Egzamin: Wstęp do Statystycznej Analizy Danych

Wydział MiI UMK, styczeń 2019, I termin.   Zestaw **A**

Wypełnij miejsca wykropkowane [".."]. Oceniane będą **tylko** 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. Wiadomo, że $\mathbb{P}(A) = 0.5$, $\mathbb{P}(B) = 0.4$ i $\mathbb{P}(A \setminus B) = 0.2$. Oblicz
   (a) $\mathbb{P}(A \cap B) =$..
   (b) $\mathbb{P}(A \cup B) =$..
   (c) $\mathbb{P}(B \setminus A) =$..
   (d) $\mathbb{P}(B|A) =$..
   (e) Czy zdarzenia losowe A i B są niezależne? (napisz TAK lub NIE) ....

2. Rzucamy dwie rzetelne kostki do gry: białą i czarną.
   (a) Oblicz prawdopodobieństwo tego, że liczba oczek na czarnej kości będzie większa niż na białej.

   Odpowiedź:...

   (b) Oblicz prawdopodobieństwo tego, że na obu kościach liczba oczek będzie taka sama.

   Odpowiedź:...

   (c) Oblicz prawdopodobieństwo otrzymania sumy oczek na obu kościach równej 11.

   Odpowiedź:...

3. Fabryki A, B i C produkują odpowiednio 50%, 30% i 20% ogólnej produkcji żarówek. Udział braków wynosi odpowiednio 5% produkcji fabryki A, 2% produkcji B i 3% produkcji C.
   (a) Oblicz prawdopodobieństwo tego, że losowo wybrana żarówka jest brakiem.

   Odpowiedź:...

   (b) Oblicz prawdopodobieństwo tego, że losowo wybrana żarówka pochodzi z fabryki A, jeśli wiadomo, że jest brakiem.

   Odpowiedź:...

   (c) Oblicz prawdopodobieństwo tego, że losowo wybrana żarówka pochodzi z fabryki A, jeśli wiadomo, że nie jest brakiem.

   Odpowiedź:...

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

$$f(x) = \begin{cases} 4x^3 & \text{dla } 0 \leqslant x \leqslant 1; \\ 0 & \text{w przeciwnym przypadku .} \end{cases}$$

   (a) $\mathbb{E}(X) =$..
   (b) $\mathrm{Var}(X) =$..
   (c) $\mathbb{P}(X \leqslant 0.5) =$..
   (d) Podaj medianę rozkładu zmiennej losowej X). Odpowiedź:..

5. Wykonujemy 6 rzutów symetryczną monetą.

   (a) Oblicz prawdopodobieństwo otrzymania dokładnie 3 orłów.

   Odpowiedź:..

   (b) Oblicz prawdopodobieństwo otrzymania ciągu wyników 'OOORRR' (w podanej kolejności).

   Odpowiedź:..

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

   Odpowiedź:..

   (d) Oblicz wartość oczekiwaną (średnią) liczby orłów.

   Odpowiedź:..

   (e) Oblicz wariancję liczby orłów.

   Odpowiedź:..

6. Zakładamy, że dzienny utarg w pewnym markecie jest zmienną losową X taką, że $\mathbb{E}X = 10$ i $\mathrm{Var}X = 10^2$ (w tysiącach PLN). Zakładamy, że zmienne $X_1, \ldots, X_{256}$ opisujące utarg w kolejnych dniach roku (rok ma 256 dni roboczych) są niezależne i mają taki sam rozkład prawdopodobieństwa. Niech $S = X_1 + \cdots + X_{256}$ będzie sumarycznym utargiem w ciągu roku, a $\bar{X} = S/256$ – średnim utargiem. Oblicz:

   (a) $\mathbb{E}(S) =$..
   (b) $\mathrm{Var}(S) =$..
   (c) $\mathbb{E}(\bar{X}) =$..
   (d) $\mathrm{Var}(\bar{X}) =$..
   (e) Oblicz w przybliżeniu $\mathbb{P}(S > 2500)$, wykorzystując Centralne Twierdzenie Graniczne.

   Odpowiedź:..

   (f) Znajdź liczbę a taką, że $\mathbb{P}(S \leqslant a) \approx 0.975$.

   Odpowiedź:..

Imię i NAZWISKO: ....

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

$$150, 245, 225, 195, 300, 170, 120, 390, 200, 190, 235$$

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:..

8. Zakładamy, że $X_1, \ldots, X_{400}$ jest próbką z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i wariancją σ^2. Obliczono średnią z próbki i nieobciążony estymator wariancji: $\bar{X} = 51.75$, $S^2 = 20^2$.

