

Zadanie 1. Wykaż, że dla dowolnych zmiennych losowych X i Y zachodzą następujące równości:

- (a) $\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}Y$ (wewnętrzna wartość oczekiwana jest po Y , zewnętrzna po X),
 (b) $\text{var}(X) = \mathbb{E}(\text{var}(X|Y)) + \text{var}(\mathbb{E}(X|Y))$.

Zadanie 2. Wykaż, że dla zmiennej X o rozkładzie $\text{Beta}(\alpha, \beta)$:

$$\mathbb{E}(X) = \frac{\alpha}{\alpha + \beta}, \quad \mathbb{M}(X) = \frac{\alpha - 1}{\alpha + \beta - 2}, \quad \text{var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)},$$

gdzie $\mathbb{M}(X)$ oznacza dominantę (modę) zmiennej X .

Zadanie 3. Wykaż, że jeśli zmienne losowe X i Y są niezależne, to ich kowariancja wynosi 0. Wskaż przykład, kiedy implikacja w drugą stronę nie zachodzi.

Zadanie 4. Wykaż, że dla dwóch zmiennych losowych X i Y zachodzi

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y).$$

Zadanie 5. Niech Σ będzie macierzą kowariancji dla pewnego wektora n zmiennych losowych. Wykaż, że macierz Σ jest symetryczna i dodatnio półokreślona, czyli

$$\Sigma = \Sigma^T \quad \text{oraz} \quad \vec{x}^T \Sigma \vec{x} \geq 0 \quad \text{dla} \quad \vec{x} \in \mathbb{R}^n.$$

Zadanie 6. Podaj przykład zmiennej losowej rzeczywistej X i funkcji wypukłej $f: \mathbb{R} \rightarrow \mathbb{R}$ takich, że $\mathbb{E}X < \infty$ oraz $\mathbb{E}(f(X)) = \infty$.

Zadanie 7. Wykaż, że dla rzeczywistej zmiennej losowej X o skończonej wartości oczekiwanej oraz różniczkowalnej wypukłej funkcji $f: X \rightarrow \mathbb{R}$ takiej, że $\mathbb{E}(f(X)) < \infty$ zachodzi nierówność *Jensena*:

$$\mathbb{E}(f(X)) \geq f(\mathbb{E}(X)).$$

Przedstaw na rysunku nierówność Jensena dla wybranej dyskretnej zmiennej losowej X i wybranej funkcji wypukłej g .

Zadanie 8. Podaj interpretację geometryczną zbiorów $\{\vec{x} \in \mathbb{R}^2: \mathcal{N}(\vec{x}|\vec{\mu}, \Sigma^2) = c\}$ dla $c \in \mathbb{R}_+$ i diagonalnej macierzy kowariancji Σ .

Zadanie 9. Rozważmy próbkę $\{x^{(1)}, \dots, x^{(m)}\}$ z rozkładu $\mathcal{N}(\mu, \sigma^2)$. Metodą największej wiarygodności wyznacz estymatory $\hat{\mu}_{MLE}$ i $\hat{\sigma}_{MLE}^2$ parametrów μ i σ^2 .

Zadanie 10. Danych jest m niezależnych zmiennych $X_1, \dots, X_m$ o tym samym rozkładzie o średniej μ i wariancji σ^2 . Sprawdź, który z poniższych estymatorów jest obciążony:

$$\hat{\mu} = \frac{1}{m} \sum_{i=1}^m X_i, \quad \hat{\sigma}^2 = \frac{1}{m} \sum_{i=1}^m (X_i - \hat{\mu})^2.$$

Zadanie 11. Rozważmy próbkę $\{x^{(1)}, \dots, x^{(m)}\}$ z rozkładu $\mathcal{N}(\mu, \sigma^2)$. Niech μ ma rozkład aprioryczny $\mathcal{N}(\mu_0, \sigma_0^2)$. Metodą maksymalnego a posteriori wyznacz estymator $\hat{\mu}_{MAP}$ parametru μ .

Zadanie 12. Chcemy dokonać zakupu na znanym portalu aukcyjnym. Jeden ze sprzedawców ma 90 ocen pozytywnych i 10 negatywnych, natomiast drugi 2 oceny pozytywne i żadnej negatywnej. Przeprowadź analizę bayesowską i na tej podstawie zdecyduj, u którego sprzedawcy dokonasz zakupu.

Wskazówka. Na podstawie ocen wyznacz prawdopodobieństwo, że jeden ze sprzedawców jest bardziej wiarygodny niż drugi. W celu wyznaczenia wartości tego prawdopodobieństwa

skorzystaj z komputera. Jako rozkład a priori przyjmij rozkład jednostajny na przedziale $[0, 1]$ (równoważnie: rozkład $\text{Beta}(1, 1)$).

