**Tadeusz Inglot**

# STATYSTYKA STOSOWANA
# 2018/2019

LITERATURA PODSTAWOWA

J. Koronacki, J. Mielniczuk, Statystyka dla studentów kierunków technicznych i przyrodniczych, WNT, Warszawa 2004.

LITERATURA UZUPEŁNIAJĄCA

1. L. Gajek, M. Kałuszka, Wnioskowanie statystyczne. Modele i metody, WNT, Warszawa 2004.
2. J. Greń, Statystyka matematyczna. Modele i zadania, PWN, Warszawa 1976.
3. T. Inglot, T. Ledwina, Z. Ławniczak, Materiały do ćwiczeń z rachunku prawdopodobieństwa i statystyki matematycznej, Oficyna Wydawnicza PWr, Wrocław 1984.
4. H. Jasiulewicz, W. Kordecki, Rachunek prawdopodobieństwa i statystyka matematyczna. Przykłady i zadania, Oficyna Wydawnicza GiS, Wrocław 2001.
5. W. Klonecki, Statystyka matematyczna, PWN, Warszawa 1999.
6. W. Kordecki, Rachunek prawdopodobieństwa i statystyka matematyczna. Definicje, twierdzenia, wzory, Oficyna Wydawnicza GiS, Wrocław 2002.
7. W. Krysicki, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach, Cz. I i II, PWN, Warszawa 2007.
8. A. Plucińska, E. Pluciński, Zadania z probabilistyki, PWN, Warszawa 1983.
9. P. Pusz, L. Zaręba, Elementy statystyki, Wydawnictwo Oświatowe FOSZE, Rzeszów 2006.

**STATYSTYKA STOSOWANA**
**Lista zadań nr 1**

1. Niech $\overline{x}$ i $s^2$ będą średnią i wariancją z próby. Wykazać, że dla dowolnego $a \in R$ zachodzi wzór $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - a)^2 - (\overline{x} - a)^2$.

2. Katalogowe zużycie paliwa (w l) na 100 km 24 modeli samochodów występujących na rynku polskim wynosi: 6.3, 8.0, 8.5, 9.3, 5.5, 5.9, 5.9, 6.5, 6.4, 6.6, 8.2, 10.1, 6.3, 6.8, 7.6, 6.7, 7.3, 7.1, 9.2, 6.9, 5.9, 7.5, 8.6, 6.0. Wyznaczyć charakterystyki dla badanej zmiennej, sporządzić wykres ramkowy i histogram. (Koronacki i Mielniczuk, zad. 1.2, str. 57)

3. Wynagrodzenie 41 pracowników pewnej firmy w tys. zł wynosi (po uporządkowaniu rosnąco): 2, 2, 2, 2, 2, 2.4, 2.4, 2.4, 2.8, 2.8, 2.8, 2.8, 3, 3, 3, 3, 3, 3, 3, 3, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 4.4, 4.4, 4.4, 4.4, 5.2, 5.2. Wyznaczyć charakterystyki, sporządzić wykres ramkowy oraz histogram.

4. Suma opadów (w mm) w Warszawie w lipcu w kolejnych latach poczynając od roku 1811 do roku 1960 wynosiła: 35, 82, 48, 75, 77, 123, 117, 75, 92, 101, 116, 113, 42, 44, 36, 71, 9, 74, 114, 49, 83, 94, 223, 28, 57, 46, 33, 86, 85, 74, 72, 104, 37, 229, 41, 50, 73, 40, 76, 100, 171, 41, 160, 120, 144, 46, 143, 105, 29, 92, 138, 44, 26, 80, 50, 84, 78, 74, 53, 51, 76, 30, 48, 6, 54, 63, 20, 74, 81, 45, 50, 174, 82, 18, 139, 31, 47, 78, 173, 71, 72, 20, 85, 19, 35, 39, 120, 92, 172, 98, 37, 77, 143, 26, 96, 13, 132, 109, 116, 132, 37, 32, 91, 101, 77, 87, 99, 181, 166, 68, 5, 122, 33, 84, 66, 64, 149, 23, 20, 115, 71, 108, 55, 166, 124, 115, 53, 71, 49, 73, 93, 76, 113, 53, 77, 37, 78, 124, 84, 44, 68, 26, 65, 136, 154, 82, 88, 38, 80, 159. Obliczyć średnią, wariancję, medianę i rozstęp międzykwartylowy. Sporządzić histogram i wykres ramkowy. Wyznaczyć średnią obciętą odrzucając po 15% skrajnych wyników. Ocenić własności rozkładu sumy opadów (jednomodalność, skośność, spłaszczenie). (Koronacki i Mielniczuk, zad. 1.3, str. 57)

