Лабораторна робота № 1: побудова довірчих інтервалів

Нехай ω_1 та ω_2 – це незалежні рівномірно розподілені на $[0,1]$ випадкові величини (в.в.). Пара незалежних в.в. (ξ_1,ξ_2), які мають стандартний нормальний розподіл (тобто $N(0,1)$), генеруються за допомогою перетворення:

$$\xi_1 = \sqrt{-2\ln\omega_1}\,\sin(2\pi\,\omega_2), \qquad \xi_2 = \sqrt{-2\ln\omega_1}\,\cos(2\pi\,\omega_2)$$

(в.в. $N(0,1)$ можна генерувати і за допомогою вбудованого в комп’ютер генератора). Позначимо $a = \mathbf{M}\xi_i = 0,\ \sigma^2 = \mathbf{D}\xi_i = 1$.

Нехай спостерігається вибірка $\overline{X} = (X_1, \boxed{?}, X_n)$, де $X_i\ \boxed{?}N(0,1)$.

Всі довірчі інтервали будувати із достовірністю $1-\gamma = 0.99$. Цьому випадку відповідає $z_\gamma = 2.575$.

Завдання 1. Кожне з наступних трьох завдань виконувати для $n = 100$, $n = 10\,000$ та $n = 1\,000\,000$. В усіх трьох випадках дослідити, чи потрапляють математичне сподівання та дисперсія у побудовані довірчі інтервали, а також оцінити, як змінюється довжина інтервалу при збільшенні n.

A. Побудувати довірчий інтервал для математичного сподівання a у припущенні, що спостерігаються в.в. $\{X_i\}$, які мають нормальний розподіл, але дисперсія σ^2 невідома.
B. Побудувати довірчий інтервал для математичного сподівання a у припущенні, що спостерігаються в.в. $\{X_i\}$, розподіл яких невідомий.
C. Побудувати довірчий інтервал для дисперсії σ^2 у припущенні, що спостерігаються в.в. $\{X_i\}$, які мають нормальний розподіл.

Завдання 2: обчислення інтегралу чотирьома способами із дослідженням швидкості збіжності. Потрібно обчислити наступний інтеграл:

$$Q = \mathbf{P}\{\xi < \eta\},$$

де ξ та η – в.в., які мають функції розподілу (ф.р.) $F(x)$ та $G(x)$ відповідно. Припустимо, що $\xi \geq 0$ та $\eta \geq 0$ з ймовірністю 1 і існують щільності $f(u) = F'(u),\ u \geq 0,$ та $g(u) = G'(u),\ u \geq 0$. Виберемо наступні ф.р.:

$$F(u) \ = \ 1 - e^{-(\alpha u)^4}, \quad G(u) \ = \ 1 - e^{-u^2}, \quad u \geq 0 .$$

Тоді ймовірність Q обчислюється за формулою:

$$Q = \mathbf{P}\{\xi < \eta\} = \int_0^\infty F(u)\, dG(u) = \int_0^\infty \left(1 - e^{-(\alpha u)^4}\right) 2u\, e^{-u^2}\, du \ \underset{\alpha \to 0}{\boxed{?}}\ 2\alpha^4 \int_0^\infty u^5 e^{-u^2}\, du \ =$$

$$\left\| v = u^2,\ dv = 2u\, du \right\| \ = \ 2\alpha^4 \int_0^\infty u^5 e^{-u^2}\, du \ = \ \alpha^4 \int_0^\infty v^2 e^{-v}\, dv \ = \alpha^4\, \Gamma(3) = 2\alpha^4 . \qquad (1)$$

Зауваження 1. Нехай $\omega_1, \omega_2, \boxed{?}$ – послідовність незалежних однаково розподілених на відрізку $[0, 1]$ в.в. (послідовність псевдовипадкових чисел). Тоді $\xi = F^{-1}(\omega)$ і $\eta = G^{-1}(\omega)$. Тобто

$$\xi_i = \frac{1}{\alpha}\left(-\ln \omega_i\right)^{\frac{1}{4}}, \quad \eta_i = \left(-\ln \omega_i\right)^{\frac{1}{2}}, \quad i = 1, 2, \boxed{?} .$$

