

ZESTAW ZADAŃ Z ZAGADNIENIA: PARAMTRYCZNE TESTY ISTOTNOŚCI

Test istotności dla średniej

1. Dzielne zużycie wody w fabryce AXA jest zmienną losową o rozkładzie normalnym $N(1000; 20)$. Na podstawie 196 dni roku stwierdzono, że średnie zużycie wody wynosi 1025 m^3 . Na poziomie istotności 0,05 zweryfikować hipotezę, iż średnie rzeczywiste dziennie zużycie wody różni się istotnie od teoretycznego. (wskazówka: σ – znane).

Jeżeli statystyka testowa należy do obszaru krytycznego, to hipotezę zerową ($H_0: m = m_0$) należy odrzucić na korzyść hipotezy alternatywnej.
Jeżeli statystyka testowa nie należy do obszaru krytycznego stwierdzamy brak podstaw do odrzucenia hipotezy zerowej na przyjętym poziomie istotności.

2. Wiadomo, że rozkład wyników pomiarów głębokości morza w pewnym rejonie jest rozkładem normalnym z odchyleniem standardowym 5 metrów. Dokonano sześciu niezależnych pomiarów głębokości morza w tym rejonie i otrzymano wyniki (w metrach): 860, 867, 876, 866, 871, 855. Na poziomie ufności 0,05 zweryfikować hipotezę, że średnia głębokość morza (wskazówka: σ – znane):
- wynosi 870 metrów;
 - jest mniejsza od 870 metrów;
 - jest większa od 870 metrów.

| | | | | | | | | | |
|--|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Poziom istotności: α | 0,01 | 0,02 | 0,03 | 0,04 | 0,05 | 0,06 | 0,07 | 0,08 | 0,1 |
| Wartość krytyczna Z_α | 2,58 | 2,33 | 2,17 | 2,05 | 1,96 | 1,88 | 1,81 | 1,75 | 1,64 |

3. W 81 zakładach pewnej gałęzi przemysłu zbadano koszty materiałowe przy produkcji pewnego wyrobu i otrzymano średnią 100 zł przy odchyleniu standardowym 28 zł. Na poziomie istotności $\alpha = 0,05$ zweryfikować hipotezę, że średnie koszty materiałowe przy produkcji tego wyrobu wynoszą 110zł (wskazówka: σ – nieznane, duża próba).
4. Wymagania dietetyczne dotyczące posiłków w stołówkach studenckich wskazują, że norma dzienna posiłków to 3400 kalorii. W wybranej stołówce uniwersyteckiej zbadano kaloryczność 144 posiłków i stwierdzono, że norma to 3350 kalorii z 7% poziomem zmienności. Czy na tej podstawie można twierdzić, że kierownictwo stołówki podaje posiłki o kaloryczności o obniżonej normie, zakładając poziom istotności 0,075? (wskazówka: σ – nieznane, duża próba).
5. Zebrano informacje dla grupy kierowców, którzy w okresie ostatnich 10 lat na terenie pewnego miasta w województwie śląskim spowodowali wypadek drogowy znajdując się pod wpływem alkoholu. Uzyskano następujące zestawienie:

| | | | | | |
|--------------------------------------|---------|---------|---------|---------|---------|
| Poziom alkoholi w krwi (w promilach) | 0,4-1,0 | 1,0-1,6 | 1,6-2,2 | 2,2-2,8 | 2,8-3,0 |
| Liczba kierowców | 15 | 100 | 170 | 120 | 30 |

Zakładając, że badana populacja ma charakter rozkładu normalnego przy poziomie istotności 0,05 zweryfikować hipotezę, że średnie stężenie alkoholu we krwi w całej populacji nietrzeźwych kierowców, którzy spowodowali wypadek drogowy, jest większe od 2,3 promila (wskazówka: σ – nieznane, duża próba).