5. Dla danych z poprzedniego zadania rozważyć oddzielnie sumy opadów z lat 1811 - 1860 oraz z lat 1911 - 1960. Wyznaczyć wykresy ramkowe i histogramy dla tych danych i ocenić zgrubnie, czy po 100 latach zmienił się rozkład sumy opadów w lipcu.

6. Przeciętna długość życia mężczyzn w 29 państwach świata wynosi: 74, 76, 77, 72, 69, 65, 72, 68, 70, 72, 72, 73, 74, 75, 72, 74, 68, 72, 75, 72, 69, 71, 75, 76, 81, 73, 69, 78, 74. Wyznaczyć charakterystyki, sporządzić wykres ramkowy i histogram. Ocenić, czy rozkład jest skośny, czy symetryczny. Wyznaczyć średnią obciętą odrzucając po 10% najmniejszych i największych wyników. (Pusz i Zaręba, zad. 10, str. 40)

7. Przeciętna długość życia kobiet w tych samych 29 państwach świata co w zadaniu poprzednim wynosi: 80, 74, 76, 77, 80, 82, 81, 80, 84, 79, 81, 75, 71, 73, 76, 78, 83, 81, 73, 74, 75, 79, 81, 75, 80, 79, 75, 77, 81. Wyznaczyć charakterystyki, sporządzić wykres ramkowy i histogram. Porównać jakościowo rozkłady obu zmiennych. Czy z powyższych danych można wnioskować, że rozkład długości życia kobiety jest przesuniętym rozkładem długości życia mężczyzny? O ile lat? (Pusz i Zaręba, zad. 9, str 40)

8. Wykonać 200 rzutów kostką do gry. Wyznaczyć charakterystyki, sporządzić wykres ramkowy i histogram. Czy wyniki przemawiają za prawidłowością kostki?

9. Dla danych z poprzedniego zadania przjąć wypadnięcie 4, 5 lub 6 jako wynik 0, wypadnięcie 1 jako wynik 1 i wypadnięcie 2 lub 3 jako wynik 2 (czyli nowa zmienna ma tylko wartości 0, 1 i 2) i powtórzyć czynności z poprzedniego zadania.

**STATYSTYKA STOSOWANA**
**Lista zadań nr 2**

1. Zmienna $X$ ma rozkład o gęstości

$$f(x) = \begin{cases} cxe^{-x/2} & , \quad x \geq 0, \\ 0 & , \quad \text{poza tym.} \end{cases}$$

Wyznaczyć $c$. Narysować wykres i na jego podstawie ocenić własności tego rozkładu. Obliczyć $P(1 \leq X \leq 6)$, $EX$, $\mathrm{Var}X$ oraz medianę. Na wykresie gęstości zaznaczyć obliczone uprzednio prawdopodobieństwo oraz $EX$ i medianę. Uzasadnić, że powyższy rozkład jest rozkładem gamma. Z jakimi parametrami?

2. Zmienna $Y$ ma rozkład beta z parametrami $p = 3$, $q = 2$. Napisać wzór i narysować gęstość. Obliczyć wartość oczekiwaną $EY$, medianę i prawdopodobieństwo $P(Y \geq \frac{2}{3})$. Obliczyć współczynnik skośności i spłaszczenia. Zinterpretować otrzymane wyniki.

3. Niech $M$ oznacza największą spośród 4 losowo wybranych liczb z odcinka [0,1]. Wykazać, że $P(M < x) = x^4$ dla $x \in [0, 1]$. Wywnioskować stąd, że zmienna $M$ ma gęstość

$$f(x) = \begin{cases} 4x^3 & \text{dla } x \in [0, 1] \\ 0 & \text{poza tym} \end{cases} .$$

Obliczyć $EM$ i $\mathrm{Var}M$ oraz współczynnik zmienności. Nazwać ten rozkład.

