Wstęp do statystyki matematycznej
Lista 4

1. Udowodnić, że rodzina rozkładów normalnych z parametrem położenia (wariancja $\sigma^2 = 1$) jest rodziną zupełną.

2. Udowodnić, że rodzina $\mathcal{P}$ rozkładów jednostajnych (dyskretnych) na zbiorach $\{1, 2, ..., k\}$, gdzie $k \geqslant 1$ jest parametrem, jest zupełna. Pokazać, że każda właściwa podrodzina $\mathcal{P}$ nie jest zupełna.

3. Sformułować definicję rodziny rozkładów ograniczenie zupełnej.
Niech $\mathcal{P} = \{p_\theta : \theta \in (0,1)\}$ będzie rodziną gęstości (względem miary liczącej) na zbiorze $\mathcal{X} = \{-1, 0, 1, 2, ...\}$, gdzie $p_\theta(-1) = \theta, \quad p_\theta(x) = (1-\theta)^2\theta^x$, dla $x = 0, 1, 2, ...$
Udowodnić, że rodzina $\mathcal{P}$ nie jest zupełna, ale jest ograniczenie zupełna.

---

4. Niech $X = (X_1, ..., X_n)$ będzie próbą z rozkładu Poissona o średniej $\lambda > 0$. Wyznaczyć rozkład warunkowy próby przy warunku $T = t$, gdzie $T(X) = \sum_{i=1}^n X_i$. Za pomocą definicji wykazać, że T jest statystyką dostateczną dla λ (w rodzinie rozkładów produktowych).

5. Losujemy bez zwracania n jednostek z partii N wyrobów, spośród których $N\theta$ jest wadliwych. Niech X_i, $i = 1, ..., n$, przyjmuje wartość 1, gdy i-ta jednostka jest wadliwa i wartość 0, gdy i-ta jednostka jest dobra. Pokazać, że statystyka $T = \sum_{i=1}^n X_i$ jest dostateczna dla parametru θ.

6. Pokazać, że jeśli T jest statystyką dostateczną dla $\mathcal{P}$ oraz $T = g(S)$ dla pewnej statystyki S i odwzorowania mierzalnego g, to S jest statystyką dostateczną dla $\mathcal{P}$.

7. Pokazać, że jeśli T jest statystyką dostateczną dla $\mathcal{P}$, to jest statystyką dostateczną dla uwypuklenia $\mathcal{P}$ tzn. rodziny wszystkich kombinacji wypukłych rozkładów z $\mathcal{P}$.

8. Niech X będzie czasem czekania na k-ty sukces w schemacie Bernoulliego z prawdopodobieństwem sukcesu $\theta \in (0,1)$, gdzie $k \geqslant 1$ ustalone (znane). Pokazać, że rozkłady X tworzą jednoparametrową rodzinę wykładniczą. Niech $X_1, ..., X_n$ będzie próbą z rozkładu z tej rodziny. Korzystając z twierdzenia o rodzinach wykładniczych, podać statystykę dostateczną dla parametru θ i wyznaczyć jej rozkład.

9. Niech $h(x)$ będzie dodatnią funkcją całkowalną na $(-\infty, \infty)$ i niech
$$p_{\theta\eta}(x) = c(\theta, \eta)h(x)\mathbf{1}_{(\theta,\eta)}(x), \ -\infty < \theta < \eta < \infty,$$
będzie rodziną gęstości. Niech $X_1, ..., X_n$ będzie próbą o rozkładzie z tej rodziny. Udowodnić, że $(X_{(1)}, X_{(n)})$ jest statystyką dostateczną dla parametru (θ, η).

10. Niech $X = (X_1, ..., X_n)$ będzie próbą z rozkładu należącego do rodziny:
    (a) Poissona $\mathcal{P}(\lambda)$;
    (b) dwumianowej ujemnej $\mathcal{NB}(r_0, p)$ z parametrem p (r_0 ustalone i znane);
    (c) normalnej $\mathcal{N}(\mu, \sigma^2)$;
    (d) gamma $\mathcal{G}(\alpha, \beta)$;
    (e) jednostajnej $\mathcal{U}(a, b)$;
    (f) Pareto $\mathcal{P}a(\theta, \alpha)$ o gęstości $p_{\theta\alpha}(x) = \alpha\theta^\alpha x^{-\alpha-1}\mathbf{1}_{(\theta,\infty)}(x), \ \theta, \alpha > 0$;
    (g) wykładniczej o gęstości $p_{\theta\lambda}(x) = \lambda^{-1}e^{-(x-\theta)/\lambda}\mathbf{1}_{(\theta,\infty)}(x), \ \theta \in R, \lambda > 0$.

