PRZYKŁADOWE PYTANIA NA EGZAMIN

1. Podać przykłady zmiennych losowych

   (a) dyskretnych,
   (b) ciągłych.

   Niech X będzie zmienną losową ciągłą i $a \in \mathbb{R}$. Ile wynosi $P(X = a)$?

2. Podać interpretację następujących rozkładów:

   - rozkładu dwumianowego,
   - rozkładu geometrycznego,
   - rozkładu ujemnego dwumianowego.

   Podać wzory na $P(X = k)$, gdzie X to zmienna losowa o rozkładzie dwumianowym, geometrycznym i ujemnym dwumianowym.

3. Zmienna losowa X ma rozkład dany tabelką:

x_k	-1	0	1
p_k	$\frac{1}{10}$	C	$\frac{1}{10}$

   .

   Wyznaczyć C. Obliczyć wartość oczekiwaną, wariancję i odchylenie standardowe X.

4. Podać parę własności rozkładu normalnego. Jak wygląda wykres jego gęstości?

5. Stosując metodę momentów, wyznaczyć estymator parametru θ na podstawie próby losowej $X_1, X_2, \ldots, X_n$

   (a) z rozkładu trzypunktowego takiego, że $P(X = -1) = \frac{\theta}{4}$, $P(X = 0) = 1 - \theta$ oraz $P(X = 2) = \frac{3\theta}{4}$, gdzie $\theta \in (0, 1)$;
   (b) z rozkładu jednostajnego na odcinku $(0, \theta)$, gdzie $\theta > 0$, czyli z rozkładu o gęstości $f_\theta(x) = \begin{cases} \frac{1}{\theta} & \text{dla } x \in (0,\theta) \\ 0 & \text{dla } x \notin (0,\theta) \end{cases}$.

6. Niech $X_1, X_2, \ldots, X_n$ będzie próbą losową z rozkładu Poissona z parametrem $\theta > 0$, tzn. z rozkładu dyskretnego o funkcji masy prawdopodobieństwa $P(X = x) = \mathrm{e}^{-\theta}\frac{\theta^x}{x!}$, $x = 0, 1, 2, \ldots$. Wyznaczyć estymator parametru θ, stosując metodę największej wiarygodności.

7. Niech $X_1, X_2, \ldots, X_n$ będzie prostą próbą losową z rozkładu Pareto o gęstości danej wzorem $f_a(x) = \begin{cases} \frac{a}{x^{a+1}} & \text{dla } x \geqslant 1 \\ 0 & \text{dla } x < 1 \end{cases}$. Wyznaczyć estymator parametru a, stosując metodę największej wiarygodności.

8. Podać definicję

   (a) błędu średniokwadratowego estymatora;
   (b) estymatora nieobciążonego;
   (c) obciążenia estymatora;
   (d) estymatora nieobciążonego o minimalnej wariancji;
   (e) asymptotycznie nieobciążonego ciągu estymatorów;
   (f) zgodnego, mocno zgodnego i zgodnego w sensie zbieżności średniokwadratowej ciągu estymatorów.

9. Niech $X_1, X_2, \ldots, X_n$ będzie próbą losową z rozkładu jednostajnego na odcinku $(0, \theta)$, gdzie $\theta > 0$. Mamy dwa estymatory parametru θ:

   $$\hat{\theta}_1 = 2\bar{X} \quad \text{i} \quad \hat{\theta}_2 = \max\{X_1, X_2, \ldots, X_n\}.$$

   Ponadto wiadomo, że $E(\hat{\theta}_1) = \theta$, $E(\hat{\theta}_2) = \frac{n\theta}{n+1}$, $MSE_{\hat{\theta}_1}(\theta) = \frac{\theta^2}{3n}$, $MSE_{\hat{\theta}_2}(\theta) = \frac{2\theta^2}{(n+1)(n+2)}$.

(a) Czy estymator $\hat{\theta}_1$ jest estymatorem

   i. nieobciążonym,
   ii. zgodnym w sensie zbieżności średniokwadratowej,
   iii. zgodnym?

(b) Czy estymator $\hat{\theta}_2$ jest estymatorem

   i. nieobciążonym,
   ii. asymptotycznie nieobciążonym,
   iii. zgodnym w sensie zbieżności średniokwadratowej,
   iv. zgodnym?

(c) Ile wynosi obciążenie estymatora $\hat{\theta}_2$?

Odpowiedzi uzasadnić powołując się na odpowiednie definicje lub fakty.

10. Zdefiniować błąd I-go rodzaju, błąd II-go rodzaju i moc testu.

11. O czym informuje współczynnik zwany *p-value*? Omówić zasadę posługiwania się tą wielkością.

12. Podczas przeprowadzania testu statystycznego weryfikującego hipotezę H_0 przeciwko hipotezie H_1, otrzymaliśmy statystykę testową równą $2,34$ i zbiór krytyczny wynoszący $W = [5, \infty)$. Jaki wyciągniemy wniosek?

