Экзамен, 5-ый билет.

Романенко Демьян, M3238

22.06.2020

1 Билет

1. Ошибки первого и второго рода и их вероятности как критерий качества критерия (теста) проверки гипотез. Подход Неймана-Пирсона.

2. Свойства ЭФР в целом. Расстояние Колмогорова, Смирнова. Теоремы Гливенко-Кантелли, Колмогорова, Мизеса - Смирнова. Построение доверительной полосы для функции распределения.

3. Пусть простейший процесс имеет вид $X(t) = Yt + c$, $Y \sim U(a,b)$. Найти корреляционную функцию и нормированную корреляционную функцию этого процесса.

2 Ответы

2.1 Первый вопрос

2.1.1 Ошибки первого и второго рода и их вероятности как критерий качества критерия (теста) проверки гипотез.

Пусть $H_0 : \theta \in H_0$ – гипотеза, $H_1 : H_0 \subset \Theta, H_1 = \bar{H}_0$ – альтернатива, $\varphi : \mathcal{X} \to \{0,1\}$ – тест.

Ошибкой I рода называют отклонение основной гипотезы, в то время как она была справедлива.

Вероятность ошибок I рода теста φ называют α такую, что: $\alpha(\varphi,\theta) := P_\theta(\mathcal{X}_{n,1})$, $\theta \in \Theta_{H_0}$.

Ошибкой II рода называют принятие основной гипотезы, в то время как она не была справедлива.

Вероятность ошибок II рода теста φ называют β такую, что: $\beta(\varphi,\theta) := P_\theta(\mathcal{X}_{n,0})$, $\theta \in \Theta_{H_1}$

Уровнем значимости теста называют верхнюю границу вероятности ошибки I рода по всем возможным наблюдаемым значениям неизвестных параметров, отвечающих основной гипотезе: $\alpha(\varphi) := \sup\limits_{\theta \in \Theta_{H_0}} \alpha(\varphi,\theta)$

Мощностью теста называют следующую величину: $\gamma(\varphi,\theta) := 1 - \beta(\varphi,\theta)$

2.1.2 Подход Неймана-Пирсона.

Будем искать лучшую оценку в классе оценок с зафиксированным максимальным значением уровня значимости. Зафиксируем $\alpha \in (0,1)$ (обычно выбирают малое значение, близкое к нулю). Будем считать это значение минимальной допустимой величиной ошибки I рода (*допустимый уровень значимости*). Рассмотрим множество всех тестов таких, что: $\bar{\Phi}_\alpha = \{\varphi = \varphi(x) \mid \alpha(\varphi) \leqslant \alpha\}$. Среди этих тестов выбирается тест с минимальным значением β. В асимптотических задачах ограничения накладываются на предельные значения.

2.2 Второй вопрос

2.2.1 Свойства ЭФР в целом.

Эмпирической функцией распределения (ЭФР) в точке $t \in \mathbb{R}$ называют следующую оценку функции распределения генеральной совокупности: $F_n(t) = \frac{1}{n}\sum_{i=1}^n 1_{(-\infty,t)}$. Иными словами, значение ЭФР в точке t равно отношению числа наблюдений, меньших t, к их общему числу n.

Свойства ЭФР:

1. ЭФР кусочно-постоянна;

2. Скачки ЭФР имеют вид $\frac{k}{n}$ для некоторого $k \in (1;n)$;

3. Область принимаемых значений: $[0;1]$;

4. $F_n(t)$ является состоятельной оценкой: $F_n(t_0) = \bar{\xi}_n : F_n(t_0) \xrightarrow[p=1]{} F_x$;

5. $F_n(t)$ является асимптотически нормальной оценкой;

6. Частота может служить как оценка функции распределения генеральной совокупности. При фиксированном $t = t_0$: $F_x(t_0) \approx F_n(t_0) = \frac{\xi_1 + \ldots + \xi_n}{n} = \frac{k_n}{n}$ – частота.

2.2.2 Расстояние Колмогорова, Смирнова.

Расстояние Колмогорова: $\rho_\infty(F_n, F_x) = \sup\limits_t |F_n(t) - F_x(t)|$
Расстояние Смирнова: $\rho_2^2(F_n, F_x) = \int\limits_{\mathbb{R}} (F_n(t) - F_x(t))^2 \, \mathrm{d}F_x(t)$.

