

KOVARIANCE A KORELACE

14.11.2024

1. V šuplíku je 6 ponožek: 2 bílé, 2 černé a 2 oranžové. Potmě náhodně vytáhneme z šuplíku 3 ponožky. Označme jako X počet vytažených bílých ponožek a Y je počet vytažených černých ponožek.

- (a) Napište tabulku rozdělení náhodného vektoru $(X, Y)^\top$.
- (b) Jsou X a Y nezávislé?
- (c) Spočítejte kovarianci a korelaci X a Y .
- (d) Spočítejte rozdělení, střední hodnotu a rozptyl náhodné veličiny $U = X - Y$.
- (e) Nechť Z značí počet černých ponožek, které zůstaly v šuplíku. Určete pravděpodobnostní rozdělení $(X, Z)^\top$ a kovarianci X a Z .

2. Náhodný vektor $(X, Y)^\top$ má spojitě rozdělení charakterizované sdruženou hustotou

$$f(x, y) = \begin{cases} cxy & \text{pro } 0 < x < 1, 0 < y < 1, \\ 0 & \text{jinak.} \end{cases}$$

- (a) Určete $c > 0$ tak, aby f byla hustota.
 - (b) Spočítejte kovarianci $\text{Cov}(X, Y)$.
 - (c) Rozhodněte, zda jsou X a Y nezávislé,
3. Náhodná veličina X má rovnoměrné rozdělení na intervalu $(-1, 1)$. Označme $Y = X^2$. Spočítejte kovarianci veličin X a Y a jejich korelační koeficient ρ_{XY} . Jsou X a Y nezávislé?
4. Náhodný vektor $(X, Y)^\top$ má rovnoměrné rozdělení na množině $M = \{(x, y) : 0 \leq y \leq 1 - x, 0 \leq x \leq 1\}$, tj. spojitě rozdělení s hustotou

$$f(x, y) = \begin{cases} c & \text{pro } 0 \leq y \leq 1 - x, 0 \leq x \leq 1, \\ 0 & \text{jinak,} \end{cases}$$

- (a) Určete konstantu $c > 0$, aby f byla hustota.
 - (b) Rozhodněte, zda jsou X a Y nezávislé.
 - (c) Spočítejte kovarianci a korelaci X a Y .
5. Řešte předchozí příklad pro množinu $M = \{(x, y) : x - 1 \leq y \leq 1 - x, 0 \leq x \leq 1\}$.
6. Uvažujme posloupnost n znaků, ve které se každý znak zkreslí s pravděpodobností $p \in (0, 1)$, a to nezávisle na ostatních znacích. Nechť X značí počet takto zkreslených znaků. Dále označme jako Y_i identifikátor toho, zda se i -tý znak zkreslí. Vyjádřete X pomocí veličin $Y_1, \dots, Y_n$ a pomocí tohoto vyjádření spočítejte $\text{E}X$ a $\text{Var} X$.

OPAKOVÁNÍ

KOVARIANCE A KORELACE: Kovariance $\text{Cov}(X, Y)$ náhodných veličin X, Y je definována jako

$$\text{Cov}(X, Y) = E(X - EX)(Y - EY) = E(XY) - (EX)(EY),$$

je-li $EX^2 < \infty$, $EY^2 < \infty$.

Korelace $\text{Corr}(X, Y)$ je definována jako

$$\rho_{X,Y} = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var } X} \sqrt{\text{Var } Y}},$$

je-li $\text{Var } X, \text{Var } Y > 0$. Platí vždy $-1 \leq \text{Corr}(X, Y) \leq 1$. Korelace je mírou lineární závislosti mezi X a Y .

UŽITEČNÉ VLASTNOSTI: Nechť X, Y jsou náhodné veličiny a $a, b, c, d \in \mathbb{R}$. Pak (za předpokladu existence daných momentů)

- $\text{Cov}(aX + c, bY + d) = ab\text{Cov}(X, Y)$,
- $E(aX + bY) = aEX + bEY$,
- $\text{Var}(aX + bY) = a^2\text{Var } X + b^2\text{Var } Y + 2ab\text{Cov}(X, Y)$.

Obecně, jsou-li $X_1, \dots, X_n$ jsou náhodné veličiny a $a_1, \dots, a_n \in \mathbb{R}$, pak platí (za předpokladu existence daných momentů)

- $E(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i EX_i$
- $\text{Var}(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^2 \text{Var } X_i + \sum \sum_{i \neq j} a_i a_j \text{Cov}(X_i, X_j)$