13. Podczas przeprowadzania testu statystycznego weryfikującego hipotezę H_0 przeciwko hipotezie H_1, otrzymaliśmy $p\text{-}value = 0,027$. Jaki wyciągniemy wniosek na poziomie istotności $0,05$?

14. Zakłada się, że napięcie wyjściowe na zasilaczu ma rozkład normalny. Przeprowadzono szesnaście niezależnych pomiarów napięcia (w woltach) i otrzymano średnie napięcie z próby równe $12,5$ wolta oraz wariancję z próby równą $2,25$ wolta^2.

   (a) Na poziomie istotności $0,01$ sprawdzić czy średnie napięcie przekracza 12 woltów.

   (b) Zakładając, że rzeczywiste średnie napięcie to 12,5 wolta, wyznaczyć prawdopodobieństwo, że test używany w pkt. (a)

      i. przyjmie hipotezę, że średnie napięcie przekracza 12 woltów;
      ii. da odpowiedź, że średnie napięcie nie przekracza 12 woltów.

   Odpowiedzi zapisać za pomocą funkcji $m(\cdot)$, gdzie $m(\cdot)$ to funkcja mocy testu używanego w pkt. (a).

15. Pewna firma szyje prześcieradła o nominalnej długości 200 cm. Jest jednak podejrzenie, że średnia długość tych prześcieradeł jest mniejsza niż 200 cm. Aby sprawdzić czy owo podejrzenie jest słuszne, zmierzono 100 losowo wybranych prześcieradeł pochodzących od tej firmy i otrzymano średnią długość 190 cm z odchyleniem standardowym długości wynoszącym 10 cm. Jakie wyciągniemy wnioski na poziomie istotności $\alpha = 0,05$?

16. Jakich metod graficznych można użyć by sprawdzić czy analizowane dane pochodzą z rozkładu normalnego?

17. Co to jest wykres normalności i do czego służy? Omówić budowę tego wykresu i sposób korzystania z niego.

18. Wymienić testy normalności. Sformuować hipotezy, które w tych testach weryfikujemy. Który z tych testów uznawany jest za najlepszy?

ODPOWIEDZI:

3. $C = 0,8$; $EX = 0$; $Var(X) = 0,2$; $\sigma_X = \sqrt{0,2}$.

5. (a) $EX = \frac{5\theta}{4} \implies \hat{\theta} = \frac{4}{5}\bar{X}$ (b) $EX = \frac{1}{2}\theta \implies \hat{\theta} = 2\bar{X}$

6. $\hat{\theta}_{NW} = \bar{X}$

7. $\hat{a}_{NW} = \frac{n}{\sum_{i=1}^{n} \ln X_i}$

9. (a) i. Tak, bo $E(\hat{\theta}_1) = \theta$.
   ii. Tak, bo $\lim_{n\to\infty} MSE_{\hat{\theta}_1}(\theta) = 0$.
   iii. Tak, bo każdy estymator zgodny w sensie zbieżności średniokwadratowej jest zgodny.
   (b) i. nie ii. tak iii. tak iv. tak
   (c) $B(\theta) = E(\hat{\theta}_2) - \theta = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1}$

12. Ponieważ statystyka testowa nie należy do W, to stwierdzamy, że nie mamy podstaw do odrzucenia H_0.

13. $p\text{-}value = 0,027 < \alpha = 0,05$, więc odrzucamy H_0.

14. Niech X to zmienna losowa opisująca napięcie wyjściowe na zasilaczu (w woltach).
Z założenia X ma rozkład normalny o nieznanej wartości średniej μ i nieznanej wariancji σ^2, więc możemy zastosować test t-Studenta.
(a) $H_0 : \mu = 12$, $H_1 : \mu > 12$; $T = \frac{4}{3} \approx 1,33 \notin W \approx \langle 2,6025; +\infty)$, zatem brak podstaw do odrzucenia H_0. Stwierdzamy, że średnie napięcie nie przekracza 12 woltów.
(b) i. $m(12,5)$ ii. $1 - m(12,5)$

15. Niech X to zmienna losowa opisująca długość prześcieradła (w cm).
Nie mamy założenia o normalności X ale dysponujemy dużą próbą ($n = 100$), więc możemy użyć test t-Studenta.
$H_0 : \mu = 200$, $H_1 : \mu < 200$; $T = -10 \in W \approx (-\infty; -1.64\rangle$, więc H_0 odrzuacamy i stwierdzamy, że rzeczywiście średnia długość tych prześcieradeł jest mniejsza niż 200 cm.