

Wykład nr

Nr indeksu:

--	--	--	--	--

Imię i nazwisko studenta

.....

Temat

Egzamin ze statystyki

Studia Licencjackie Stacjonarne

Zadanie		Pkt.
1	a	
	b	
	c	
	d	
	e	
	f	
2	a	
	b	
	c	
3	a	
	b	
4		
5	a	
	b	
	c	
	d	
6	a	
	b	
7	(TEST)	
Razem		

Uwaga. W rozwiązaniach – tam gdzie to konieczne – przyjmij poziom istotności 0,01 i współczynnik ufności 0,99.

Zadanie 1. [10,5] Do urzędu miasta A wpłynęło wiele skarg na opieszałość urzędu w udzielaniu odpowiedzi na pisma mieszkańców. Dla zbadania zasadności tych skarg wylosowano 200 pism spośród wszystkich, na które udzielono odpowiedzi w ubiegłym roku i uzyskano wyniki:

Czas oczekiwania na odpowiedź (w dniach)	Liczba pism
9-21	20
22-28	40
29-35	80
36-42	60
Razem	200

Ponadto obliczono:

$$\sum (\dot{x}_i - \bar{x})^2 n_i = 10480$$

$$\sum (\dot{x}_i - \bar{x})^3 n_i = - 59760$$

- a) Oblicz średni czas oczekiwania na odpowiedź oraz uzasadnij, czy wynik mógłby być inny, gdybyśmy zamiast przedziałów mieli informacje o czasie oczekiwania na odpowiedź na każde z pism. [1]
- b) Oblicz i zinterpretuj kwartył trzeci w tym rozkładzie. [1,5]
- c) Jak długo czekał na odpowiedź na swoje pismo jeden z badanych, skoro po wystandaryzowaniu jego czasu oczekiwania wynik wyniósł 0? [1]
- d) Scharakteryzuj kierunek i siłę asymetrii rozkładu czasu oczekiwania na odpowiedź, obliczając odpowiednią miarę.[2]

- e) Czy wylosowana próba (200 pism) jest dostatecznie liczna, aby z błędem 5% oszacować przedziałowo udział pism, na które w tym urzędzie odpowiedź udzielana jest z opóźnieniem? Z wcześniejszych badań wynika, że takich pism może być około 40%. [1,5]
- f) Postanowiono sprawdzić, czy średni czas oczekiwania na odpowiedź przekracza czas określony przepisami, tzn. 30 dni. Wyznacz krytyczny poziom istotności i w oparciu o jego poziom zweryfikuj odpowiednią hipotezę. [3,5]

Zadanie 2. [4,5] Postanowiono sprawdzić, czy czas oczekiwania na odpowiedź jest niezależny stochastycznie od wydziału, do którego kierowane jest pismo. Dla wylosowanej próby pism (dane z zad.1) i trzech wydziałów otrzymano:

Czas oczekiwania na odpowiedź (w dniach)	Wydział Nieruchomości (WN)	Wydział Komunikacji (WK)	Wydział Oświaty (WO)	Razem
9-21	-	10	10	20
22-28	20	10	10	40
29-35	60	10	10	80
36-42	60	-	-	60
Razem	140	30	30	200

- a) Oblicz liczebność teoretyczną \hat{n}_{43} (dla wariantu: 36-42 dni / Wydział Oświaty) i wytłumacz co oznacza ta wartość. [1]
- b) Oblicz wartość statystyki chi-kwadrat wiedząc, że jej wartość wyznaczona dla wszystkich składników poza ostatnim (36-42 dni/ WO) wyniosła 71,95. Czy zależność w próbie między czasem oczekiwania na odpowiedź i wydziałem, do którego było ono skierowane można uznać za silną? [1,5]
- c) Czy można uznać, że powyższa zależność dotyczy 200 wylosowanych pism, czy wszystkich trafiających do tego urzędu? Zweryfikuj odpowiednią hipotezę. [2]

Zadanie 3. [4] Losowo wybranych pracowników trzech wydziałów urzędu miasta (WN, WK, WO) zapytano o ich zarobki [tys. zł]. Odpowiedzi udzieliło 12 pracowników Wydziału Nieruchomości, 16 Wydziału Komunikacji i 20 Wydziału Oświaty. Suma kwadratów odchyleń wewnątrzgrupowych wyniosła 7,65, natomiast średnie zarobki badanych osób w poszczególnych wydziałach wyniosły: 3,7 w WN; 3,6 w WK; 3,3 w WO.

- a) Na jakie pytanie da nam odpowiedź zastosowanie analizy wariancji? [0,5]

b) Jaka to będzie odpowiedź – przeprowadź całą procedurę weryfikacyjną. [3,5]

Zadanie 4 [3] Dobowy czas snu studentów pewnej uczelni ekonomicznej w czasie sesji egzaminacyjnej charakteryzuje się średnią równą 6 godzin i odchyleniem standardowym 1 godzina. Podaj jaki przybliżony rozkład i o jakich parametrach ma średni czas snu w czasie sesji egzaminacyjnej losowo wybranych 120 studentów tej uczelni. Oblicz prawdopodobieństwo, że średni czas snu tych 120 studentów znajdzie się w przedziale od 6 do 6,2 godziny. Uzasadnij metodę rozwiązania zadania, powołując się na stosowne twierdzenie.