4. Wybrano losowo liczbę z odcinka [0,1]. Niech $V$ oznacza iloraz długości krótszego z otrzymanych dwóch odcinków przez długość dłuższego. Wyznaczyć i narysować gęstość zmiennej $V$. Obliczyć jej wartość oczekiwaną i wariancję.

5. Cząstka wykonuje ruch chaotyczny na płaszczyźnie. Niech $X$ oznacza odległość cząstki od położenia początkowego po czasie $T$. Gęstość $T$ ma postać

$$f(x) \begin{cases} Cxe^{-\lambda x^2}, & x \geqslant 0 \\ 0, & x < 0 \end{cases} ,$$

gdzie $\lambda > 0$ jest parametrem zależnym od $T$ i od środowiska, w którym odbywa się ruch. Obliczyć stałą $C$, wartość oczekiwaną i wariancję $X$ oraz prawdopodobieństwo $P(X > \lambda)$.

**STATYSTYKA STOSOWANA**
**Lista zadań nr 3**

1. Prawdopodobieństwo trafienia w dziesiątkę wynosi 0.7, a w dziewiątkę wynosi 0.3. Niech S oznacza liczbę punktów zdobytych w 10 strzałach. Wyznaczyć rozkład S. Obliczyć prawdopodobieństwo, że strzelec zdobędzie co najmniej 98 punktów oraz wartość oczekiwaną i wariancję S.

2. W pewnym sklepie znajdują się cztery jednakowe, pracujące niezależnie stoiska, każde obsługiwane przez jedną ekspedientkę. Przeciętnie ekspedientka jest zajęta obsługą klientów przez 1/3 część czasu pracy. Obliczyć prawdopodobieństwo tego, że w momencie naszego przyjścia zastaniemy co najmniej jedno wolne stoisko, prawdopodobieństwo, że wszystkie stoiska będą wolne i średnią liczbę zajętych stoisk.

3. Czas poprawnej pracy pewnego urządzenia ma rozkład wykładniczy o średniej 10 000 godzin. Włączono 10 takich urządzeń. Niech $X$ oznacza liczbę poprawnie pracujących urządzeń po 5000 godzin. Określić rozkład $X$. Obliczyć $EX$ oraz prawdopodobieństwo tego, że po 5000 godzin pracuje poprawnie co najmniej 7 urządzeń.

4. Punkt startuje z początku układu współrzędnych i porusza się po prostej: przesuwa się o jednostkę w prawo z prawdopodobieństwem 0.5 i o jednostkę w lewo z prawdopodobieństwem 0.5. Przyjmując, że poszczególne przesunięcia są niezależne, wyznaczyć rozkład zmiennej losowej $D_6$ oznaczjącej położenie punktu po 6 przesunięciach. Obliczyć prawdopodobieństwo, że po 10 przesunięciach punkt znajdzie się w przedziale $[-2, 2]$. Zaproponować interpretację fizyczną opisanego zjawiska losowego i otrzymanego wyniku.

5. Aparatura zawiera 2000 jednakowo niezawodnych elementów. Prawdopodobieństwo zepsucia się każdego z nich wynosi $p = 0.001$. Jakie jest prawdopodobieństwo, że aparatura przestanie działać, jeśli awaria następuje przy uszkodzeniu a) co najmniej jednego elementu; b) co najmniej trzech elementów ? Obliczyć wartość oczekiwaną liczby uszkodzonych elementów.

6. Centrala telefoniczna obsługuje 300 abonentów. Każdy abonent, niezależnie od pozostałych abonentów, może z prawdopodobieństwem 0.02 zamówić połączenie zewnę-trzne. Jaka powinna być minimalna liczba łączy zewnętrznych do tej centrali, aby z prawdopodobieństwem 0.9 były zrealizowane wszystkie zamówienia abonentów na połączenia zewnętrzne?