Zadanie 13. Rozważmy zmienną losową X o dwumianowym rozkładzie z parametrem μ , dla którego rozkład a priori jest rozkładem beta z parametrami α i β . Zaobserwowaliśmy m zdarzeń $X = 1$ oraz ℓ zdarzeń $X = 0$. Wykaż, że wartość oczekiwana a posteriori wynosi

$$\lambda \mathbb{E}(\mu) + (1 - \lambda) \hat{\mu}_{MLE}$$

dla pewnego $\lambda \in [0, 1]$. Wyznacz wartość λ (w zależności od α , β , m oraz ℓ).

WARTOŚĆ OCZEKIWANA, WARIANCJA, KOWARIANCJA

Wartość oczekiwana rzeczywistej zmiennej losowej X :

(zmienna dyskretna)
$$\mathbb{E}X = \sum_x xP(X = x)$$

(zmienna ciągła)
$$\mathbb{E}X = \int_x xp(x)dx$$

Wariancja zmiennej X :

$$\text{var}X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X^2) - (\mathbb{E}X)^2$$

Kowariancja określa zależność liniową między zmiennymi X i Y i zadana jest wzorem

$$\text{cov}(X, Y) = \mathbb{E}_{X,Y}((X - \mathbb{E}X)(Y - \mathbb{E}Y)) = \mathbb{E}_{X,Y}(XY) - \mathbb{E}X \cdot \mathbb{E}Y.$$

Dla wektorowych zmiennych $\vec{X}$ i $\vec{Y}$ kowariancja jest macierzą

$$\text{cov}(\vec{X}, \vec{Y}) = \mathbb{E}_{\vec{X}, \vec{Y}}(\vec{X} - \mathbb{E}\vec{X})(\vec{Y}^T - \mathbb{E}\vec{Y}^T) = \mathbb{E}_{\vec{X}, \vec{Y}}(\vec{X}\vec{Y}^T) - \mathbb{E}\vec{X} \cdot \mathbb{E}\vec{Y}^T.$$

Kiedy rozważamy kowariancję wektora $\vec{X}$ z samym sobą, to piszemy $\text{cov}\vec{X} = \text{cov}(\vec{X}, \vec{X})$.

ROZKŁAD NORMALNY

Dla zmiennej rzeczywistej x funkcja gęstości prawdopodobieństwa rozkładu normalnego o średniej μ i wariancji σ^2 jest zadana wzorem

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right).$$

Funkcja gęstości prawdopodobieństwa D -wymiarowego rozkładu normalnego wektora losowego $\vec{x}$ o wektorze wartości oczekiwanych $\boldsymbol{\mu} = [\mu_1, \dots, \mu_D]^T$ i macierzy kowariancji Σ dana jest wzorem

$$\mathcal{N}(\vec{x}|\vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^T \Sigma^{-1}(\vec{x} - \vec{\mu})\right).$$

ROZKŁAD BETA

Rozkład beta to ciągły rozkład prawdopodobieństwa dany funkcją gęstości zdefiniowaną na przedziale $[0, 1]$ wzorem

$$\text{Beta}(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1},$$

gdzie $\alpha, \beta > 0$ są parametrami rozkładu, a Γ to funkcja gamma Eulera:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt.$$

Funkcja gamma rozszerza pojęcie silni na zbiór liczb zespolonych o dodatniej części rzeczywistej. W szczególności

$$\Gamma(z+1) = z\Gamma(z) \text{ dla } z \in \mathbb{C} \text{ t.j. } \text{Re}(z) > 0 \quad \text{oraz} \quad \Gamma(n+1) = n! \text{ dla } n \in \mathbb{N}.$$

METODA NAJWIĘKSZEJ WIARYGODNOŚCI

Funkcja wiarygodności względem parametru θ dla obserwacji D zadana jest wzorem

$$L(\theta) = p(D|\theta).$$

Estymator największej wiarygodności $\hat{\theta}_{MLE}$ parametru θ to argument maksymalizujący wartość funkcji wiarygodności $L(\theta)$:

$$\hat{\theta}_{MLE} := \arg \max_{\theta \in \Theta} L(\theta) = \arg \max_{\theta \in \Theta} \log L(\theta).$$

METODA MAKSYMALNEGO A POSTERIORI

Metodą maksymalnego a posteriori maksymalizujemy prawdopodobieństwo a posteriori $p(\theta|D)$ uwzględniające nasze wcześniejsze przekonania oraz obserwacje D . *Estymator maksymalnego a posteriori* $\hat{\theta}_{MAP}$ parametru θ to wartość, który maksymalizuje wartość funkcji $p(\theta|D)$:

$$\hat{\theta}_{MAP} := \arg \max_{\theta \in \Theta} p(\theta|D) = \arg \max_{\theta \in \Theta} \frac{p(D|\theta)p(\theta)}{p(D)} = \arg \max_{\theta \in \Theta} p(D|\theta)p(\theta).$$

WNISKOWANIE BAYESOWSKIE

Wnioskowanie bayesowskie polega na wyznaczeniu pełnego rozkładu a posteriori $p(\theta|D)$ uwzględniającego nasze wcześniejsze przekonania oraz obserwacje D :

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}.$$

Warto korzystać z faktu, że $p(D) = \int_{\theta \in \Theta} p(D|\theta)p(\theta)d\theta$.