Zadanie 5. [9,5] Pewien mieszkaniec miasta A twierdzi, że liczba spraw załatwianych po ustawowo przewidzianym terminie przez dany wydział jest stochastycznie liniowo zależna od ilości kawy, kupowanej na potrzeby pracowników tego wydziału. Zgromadził on dane z ostatniego miesiąca dla ośmiu wydziałów, które przedstawione są w tabeli poniżej. Ponadto, mieszkaniec obliczył, że średnia ilość zamawianej kawy wyniosła **17,375** kg, a odchylenie standardowe **10,53** kg (liczone z wariancji nieobciążonej). Podobnie średnia dla liczby spraw załatwionych po terminie w pojedynczym wydziale wyniosła **28**, a odchylenie standardowe **13,27**.

Nr wydziału	1	2	3	4	5	6	7	8
Ilość zamawianej kawy w miesiącu (X w kg)	10	15	14	34	5	33	12	16
Liczba spraw załatwiona po terminie (Y)	15	22	29	39	10	51	24	34

- a) Oblicz i zinterpretuj wartość współczynnika korelacji liniowej pomiędzy ilością zamawianej kawy i liczbą spraw załatwionych po terminie. Wiadomo, że $\sum_{i=1}^8 x_i y_i = 4777$. [2,5]

- b) Oblicz i zinterpretuj oszacowania parametrów strukturalnych liniowej funkcji regresji liczby spraw załatwionych po terminie względem ilości zamawianej kawy. [2]

c) Odpowiedz na pytanie, czy oszacowanie współczynnika regresji jest statystycznie istotne w kontekście tezy postawionej przez mieszkańca. Suma kwadratów reszt w modelu wynosi **222,53**. [3]

d) Część mieszkańców uważa, że jedynie całkowity zakaz picia kawy może uzdrowić sytuację w lokalnym urzędzie. Dokonaj prognozy liczby przeterminowanych spraw w sytuacji, gdyby zakupy kawy w hipotetycznym wydziale wyniosły 0 kg. Oszacuj wartość standardowego błędu tej prognozy.[2]

Zadanie 6 [4,5]

a) Ceny detaliczne paliwa Pb98 (na koniec miesiąca) w kolejnych pięciu miesiącach od stycznia do maja na jednej ze stacji wynosiły w (zł/litr): 4,80; 5,00; 5,10; 5,22; 5,30. Oblicz i zinterpretuj średnie tempo zmian cen tego paliwa w analizowanym okresie. Następnie wyznacz prognozę ceny paliwa na koniec lipca tego roku przy założeniu, że średnie tempo zmian utrzyma się przez kolejne dwa miesiące. [2]

b) Wartość sprzedaży oraz dynamika ilości sprzedaży dwóch rodzajów paliwa na jednej ze stacji kształtowały się następująco:



Paliwo	Indeksy ilości sprzedaży	Wartość sprzedaży w maju w mln zł
	maj/styczeń	
Pb95	1,15	2,5
ON (olej napędowy)	1,1	1,0

Wiedząc, że wartość sprzedaży obu rodzajów paliwa łącznie wyniosła w styczniu 2,0 mln zł, określ wpływ zmian cen jednostkowych na dynamikę wartości sprzedaży tych paliw w analizowanym okresie. [2,5]

Zadanie 7 Proszę zaznaczyć w każdym przypadku odpowiedź **T-tak** lub **N-nie**

Punktacja: odpowiedź poprawna = 1 pkt; brak odpowiedzi = 0 pkt; odpowiedź błędna = -1 pkt;

Jeżeli całkowita suma punktów z części testowej będzie ujemna, jako wynik części testowej zostanie przyjęte 0 pkt.

1. Przy współczynniku ufności 0,95 wyznaczono przedział ufności dla przeciętnej wartości oszczędności Polaków. Oznacza to, że:

- | | | |
|---|----------|----------|
| a) prawdopodobieństwo, że w wyznaczonym przedziale znajduje się punktowe oszacowanie parametru wynosi 0,95 | T | N |
| b) z prawdopodobieństwem 0,95 przeciętne oszczędności w próbie są równe przeciętnym oszczędnościom wszystkich Polaków | T | N |
| c) przedział ten jest jednym z tych, które z prawdopodobieństwem 0,95 obejmują przeciętne oszczędności Polaków | T | N |

2. Podana niżej miara może przyjąć wartości niższe niż -1

- | | | |
|--------------------------------|----------|----------|
| a) współczynnik asymetrii | T | N |
| b) kowariancja | T | N |
| c) współczynnik korelacji rang | T | N |

3. Rozkład t-Studenta

- | | | |
|--|----------|----------|
| a) ma wariancję wyższą niż wariancja w rozkładzie normalnym standardowym | T | N |
| b) ma symetryczną funkcję gęstości prawdopodobieństwa | T | N |
| c) ma dystrybuantę, która dla wartości 0 wynosi 0,5 | T | N |