7. Jaka jest najbardziej prawdopodobna liczba studentów urodzonych w niedzielę wśród 60 obecnych na wykładzie? Ile wynosi to prawdopodobieństwo? Porównać ze stanem faktycznym.

8. Z talii 52 kart wylosowano 6. Niech $D$ oznacza liczbę pików wśród wylosowanych kart. Znaleźć rozkład zmiennej $D$. Obliczyć prawdopodobieństwo, że otrzymano co najmniej 3 piki.

9. Student zna odpowiedź na 20 spośród 30 pytań. Losuje 5 pytań. Określić rozkład zmiennej $Y$ oznaczającej liczbę pytań, na które student udzieli poprawnej odpowiedzi. Obliczyć prawdopodobieństwo, że student udzieli poprawnej odpowiedzi na co najmniej 3 pytania. Obliczyć wartości oczekiwaną oraz wariancję tej zmiennej.

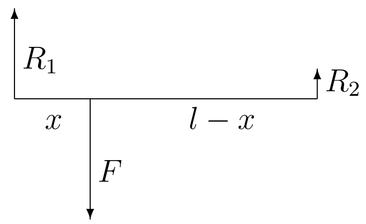
## STATYSTYKA STOSOWANA
## Lista zadań nr 4

1. Czas bezawaryjnej pracy (w godz.) pewnego typu drukarki ma rozkład normalny $N(1500, 200)$. Jaki procent drukarek ulegnie awarii przed upływem 1000 godzin? Jaki powinien być okres gwarancji, aby najwyżej 5% drukarek uległo awarii w okresie gwarancyjnym? Przyjąć, że drukarka pracuje 8 godzin dziennie.

2. Długość produkowanych detali (w mm) ma rozkład normalny $N(9, 0.03)$. Norma przewiduje wyroby o wymiarach $9 \pm 0.05$ mm. Jaki procent produkowanych detali nie spełnia wymogów normy? Jakie może być dopuszczalne $\sigma$, aby procent detali nie spełniających wymagań normy nie przekroczył 0.1% ?

3. Temperatura $T$ gazu (w $^0C$) znajdującego się w zamkniętym naczyniu o stałej objętości $v = 0.01$ m$^3$ jest zmienną o rozkładzie normalnym N(20, 0.5). Wyznaczyć rozkład ciśnienia $P$ tego gazu, korzystając z równania Clapeyrona: $\frac{Pv}{t} = nR$, gdzie $R = 8.31$ J/K mol, a $n = 1$ mol gazu oraz $t = T + 273.15$. Obliczyć prawdopodobieństwo, że ciśnienie nie przekroczy $2.45 \cdot 10^5$ Pa.

4. Na belce żelbetowej długości $l = 4\,$m (rysunek) umieszczono losowe obciążenie $F$ (w N) w odległości $x = 1\,$m od lewego końca. $F$ ma rozkład normalny $N(2000, 500)$.




   Wyznaczyć rozkłady reakcji $R_1$ i $R_2$ na podporach ($R_1 l = (l-x)F,\ R_2 l = xF$). Obliczyć prawdopodobieństwo tego, że $R_1$ przekroczy 2000 N.

5. Dwie piekarnie dostarczjące pieczywo do sklepu w proporcjach 1:2 wypiekają chleby o nominalnej wadze 0.5 kg. W rzeczywistości waga chleba z pierwszej piekarni ma rozkład normalny $N(0.48, 0.02)$, a waga chleba z drugiej piekarni $N(0.51, 0.15)$. Wyznaczyć gęstość zmiennej oznaczającej wagę chleba sprzedawanego w tym sklepie. Obliczyć jej wartość oczekiwaną i prawdopodobieństwo, że kupiony bochenek będzie miał wagę co najmniej 0.5 kg.

6. Czas życia żarówki ma rozkład wykładniczy z parametrem $1/\lambda = 400$ godzin. Stosując nierówność Czebyszewa oszacować prawdopodobieństwo, że żarówka będzie świeciła co najmniej 1000 godzin? Porównać to oszacowanie z wartością dokładną.

7. Rzucono sto razy kostką do gry. Stosując nierówność Czebyszewa oszacować prawdopodobieństwo tego, że liczba szóstek wyniesie co najmniej 10 oraz prawdopodobieństwo tego, że suma wyrzuconych oczek znajdzie się w przedziale od 300 do 400.

