

Permutacijski i randomizacijski testovi

STATISTIČKI PRAKTIKUM 2

Randomizacijski testovi čine grupu statističkih testova u kojima se distribucija testne statistike određuje permutacijom dobivenih podataka.

Prednosti:

- ▶ ne ovise o distribuciji iz koje dolaze sami podaci
- ▶ ne zahtijevaju dovoljno veliki uzorak (čak i za male uzorke, vjerojatnost pogreške ne prelazi 0.05)
- ▶ p-vrijednosti se (u slučaju permutacijskih testova) računaju egzaktno
- ▶ ne ovise o uzimanju poduzoraka iz neke prilagođene populacije
- ▶ lako se poopćuju do testova s nekoliko nezavisnih djelujućih tretmana

Poseban slučaj: *permutacijski testovi*, kada se distribucija testne statistike i pripadne p vrijednosti računaju egzaktno (budući da su sve permutacije početnih podataka jednako vjerojatne).
Za velike skupove podataka koristimo aproksimativne metode.

Primjer

Bila jednom jedna britanska gospođa koja je tvrdila da po okusu može prepoznati što je bilo prvo uliveno u šalicu, čaj ili mlijeko ... Fisher je napravio eksperiment da bi njezine tvrdnje testirao:

- ▶ od 8 šalica, u 4 je prvo uliveno mlijeko, a u 4 čaj
- ▶ šalice su prezentirane po slučajnom redoslijedu
- ▶ gospođa ih je probala i na temelju okusa dala ocjenu za svaku šalicu (pritom je znala da od svakog okusa ima 4 šalice)

Formiranje nulte hipoteze i testiranje

H_0 : gospođa nema sposobnost raspoznavanja razlike u okusima

Ako je H_0 istinita, svi mogući rasporedi njezinih odgovora (sve permutacije) u odnosu na istinu o šalicama moraju biti jednako vjerojatni (naime, raspored prezentiranja šalica je odabran na slučajan način).

Testna statistika:

T = broj pogođenih okusa ili broj točnih odgovora za šalice s mlijekom

Distribucija testne statistike

U uvjetima H_0 , razlog za njezinu odluku nije poznat, osim što nema veze s istinom o šalicama. Taj niz fiksiramo. Na koliko načina možemo odabrati raspored za prezentaciju šalica? Na $8!$ načina. Međutim, nama je samo bitno koji se okus, od dva moguća, nalazi u pojedinoj šalici. Stoga imamo onoliko rasporeda koliko ima načina da odaberemo 4 šalice u koje ćemo prvo uliti čaj:

$$\binom{8}{4} = 70.$$

Svi rasporedi su jednako vjerojatni. Samo jedan raspored je u potpunosti identičan kao i gospođini odgovori, a njegova je vjerojatnost pod uvjetima H_0 jednaka $\frac{1}{70} = 0.0413 < 0.05$. Dakle, ako je gospođa u potpunosti pogodila okuse, H_0 bismo odbacili na nivou značajnosti $\alpha = 0.05$.

U ostalim slučajevima:

$$\mathbb{P}(T = t) = \frac{\binom{4}{t} \binom{4}{4-t}}{\binom{8}{4}}, \quad t = 0, \dots, 4.$$

	t	P(T=t)
[1,]	0	0.01428571
[2,]	1	0.22857143
[3,]	2	0.51428571
[4,]	3	0.22857143
[5,]	4	0.01428571

Da je gospođa ukupno probala 10 šalica, mogli bismo odbaciti H_0 i u slučaju da nije savršeno pogodila okuse.

Permutacijski testovi

Općenito, postupak formiranja testa je sljedeći:

1. definirati testnu statistiku T
2. odrediti skup mogućih vrijednosti od T (dobivenih permutacijom podataka)
3. pod uvjetima H_0 , svi mogući rasporedi (permutacije) podataka su jednako vjerojatni
4. odrediti distribuciju statistike T : $\mathbb{P}(T = t)$ je proporcija onih rasporeda kojima se dobila vrijednost t
5. p -vrijednost testa: vjerojatnost da se dobije vrijednost statistike T koja je jednako ili više "ekstremna" od one opažene (trenutne), tj. koja jednako ili još više od opažene potvrđuje alternativnu hipotezu

Nedostaci

Glavni nedostatak – **teško je izračunati egzaktnu distribuciju** čim broj podataka malo naraste. S pojavom računala to je djelomično olakšano. Ipak, čak i uz pomoć računala, računanje svih permutacija može biti nemoguće ili neefikasno.

Rješenja

- ▶ umjesto podataka, promatrati njihove rangove
 - ▶ današnji neparametarski testovi s rangovima su u principu permutacijski testovi napravljeni nad rangovima
- ▶ promatrati samo dio permutacija (**randomizacijski testovi**)
 - ▶ na slučajan način permutirati podatke i izračunati testnu statistiku
 - ▶ ponoviti puno puta: dobivamo uzorak $t_1, \dots, t_B$
 - ▶ Zakon velikih brojeva: p -vrijednost aproksimiramo proporcijom slučajnih vrijednosti statistike T koje su jednake ili premašuju opaženu vrijednost t_0 (početnu realizaciju)

Primjer: nekategorijski podaci

Imamo dvije populacije, iz svake su uzeta 4 podatka (podaci se mogu kombinirati). Dobiveni brojevi su:

grupa 1	grupa 2
11	2
14	9
7	0
8	5
aritm. sredina 10	aritm. sredina 4

T = razlika između aritm. sredina
realizacija: $t = \mathbf{6}$

Permutacija podataka: na koliko se načina može podijeliti ovih 8 brojeva u dvije skupine po 4? $\binom{8}{4} = 70$

H_0 : nema razlike među populacijama

H_a : Grupa 1 ima veće očekivanje

p -vrijednost: proporcija onih rasporeda za koje je razlika u aritmetičkim sredinama jednaka ili veća od trenutne (tj. one koja se dobila na danim podacima).

Veću razliku od 6 možemo dobiti u još samo dva rasporeda (zamjena brojeva 9 i 7 ili 9 i 8). Dakle:

$$p - vr. = \frac{3}{70} = 0.0429.$$

Za dvostrani test uzeli bismo u obzir i slučajeve gdje razlika ide u korist druge grupe pa je

$$p - vr. = \frac{6}{70} = 0.0857.$$