

Egzamin: Wstęp do Statystycznej Analizy Danych

UMK, luty 2016, II termin.

Wypełnij miejsca wykropkowane [“.....”]. 1 pytanie = 1 punkt; maksimum = 30 punktów Oceniane będą **wyłącznie** odpowiedzi wpisane w wykropkowane miejsca na tej kartce, bez uzasadnień i rachunków. Wpisz odpowiedzi dopiero **po ostatecznym ich sprawdzeniu**; odpowiedzi pokreślone lub nieczytelne będą traktowane jako **błędne**! Możesz korzystać z kalkulatora, tablic statystycznych, notatek, książek. Nie wolno korzystać z komputerów, telefonów ani z tabletów. **Nie wolno się porozumiewać.** Powodzenia!

Imię i NAZWISKO:

1. Rzucamy dwie kostki do gry: białą i czarną.

(a) Oblicz prawdopodobieństwo tego, że liczba oczek na czarnej kości będzie mniejsza niż na białej.

Odpowiedź:.....

(b) Oblicz prawdopodobieństwo otrzymania sumy oczek na obu kościach równej 11 lub więcej.

Odpowiedź:.....

2. Zmienna losowa X ma rozkład prawdopodobieństwa o gęstości danej wzorem

$$f(x) = \begin{cases} 4x^{-5} & \text{dla } x \geq 1; \\ 0 & \text{dla } x < 1. \end{cases}$$

(a) Oblicz $\mathbb{E}(X)$. Odpowiedź:.....

(b) Oblicz $\text{Var}(X)$. Odpowiedź:.....

(c) Oblicz $\mathbb{P}(X \leq 2)$. Odpowiedź:.....

(d) Oblicz $\mathbb{P}(X < 2)$. Odpowiedź:.....

3. Wykonujemy 7 rzutów monetą.

(a) Oblicz prawdopodobieństwo otrzymania dokładnie 5 orłów.

Odpowiedź:.....

(b) Oblicz prawdopodobieństwo otrzymania następującego ciągu wyników: OOOOORR (w tej kolejności).

Odpowiedź:.....

- (c) Oblicz prawdopodobieństwo tego, że liczba orłów jest niemniejsza od liczby reszek.

Odpowiedź:.....

- (d) Oblicz wartość oczekiwaną (średnią) liczby reszek.

Odpowiedź:.....

- (e) Oblicz wariancję liczby reszek.

Odpowiedź:.....

4. Zakładamy, że dzienny utarg w pewnym markecie jest zmienną losową X o rozkładzie normalnym $N(50, 10^2)$ (w tysiącach PLN). Zakładamy ponadto, że zmienne X_1, \dots, X_{25} opisujące utarg w kolejnych dniach miesiąca (miesiąc ma 25 dni roboczych) są niezależne i mają taki sam rozkład $N(50, 10^2)$. Niech $S = X_1 + \dots + X_{25}$ będzie sumarycznym utargiem w ciągu miesiąca, a $\bar{X} = S/25$ – średnim utargiem.

- (a) Oblicz $\mathbb{E}(\bar{X})$. Odpowiedź:.....

- (b) Oblicz odchylenie standardowe średniej $D(\bar{X})$. Odpowiedź:.....

- (c) Znajdź liczbę a taką, że $\mathbb{P}(S \leq a) = 0,975$. Odpowiedź:.....

- (d) Oblicz $\mathbb{P}(S > 1300)$. Odpowiedź:.....

5. Wysokości cen 11 mieszkań sprzedanych przez pewnego pośrednika były następujące:

245, 160, 255, 235, 205, 310, 180, 130, 400, 210, 200

Na podstawie tych danych należy obliczyć następujące wielkości.

- (a) Oblicz wartość średnią ceny mieszkania. Odpowiedź:.....

