Решение КР-2020 по временным рядам

1. Рассмотрим уравнение $y_t = 3 + 0.4y_{t-1} + u_t$, где u_t независимы и нормальны $\mathcal{N}(0{,}9)$. Я не спрашиваю, есть ли у уравнения стационарное решение и сколько их. Скажу прямо: оно есть! Верь мне! И даже добавлю, что в нём y_t представим в виде

$$y_t = c + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \dots$$

   (a) Найди c и все α_k.

   (b) Найди $\mathrm{E}(y_t), \mathrm{Var}(y_t)$ и первые два значения автокорреляционной функции.

   Дополнительно известно, что $y_{100} = 5$

   (c) Найди 95%-й предиктивный интервал для y_{101}.

   (d) Найди 95%-й долгосрочный предиктивный интервал для y_{100+h}, где h $\to \infty$. Зависит ли он от y_{100}?

   Решение:

   (a) Рассмотрим временной ряд $y_t = 3 + 0.4y_{t-1} + u_t$

$$y_t = 3 + 0.4y_{t-1} + u_t$$
$$(1 - 0.4L)y_t = 3 + u_t$$
$$y_t = \frac{3+u_t}{1-0.4L}$$
$$y_t = \frac{30}{6} + \sum_{i=0}^{n} 0.4^i L^i u_t = 5 + \sum_{i=1}^{n} 0.4^i u_{t-i}$$

   Таким образом, получаем что $c = 5$, $\alpha_k = 0.4^k$

   (b) Для решения пункта b рассмотрим $y_t = 5 + \sum_{i=0}^{n} 0.4^i u_{t-i}$

$$\mathrm{E}(y_t) = \mathrm{E}(5 + \sum_{i=0}^{n} 0.4^i u_{t-i}) = 5$$
$$\mathrm{Var}(y_t) = \mathrm{Var}(5 + \sum_{i=0}^{n} 0.4^i u_{t-i}) = \sum_{i=0}^{n} 0.4^{2i}\mathrm{Var}(u_t) = \sum_{i=0}^{n} 0.4^{2i} * 9 = \frac{9}{0.84}$$

   Автокорреляционная функция $\rho_k = \frac{\gamma_k}{\gamma_0}$, где $\gamma_k = \mathrm{cov}(y_t, y_{t+k})$

$$\gamma_0 = \mathrm{Var}(y_t) = \frac{9}{0.84}$$
$$\gamma_1 = \mathrm{cov}(y_t, y_{t+1}) = \mathrm{cov}(y_t, 3 + 0.4y_t + u_{t+1}) = \mathrm{cov}(y_t, u_{t+1}) + 0.4\mathrm{Var}(y_t) = 0.4\mathrm{Var}(y_t)$$

   Получаем ρ_0, ρ_1:

$$\rho_0 = \frac{\mathrm{Var}(y_t)}{\mathrm{Var}(y_t)} = 1$$
$$\rho_1 = \frac{0.4\mathrm{Var}(y_t)}{\mathrm{Var}(y_t)} = 0.4$$

(c) Найди 95%-й предиктивный интервал для y_{101}.

$$y_{101|100} = \mathrm{E}(y_{100} * 0.4 + u_{101} + 3|\mathcal{F}_{100}) = 0.4y_{100} + 3 = 5$$
$$\mathrm{Var}(y_{101|100}) = \mathrm{Var}(y_{100} * 0.4 + u_{101} + 3|\mathcal{F}_{100}) = 9$$

Тогда доверительный интервал принимает вид: $5 \pm 1.96 * 3$

(d) Найди 95%-й долгосрочный предиктивный интервал для y_{100+h}, где h $\to \infty$. Зависит ли он от y_{100}?

Рассмотрим для начала $y_{101|100}, y_{102|100}, y_{103|100}$:

$$y_{101|100} = 0.4y_{100} + 3(\text{знаем из предыдущего пункта})$$
$$y_{102|100} = \mathrm{E}(0.4y_{101|100} + u_{102} + 3|\mathcal{F}_{100}) = 0.4^2y_{100} + 0.4 * 3 + 3$$
$$y_{103|100} = \mathrm{E}(0.4y_{102|100} + u_{103} + 3|\mathcal{F}_{100}) = 0.4(0.4^2y_{100} + 0.4 * 3 + 3) + 3$$

Получаем, что формула для $y_{100+h|100}$ выглядит следующим образом:

$$0.4^h y_{100} + 3(1 + 0.4^2 + ...0.4^{h-1}) = 5$$

2. Временной ряд порождается MA(2) процессом $y_t = 3 + u_t + 0.5u_{t-1} + 0.2u_{t-2}$. Однако Винни-Пух строит регрессию $\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 y_{t-1}$ с помощью МНК.

