

Statistično sklepanje

Preverjanje hipotez

dr. Nino Rode

1

Sklepanje na podlagi vzorcev

- metoda za **statistično odločanje** na podlagi empiričnih podatkov
- Na podlagi vzorca nikoli **ne moremo dokončno potrditi** hipoteze, lahko pa jo **zavrremo**
- Ugotavljamo, če rezultati vsebujejo **dovolj informacije**, da lahko utemeljeno **podvomimo** v takšno hipotezo
- kako močno so dobljeni podatki **podobni** podatkom, ki bi jih dobili, če bi delovali **samo slučajni dejavniki**

2

Sklepanje na podlagi vzorcev:

Primer aritmetične sredine

- Statistike (izračuni) iz vzorca so mešanica parametra in napake vzorčenja

$$\bar{x} = \mu + e$$

- Razlika med statistikama je odvisna od razlike parametrov in napake

$$\bar{x}_1 - \bar{x}_2 = (\mu_1 - \mu_2) + (e_1 - e_2)$$

3

Sklepanje na podlagi vzorcev:

Primer aritmetične sredine

- Za preverjanje hipoteze **predpostavimo**, da razlik v populaciji ni (ničelna hipoteza)

$$\mu_1 - \mu_2 = 0$$

- Razlika med statistikama je tako odvisna **samo od napake vzorčenja**, ki pa se porazdeljuje normalno okoli 0

$$\bar{x}_1 - \bar{x}_2 = e_1 - e_2 = e$$

4

Sklepanje na podlagi vzorcev:

Primer aritmetične sredine

- Razlike med statistikama se **ob ničelni hipotezi** porazdeljuje normalno okoli 0

$$(\bar{x}_1 - \bar{x}_2) = 0$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2}$$

- Če dobimo absolutno razliko dosti večjo od 0, sledi, da razlika med statistikama ni le posledica napak vzorčenja:

RAZLIKE SO ŽE V POPULACIJI

5

Sklepanje na podlagi vzorcev:

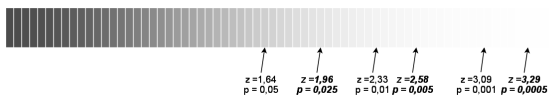
Logika preverjanja hipotez

- Kdaj je razlika „dovolj“ večja od 0? Postaviti moramo mejo: kritično vrednost
- **Kritična vrednost**: Prag ko sprejmemo odločitev (določimo ga glede na naše namene)
- **Primeri iz vsakdanjega življenja**:
 - Glavobol: kdaj vzamemo tableto?
 - Menjava obleke jeseni: kdaj so jutra dovolj hladna?
 - Merjenje telesne temperature: kdaj imamo vročino?
 - Sodstvo: Kdaj lahko rečemo, da je dovolj dokazov, da je obdolženi kriv?

6

Sklepanje na podlagi vzorcev: Logika preverjanja hipotez

- Intenzivnost rdeče barve je sorazmerna z verjetnostjo, da dobimo dani z ali večjega



- Kritične Meje so v statistiki postavljene zelo konzervativno (previdno)

7

Sklepanje na podlagi vzorcev: Logika preverjanja hipotez

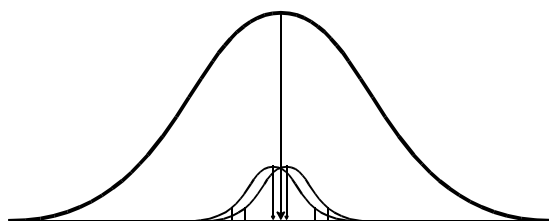
- Razlika med dvema populacijama (npr. moškimi in ženskami) na podlagi vzorcev

Med populacijama	Med vzorcema	
	Ni razlik	So razlike
Ni razlik	Ni napake ocene vzorca omogočata pravi sklep	Napaka I. tipa (α) iz vzorcev je videti, da razlike so, čeprav jih v populaciji ni
So razlike	Napaka II. tipa (β) iz vzorcev je videti, da razlik ni, čeprav v populaciji so	Ni napake ocene vzorca omogočata pravi sklep

8

Sklepanje na podlagi vzorcev: Logika preverjanja hipotez

- Če vzorci prihajajo iz iste populacije, so razlike med statistikami izračunanimi iz njih najverjetneje **majhne**:
RAZLIKE NI