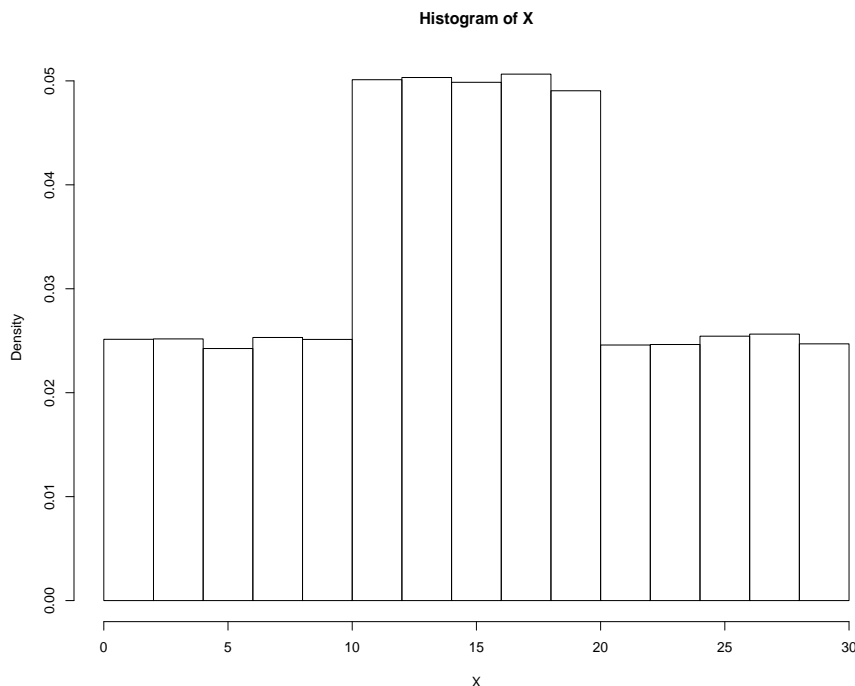
- (b) Oblicz medianę ceny mieszkania. Odpowiedź:.....

- (c) Oblicz wariancję ceny mieszkania. Odpowiedź:.....

- (d) Oblicz odchylenie standardowe ceny mieszkania. Odpowiedź:.....

Imię i NAZWISKO:

6. Na podstawie 20000 pomiarów cechy x sporządzono następujący histogram.



Odczytaj z rysunku przybliżone wartości pewnych charakterystyk liczbowych cechy x .

(a) Podaj medianę $\text{med}(x_1, x_2, \dots, x_{20000})$. Odpowiedź:.....

(b) Podaj kwartyle Q_1 i Q_3 (kwantyle rzędu $1/4$ i $3/4$). Odpowiedź:.....

(c) Podaj rozstęp międzykwartylowy. Odpowiedź:.....

(d) Podaj w przybliżeniu ile pomiarów dało wynik mniejszy, niż 5. Odpowiedź:.....

Wskazówka: Dokładne wartości wymienionych w pytaniach charakterystyk są niemożliwe do obliczenia bez znajomości danych. Niemniej, na podstawie rysunku można podać bardzo rozsądne przybliżenia. Pamiętaj, że histogram dla dużej liczby danych jest przybliżeniem gęstości (pola pod histogramem są bliskie odpowiednim prawdopodobieństwom).

7. Zakładamy, że X_1, \dots, X_{400} jest próbka z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i wariancją σ^2 . Obliczono średnią z próbki i nieobciążony estymator wariancji: $\bar{X} = 50, 75, S^2 = 10^2$.

(a) Przeprowadź test hipotezy $H_0 : \mu = 50$ przeciw alternatywie $H_1 : \mu > 50$ na poziomie istotności $\alpha = 0,05$. Oblicz wartość statystyki T (t-Studenta), porównaj z odpowiednim kwantylem i podejmij decyzję: odrzucamy H_0 czy nie?

Odpowiedź:.....

- (b) Przeprowadź test hipotezy $H_0 : \mu = 50$ przeciw alternatywie $H_1 : \mu \neq 50$ na poziomie istotności $\alpha = 0,05$. Oblicz wartość statystyki T (t-Studenta), porównaj z odpowiednim kwantylem i podejmij decyzję: odrzucamy H_0 czy nie?

Odpowiedź:.....

- (c) Oblicz przedział ufności dla μ na poziomie $1 - \alpha = 0,95$.

Odpowiedź:.....

Uwaga: Kwantyle rozkładu t-Studenta z 399 stopniami swobody są w przybliżeniu równe odpowiednim kwantylom rozkładu normalnego.

8. W losowo wybranej próbie $n = 400$ studentów, znalazło się $x = 144$ palących. Interesuje nas frakcja p palących w populacji studentów.

- (a) Zbuduj przedział ufności dla p na poziomie ufności 0,95.

Odpowiedź:.....

- (b) Przeprowadź test hipotezy $H_0 : p = 0,20$ przeciw $H_1 : p > 0,20$. Przyjmij poziom istotności 0,05.

Odpowiedź: Odrzucamy H_0 na rzecz H_1 czy nie?.....

9. Zmienna losowa X ma rozkład prawdopodobieństwa o gęstości danej wzorem

$$f(x) = \begin{cases} 4x^3 & \text{dla } 0 \leq x \leq 1; \\ 0 & \text{w pozostałych przypadkach.} \end{cases}$$

- (a) Podaj wzór na dystrybuantę $F_X(x) = \mathbb{P}(X \leq x)$.

Odpowiedź:.....

- (b) Podaj medianę zmiennej losowej X .

Odpowiedź:.....