8. Roczny opad deszczu (w cm) w pewnym regionie ma rozkład normalny N(100, 10). Jakie jest prawdopodobieństwo, że w roku bieżącym oraz dwu kolejnych latach roczny opad nie przekroczy 112 cm? Wygenerować 5 obserwacji rocznych opadów w trzech kolejnych latach. Obliczyć częstość opisanego wyżej zdarzenia.

9. Roczny dochód $Z$ pewnego przedsiębiorstwa ma rozkład Pareto o gęstości $f(x) = 3(x+1)^{-4}$ dla $x > 0$ i zero poza tym (w mln. zł). Obliczyć $P(Z > \frac{1}{2})$ oraz $EZ$ i $\operatorname{Var} Z$. Wygenerować próbę 30-elementową z tego rozkładu, obliczyć częstość zdarzenia $\{Z > \frac{1}{2}\}$, średnią oraz wariancję z próby i porównać wyniki z wartościami dokładnymi (teoretycznymi).

**STATYSTYKA STOSOWANA**
**Lista zadań nr 5**

1. Zbiór składający się z 1000 znaków jest przesyłany między dwoma komputerami. Prawdopodobieństwo błędnej transmisji jednego znaku wynosi 0,02. Obliczyć prawdopodobieństwo tego, że podczas transmisji liczba błędów zmieści się w granicach od 10 do 20. Użyć centralnego twierdzenia granicznego.

2. 64% studentów zdaje egzaminy w pierwszym terminie. Obliczyć prawdopodobieństwo tego, że w grupie 5000 studentów w pierwszym terminie zda od 3000 do 3500 studentów. Użyć centralnego twierdzenia granicznego.

3. Rozwiązać zad. 7 z listy 4., stosując centralne twierdzenie graniczne.

4. W pewnej loterii wygrana $W$ wynosi 1, 2 i 10 złotych z prawdopodobieństwem $p_1 = 0.4$, $p_2 = 0.2$ oraz $p_{10} = 0.1$. W pozostałych wypadkach los jest pusty. Podać symulowane wartości wygranych dla 36 losów oraz częstość wygrania 1 zł. Stosując centralne twierdzenie graniczne, obliczyć prawdopodobieństwo tego, że łączna wygrana 36 losów znajdzie się w granicach 50 do 80 zł. Jaka jest łączna wygrana dla wygenerowanej próby?

5. Liczba samochodów przyjeżdżających w ciągu 5 sekund na skrzyżowanie z danego kierunku ma rozkład Poissona z parametrem $\lambda = 4$. Samochody przyjeżdżają niezależnie od siebie. Po zapaleniu się świateł zielonych pierwsze 5 sekund stanowi czas martwy, po czym samochody opuszczają skrzyżowanie w sposób nielosowy po jednym pojeździe na sekundę z każdego z dwóch pasów ruchu. Ile minimalnie sekund w cyklu dwuminutowym należy przeznaczyć na światła zielone dla rozważanego kierunku ruchu, aby z prawdopodobieństwem co najmniej 0.9 wszystkie samochody, które przyjechały na skrzyżowanie w ciągu jednego cyklu, zdążyły opuścić skrzyżowanie?

6. Żywotność nowej baterii ma nieznany rozkład o wartości oczekiwanej $\mu = 100$ godz. i odchyleniu standardowym $\sigma = 20$ godz. Obliczyć prawdopodobieństwo tego, że średnia żywotność baterii z opakowania zawierającego 64 sztuk wyniesie co najmniej 98 godzin. Użyć centralnego twierdzenia granicznego.

**STATYSTYKA STOSOWANA**
**Lista zadań nr 6**

1. Zmierzono pięciokrotnie napięcie w pewnym obwodzie elektrycznym i otrzymano wyniki (w $V$): 12.6 12.3 12.7 12.55 12.4. Oszacować przedziałowo mierzone napięcie oraz odchylenie standardowe błędu pomiaru na poziomie ufności 0.9, zakładając, że napięcie ma rozkład normalny.