6. W stołówce studenckiej przeprowadzono wrywkową kontrolę masy porcji obiadowej mięsa, która nominalnie powinna wynosić 120g. Losowo wybrano a następnie zważono 10 porcji, uzyskując następujące wyniki pomiarów w gramach: 122, 118, 116, 115, 123, 125, 116, 114, 120, 121. Na poziomie istotności $\alpha = 0,05$ sprawdzić hipotezę, że studenci w badanej stołówce są żywieni zgodnie z recepturą, czyli otrzymywana przez nich waga mięsa wynosi co najmniej 120g. Zakłada się, że rozkład masy porcji mięsa w całej populacji jest rozkładem normalnym (wskazówka: σ – nieznane, małą próbą).
7. Okresowe badania lekarskie przeprowadzone wśród dziennikarzy dwóch regionalnych gazet dotyczyły między innymi ilości wypalanych przez nich dziennie papierosów. Informacje dotyczące gazety GW przedstawia tabela:

| | | | | | |
|-------------------|-----|------|-------|-------|-------|
| Liczba papierosów | 1-9 | 9-19 | 19-29 | 29-39 | 39-49 |
| Liczba osób | 15 | 35 | 70 | 20 | 10 |

Dla 100 dziennikarzy gazety WZ ustalono, że średnia ilość wypalonych przez nich papierosów wynosi 28 sztuk/osobę, przy odchyleniu standardowym 7 sztuk.

Sprawdzić hipotezę, że średnia liczba wypalonych papierosów przez dziennikarzy obu gazet wynosi co najwyżej 18 sztuk dziennie/osobę. Przyjąć poziom istotności 0,05.

Test istotności dla dwóch średnich

8. Rozkład tygodniowego czasu poświęconego na naukę poza uczelnią studentów I roku UE jest rozkładem normalnym $N(m, 5)$, natomiast studentów II roku rozkład czasu nauki poza uczelnią jest rozkładem normalnym $N(m, 6)$. Pobrane losowo próby 10 studentów I roku i 24 studentów z II roku pozwoliły na wyznaczenie średniego czasu nauki równego odpowiednio: 20 godzin i 15 godzin. Czy na podstawie tych informacji możemy twierdzić, że studenci I roku uczą się przeciętnie dłużej niż studenci II roku? Przyjąć poziom istotności 0,01. (Wskazówka: σ_1, σ_2 – wartości znane).
9. Zaistniało podejrzenie, że pierwsza z dwóch obrabiarek pracujących na wydziale R dostarcza elementy zawyżonej długości. Wybrano losowo i zmierzono 200 elementów wytworzonych na obrabiarce pierwszej i 100 elementów otrzymanych z drugiej obrabiarki. Po przeprowadzeniu obliczeń uzyskano: $\bar{X}_1 = 20,9mm$, $\bar{X}_2 = 20,1mm$, $S^2(X_1) = 0,8mm^2$. $S^2(X_2) = 0,2mm^2$. Na poziomie istotności $\alpha = 0,05$ zweryfikować hipotezę, że obie obrabiarki wytwarzają elementy o tej samej długości. (Wskazówka: σ_1, σ_2 – wartości nieznane, duża próba).
10. Badając próby losowe liczące po 150 pracowników zakładów A i B, otrzymano następujące wyniki:

- Dla zakładu A: poniższa tabela

| | | | | | | |
|--------------------|-----|-----|------|-------|-------|-------|
| Staż pracy | 0-4 | 4-8 | 8-12 | 12-16 | 16-20 | 20-24 |
| Liczba pracowników | 6 | 15 | 60 | 45 | 15 | 9 |

- Dla zakładu B: $Me = 13$ lat, $D = 11$ lat, $V(x) = 40\%$.

Sprawdzić, czy średnie stażu pracy zatrudnionych w tych dwóch zakładach różnią się od siebie istotnie. (Wskazówka: σ_1, σ_2 – wartości nieznane, duża próba).

$$\bar{x} - D = 3(\bar{x} - Me); AS = \frac{\bar{x} - D}{S(x)}; V(x) = \frac{S(x)}{\bar{x}}$$

11. Wylosowano 100 pracowników pewnego przedsiębiorstwa (60 mężczyzn i 40 kobiet) i zbadano ich pod względem wieku. Wyniki są następujące:
- MĘŻCZYŹNI:
 - Średnia arytmetyczna: 40;
 - Dominanta: 35;
 - Współczynnik skośności Pearsona: +0,5
 - KOBIECY:
 - Średnia arytmetyczna: 30;
 - Dominanta: 35;
 - Współczynnik skośności Pearsona: -0,5

Czy na poziomie istotności 0,05 można twierdzić, że kobiety w tym przedsiębiorstwie są, przeciętnie rzecz biorąc, młodsze od mężczyzn? (Wskazówka: σ_1, σ_2 – wartości nieznane, duża próba).

