§ Контрольна робота 2, Варіант 5 §

Задача 1: Кореляція і умовний розподіл

Умова. Дано щільність розподілу двовимірного випадкового вектору (ξ, η).

$$f_{(\xi,\eta)}(x,y) = \alpha x^2 y^4 \cdot \mathbb{1}_{[0,1]\times[0,2]}(x,y)$$

(а) Знайти значення сталого параметру α, умовну щільність розподілу ξ за умови $\eta = y$.

(б) Знайти коефіцієнт кореляції випадкових величин ξ, η. Чи незалежні ці випадкові величини?

Розв'язання.

Пункт а. Для знаходження коефіцієнту α скористаємося умовою нормування, тобто $\iint_{\mathbb{R}\times\mathbb{R}} f_{(\xi,\eta)}(x,y)dxdy = 1$. Підставимо нашу щільність:

$$\iint_{\mathbb{R}\times\mathbb{R}} f_{(\xi,\eta)}(x,y)dxdy = \iint_{\mathbb{R}\times\mathbb{R}} \alpha x^2 y^4 \cdot \mathbb{1}_{[0,1]\times[0,2]}(x,y)dxdy \tag{1.1}$$

Якщо врахувати, що щільність зануляється за прямокутником $[0,1]\times[0,2]$, то маємо умову

$$\alpha \iint_{[0,1]\times[0,2]} x^2y^4dxdy = 1 \implies \alpha = \left(\iint_{[0,1]\times[0,2]} x^2y^4dxdy\right)^{-1} =: I^{-1} \tag{1.2}$$

за умови, що інтеграл I має ненульове значення. Отже, обраховуємо сам інтеграл. За теоремою Фубіні, можемо записати:

$$I = \int_0^1 \left(\int_0^2 x^2y^4dy\right)dx = \int_0^1 \left(\frac{x^2y^5}{5}\Big|_{y=0}^{y=2}\right)dx = \frac{32}{5}\int_0^1 x^2dx = \frac{32}{15} \tag{1.3}$$

Звідси остаточно $\boxed{\alpha = \frac{15}{32}}$ і наша щільність розподілу вектору (ξ, η) тому має вигляд

$$f_{(\xi,\eta)}(x,y) = \frac{15}{32} \cdot x^2y^4 \cdot \mathbb{1}_{[0,1]\times[0,2]}(x,y) \tag{1.4}$$

Тепер потрібно знайти щільність розподілу ξ за умови $\eta = y$. За означенням, маємо

$$f_{\xi|\eta}(x \mid y) = \begin{cases} f_{(\xi,\eta)}(x,y)/f_\eta(y), & f_\eta(y) \neq 0 \\ 0, & f_\eta(y) = 0 \end{cases} \tag{1.5}$$

Отже бачимо, що нам потрібно знайти маргінальний розподіл $f_\eta(y)$. Більш того, буде зручно для наступного пункту також знайти маргінальний розподіл $f_\xi(x)$, тому пропоную це зробити одразу. Знову ж таки за означенням

$$f_\xi(x) = \int_{\mathbb{R}} f_{(\xi,\eta)}(x,y)dy = \begin{cases} \int_0^2 \frac{15}{32} \cdot x^2y^4dy, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases} \tag{1.6}$$

Тут значення інтегралу $\int_0^2 \frac{15}{32} \cdot x^2y^4dy = \frac{15x^2}{32} \cdot \frac{y^5}{5}\Big|_{y=0}^{y=2} = 3x^2$. Для $f_\eta(y)$ маємо

$$f_\eta(y) = \int_{\mathbb{R}} f_{(\xi,\eta)}(x,y)dx = \begin{cases} \int_0^1 \frac{15}{32} \cdot x^2y^4dx, & y \in [0,2] \\ 0, & y \notin [0,2] \end{cases} \tag{1.7}$$

Тут інтеграл $\int_0^1 \frac{15}{32} \cdot x^2y^4dx = \frac{15y^4}{32} \cdot \frac{x^3}{3}\Big|_{x=0}^{x=1} = \frac{5y^4}{32}$. Отже, остаточно:

$$f_\xi(x) = 3x^2 \cdot \mathbb{1}_{[0,1]}(x),\ f_\eta(y) = \frac{5y^4}{32} \cdot \mathbb{1}_{[0,2]}(y) \tag{1.8}$$

Проте, помітимо, що

$$f_\xi(x)f_\eta(y) = \frac{15x^2y^4}{32} \cdot \mathbb{1}_{[0,1]}(x) \cdot \mathbb{1}_{[0,2]}(y) = f_{(\xi,\eta)}(x,y) \tag{1.9}$$

Отже, ξ, η є незалежними випадковими величинами. Тому, для всіх значень $y \in [0,2]$, маємо $f_{\xi|\eta}(x \mid y) \equiv f_\xi(x) = 3x^2\mathbb{1}_{[0,1]}(x)$, відповідно для $y \notin [0,2]$, $f_{\xi|\eta}(x \mid y) \equiv 0$.