2. Z rozważań teoretycznych wynika, że siła powodująca uplastycznienie prętów zbrojeniowych ze stali miękkiej ma rozkład normalny. Zmierzono siłę powodującą uplas-tycznienie 27 wylosowanych prętów o średnicy 12 mm i uzyskano wyniki (w kg): 4280, 4350, 4210, 4290, 4230, 4220, 4230, 4220, 4275, 4225, 4310, 4290, 4325, 4230, 4250, 4210, 4250, 4310, 4220, 4210, 4210, 4205, 4350, 4180, 4155, 4165, 4255. Wyznaczyć przedziały ufności dla parametrów $m$ i $\sigma$ na poziomie ufności 0.95.

3. W badaniu rozrzutu średnic produkowanych wałów wylosowano 150 wałów i zmierzono odchylenia ich średnicy od nominalnej. Otrzymano $\overline{x} = 20\ \mu m$, $\ s^2 = 46\ \mu m^2$. Oszacować parametr $\sigma$ na poziomie ufności 0.98.

4. W celu zbadania wieku osób zatrudnionych w firmie komputerowej wylosowano 36 pracowników i otrzymano

| wiek w latach | 18-20 | 21-22 | 23-24 | 25-26 | 27-28 | 29-30 |
|---|---|---|---|---|---|---|
| liczba osób | 3 | 7 | 10 | 9 | 5 | 2 |

   Narysować histogram. Zakładając, że wiek pracowników ma rozkład normalny, wyznaczyć 96% przedział ufności dla średniego wieku pracownika tej firmy.

5. Wytrzymałość (w $10^5$ N/m$^2$) pewnego materiału budowlanego ma rozkład normalny. Wygenerować 10 prób o liczebności $n = 5$ każda z rozkładu $N(20; 0,6)$ i w oparciu o te próby zbudować 10 przedziałów ufności dla parametru $\mu$ na poziomie ufności 0.7. Ile jest przedziałów, które nie zawierają rzeczywistej wartości parametru $\mu$ tj. 20? Jak to się ma do poziomu ufności 0.7?

6. W celu zbadania niezawodności działania pewnego urządzenia sygnalizującego awarię wykonano 400 niezależnych obserwacji i stwierdzono, że w 330 przypadkach urządzenie pracowało poprawnie. Oszacować przedziałowo prawdopodobieństwo $p$ poprawnego działania urządzenia na poziomie ufności 0.98.

7. W badaniach ankietowych pewna firma ustaliła, że spośród 200 ankietowanych osób 50 byłoby skłonnych skorzystać z jej usług. Z wiarogodnością 95% oszacować odsetek ludzi skłonnych do korzystania z usług tej firmy.

8. Dostawca drewnianych belek twierdzi, że ich średnia wytrzymałość jest równa 40 kg/cm$^2$. Wybrano losowo 25 belek, zmierzono ich wytrzymałość i otrzymano $\overline{x} = 39$, $\ s = 2,2$. Zweryfikować hipotezę o średniej wytrzymałości wynoszącej 40 kg/cm$^2$ przeciwko hipotezie, że wytrzymałość jest mniejsza na poziomie istotności 0.1.

9. Ocenia się, że w pewnym województwie skorzystało bezprawnie z ulgi podatkowej 10% podatników. Istnieje obawa, że zmiana przepisów podatkowych mogła zwiększyć podany odsetek osób nieprawidłowo obliczających płacony przez nich podatek. Wylosowano 150 podatników i wykazano, że 21 z nich niesłusznie skorzystało ze wspomnianej ulgi. Zweryfikować postawioną uprzednio hipotezę na poziomie istotności 0.05.

10. Dla danych z zad.8 zweryfikować hipotezę, że $\sigma \leqslant 2$ przeciwko $\sigma > 2$ na poziomie istotności 0.1.

**STATYSTYKA STOSOWANA**
**Lista zadań nr 7**

1. Wysunięto hipotezę, że użycie innego typu noża tokarskiego skróci czas obróbki pewnego detalu. Wykonano 10 pomiarów czasu obróbki detalu i otrzymano dla nowego noża (w min.): 57, 55, 63, 24, 67, 43, 33, 68, 56, 54. Dla starego noża wyniki były następujące: 58, 58, 56, 38, 70, 42, 75. Na poziomie istotności 0.05 zweryfikować hipotezę o równości średnich czasów obróbki detalu przy użyciu obu noży tokarskich.