(a) Найди $\mathrm{E}(y_t), \mathrm{Var}(y_t), \mathrm{Cov}(y_t, y_s)$.

(b) Какие коэффициенты примерно получит Винни-Пух, если у него много наблюдений?

Решение:

(a) Найди $\mathrm{E}(y_t), \mathrm{Var}(y_t), \mathrm{Cov}(y_t, y_s)$.

$$\mathrm{E}(y_t) = \mathrm{E}(3 + u_t + 0.5u_{t-1} + 0.2u_{t-2}) = 3$$
$$\mathrm{Var}(y_t) = \mathrm{Var}(3 + u_t + 0.5u_{t-1} + 0.2u_{t-2}) = 0.5^2\sigma^2 + 0.2^2\sigma^2 + \sigma^2 = 1.29\sigma^2$$
$$\mathrm{Cov}(y_t, y_s) = \mathrm{Cov}(3 + u_t + 0.5u_{t-1} + 0.2u_{t-2}, 3 + u_s + 0.5u_{s-1} + 0.2u_{s-2}) =$$
$$\sigma^2\mathbb{1}[t = s] + 0.5\sigma^2\mathbb{1}[t = s - 1] + 0.2\sigma^2\mathbb{1}[t = s - 2] + 0.5\sigma^2\mathbb{1}[t - 1 = s] + 0.5^2\sigma^2\mathbb{1}[t = s] + 0.5 *$$
$$0.2\sigma^2\mathbb{1}[t - 1 = s - 2] + 0.2\sigma^2\mathbb{1}[t - 2 = s] + 0.2 * 0.5\sigma^2\mathbb{1}[t - 2 = s - 1]$$

(b) Какие коэффициенты примерно получит Винни-Пух, если у него много наблюдений?

Так как Винни-Пух строит регрессию с помощью МНК то верны следующие формулы:

$$\hat{\beta}_2 = \frac{\mathrm{Cov}(y_t, y_{t-1})}{\mathrm{Var}(y_t)} = \frac{0.5\sigma^2 + 0.2 * 0.5\sigma^2}{\sigma^2(1+0.5^2+0.2^2)} = \frac{0.6}{1.29} \approx 0.47$$
$$\hat{\beta}_1 = E(y_t) - \hat{\beta}_2 E(y_{t-1}) = 3(1-\hat{\beta}_2) = 3(1-0.47) \approx 1.69$$

3. Рассмотрим процесс $y_t = u_1 sin(5t) + u_2 cos(5t)$, где u_t — белый шум.

   (a) Является ли данный процесс стационарным?

   (b) Можно ли представить данный процесс в виде MA(∞)? На всякий случай, чтобы не гуглить, я напомню, MA(∞) - процесс имеет вид:

   $$y_t = c + \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \dots,$$

   где ϵ_t — белый шум. И да, обращу внимание, что шум ϵ_t не обязательно совпадает с шумом u_t

Решение:

(a) Является ли данный процесс стационарным?

$$\mathrm{E}(y_t) = 0$$
$$\mathrm{Cov}(y_t, y_s) = sin(5t)sin(5s) + cos(5t)cos(5s) = cos(5(t-s))$$

Процесс является стационарным, так как выполнены условия.

(b) Можно ли представить данный процесс в виде MA(∞)?