2.2.3 Теоремы Гливенко-Кантелли, Колмогорова, Мизеса - Смирнова

Теорема Гливенко-Кантелли Пусть $\mathcal{F}$ – множество функций распределения. Тогда $\forall F_x(t) \in \mathcal{F}$ с вероятностью 1 справедливо предельное неравенство: $\rho_\infty(F_n, F_x) \xrightarrow[n\to\infty]{} 0$. То же верно для ρ_2, так как $\rho_2 \leqslant \rho_\infty$. $F_n(t)$ – состоятельная оценка $F_x(t)$ в расстояниях Колмогорова и Смирнова.

Теорема Колмогорова Пусть $\mathcal{F}_c$ – множество всех непрерывных функций распределения, $F_x \in \mathcal{F}_c$. Тогда
$$P_{n,F}(\sqrt{n}\rho_\infty(F_n, F_x) < u) \xrightarrow[n\to\infty]{} \mathcal{K}(u) = \begin{cases} 0, \; u = 0, \\ \sum\limits_{j=-\infty}^{+\infty} (-1)^j e^{-2(ju)^2}, \; u > 0. \end{cases}$$

Теорема Мизеса - Смирнова $P_{n,F}(\sqrt{n}\rho_2^2(F_n, F_x) < u) \xrightarrow[n\to\infty]{} \mathcal{S}(u)$, где $\mathcal{S}(u)$ есть функция распределения следующей случайной величины: $\mathcal{U} = \sum\limits_{j=1}^{\infty} \frac{\xi_j^2}{j^2\pi^2}, \; \xi_j \sim N(0,1), \; \xi_j$независимые.

2.2.4 Построение доверительной полосы для функции распределения.

Используя теорему Колмогорова, можно построить доверительную полосу для неизвестной функции распределения. *Доверительной полосой* γ называют часть плоскости, в которую с надежностью γ попадает функция распределения генеральной совокупности.

Теорема Доверительная полоса задаётся функциями: $F_n^-(t) = \max\left(0, F_n(t) - \frac{u_\gamma}{\sqrt{n}}\right)$, $F_n^+(t) = \min\left(1, F_n(t) + \frac{u_\gamma}{\sqrt{n}}\right)$, где u_γ определяется из условия $\mathcal{K}(u_\gamma) = \gamma$.

$\square$

Достаточно проверить, что $P_x(F_n^-(t) \leqslant F_x(t) \leqslant F_n^+(t)) \xrightarrow[n\to\infty]{} \gamma$:
$P_x(F_n^-(t) \leqslant F_x(t) \leqslant F_n^+(t)) = P_x\left(F_x(t) - \frac{u_\gamma}{\sqrt{n}} \leqslant F_x(t) \leqslant F_n(t) + \frac{u_\gamma}{\sqrt{n}}\right) = P_x\left(\sqrt{n}|F_x(t) - F_n(t)| \leqslant u_\gamma\right) =$
$= P_x\left(\sqrt{n}\sup\limits_{t\in\mathbb{R}} |F_x(t) - F_n(t)| \leqslant u_\gamma\right) \xrightarrow[n\to\infty]{} \mathcal{K}(u_\gamma) = \gamma$

$\blacksquare$

3 Задача

Формулировка Пусть простейший процесс имеет вид $X(t) = Yt + c$, $Y \sim U(a,b)$. Найти корреляционную функцию и нормированную корреляционную функцию этого процесса.

Решение

1. Найдём матожидание процесса: $E_X(t) = E(Yt + c) = tE(y) + c = t\frac{a+b}{2} + c$;

2. Найдем вид центрированного сечения: $\tilde{X}(t) = X(t) + E_X(T) = Yt + c - t\frac{a+b}{2}$;

3. Найдём саму корреляционную функцию этого процесса: $K_X(t, t') = E(\tilde{X}(t) * \tilde{X}(t')) = tt'E(Y - \frac{a+b}{2})^2 =$
$= tt'D(Y) = tt'\frac{(b-a)^2}{12}$;

4. Найдём дисперсию сечений: $D_X(t) = D(Yt + c) = t^2D(Y) = t^2\frac{(b-a)^2}{2}$;

5. Найдём саму нормированную корреляционную функцию этого процесса: $\rho_X(t, t') = \frac{K_X(t,t')}{\sigma_X(t)\sigma_X(t')} =$
$= tt'\frac{(b-a)^2}{12}\sqrt{\frac{12}{t^2(b-a)^2}}\sqrt{\frac{12}{t'^2(b-a)^2}} = 1$;

Проверка: из определения $X(t)$ зависимость между сечениями должна быть линейной, какой она и является по результату вычислений.