ся

Упр. (y_t) - стац. процесс

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

долг. дисп. $\lambda^2 = ?\ \gamma_0 + 2\gamma_1 + 2\gamma_2 + \ldots$

Все $|Corr(y_t, y_{t-k})| < 1$

$$\gamma_k = \gamma_0 = Cov(y_t, y_{t-k})$$

$$\lambda^2 = \gamma_0 + 2\gamma_0 + 2\gamma_0 + 2\gamma_0 + \ldots = +\infty$$

$$Var\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} \cdot \sigma^2$$

$$\det \begin{bmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{bmatrix} \geq 0$$

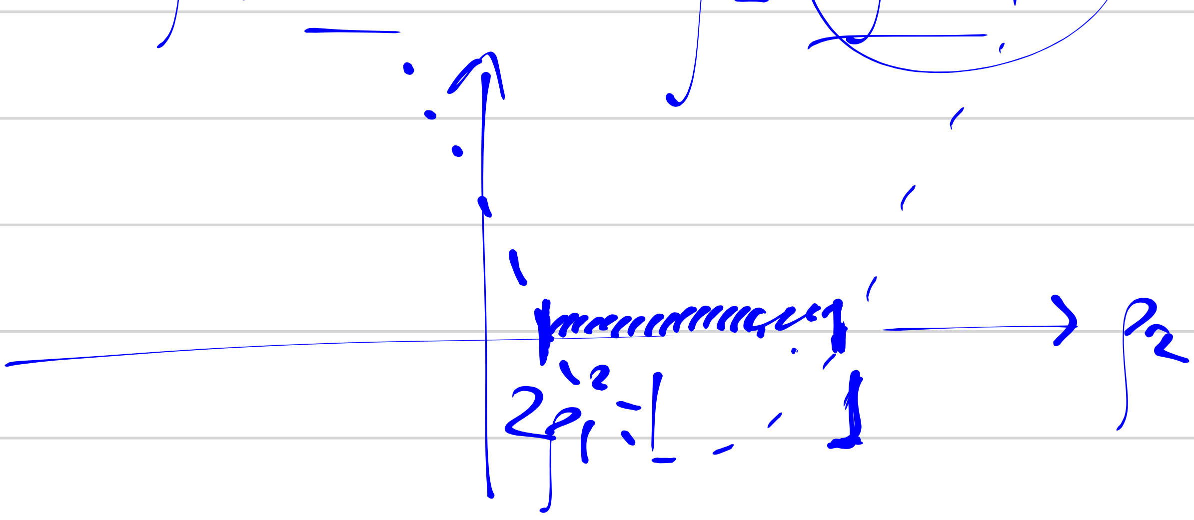
$$1 - \rho_2^2 \geq 0$$

$$\det \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} = 1 + \rho_1^2 \cdot \rho_2 + \rho_1^2 \cdot \rho_2 - \rho_2^2 - \rho_1^2 - \rho_1^2 \geq 0$$

$$1 + 2\rho_1^2 \cdot \rho_2 - 2\rho_1^2 - \rho_2^2 \geq 0$$

$$\rho_2^2 - \rho_2 \cdot 2\rho_1^2 + (2\rho_1^2 - 1) \leq 0$$

$$\rho_2^* = 1 \qquad \rho_2^* = 2\rho_1^2 - 1$$



y_1, y_2, y_1, y_2, y_1, y_2, ⋮

? достат-но ли для ю-сти λ^2?

$$|\mathrm{corr}(y_t, y_s)| < 1$$

$$y_1$$

$$0{,}9 \cdot y_1 + \sqrt{0{,}1} \cdot u_2$$

(u_t) - б. шум.
y — один. сл.

$$y_t = \sqrt{1-0{,}1^{t}} \cdot y + \sqrt{0{,}1^{t}} \cdot u_t$$

Упр. $\boxed{\text{Придумайте}}$ стац. процесс с $|\mathrm{corr}(y_t, y_s)| < 1$ и $\forall \varepsilon > 0 \ \exists t, s \ |\mathrm{corr}(y_t, y_s)| > 1-\varepsilon$

[или док-ть, что нет].