Зауваження 2. Загальна схема обчислення ймовірності Q виглядає наступним чином. Нехай $\hat{q}_1, \hat{q}_2, \boxed{?}$ – незміщені оцінки ймовірності Q. Позначимо

$$\hat{Q}_n = \frac{1}{n} \sum_{i=1}^{n} \hat{q}_i, \quad \hat{\sigma}_n^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} \hat{q}_i^2 - n\, \hat{Q}_n^2 \right).$$

Кількість реалізацій n^* алгоритму, які потрібно здійснити для обчислення ймовірності Q із заданою *достовірністю* $1 - \gamma$ та *відносною похибкою* ε обчислюється за формулою:

$$n^* = \min \left\{ n \geq n_0 :\ n \geq \frac{z_\gamma^2\, \hat{\sigma}_n^2}{\varepsilon^2\, \hat{Q}_n^2} \right\},$$

де n_0 – початкова кількість реалізацій, яка потрібна для “стабілізації” дисперсії, а z_γ – це коефіцієнт, який знаходиться з рівняння $2\Phi(z) = 1 - \gamma$ ($\Phi(z)$ – це функція Лапласа).

В усіх наведених вище випадках обчислення вести із достовірністю 0.99 та відносною похибкою 1%, тобто $z_\gamma = 2.575$ і $\varepsilon = 0.01$. Розглядаються три можливі значення параметра α : 1, 0.1 та 0.01. Потрібно виконати наступні завдання.

A. При кожному $\alpha = 1;\ 0.1;\ 0.01$ обчислити точне значення ймовірності $Q = Q(\alpha)$.

B. Стандартний метод Монте-Карло (*метод 1*):

$Q(\alpha) = \mathbf{M}\, I(\xi < \eta)$, тобто $\hat{q}_i = I(\xi_i < \eta_i)$, де $I(\cdot)$ – індикаторна функція.

C. Метод 2: $Q(\alpha) = \int_0^\infty [1 - G(u)]\, dF(u) = \mathbf{M}[1 - G(\xi)]$, тобто $\hat{q}_i = 1 - G(\xi_i) = e^{-\xi_i^2}$.

D. Метод 3: $Q(\alpha) = \int_0^\infty F(u)\, dG(u) = \mathbf{M}\, F(\eta)$, тобто $\hat{q}_i = F(\eta_i) = 1 - e^{-(\alpha\eta_i)^4}$.

E. Метод 4:

$$Q(\alpha) = \int_0^\infty F(u)\, dG(u) = \int_0^\infty F(u)\, g(u)\, du = \int_0^\infty F(u)\, \frac{g(u)}{h(u)}\, h(u)\, du = \mathbf{M}\left[F(\beta)\, \frac{g(\beta)}{h(\beta)}\right],$$

де β – невід’ємна в.в. із щільністю $h(u),\ u \geq 0$. Використовуючи співвідношення (1), маємо

$$F(u)\, g(u) = \left(1 - e^{-(\alpha u)^4}\right) 2u\, e^{-u^2} \underset{\alpha \to 0}{\boxed{?}} 2\alpha^4 u^5\, e^{-u^2}.$$

У співвідношенні (1) було показано, що $\int_0^\infty u^5\, e^{-u^2}\, du = 1$. Тому як щільність $h(u),\ u \geq 0$, раціонально вибрати $h(u) = u^5\, e^{-u^2},\ u \geq 0$. Легко показати, що саме таку щільність має в.в. $\beta = \sqrt{\theta^{(1)} + \theta^{(2)} + \theta^{(3)}}$, де $\theta^{(1)}$, $\theta^{(2)}$ та $\theta^{(3)}$ – незалежні в.в., які мають показниковий розподіл з параметром 1, тобто $\theta^{(i)} = -\ln\omega_i$. Звідси випливає, що

$$\hat{q}_i = \left[1 - e^{-(\alpha\beta_i)^4}\right] \frac{2\beta_i\, e^{-\beta_i^2}}{\beta_i^5\, e^{-\beta_i^2}} = \frac{2}{\beta_i^4}\left[1 - e^{-(\alpha\beta_i)^4}\right],$$

де $\beta = \sqrt{\theta_i^{(1)} + \theta_i^{(2)} + \theta_i^{(3)}}$, причому $\theta_i^{(1)} = -\ln\omega_{3i-2},\ \theta_i^{(2)} = -\ln\omega_{3i-1},\ \theta_i^{(3)} = -\ln\omega_{3i}$, $i = 1, 2, \boxed{?}$.