Пункт б. Оскільки випадкові величини є незалежними, то і кореляція дорівнює 0, тобто $\mathrm{corr}[\xi,\eta] = 0$ (обернене, проте, не завжди справедливе). Тим не менш, про всяк випадок перевіримо це, показавши, що $\mathrm{cov}[\xi,\eta] = 0$. За означенням:

$$\mathrm{cov}[\xi,\eta] = \mathbb{E}[\xi\eta] - \mathbb{E}[\xi]\mathbb{E}[\eta] \tag{1.10}$$

Знайдемо математичні сподівання:

$$\mathbb{E}[\xi\eta] = \int_{\mathbb{R}\times\mathbb{R}} xyf_{(\xi,\eta)}(x,y)dxdy = \int_{[0,1]\times[0,2]} \frac{15}{32}x^3y^5dxdy = \frac{5}{4} \tag{1.11}$$

$$\mathbb{E}[\xi] = \int_{\mathbb{R}} xf_\xi(x)dx = \int_0^1 3x^3dx = \frac{3}{4} \tag{1.12}$$

$$\mathbb{E}[\eta] = \int_{\mathbb{R}} yf_\eta(y)dy = \int_0^2 \frac{5y^5}{32}dy = \frac{5}{3} \tag{1.13}$$

Отже, коваріація:

$$\mathrm{cov}[\xi,\eta] = \mathbb{E}[\xi\eta] - \mathbb{E}[\xi]\mathbb{E}[\eta] = \frac{5}{4} - \frac{3}{4}\cdot\frac{5}{3} = 0 \tag{1.14}$$

Звідси випливає, що $\mathrm{corr}[\xi,\eta] = 0$.

Відповідь. *Пункт а.* $\alpha = \frac{15}{32}$, $f_{\xi|\eta}(x \mid y) = 3x^2\mathbb{1}_{0,1}(x)$ для $y \in [0,2]$ і 0 в іншому випадку. *Пункт б.* Кореляція 0, оскільки ξ,η – незалежні.

Задача 2: Характеристична функція

Умова. За заданою характеристичною функцією

$$\varphi_\xi(t) = \frac{1}{3}\cos 3t + \frac{1}{3}\cos 2t + \frac{1}{3}e^{-4it}$$

відновити закон розподілу випадкової величини ξ.

Розв'язання. За видом характеристичної функції можна вгадати, що випадкова величина ξ є дискретною зі скінченною кількістю можливих значень. Тобто, розподіл має вигляд

$$\Pr[\xi = x_k] = p_k,\ k \in \{1,\dots,n\} \tag{2.1}$$

За означенням, характеристична функція:

$$\varphi_\xi(t) = \int_{-\infty}^{+\infty} e^{itx}dF_\xi(x) = \sum_{k=1}^{n} e^{itx_k}p_k \tag{2.2}$$

Дійсно, задана $\varphi_\xi(t)$ є лінійною комбінацією експонент виду e^{itx_k}. Знайдемо цей вигляд конкретно, скориставшись тим фактом, що $\cos z = \frac{e^{iz}+e^{-iz}}{2}$:

$$\varphi_\xi(t) = \frac{1}{6}e^{3it} + \frac{1}{6}e^{-3it} + \frac{1}{6}e^{2it} + \frac{1}{6}e^{-2it} + \frac{1}{3}e^{-4it} \tag{2.3}$$

Отже, наш розподіл має вигляд:

$$\boxed{\Pr[\xi = \pm 3] = \Pr[\xi = \pm 2] = \frac{1}{6},\ \Pr[\xi = -4] = \frac{1}{3}} \tag{2.4}$$

Задача 3: Закон великих чисел

Умова. Дана послідовність незалежних випадкових величин $\{\xi_n\}_{n\in\mathbb{N}}$. Випадкова величина ξ_n $(n \in \mathbb{N})$ приймає значення $\pm 1, 0$, причому

$$\Pr[\xi_n = \pm 1] = \frac{5n}{10n+2}.$$

Довести, що до цієї послідовності застосовний закон великих чисел.