$KPSS_c$ $\qquad\qquad$ $KPSS_{ct}$

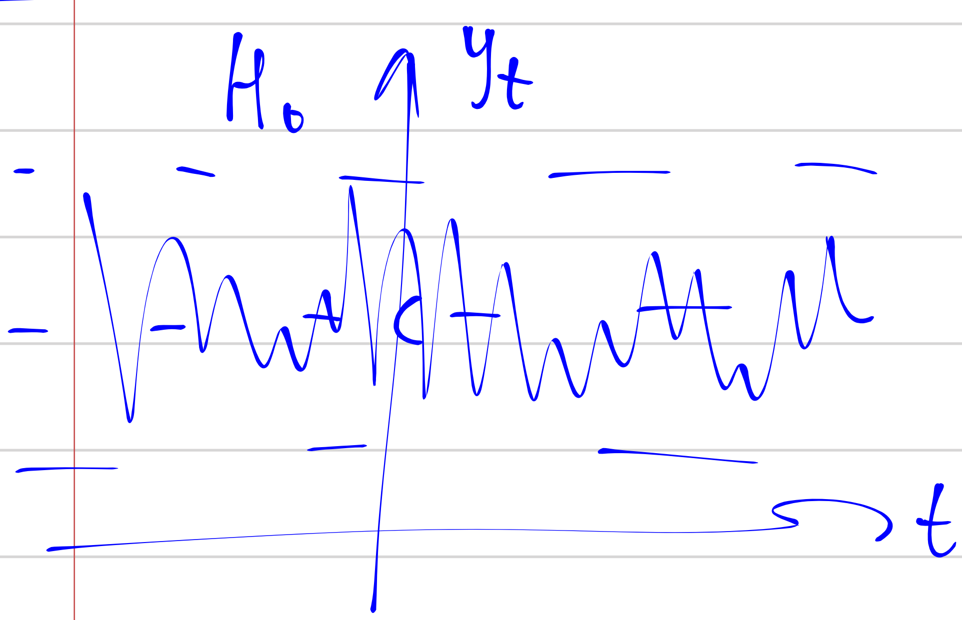
$$y_t = c + z_t + x_t + [\beta \cdot t]$$

(x_t) – стац. процесс с $E(x_t)=0$.

$H_0: z_t = 0$ $\qquad$ $y_t = c + x_t + [\beta \cdot t]$

$H_A: \begin{cases} z_t = z_{t-1} + u_t \\ z_0 = 0 \end{cases}$ $\qquad$ $y_t = c + u_1 + \ldots + u_t + x_t + [\beta \cdot t]$

(u_t) – б. шум не зав с (x_t).



$$\hat{c} = \bar{y} \qquad \hat{u}_t = y_t - \hat{c} = y_t - \bar{y}$$

$$S_t = \sum_{i=1}^{t} \hat{u}_i$$

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\lambda}^2 \cdot T^2}$$

$\hat{\lambda}^2$ (сост) – оценка долг. дисперсии.

упр.

При H_0: $KPSS \xrightarrow[n \to \infty]{dist} KPSS_c \to KPSS_{ct}$ $\qquad$ $\hat{\lambda}^2 = \hat{\gamma}_0 + 2\hat{\gamma}_1 + \ldots + 2\hat{\gamma}_{T/2}$

Если $KPSS > KPSS_c^{crit}$, то H_0 отвергается.

Алгоритм Хайндмана-Хандакара (автом-ий подбор ARIMA)

Шаг 1. (для сезонных рядов)
Надо ли $(y_t) \to (\Delta_{12} y_t)$?

раньше брали тест Canova-Hansen

STL разложение. $y_t = trend_t + season_t + res_t$

$$\text{Сила сезонности} = 1 - \frac{sVar(residual)}{sVar(residual + season)}$$

Если Сила сезонности > 0,65 (порог), то переходим от (y_t) к $(\Delta_{12} y_t)$.

Шаг 2

$$(y_t) \xrightarrow{KPSS_c} (\Delta y_t) \xrightarrow{KPSS_c} (\Delta^2 y_t)$$

Шаг 3 перебор SARMA(p, q)(P, Q)

$p+q \leq 5$
$P+Q \leq 5$

и выбор опт-ой по AIC.

Упр. скачать ряд заканчив-ся автоARIMA в питоне. (без пропусков)