    Korzystając z kryterium faktoryzacji, wyznaczyć statystykę dostateczną dla parametru tej rodziny.

Wstęp do statystyki matematycznej
Lista 5

1. Niech $X_1, ..., X_n$ będzie próbą z rozkładu beta z parametrami (α, α), $\alpha \in (0, \infty)$. Wyznaczyć minimalną statystykę dostateczną dla parametru α. Uzasadnić, że jest to statystyka zupełna.

2. Niech $X_1, ..., X_n$ będzie próbą z rozkładu normalnego $\mathcal{N}(\sigma, \sigma^2)$, $\sigma > 0$. Wyznaczyć minimalną statystykę dostateczną dla parametru σ. Pokazać, że nie jest ona zupełna.

3. Niech $X = (X_1, ..., X_n)$ będzie próbą z rozkładu:

   (a) jednostajnego na odcinku $[0, \theta]$, $\theta \in (0, \infty)$. Pokazać, że $T(X) = X_{(n)}$ jest minimalną, zupełną statystyką dostateczną.

   (b) jednostajnego na odcinku $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$, $\theta \in R$. Udowodnić, że $(X_{(1)}, X_{(n)})$ jest minimalną statystyką dostateczną dla parametru θ, ale nie jest statystyką zupełną.

4. Rozważmy rodzinę rozkładów wykładniczych $\mathcal{E}(\theta, \lambda)$ (por. 10 (g) z listy 4). Niech $X_1, ..., X_n$ będzie próbą z tego rozkładu. Wykazać, że statystyka $(X_{(1)}, \sum_i (X_i - X_{(1)}))$ jest minimalną statystyką dostateczną.

5. Wykazać, że w rodzinie rozkładów logistycznych z parametrem położenia o gęstości

   $$p_m(x) = \frac{e^{-(x-m)}}{[1 + e^{-(x-m)}]^2}, \; m \in R,$$

   wektor statystyk pozycyjnych jest minimalną statystyką dostateczną.

6. Niech $\mathcal{P}$ będzie rodziną wszystkich rozkładów absolutnie ciągłych (lub tylko bezatomowych) na prostej. Niech $X_1, ..., X_n$ będzie próbą z rozkładu z tej rodziny. Wykazać, że wektor statystyk pozycyjnych (a więc także dystrybuanta empiryczna) jest statystyką dostateczną i zupełną, a więc również minimalną dostateczną.

7. Niech $(X_1, Y_1), ..., (X_n, Y_n)$ będzie próbą z rozkładu dwuwymiarowego normalnego o średnich 0, wariancjach 1 i kowariancji $\theta \in (-1, 1)$. Znaleźć minimalną statystykę dostateczną dla parametru θ i zbadać, czy jest zupełna. Pokazać, że statystyki $T_1 = \sum_i X_i^2$ i $T_2 = \sum_i Y_i^2$ są swobodne.

8. Niech $X_1, ..., X_n$ będzie próbą z rozkładu gamma $\mathcal{G}(\alpha, \beta)$. Korzystając z twierdzenia Basu, wykazać, że statystyki $\sum_i X_i$ oraz $\sum_i (\log X_i - \log X_{(1)})$ są niezależne.

9. Niech X i Y będą niezależnymi próbami rozmiarów n każda z rozkładów normalnych odpowiednio $\mathcal{N}(\mu_X, \sigma_X^2)$ i $\mathcal{N}(\mu_Y, \sigma_Y^2)$.

   (a) Udowodnić, że statystyka $T = (\overline{X}, \sum_{i=1}^{n} X_i^2, \overline{Y}, \sum_{i=1}^{n} Y_i^2)$ jest zupełną statystyką dostateczną dla parametru $(\mu_X, \sigma_X^2, \mu_Y, \sigma_Y^2)$.

   (b) Wykazać, że statystyki

   $$V = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\left[\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i=1}^{n} (Y_i - \overline{Y})^2\right]^{1/2}}$$

   i T są niezależne.