Пользуясь теоремой Юла-Волкера:

$$y_t = \alpha + \phi_{11} y_{t-1} + w_t$$
$$cov(y_{t-1}, y_t - \phi_{11} y_{t-1}) = \gamma_1 - \phi_{11}\gamma_0 = 0 \mid \ : \gamma_0$$
$$\rho_1 - \phi_{11} = 0 \Rightarrow \phi_{11} = \rho_1 = \cos(5)$$

Получаем, что наш исходный процесс представим в виде: $y_t = \alpha + \cos(5) y_{t-1} + w_t$

$$(1 - \cos(5)L) y_t = \alpha + w_t$$
$$y_t = \frac{\alpha}{1-\cos(5)} + \frac{w_t}{1-\cos(5)L} = \frac{\alpha}{1-\cos(5)} + \sum_{i=0}^{n} \cos(5)^i L^i w_t = \frac{\alpha}{1-\cos(5)} + \sum_{i=0}^{n} \cos(5)^i w_{t-i}$$

Так как $\cos(5) \approx 0.28$, то наш процесс представим в виде MA(∞)

4. У стационарного процесса y_t первые две обычные корреляции равны $\rho_1 = 0.5, \rho_2 = 0.2$, а ожидание равно $\mathrm{E}(y_t) = 20$. Известно, что $y_{100} = 25, y_{99} = 22$. Найди наилучший

точечный прогноз для y_{101}. Псссст, парень! Это была задача про частные корреляции!

Решение: Воспользуемся теоремой Юла-Волкера(так как процесс стационарен, то теоремой можно пользоваться): $y_t = \alpha + \phi_{21} y_{t-1} + \phi_{22} y_{t-2} + w_t$

$$\begin{cases} \operatorname{cov}(\mathrm{y}_t - \alpha - \phi_{21} y_{t-1} - \phi_{22} y_{t-2}, y_{t-1}) = 0 \\ \operatorname{cov}(\mathrm{y}_t - \alpha - \phi_{21} y_{t-1} - \phi_{22} y_{t-2}, y_{t-1}) = 0 \end{cases} \to \begin{cases} \varphi(1) - \phi_{21}\varphi(0) - \phi_{22}\varphi(1) = 0 \\ \varphi(2) - \phi_{21}\varphi(1) - \phi_{22}\varphi(0) = 0 \end{cases}$$

Делим на $\varphi(0)$: $\begin{cases} \rho_1 - \phi_{21} - \phi_{22}\rho_1 = 0 \\ \rho_2 - \phi_{21}\rho_1 - \phi_{22} = 0 \end{cases} \to \left[\begin{array}{c} \phi_{21} = \frac{\rho_1 - \rho_1\rho_2}{1-\rho_1^2} \approx 0.53 \\ \phi_{22} = \rho_2 - \rho_1 \frac{\rho_1 - \rho_1\rho_2}{1-\rho_1^2} \approx -0.07 \end{array} \right.$

Получаем, что $y_t = \alpha + 0.53 y_{t-1} - 0.07 y_{t-2} + w_t$

Так как процесс стационарен, то $\forall t : \mathrm{E}(y_t) = const$. Отсюда найдем α:

$$\mathrm{E}(y_t) = \alpha + 0.53\mathrm{E}(y_t) - 0.07\mathrm{E}(y_t)$$

$$\alpha = 10.8$$

Теперь найдем прогноз:

$$y_{101|100} = E(10.8 + 0.53 y_{100} - 0.07 y_{99} + w_{101} | \mathcal{F}_{100}) = 10.8 + 0.53 * 25 - 0.07 * 22 = 22.51$$

5. Вспомни $ETS(AAN)$ модель, а я тебе даже уравнения напишу:

$$\begin{cases} y_t = \ell_{t-1} + b_{t-1} + u_t \\ \ell_t = \ell_{t-1} + b_{t-1} + \alpha u_t \\ b_t = b_{t-1} + \beta u_t \\ u_t \sim \mathcal{N}(0; \sigma^2) \end{cases} \tag{1}$$

a) Ты вчера чатик читал? Помнишь там вопрос был? Ага! Докажи, что ни при каких l_0 и b_0 этот процесс не будет стационарным. Или опровергни и приведи пример, при каких будет.

   Константы α, β лежат в интервале (0;1).

b) При $l_{100} = 20$, $b_{100} = 2, \alpha = 0.2, \beta = 0.3, \sigma^2 = 16$ построй интервальный прогноз на один и два шага вперёд.

Решение:

a) Ты вчера чатик читал? Помнишь там вопрос был? Ага! Докажи, что ни при каких l_0 и b_0 этот процесс не будет стационарным. Или опровергни и приведи пример, при каких будет.