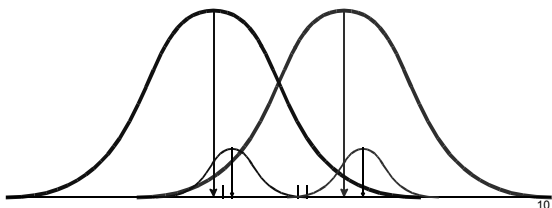
9

Sklepanje na podlagi vzorcev:

Logika preverjanja hipotez

- Če vzorci prihajajo iz **različnih populacij**, so **razlike** med statistikami izračunanimi iz njih najverjetneje **velike** (podobne razlikam med populacijama)

RAZLIKA JE



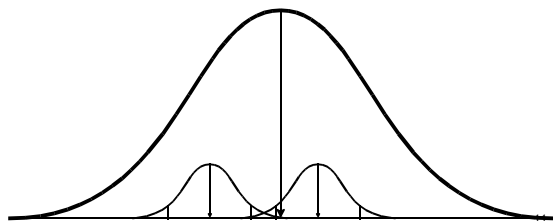
10

Sklepanje na podlagi vzorcev:

Možne napake

- **Napaka prvega tipa (α)**: vzorci prihajajo iz **iste populacije**, a razlike med statistikami izračunanimi iz njih so **velike**:

NAVIDEZNA RAZLIKA

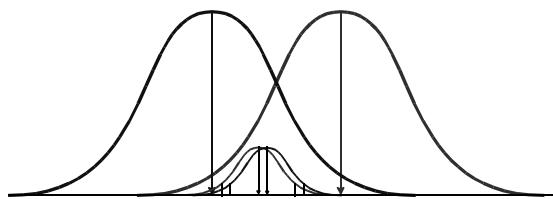


Sklepanje na podlagi vzorcev:

Možne napake

- **Napaka drugega tipa (β)**: vzorci prihajajo iz **različnih populacij**, a razlike med statistikami izračunanimi iz njih so **majhne**:

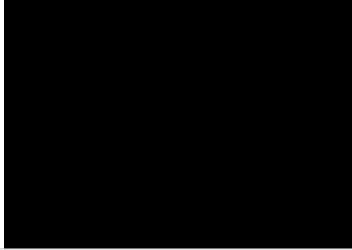
RAZLIKE NE MOREMO DOKAZATI



12

Sklepanje na podlagi vzorcev

- Če je razlika dovolj velika, je to zato, ker je razlika najverjetneje že v populaciji (tvegamo le α da nimamo prav)
- Če razlika ni statistično značilna, še ne pomeni, da ni razlik v



13

Sklepanje na podlagi vzorcev:

Primer

- Primerjali so znanje tujega jezika pri učencih dveh osnovnih šol in dobili naslednje rezultate:

	1.šola	2.šola
n	210	191
M	27,92	23,56
s ²	112,42	94,24
s	10,6	9,71

Kaže, da so učenci prve šole pokazali boljše znanje, vendar so razlike med učenci znotraj šole velike.

Je **zares možno sklepati**, da učenci iz prve šole znajo bolje tuji jezik? (ali pa je razlika zgolj slučajna)

14

Sklepanje na podlagi vzorcev:

Primer

- Poglejmo izračune:
- Razlika med povprečnima ocenama je:
- Standardna napaka za to razliko je:

$$s_{\bar{x}_1}^2 = \frac{s_1^2}{n_1} = \frac{112,42}{210} = 0,54 \quad s_{\bar{x}_2}^2 = \frac{s_2^2}{n_2} = \frac{94,24}{191} = 0,49$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2} = \sqrt{0,54 + 0,49} = \sqrt{1,03} = 1,01$$

Sklepanje na podlagi vzorcev:

Primer

- z-vrednost je razmerje med razliko povprečij in standardno napako:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{4,36}{1,01} = 4,30$$

- Ta razlika je večja kot 1,96 (kritična meja za 5% tveganje) in tudi od 3,29 (kritična meja za 0,1% tveganje)

$$z = 4,30 > 3,29 = z_{0,001} > 1,96 = z_{0,05}$$

16

Sklepanje na podlagi vzorcev:

Primer

ODGOVOR:

- S tveganjem manjšim kot 0,1% lahko trdimo da učenci iz prve šole znajo tuji jezik bolje od učencev druge šole

ALI:

- Razlika v znanju tujega jezika med učenci prve in druge šole je statistično značilna pri stopnji tveganja 0,1%

17