12. W pewnym przedsiębiorstwie zbadano pracowników dwóch zmian ze względu na czas obróbki detalu A. Dla losowo wybranych 10 pracowników zmiany I i 12 pracowników zmiany II otrzymano następujące wyniki (w minutach):
- Zmiana I: $S^2(x) = 10,24, V(x) = 35,6\%$.
 - Zmiana II: 10, 7, 5, 4, 9, 6, 11, 13, 12, 6, 7, 8

Czy można twierdzić, że zmienowość ma istotny wpływ na czas obróbki detalu A w tym przedsiębiorstwie? Przyjąć poziom istotności 0,05. (Wskazówka: σ_1, σ_2 – wartości nieznane, mała próba, $n_1 \neq n_2$).

Test istotności dla wariancji

13. Pewnego dnia dokonano 120 pomiarów opóźnień pociągów kursujących do Stacji Kraków, na podstawie otrzymanych wyników ustalono: $\bar{X} = 25$ i $S(X) = 26$. Zakłada się, że czas opóźnień pociągów ma rozkład normalny $N(m, \sigma)$. Sprawdzić na poziomie istotności $\alpha = 0,05$ hipotezę, że wariancja opóźnień wynosi 400.
14. Analizując wydajność pracy w szt/h w pewnym zakładzie dla 50 losowo wybranych pracowników, uzyskano odchylenie standardowe $S = 3,78$ szt/h zweryfikować na poziomie istotności $\alpha = 0,01$, czy wariancja wydajności przekracza 12sz/h.
15. Pewnego dnia dokonano 15 pomiarów opóźnień pociągów kursujących do Stacji Kraków, otrzymano w minutach: 20, 25, 18, 28, 14, 19, 12, 14, 19, 17, 15, 12, 21, 18. Zakłada się, że czas opóźnień pociągów ma rozkład normalny $N(m, \sigma)$. Sprawdzić na poziomie istotności $\alpha = 0,05$ hipotezę, że wariancja opóźnień wynosi 400.
16. W 7 losowo wybranych ulach pasiek nowosądeckich zmierzono ilość wybieranego jednorazowo miodu (w litrach): 5,36; 5,2; 5,28; 5,16; 5,3; 5,08; 5,23. Przyjmując poziom istotności $\alpha = 0,05$ zweryfikować sąd, że wariancja ilości wybieranego miodu przekracza 0,01 litrów².
17. Analiza stażu pracy pracowników w przedsiębiorstwie FERM przedstawia się następująco: 6 pracowników o stażu pracy 0-5 lat, 8 pracowników o stażu pracy 5-10 lat, 16 pracowników o stażu 10-15 lat, 14 pracowników o stażu pracy 15-20 lat. Przyjmując poziom istotności $\alpha = 0,01$ zweryfikować hipotezę, że odchylenie standardowe stażu pracy jest równe 4 lata.
18. Z dwóch kompanii wojskowych wybrano losowo po 12 szeregowców. Odchylenie standardowe wzrostu wylosowanych szeregowców w pierwszej kompanii wynosi 2,5 cm a w drugiej kompanii 2,3 cm. Zakładając, że rozkład wzrostu szeregowców w obu kompaniach jest rozkładem normalnym, zweryfikować hipotezę, że wariancje wzrostu szeregowców w obu kompaniach są jednakowe. Przyjąć poziom istotności 0,04. (Wskazówka: σ_1, σ_2 – wartości nieznane, mała próba, $n_1 = n_2$).