2. Dla dwustu próbek pewnego materiału przeprowadzono badanie wytrzymałości na ściskanie i uzyskano wyniki:

| wytrzymałość w MPa | 19-20 | 20-21 | 21-22 | 22-23 | 23-24 | 24-25 |
|---|---|---|---|---|---|---|
| liczba próbek | 10 | 26 | 56 | 64 | 30 | 14 |

   Za pomocą testu zgodności chi-kwadrat zweryfikować hipotezę o normalności rozkładu wytrzymałości na poziomie istotności 0.01.

3. Przeprowadzono obserwacje dziennej liczby wypadków drogowych w pewnym mieście dla losowo wybranych 300 dni i otrzymano wyniki:

| dzienna liczba wypadków | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| liczba dni | 50 | 100 | 80 | 40 | 20 | 10 |

   Za pomocą testu zgodności chi-kwadrat zweryfikować na poziomie istotności 0.02 hipotezę, że dzienna liczba wypadków drogowych w tym mieście ma rozkład Poissona.

4. W ciągu trzech miesięcy zaobserwowano 145 awarii maszyn. Poniżej podano liczby awarii poszczególnych maszyn w czasie każdej zmiany.

| zmiana | maszyna | | | |
|---|---|---|---|---|
| | A | B | C | D |
| 1 | 9 | 5 | 11 | 12 |
| 2 | 9 | 11 | 18 | 20 |
| 3 | 12 | 9 | 12 | 17 |

   Zweryfikować hipotezę, że liczba awarii zależy od zmiany na poziomie istotności 0.05.

5. Postawiono hipotezę, że dzienne zużycie energii elektrycznej w gospodarstwach domowych w dni robocze nie zależy od dnia tygodnia. Dla zweryfikowania tej hipotezy wylosowano 5 tygodni oraz 40 gospodarstw domowych w badanym rejonie i zanotowano zużycie energii elektrycznej w poszczególnych dniach tygodnia. Otrzymano wyniki

| kWh | poniedziałek | wtorek | środa | czwartek | piątek |
|---|---|---|---|---|---|
| 0–6 | 14 | 15 | 9 | 16 | 5 |
| 6–10 | 18 | 16 | 21 | 18 | 25 |
| powyżej 10 | 8 | 9 | 10 | 6 | 10 |

   Zweryfikować postawioną hipotezę na poziomie 0.1.

6. Firma sprzedająca samochody pięciu marek notowała reklamacje, zgłaszane przez kupujących, w ciągu miesiąca od ich sprzedaży. Chce ona wiedzieć, czy rezygnacja ze sprzedaży niektórych marek zmniejszy liczbę zgłaszanych reklamacji w stosunku do liczby sprzedanych samochodów. Wyniki zanotowane przez firmę są następujące:

| marka | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| liczba sprzedanych samochodów | 20 | 30 | 26 | 50 | 18 |
| liczba reklamacji | 4 | 13 | 8 | 8 | 2 |

   Zweryfikować hipotezę, że liczba reklamacji nie zależy od marki samochodu. Przyjąć poziom istotności 0.05. ($X-$ marka samochodu, $Y = 0-$ brak reklamacji, $Y = 1-$ zgłoszenie reklamacji.)

**STATYSTYKA STOSOWANA**
**Lista zadań nr 8**

1. Wyniki kolokwium (w skali 0 - 25 pkt.) oraz egzaminu (w skali 0 -50 pkt) dla grupy 19 studentów wyglądały następująco

| nr studenta | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| kolokwium | 7 | 11 | 12 | 14 | 17 | 15 | 21 | 22 | 19 | 13 | 5 | 12 | 16 | 14 | 21 | 20 | 17 | 10 | 17 |
| egzamin | 20 | 24 | 25 | 30 | 35 | 30 | 43 | 42 | 41 | 24 | 14 | 27 | 35 | 28 | 42 | 38 | 36 | 23 | 40 |

Wyznaczyć współczynniki prostej regresji $y = ax + b$ podającej związek wyniku egzaminu $y$ od wyniku kolokwium $x$. Obliczyć współczynnik determinacji. Wyznaczyć przedział ufności dla parametru $a$ na poziomie 0,9. Jaki jest przewidywany wynik egzaminu dla studenta, który uzyskał 18 pkt. na kolokwium? Wyznaczyć przedział ufności dla Y(18) na poziomie 0.95.

