

INFORMATYKA

8. Test dobroci dopasowania i niezależności chi-kwadrat

1. Stowarzyszenie Russela Reynolda przeprowadziło ankietę wśród emerytowanych menedżerów wyższego szczebla, którzy wrócili do pracy. Ustalili, że po powrocie do pracy 38% było zatrudnionych w innej organizacji, 32% samozatrudnionych, 23% było freelancerami lub konsultantami, a 7% założyło własne firmy. Aby sprawdzić, czy te wartości procentowe są zgodne z odsetkami mieszkańców hrabstwa Allegheny, lokalny badacz przeprowadził ankietę wśród 300 emerytowanych menedżerów, którzy wrócili do pracy, i odkrył, że 122 pracowało dla innej firmy, 85 prowadziło działalność na własny rachunek, 76 pracowało jako freelancer lub doradzało, a 17 założyło własne firmy. Czy przy istotności 10% dane potwierdzają, że odsetki poszczególnych zatrudnionych w hrabstwie Allegheny różnią się od ich odpowiedników w skali całego kraju?
 $\chi^2=3.2939$, $p\text{-val}=0.3485$, brak podstaw do odrzucenia H_0

2. Badacz przeczytał w artykule, że liczba zgonów związanych z bronią palną wśród osób w wieku od 1 do 18 lat rozkładała się następująco: 74% to wypadki, 16% to zabójstwa, a 10% to samobójstwa. W jej okręgu w ubiegłym roku doszło do 68 wypadków śmiertelnych, 27 zabójstw i 5 samobójstw. Czy na poziomie istotności 10% dane potwierdzają, że odsetki poszczególnych zgonów różnią się od przedstawionych w artykule?
 $\chi^2=10.549$, $p\text{-val}=0.005$, odrzucamy H_0

3. M&M/Mars, producent cukierków Skittles, twierdzi, że mieszanka smakowa wynosi 20% dla każdego smaku. Skittles to połączenie cukierków o smaku cytrynowym, limonkowym, pomarańczowym, truskawkowym i winogronowym. Poniższe dane przedstawiają wyniki czterech losowo wybranych torebek Skittles i ich mieszanek smakowych.

Torebka	Smaki				
	Malinowy	Truskawkowy	Cytrynowy	Jagodowy	Pomarańczowy
1	7	20	10	7	14
2	20	5	5	13	17
3	4	16	13	21	4
4	12	9	16	3	17

Wykonaj test, aby porównać zaobserwowane (całkowite) wartości z wartościami oczekiwanymi. Załóżmy poziom istotności 0,05.

$$\chi^2=1.4421, p\text{-val}=0.837, \text{ brak podstaw do odrzucenia } H_0$$

4. Poniższe przykładowe dane przedstawiają stężenie ozonu (mierzone w częściach na 100 milionów) w powietrzu w centrum miasta przez 78 kolejnych letnich dni w 2004 roku:

3.5	1.7	3.1	4.5	3.0	3.7	4.1	9.4	2.5	5.7	1.4	4.7	6.1
6.8	1.1	5.8	4.2	6.0	7.6	3.5	5.3	3.0	4.4	3.9	1.6	8.1
2.4	7.5	4.7	5.4	1.4	6.6	5.9	4.7	5.1	4.7	4.4	6.8	2.0
6.8	5.8	5.7	6.5	2.8	4.1	6.0	6.7	6.2	5.8	3.4	5.4	6.2
5.5	3.4	6.0	7.4	2.5	3.7	5.6	1.4	7.6	3.3	9.4	5.6	5.6
6.2	3.1	4.4	5.5	3.7	5.8	6.6	6.6	3.8	5.3	6.6	11.7	4.0

Powyższe dane można pogrupować w przedziały w następujący sposób:

stężenie ozonu	0-2	2-4	4-6	6-8	8-10	10-12
częstotliwość	7	19	31	17	3	1

Należy pamiętać, że do ostatnich klas kwalifikuje się mniej niż 5 pomiarów. Na poziomie istotności 0,05 sprawdź, czy stężenie ozonu ma rozkład normalny.

$$\chi^2=0.93, p\text{-val}=0.92, \text{ brak podstaw do odrzucenia } H_0$$

5. Poniższe dane dotyczą objętości guza (w mm³) w grupie losowo wybranych 100 myszy:

Objętość guza	(0; 2]	(2; 4]	(4; 6]	(6; 8]	(8; 10]
Numer myszy	10	25	35	20	10

Na poziomie istotności 0.1 zweryfikuj czy objętość guza ma rozkład normalny.

$$\chi^2=0.74442, p\text{-val}=0.9457, \text{ brak podstaw do odrzucenia } H_0$$

6. Poniższe dane przedstawiają liczbę punktów uzyskanych przez grupę studentów na koniec semestru:

23	60	79	32	57	74	52	70	82	36	80	77	81	95	41
65	92	85	55	76	52	10	64	75	78	25	80	98	81	67
41	71	83	54	64	72	88	62	74	43	60	78	89	76	84
48	90	15	79	84	34	67	17	82	69	74	63	80	85	61

Na poziomie istotności 0,01 sprawdź, czy oceny tej grupy uczniów mają rozkład normalny.

7. Socjolog pragnie sprawdzić, czy liczba lat nauki danej osoby ma związek z jej miejscem zamieszkania. Wybrano próbę 88 osób i sklasyfikowano, jak pokazano.

Obszar	Brak uczelni	4-letnie studia	Stopień zaawansowany
Miejski	15	12	8
Podmiejski	8	15	9
Wiejski	6	8	7

Czy na poziomie istotności 0,05 socjolog może stwierdzić, że miejsce zamieszkania danej osoby zależy od liczby lat studiów?

$$\chi^2=3.0055, p\text{-val}=0.557, \text{ brak podstaw do odrzucenia } H_0$$

8. Badacz wybrał 100 pasażerów każdej z 3 linii lotniczych i zapytał ich, czy linia lotnicza zgubiła ich bagaż podczas ostatniego lotu. Dane przedstawiono w tabeli. Czy na poziomie istotności 0,05 dane potwierdzają, że odsetek pasażerów, którzy zagubili bagaż w trakcie lotu, zależy od linii lotniczej?

	Linia lotnicza 1	Linia lotnicza 2	Linia lotnicza 3
Tak	10	7	4
Nie	90	93	96

$$\chi^2=2.765, p\text{-val}=0.251, \text{ brak podstaw do odrzucenia } H_0$$

9. W Senacie planowane jest głosowanie nad projektem ustawy zezwalającej na instalowanie anten satelitarnych dowolnej wielkości na obszarach objętych ograniczeniami wykonawczymi. Podobną ustawę przyjęła Izba. Przeprowadzono badanie opinii publicznej, aby sprawdzić, czy odczucia danej osoby w związku z ograniczeniami dotyczącymi anten satelitarnych są powiązane z jej wiekiem. Uzyskano następujące dane:

	18–29	30–49	50–64	65 i więcej
Za	96	96	90	36
Przeciw	201	189	195	234
Nie wiem	3	15	15	30

Czy dane potwierdzają, że opinia zależy od wieku? Załóżmy poziom istotności 0,05.

$$\chi^2=61.247, p\text{-val}=0, \text{ odrzucamy } H_0$$