Test istotności dla wskaźnika struktury

19. W AR w Krakowie studiuje 4000 studentów. Niedawno powstał projekt przeznaczenia dla palących specjalnego pomieszczenia, o ile okaże się, że jest ich więcej niż niepalących. Celem sprawdzenia hipotezy, że frakcja palących przekracza 50% wylosowano próbę złożoną z 400 studentów z wylosowanych osób 240 oświadczyło, że stale lub sporadycznie pali papierosy. Przeprowadzić postępowanie testowe celem podjęcia właściwej decyzji przyjmując poziom istotności $\alpha = 0,02$.
20. W sezonie wakacyjny „LATO 2019” przeprowadzono w wybranym dniu ankietę wśród turystów w Szczawnicy na temat miejsca zakwaterowania. W grupie ankietowanych turystów: 105 turystów mieszkało w sanatoriach, 100 osób mieszkało na kwaterach prywatnych, 85 osób w pensjonatach, a 120 turystów było w Szczawnicy przejazdem. Przyjmując poziom istotności 0,01, zweryfikować hipotezę, że ponad połowa turystów zamieszkała w pensjonatach lub sanatoriach.
21. Celem sprawdzenia hipotezy, że zachorowalność na miażdżycę jest taka sama w środowisku miejskim, jak i wiejskim pobrano dwie próby losowe. W grupie 800 osób wylosowanych z miasta stwierdzono 60 przypadków zachorowań na miażdżycę, wśród 1200 wylosowanych osób mieszkających na wsi odnotowano również 60 przypadków tej choroby. Przyjmując poziom istotności $\alpha = 0,05$ należy zweryfikować postawioną na wstępie hipotezę.

Test istotności dla współczynnika korelacji

22. Z populacji studentów kończących Akademię Ekonomiczną wylosowano niezależnie 10 osób. Następnie odnotowano liczbę punktów jaką każdy z nich uzyskał na egzaminie wstępnym (X) oraz średnią ocenę z wszystkich egzaminów na studiach (Y). Dane te pozwoliły wyznaczyć wartość współczynnika korelacji liniowej między cechami $r_{xy} = 0,462$. Zweryfikować na poziomie istotności $\alpha = 0,05$ hipotezę, że istnieje zależność między wynikami egzaminu wstępnego a ocenami osiągniętymi w toku studiów.
23. Wylosowano 9 par zawierających związek małżeński i otrzymano dla nich następujące dane o wieku nowożeńców w latach:

| | | | | | | | | | | |
|---------------|----|----|----|----|----|----|----|----|----|----|
| Wiek kobiet | 23 | 29 | 33 | 19 | 21 | 24 | 27 | 29 | 22 | 30 |
| Wiek mężczyzn | 27 | 30 | 34 | 22 | 26 | 28 | 30 | 30 | 25 | 31 |

Zweryfikować na poziomie istotności $\alpha = 0,05$ hipotezę, że istnieje dodatnia korelacja między wiekiem osób zawierających związek małżeński.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Interpretacja:

$0,0 < |r_{xy}| < 0,2$ - zależność nieistotna statystycznie,

$0,0 < |r_{xy}| < 0,3$ - słaba zależność,

$0,3 \leq |r_{xy}| < 0,7$ - umiarkowana zależność,

$0,7 < |r_{xy}| < 1,0$ - silna zależność.

24. W losowej próbie 24 gospodarstw domowych woj. Śląskiego przeprowadzono badanie zależności między miesięcznym spożyciem ziemniaków (y) a miesięcznym dochodem (x). Otrzymane wyniki są następujące:

$$\bar{x} = 77, S(x) = 46,2, \bar{y} = 15, V(y) = 30\%, a_1(y) = -0,08.$$

Należy sprawdzić, czy współczynnik korelacji między tymi zmiennymi w całej populacji gospodarstw domowych woj. Śląskiego jest statystycznie istotny (przyjąć $\alpha = 0,01$).

$$\hat{y} = a_1x + a_0$$

gdzie:

$$a_1 = \frac{cov(x,y)}{S^2(x)} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = r_{xy} \frac{S(y)}{S(x)}$$

$$a_0 = \bar{y} - a_1 \cdot \bar{x}$$

$$\hat{x} = b_1y + b_0$$

gdzie:

$$b_1 = \frac{cov(x,y)}{S^2(y)} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} = r_{xy} \frac{S(x)}{S(y)}$$

$$b_0 = \bar{x} - b_1 \cdot \bar{y}$$

$$V(x) = \frac{S(x)}{\bar{x}} 100\%; V(y) = \frac{S(y)}{\bar{y}} 100\%; r_{xy} = \frac{cov(x,y)}{S(x) \cdot S(y)};$$

25. Za zbiorowości uczniów szkół ponadpodstawowych wybrano losowo 150 osób i ustalono dla nich zależność pomiędzy wysokością dochodów rodziców a wysokością tygodniowego kieszonkowego u dziecka. Obliczony współczynnik korelacji liniowej był równy 0,62. Zweryfikować hipotezę, że dla ogółu uczniów szkół ponadpodstawowych zależność pomiędzy wysokością dochodów rodziców a wysokością tygodniowego kieszonkowego przekracza 43% (mierzona współczynnikiem determinacji). (Wskazówka $d = (r_{xy})^2$).

