

Istotność statystyczna.

Chcemy, na podstawie danych, sprawdzić hipotezę H . Rozstrzygnięcie, czy jest ona prawdziwa, zależy od wartości pewnego kryterium¹: im wyższa wartość tego kryterium tym bardziej prawdziwa hipoteza H . Istotność H to prawdopodobieństwo², że wartość kryterium jest co najmniej równa wartości zaobserwowanej w próbce, obliczone przy założeniu, że H nie jest prawdziwa. W praktyce przyjmuje się, że dane potwierdzają H , gdy jej zaprzeczenie jest mało prawdopodobne³.

1. Rzuca się monetą tak długo dopóki nie pojawi się strona monety przeciwna do tej, która wypadła za pierwszym razem. Liczba rzutów jest kryterium dla hipotezy: *Moneta jest niesymetryczna*. Na przykład, gdy wypadnie RRRO to wartość kryterium wynosi 4. Jaka najmniejsza wartość tego kryterium wystarczy, aby przyjąć na poziomie 0,05 hipotezę o braku symetrii? A jaka, aby poziom istotności był mniejszy niż 0,01?

2. Rzucono 10 razy monetą. Wypadło 7 orłów. Sprawdź hipotezę, że moneta jest niesymetryczna. Jako kryterium przyjmij moduł różnicy między liczbą orłów a liczbą reszek.

- Jaki jest poziom istotności tej hipotezy?
- Jaką podejmiesz decyzję: czy moneta jest symetryczna, czy niesymetryczna?
- Jaka liczba orłów musiałaby wypaść, aby poziom istotności był mniejszy niż 0,05?
- Jaka, aby poziom istotności był mniejszy niż 0,01?

3. Zawartość nikotyny w pięciu próbkach papierosów gatunku A i B wynosiła:

A: 1,1 0,8 1,0 0,9 0,8 i B: 1,1 1,7 1,7 1,1 1,1

- Czy uważasz, że papierosy gatunku B zawierają więcej nikotyny? Dlaczego?
- Zweryfikuj, na poziomie 0,05 tę hipotezę, przyjmując jako test różnicę między średnią zawartością nikotyny dla gatunku B a zawartością nikotyny dla gatunku A (patrz ramka).

Sposób obliczenia poziomu istotności, przedstawiony poniżej nazywa się **testem permutacyjnym**.

- Wyniki pomiarów dla gatunku A zapisz na 5 kartkach, podobnie na 5 kartkach zapisz wyniki pomiarów dla gatunku B.
- Gdyby hipoteza nie była prawdziwa, tzn. gdyby nie było różnicy między gatunkami papierosów to 10 kartek można swobodnie potasować. Zrób to.
- Wylosuj 5 kartek, reprezentujących gatunek A. Pozostałe kartki niech reprezentują gatunek B. Oblicz różnicę wartości średnich grupie B i A. Czy jest ona co najmniej równa 0,52⁴?
- Powtórz kilkakrotnie to losowanie. Czy często zdarza się, że różnica średnich w grupie B i A jest co najmniej równa 0,52? Czy na podstawie twoich obserwacji tego losowania można podejrzewać, że pod względem zawartości nikotyny nie ma różnicy między obu gatunkami?
- Oblicz ile jest możliwych losowań sprzyjających zdarzeniu, że różnica średnich na kartkach, reprezentujących grupy B i A jest co najmniej równa 0,52.
- Oblicz, ile jest różnych podziałów 10 kartek na dwie grupy po 5 kartek.
- Oblicz prawdopodobieństwo, że różnica średnich w grupie B i A jest co najmniej równa 0,52. Liczba ta jest poziomem istotności dla badanej hipotezy.

¹ Takie kryterium nosi nazwę testu hipotezy H .

² Nazywane ***p*-wartością** lub ***poziomem istotności testu hipotezy H*** (w skrócie poziomem istotności hipotezy H)

³ Za małe wartości przyjmuje się mniejsze od 0,05 a bardzo małe mniejsze od 0,01

⁴ Jest to wartość testu dla naszej próbki.