2. W celu zbadania zależności między kątami nachylenia terenu ($x$ w radianach) a wielkością błędów wysokościowych ($y$ w $m$) popełnianych w pewnej metodzie aerotriangulacji, wykonano 9 pomiarów otrzymując wyniki:

| $x$ | $0,157$ | $0,140$ | $0,122$ | $0,105$ | $0,087$ | $0,070$ | $0,052$ | $0,035$ | $0,017$ |
|---|---|---|---|---|---|---|---|---|---|
| $y$ | $0,111$ | $0,097$ | $0,183$ | $0,215$ | $0,214$ | $0,209$ | $0,200$ | $0,178$ | $0,225$ |

Narysować wykres rozproszenia. Przyjmując, że błąd $y$ zależy liniowo od kąta $x$ wyznaczyć oszacowanie współczynników regresji $a$ i $b$. Wyznaczyć przedział ufności dla współczynnika nachylenia $a$ na poziomie ufności 0,95. Jaka jest przewidywana wartość błędu dla kąta nachylenia terenu $7,5^0$ ? Wyznaczyć przedział ufności wartości tego błędu na poziomie ufności 0,9.

3. Zbadano zależność między długością pędów żyta ($x$ w cm) a długością kłosa ($y$ w cm) i otrzymano wyniki

| $x$ | 125 | 145 | 130 | 145 | 150 | 155 | 160 | 175 | 170 | 175 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y$ | 5.0 | 5.5 | 6.0 | 6.5 | 7.0 | 7.5 | 8.0 | 9.0 | 9.5 | 10.0 |

Wyznaczyć równanie prostej regresji zależności $y$ od $x$. Obliczyć współczynnik determinacji oraz wariancję resztkową. Zweryfikować hipotezę o braku wpływu długości pędu na długość kłosa na poziomie istotności 0.05.

4. Zbadano zależność liczby wypadków samochodowych od liczby zarejestrowanych samochodów. Wylosowano 8 powiatów i otrzymano ($x$ liczba zarejestrowanych samochodów w tys., $y$ liczba wypadków w ciągu roku)

| $x$ | 29 | 23 | 18 | 20 | 8 | 27 | 12 | 10 |
|---|---|---|---|---|---|---|---|---|
| $y$ | 110 | 70 | 96 | 83 | 40 | 100 | 65 | 55 |

Wyznaczyć prostę regresji $y = ax + b$. Narysować wykres rozproszenia i zaznaczyć wyznaczoną prostą. Obliczyć współczynnik determinacji oraz aweryfikować hipotezę, że współczynnik regresji $a$ jest co najmniej 3.0 na poziomie istotności 0.05. Oszacować przedziałowo $b$ na poziomie ufności 0.9.

5. Wykonano 8 pomiarów szybkości $y$ rozpuszczania się powłoki srebrnej (w $\mu$m/s) w różnych temperaturach $t$ (w $^o$C) roztworu i otrzymano wyniki

| $t$ | 14 | 15 | 16 | 18 | 20 | 21 | 22 | 24 |
|---|---|---|---|---|---|---|---|---|
| $y$ | 0.31 | 0.35 | 0.35 | 0.38 | 0.40 | 0.42 | 0.43 | 0.44 |

Przyjmując zależność potęgową $y = b\,t^a$ i sprowadzając ją do zależności liniowej przez zlogarytmowanie, znaleźć prostą regresji $\ln y = a \ln t + \ln b$. Sporządzić wykres rozproszenia i nanieść znalezioną krzywą. Zweryfikować hipotezę, że $a = 0.5$ na poziomie 0.1. Przyjąć, że rozkład $\ln Y$ jest normalny.