26. Restauracja „WIERZYNEK” urządzająca przyjęcia weselne podała dane dotyczące kosztu przyjęcia i czasu jego trwania dla przyjęć organizowanych w I kwartale 2019 roku. Przedstawia je tabela.

| Czas trwania w godzinach | Koszt przyjęcia w tys. zł | | | | | |
|--------------------------|---------------------------|-----|------|-------|-------|-------|
| | Do 4 | 4-8 | 8-12 | 12-16 | 16-20 | 20-24 |
| 0-2 | 2 | - | 3 | - | - | - |
| 2-4 | 3 | - | - | - | - | - |
| 4-6 | - | 8 | 10 | 11 | - | 3 |
| 6-8 | - | 6 | 12 | 16 | 4 | - |
| 8-10 | - | - | - | 7 | 2 | 3 |

Na poziomie istotności 0,05 zweryfikować hipotezę, że współczynnik korelacji liniowej pomiędzy czasem trwania przyjęcia weselnego i jego kosztem przekracza 0,5.

$$r_{xy} = \frac{\text{cov}(x, y)}{S(x) \cdot S(y)}; \quad r_{xy} \in \langle -1, 1 \rangle$$

gdzie:

- Dla danych indywidualnych (w postaci szeregu korelacyjnego):

$$\text{cov}(x, y) = \text{cov}(y, x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}),$$

$$\text{cov}(x, y) = \text{cov}(y, x) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}$$

- Dla danych ujętych w formie tablicy korelacyjnej:

$$\text{cov}(x, y) = \text{cov}(y, x) = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^r (x_i - \bar{x})(y_j - \bar{y})n_{ij},$$

$$\text{cov}(x, y) = \text{cov}(y, x) = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^r x_i y_j n_{ij} - \bar{x} \cdot \bar{y}$$

Sprawdź czy umiesz

1. Przeprowadzono badanie statystyczne, gdzie dla dwóch losowo wybranych prób w pewnym zakładzie uzyskano informacje dotyczące wynagrodzenia w dwóch resortach. Wyniki przedstawiono w tabeli.

| Płace księgowych (w tys. zł) | 1,6-2,0 | 2,0-2,4 | 2,4-2,8 | 2,8-3,2 | 3,2-3,6 | 3,6-4,0 | 4,0-4,6 |
|------------------------------|---------|---------|---------|---------|---------|---------|---------|
| Resort I | 10 | 30 | 40 | 100 | 200 | 100 | 20 |
| Resort II | 40 | 60 | 150 | 100 | 40 | 10 | - |

Na poziomie istotności 0,01 zweryfikować hipotezę o tym, że

- średnia płaca księgowych jest taka sama bez względu na resort, w którym pracują. (Wskazówka: test dla dwóch średnich, σ_1, σ_2 – wartości nieznane, duża próba).
- odsetek osób zarabiających co najmniej 3200zł jest większy w resorcie I niż w resorcie II.

Literatura do zajęć

- 1) Maria Balcerowicz-Szkutnik, Elżbieta Sojka, Włodzimierz Szkutnik, (2016), *Wnioskowanie statystyczne w przykładach i zadaniach*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
- 2) Maria Balcerowicz-Szkutnik, Elżbieta Sojka, Włodzimierz Szkutnik, (2014), *Statystyka opisowa dla ekonomistów. Przykłady i zadania*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
- 3) S. Ostasiewicz, Z. Rusnak, U. Siedlecka (2011), *Statystyka - elementy teorii i zadania*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu.
- 4) Mieczysław Sobczyk, *Statystyka*, (2019), Wydawnictwo Naukowe PWN.