- Przeprowadź test hipotezy $H_0 : \mu = 50$ przeciw alternatywie $H_1 : \mu > 50$ na poziomie istotności $\alpha = 0,05$.

(a) Oblicz wartość statystyki T (t-Studenta): $T =$...

(b) Oblicz p-wartość testu: $P =$...

(c) Podejmij decyzję: odrzucamy H_0? (napisz TAK lub NIE) ....

- Przeprowadź test hipotezy $H_0 : \mu = 50$ przeciw alternatywie $H_1 : \mu \neq 50$ na poziomie istotności $\alpha = 0,05$.

(d) Oblicz p-wartość testu: $P =$...

(e) Podejmij decyzję: odrzucamy H_0? (napisz TAK lub NIE) ....

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

Odpowiedź:..

Uwaga: Rozkład t-Studenta z 399 stopniami swobody jest przybliżeniu równy standardowemu rozkładowi normalnemu $N(0,1)$. W szczególności można wykorzystać następujące przybliżone wartości kwantyli: $t_{0,95}(399) \approx z_{0,95} = 1.65$ i $t_{0,975}(399) \approx z_{0,975} = 1.96 \approx 2$. Do obliczania p-wartości można użyć tablic rozkładu normalnego.

9. W losowo wybranej próbce $n = 400$ studentów, znalazło się $x = 80$ palących. Interesuje nas frakcja p palących w populacji studentów.

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

Odpowiedź:..

- Przeprowadź test hipotezy $H_0 : p = 0.25$ przeciw $H_1 : p \neq 0.25$. Przyjmij poziom istotności $0,05$.

(b) Oblicz wartość statystyki Z: $Z =$...

(c) Oblicz p-wartość testu: $P =$...

(d) Podejmij decyzję: odrzucamy H_0? (napisz TAK lub NIE) ....

Uwaga: Możesz przyjąć przybliżoną wartość kwantyla rozkładu normalnego: $z_{0,975} \approx 2$.

10. Zmienna losowa X ma rozkład wykładniczy o dystrybuancie

$$F_X(x) = \mathbb{P}(X \leqslant x) = \begin{cases} 1 - e^{-3x} & \text{dla } x \geqslant 0; \\ 0 & \text{w przeciwnym przypadku} . \end{cases}$$

Zmienną losową Y definiujemy wzorem

$$Y = e^{2X}.$$

Należy znaleźć rozkład prawdopodobieństwa zmiennej losowej Y.

(a) Oblicz dystrybuantę zmiennej losowej Y: $F_Y(y) = \mathbb{P}(Y \leqslant y) =$...

(b) Oblicz gęstość zmiennej losowej Y: $f_Y(y) =$...

(c) Oblicz wartość oczekiwaną zmiennej losowej Y: $\mathbb{E}(Y) =$...

Uwaga: Pamiętaj, że mamy $X \geqslant 0$, a zatem $Y = e^{2X} \geqslant 1$.

11. W celu zbadania, czy trzy obecne na rynku marki golarek elektrycznych: A,B i C są jednakowo popularne, zanotowano jaką markę wybrało 90 klientów kupujących golarki. Wyniki przedstawia następująca tabelka:

marka	A	B	C
liczba klientów	20	30	40

Czy mamy podstawy do odrzucenia hipotezy H_0 mówiącej, że klienci wybierają każdą z 3 marek z jednakowym prawdopodobieństwem $1/3$?

Przeprowadź odpowiedni test zgodności χ^2 (chi-kwadrat) na poziomie istotności $\alpha = 0.05$. Oblicz wartość statystyki testowej χ^2, podaj odpowiedni kwantyl i podejmij decyzję: odrzucamy H_0 czy nie?

(a) Statystyka $\chi^2 =$...

(b) Kwantyl $\chi^2_{0.95}(2) =$...

(c) Decyzja: odrzucamy H_0? (napisz TAK lub NIE) ....

(d) p-wartość testu: $P =$...

Wskazówka: Skorzystaj z następującego faktu. Rozkład $\chi^2(2)$ (chi-kwadrat z dwoma stopniami swobody) ma dystrybuantę F daną wzorem

$$F(x) = \begin{cases} 1 - e^{-x/2} & \text{dla } x \geqslant 0; \\ 0 & \text{w przeciwnym przypadku} . \end{cases}$$

(Jest to po prostu rozkład wykładniczy z parametrem $1/2$.) Jeśli nie masz kalkulatora „naukowego” możesz przyjąć, że $e^{-3} \approx 0.05$.