   Так как мы имеем ETS(AAN), то есть ряд y_t с аддитивным трендом, то процесс не будет стационарным ни для каких l_0, b_0, так как $E(y_t) = f(t) \neq const$.

b) При $l_{100} = 20$, $b_{100} = 2, \alpha = 0.2, \beta = 0.3, \sigma^2 = 16$ построй интервальный прогноз на один и два шага вперёд.

$y_{101|100} = E(l_{100} + b_{100} + u_{101}|\mathcal{F}_{100}) = l_{100} + b_{100} = 22$

$y_{102|100} = E(l_{101} + b_{101} + u_{102}|\mathcal{F}_{100}) = E(l_{100} + b_{100} + \alpha u_{101} + b_{100} + \beta u_{101} + u_{102}) = l_{100} + 2b_{100} = 24$

6. Величины x_t равновероятно равны 0 или 1, а величины u_t нормальны $\mathcal{N}(0, 1)$. Все упомянутые величины независимы. Рассмотрим процесс $z_t = x_t^2(1 - x_{t-1})u_t$.

a) Найди $Cov(z_t, z_s)$. Стационарен ли процесс z_t?

b) Скажу тебе по секрету, что $z_{100} = 2.3$. Построй точечный и 95%-й интервальный прогноз на один и два шага вперёд. Чем интервальные прогнозы в этой задаче особенные?

Решение:

a) Найди $Cov(z_t, z_s)$. Стационарен ли процесс z_t?

$E(1 - x_{t-1}) = 1 - \frac{1}{2} = \frac{1}{2}$

$E(x_t^2) = \frac{1}{2}, E(u_t) = 0$

$cov(z_t, z_s) = cov(x_t^2(1 - x_{t-1})u_t, x_s^2(1 - x_{s-1})u_s) = E(x_t^2(1 - x_{t-1})u_t x_s^2(1 - x_{s-1})u_s) - E(x_t^2(1 - x_{t-1})u_t)E(x_s^2(1 - x_{s-1})u_s) = |\text{так как все случайные величины независимы}| = 0$

$E(z_t) = E(x_t^2(1 - x_{t-1})u_t) = 0$

Так как $E(z_t) = const, cov(z_t, z_s) = (t - s) - (t - s) = f(t - s)$, то процесс является стационарным.

b) Скажу тебе по секрету, что $z_{100} = 2.3$. Построй точечный и 95%-й интервальный прогноз на один и два шага вперёд. Чем интервальные прогнозы в этой задаче особенные?

Рассмотрим значение $z_{100} = 2.3$. $z_{100} \neq 0$ только в случае, если $x_{100} = 1, x_{99} = 0$.

$z_{101|100} = E(x_{101}^2(1 - x_{100})u_{101}|\mathcal{F}_{100}) = |\text{Так как } z_{100} = 2.3\text{, а значит } x_{100} = 1| = 0$

$z_{102|100} = E(x_{102}^2(1 - x_{101})u_{102}|\mathcal{F}_{100}) = |\text{Пользуясь независимостью}| = 0$

Теперь рассмотрим доверительные интервалы:

$\hat{z}_{101|100} \pm 1.96 * \hat{Var}(z_{101|100})$

$Var(z_{101|100}) = Var(x_{101}^2(1-x_{100})u_{101}|\mathcal{F}_{100}) = |\text{Так как } x_{100} \text{ известно}| = (1-x_{100})^2 Var(x_{101}^2 u_{101}|\mathcal{F}_{100}) = 0$

Итоговый доверительный интервал:

$0 \pm 1.96 * 0$

$\hat{z}_{102|100} \pm 1.96 * \hat{Var}(z_{102|100})$

$Var(z_{102|100}) = Var(x_{102}^2(1-x_{101})u_{101}|\mathcal{F}_{100}) = E(x_{102}^4(1-x_{101})^2 u_{101}^2|\mathcal{F}_{100}) - E^2(x_{102}^2(1-x_{101})u_{101}|\mathcal{F}_{100}) = \frac{1}{4} - \frac{1}{8} = \frac{1}{8}$ Итоговый доверительный интервал: $0 \pm 1.96 * \frac{1}{8}$

Особенность интервальных прогнозов в данной задаче состоит в том, что прогноз на следующий период сильно зависит от того, какое принимает значение z_t(0 или другое число), так как в таком случае можно однозначно определить значение x_t в z_{t+1}