Розв'язання. Знайдемо математичне сподівання та дисперсію кожної випадкової величини ξ_n $(n \in \mathbb{N})$. Випадкова величина є симетричною відносно 0, тому $\mathbb{E}[\xi_n] = 0$, а ось дисперсія

$$\mathrm{Var}[\xi_n] = 1^2 \cdot \frac{5n}{10n+2} + (-1)^2 \cdot \frac{5n}{10n+2} = \frac{10n}{10n+2} = \frac{1}{1+\frac{1}{5n}} \tag{3.1}$$

Тепер ми скористаємося **теоремою Чебишева**: якщо $\exists \rho > 0$ таке, що $\forall n \in \mathbb{N} : \mathrm{Var}[\xi_n] \leq \rho$, то закон великих чисел застосовний. Оскільки границя $\lim_{n\to\infty} \mathrm{Var}[\xi_n] = 1$, то послідовність $\{\mathrm{Var}[\xi_n]\}_{n\in\mathbb{N}}$ збіжна, а отже обмежена, тому таке ρ знайдеться[1]. Отже, за теоремою Чебишева, закон великих чисел застосовний.

Задача 4: Центральна гранична теорема

Умова. Випадкова величина ξ є середнім арифметичним незалежних однаково розподіленних випадкових величин, середнє квадратичне відхилення кожної з яких дорівнює $\sigma^2 = 3$. Скільки потрібно узяти таких випадкових величин, щоб з ймовірністю, не меншою за $p = 0.96$, випадкова величина ξ відхилялась за абсолютною величиною від свого математичесного сподівання не більш ніж на $\delta = 0.02$?

Розв'язання. Нехай ми маємо набір незалежних випадкових величин $\{\zeta_k\}_{k=1}^n$, а також за умовою $\xi = \frac{1}{n}\sum_{k=1}^n \zeta_k$. Ми знаємо, що $\mathrm{Var}[\zeta_k] = \sigma^2 = 3$, а також позначимо $\mu := \mathbb{E}[\zeta_k]$. За умовою, треба знайти таке мінімальне n, що

$$\Pr[|\xi - \mathbb{E}[\xi]| \leq \delta] \geq p, \;\; \delta = 0.02, \; p = 0.96 \tag{4.1}$$

Оскільки $\mathbb{E}[\xi] = \mathbb{E}[\frac{1}{n}\sum_{k=1}^n \zeta_k] = \frac{1}{n}\sum_{k=1}^n \mathbb{E}[\zeta_k] = \mu$, то задача еквівалентна знаходженню такого мінімального n, що

$$\Pr[|\xi - \mu| \leq \delta] \geq p, \;\; \delta = 0.02, \; p = 0.96 \tag{4.2}$$

[1] Насправді для цієї послідовності легко знайти ρ: $\mathrm{Var}[\xi_n] = \frac{1}{1+\frac{1}{5n}} < 1 =: \rho$.

Далі скористаємось практичним наслідком центральної граничної теореми. Будемо вважати, що n велике (в кінці розв'язку, коли ми отримаємо конкретне значення n, ми це перевіримо) і будемо розглядати випадкову величину $S_n := \sum_{k=1}^{n} \zeta_k$. Тоді, приблизно $S_n \sim \mathcal{N}(n\mu, n\sigma^2)$. Оскільки $\xi = \frac{S_n}{n}$, то за формулою перетворення нормального розподілу[2], $\xi \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. Звідси випливає, що випадкова величина $\eta := (\xi - \mu)\frac{\sqrt{n}}{\sigma} \sim \mathcal{N}(0, 1)$. Отже, тоді ми можемо порахувати шукану ймовірність наступним чином:

$$\Pr[|\xi - \mu| < \delta] = \Pr\left[-\delta \cdot \frac{\sqrt{n}}{\sigma} < \underbrace{(\xi - \mu)\frac{\sqrt{n}}{\sigma}}_{\eta} < \delta \cdot \frac{\sqrt{n}}{\sigma}\right] = 2\Phi_0\left(\frac{\delta\sqrt{n}}{\sigma}\right), \quad (4.3)$$

де $\Phi_0(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\tau^2/2} d\tau$ – функція Лапласа. Отже, наша задача звелась до того, що знайти мінімальне n, для якого

$$2\Phi_0\left(\frac{\delta\sqrt{n}}{\sigma}\right) \geq p \quad (4.4)$$

Через $\Phi_0^{-1}(x)$ позначимо обернену функцію Лапласа. Тоді

$$\frac{\delta\sqrt{n}}{\sigma} \geq \Phi_0^{-1}\left(\frac{p}{2}\right) \implies \boxed{n \geq \frac{\sigma^2}{\delta^2} \cdot \Phi_0^{-1}\left(\frac{p}{2}\right)^2} \quad (4.5)$$

Якщо порахувати, то $n \geq \frac{3}{0.02^2}\Phi_0^{-1}(0.48)^2 \approx \frac{3 \cdot 2.06^2}{0.0004} = 31827$. Отже, потрібно взяти близько **31830** випадкових величин.

Відповідь. Потрібно взяти близько 31830 випадкових величин.

[2] Якщо $\zeta \sim \mathcal{N}(\mu, \sigma^2)$, тоді $a\zeta + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.