Wstęp do statystyki matematycznej
Lista 6

1. Niech $X_1, ..., X_n$ będzie próbą z rozkładu wykładniczego z parametrem $\lambda \in (0, \infty)$ (tj. $p_\lambda(x) = \frac{1}{\lambda} e^{-x/\lambda} \mathbf{1}_{(0,\infty)}(x)$). Wykazać, że nie istnieje estymator nieobciążony funkcji parametrycznej $g(\lambda) = e^\lambda$.

2. (a) Pokazać, że estymator $\overline{X}$ parametru μ w rodzinie $\mathcal{N}(\mu, 1)$ jest lepszy niż estymator $T = \frac{1}{2}(X_{(1)} + X_{(n)})$ tzn. T jest niedopuszczalny;

   (b) Pokazać, że w rodzinie rozkładów jednostajnych na odcinku $[\mu - \frac{1}{2}, \mu + \frac{1}{2}]$ jest na odwrót tzn. $\overline{X}$ jest niedopuszczalny.

3. Niech $X_1, ..., X_n$ będzie próbą z rozkładu o dystrybuancie F i dystrybuancie empirycznej F_n. Udowodnić, że dla każdego $x \in R$, $F_n(x)$ jest nieobciążonym estymatorem $F(x)$.

4. Niech $X_1, ..., X_n$ będzie próbą z rozkładu o wartości oczekiwanej 0, wariancji $\sigma^2 = EX_1^2$ i skończonym czwartym momencie $\mu_4 = EX_1^4$. Wykazać, że
$$\operatorname{Var} S^2 = \frac{(n-1)^2}{n^3}\mu_4 - \frac{(n-1)(n-3)}{n^3}\sigma^4,$$
gdzie S^2 jest wariancją z próby. **Wsk.** Najpierw sprawdzić, że $\operatorname{Var}(\sum_i X_i)^2 = n\mu_4 + n(2n-3)\sigma^4$.

5. Niech $X_1, ..., X_n$ będzie próbą z rozkładu jednostajnego (dyskretnego!) na zbiorze $\{\theta - 1, \theta, \theta + 1\}$, gdzie $\theta \in Z$. Wykazać, że dla żadnej funkcji niestałej $g(\theta)$ nie istnieje estymator UMVU parametru $g(\theta)$, chociaż estymatory nieobciążone istnieją.

6. Niech $X = (X_1, ..., X_n)$ będzie próbą z rozkładu jednostajnego na odcinku $[0,\ \theta]$, $\theta > 0$. Wykazać, że w klasie wszystkich estymatorów estymator
$$T(X) = \frac{n+1}{n} X_{(n)}$$
jest niedopuszczalny, chociaż jest estymatorem UMVU. **Wsk.** Rozważyć estymator $T_1(X) = \frac{n+2}{n+1} X_{(n)}$.

7. Niech $X_1, ..., X_n$ będzie próbą z rozkładu dyskretnego
$$p_\theta(-1) = 2\theta(1-\theta),\ \ p_\theta(k) = \theta^k(1-\theta)^{3-k},\ \ k = 0, 1, 2, 3,\ \theta \in (0, 1).$$
   (a) Sprawdzić, czy istnieje estymator UMVU parametru θ.

   (b) Sprawdzić, czy istnieje estymator UMVU funkcji parametrycznej $g(\theta) = \theta(1-\theta)$.

8. Niech $X_1, ..., X_n$ będzie próbą z rozkładu Pareto z parametrami (θ, α). Wyznaczyć estymator UMVU (a) parametru (θ, α); (b) parametru α, gdy $\theta = 1$; (c) parametru θ, gdy α ustalone (znane).

9. Niech $X_1, ..., X_m, Y_1, ..., Y_n$ będą niezależnymi próbami z rozkładów odpowiednio $\mathcal{U}(0, \theta)$ i $\mathcal{U}(0, \vartheta)$. Wyznaczyć estymator UMVU ilorazu $g(\theta, \vartheta) = \theta/\vartheta$, gdy $n > 1$.

10. Niech $X_1, ..., X_n$ będzie próbą z rozkładu $\mathcal{N}(\mu, 1)$. Wykazać, że $\Phi\left(\sqrt{\frac{n}{n-1}}(c - \overline{X})\right)$ jest estymatorem UMVU funkcji parametrycznej $g_c(\mu) = P_\mu(X_1 \leqslant c) = \Phi(c - \mu)$ dla dowolnego ustalonego c. Zbadać asymptotyczne zachowanie się tego estymatora (zgodność, $\sqrt{n}$-zgodność, asymptotyczna normalność).