

## Probleemid hüpoteeside testimisega: Mitmese testimise probleem, põhjusliku mõju testimisest

Märt Möls

1

## Mitmene testimine

Meil on 100 „kahtlusalust“ riskifaktorit, mis võivad olla seotud meid huvitava näitajaga (milline toit tõstab vererõhku?). Ja siis, milles on probleem? Kogu lihtsalt andmed ja hakka testima riskifaktorite kandidaate kasutades olulisuse nivood 0,05?

No nii ei lähe mitte...

Kujuta hetkeks ette, et ükski neist sajast riskifaktorist ei seostu meid huvitava näitajaga. Mitu statistilist testi siis „leiavad“ statistiliselt olulise tulemuse (mitu valepositiivset, eksitavat vastust me saame)?

Milline on oodatav valepositiivsete tulemuste arv?

Vastus: 5!

Mis on tõenäosus saada vähemalt ühte (või enamat) valepositiivset tulemust (100 sõltumatu testi korral)?

$$1 - (1 - 0,05)^{100} = 0,994$$

Seega testides 100 riskifaktorit leiame peaaegu kindlalt mõne, mis justnagu näiks olevat seotud meid huvitava näitajaga!

### Kus kerkib esile mitmese testimise probleem?

„Pühendunud teadlane“ kordab oma eksperimenti 50 korda saamaks soovitud andmeid;

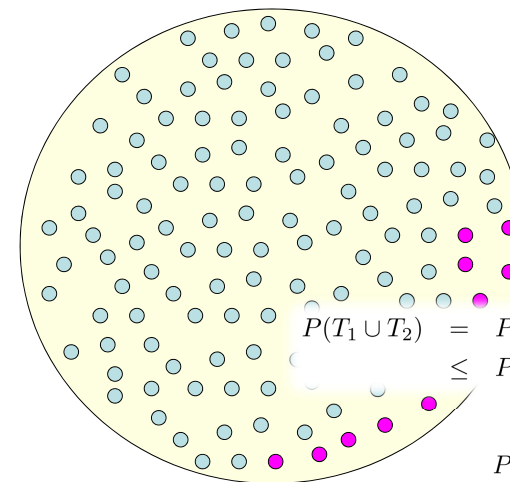
„Teaduslik“ ajakiri avaldab vaid neid statistiliselt olulisi tulemusi sisaldavaid artikleid (nn *publication bias*);

Nõutu ja ideedetü teadlane teeb lihtsalt 100 eksperimenti – “miski võib ju välja tulla”;

Hoolikas doktorant kogub veidi andmeid ja katsetab siis, kas õnnestub tõestada teda huvitavat hüpoteesi. Kui alternatiivse hüpoteesi tõestamine ebaõnnestub, siis teeb ta täiendavalt mõned mõõtmised. Ikka veel pole võimalik saada statistiliselt olulist tulemust? Pole probleemi! Lisame veel mõned vaatlused ja testime uuesti!

Uuritud on 5 raviviisi, kas mõni neist annab parema tulemuse kui mõni teine?

## Lahendused I – Bonferroni meetod



Kui 1. test teeb vea tõenäosusega  $\alpha$  ja teine test eksib ka tõenäosusega  $\alpha$  siis neid valimeid kus üks kahest testist eksib ei esine rohkem kui tõenäosusega  $2\alpha$  (aga võib olla vähem).

$$P(T_1 \cup T_2) = P(T_1) + P(T_2) - P(T_1 \cap T_2) \leq P(T_1) + P(T_2)$$

$$P\left(\bigcup_i^k T_i\right) \leq \sum_i^k P(T_i)$$

## Lahendused I – Bonferroni meetod

Tehes 10 testi, peaksime iga üksiktesti tegemisel kasutama olulisuse nivood

$$\alpha_{\beta} = 0.05 / 10 = 0.005.$$

Seega juhul, kui kasutame iga üksiktesti tegemisel olulisuse nivood 0.005 võime olla kindlad, et vale testitulemuse – ekslikult tõestatud alternatiivse hüpoteesi – saamise tõenäosus 10 testi peale kokku on 5% või väiksem.

## Lahendused I – Bonferroni meetod

Üldjuhul:

Kui teed  $k$  erinevat testi ja soovid, et nende testide peale kokku I-liiki vea tegemise tõenäosus poleks suurem kui  $\alpha$ , siis pead iga üksikul testil kasutama olulisuse nivood

$$\alpha_{\text{Bonferroni}} = \alpha / k.$$

Seega peaksid otsustama  $H_1$  kasuks ( $H_0$  valeks lugema) vaid siis, kui

$$p\text{-väärtus} < \alpha / k.$$

## Lahendused I – Bonferroni meetod

### Bonferroni meetodil korrigeeritud p-väärtus

Bonferroni meetodi otsustuskriteerium:

$$\text{Kui } p\text{-väärtus} < \alpha / k \text{ siis } H_1.$$

Mis on muidugi samaväärne otsustuskriteeriumiga:

$$\text{Kui } p\text{-väärtus} * k < \alpha \text{ siis } H_1.$$

Suurust

$$p_{\text{korrigeeritud}} = \min(p\text{-väärtus} * k, 1)$$

Tuntakse ka kui Bonferroni meetodil korrigeeritud p-väärtust

## Bonferroni meetodi modifikatsioonid

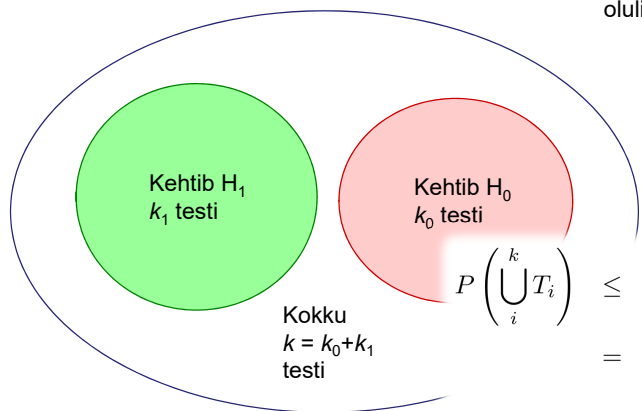
Järjestikusel testimisel võime „kulutada“ oma olulisuse nivood ebaühtlaselt. Näiteks võime teha esimese statistilise testi peale 10 vaatluse tegemist, kuid kasutame siis olulisuse tõenäosust 0,001 – kui vastus on ilmne (olulisuse nivool 0,001), siis lõpetame andmete edasise kogumise ja loeme  $H_1$  tõestatuks. Kui peale 10 vaatluse tegemist pidime jääma nullhüpoteesi juurde, siis võime teha veel 100 vaatlust ja testida siis (kasutades olulisuse nivood 0,004), kas võime juba andmete kogumise lõpetada. Kui ei õnnestunud ikka nullhüpoteesi kummutada, võime koguda veel täiendavalt andmeid 400 objekti kohta ja teha siis viimane statistiline test olulisuse nivool 0,045. Sellise testprotseduuri korral (kus võime teha kuni kolm testi) ei kerki summaarne I-liiki vea tegemise tõenäosus suuremaks kui 0,05, sest

$$0,001 + 0,004 + 0,045 = 0,05$$

# Bonferron-Holm'i meetod

Piisab, kui kasutaksime olulisuse nivood  $\alpha/k_0$ .

Paraku aga ei tea me  $k_0$  väärtust.



$$\begin{aligned}
 P\left(\bigcup_i^k T_i\right) &\leq \sum_i^k P(T_i) \\
 &= \sum_{i \in H_1} P(T_i) + \sum_{i \in H_0} P(T_i) \\
 &= \sum_{i \in H_1} 0 + \sum_{i \in H_0} \alpha/k_0 \\
 &= \alpha
 \end{aligned}$$

# Bonferron-Holm'i meetod

Järjestame olulisustõenäosused ehk p-väärtused kasvavalt:

$$p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(j)} \leq \dots \leq p_{(k)}$$

Võrdleme järjestatud p-väärtuseid suurustega (ol. nivooodega)

$$\alpha/k \quad \alpha/(k-1) \quad \dots \quad \alpha/(k-j+1) \quad \dots \quad \alpha$$

kuni leiame esimese p-väärtuse, mis on suurem kui vastav olulisuse nivoo. Kõigi ülejäänud testide (mille p-väärtus on suurem) puhul jääme  $H_0$  juurde.

# Bonferron-Holm'i meetod

Järjestame olulisustõenäosused ehk p-väärtused kasvavalt:

$$p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(j)} \leq \dots \leq p_{(k)}$$

Väikseima p-väärtusega test, mille puhul kehtib  $H_0$ .

Teeme I-liiki vea kui  $p_{(j)} < \alpha/(k-j+1)$ .

Aga  $j \leq k_1 + 1$  sest maksimaalselt saab olla  $k_1$  väiksemat p-väärtust  
 $j \leq k - k_0 + 1$   
 $k_0 \leq k - j + 1$

ja seega:

$$\alpha/(k-j+1) \leq \alpha/k_0$$

# Bonferron-Holm'i meetod

Aga kui  $\alpha/(k-j+1) \leq \alpha/k_0$

siis on I-liiki vea tegemise tõenäosus  $j$ . testil:

$$P(p_{(j)} < \alpha/(k-j+1)) \leq P(p_{(j)} < \alpha/k_0)$$

ja seega on summaarne I-liiki vea tegemise tõenäosus nende  $k_0$  testi puhul, kus üldse saame teha I-liiki viga väiksem (või võrdne)  $\alpha$ -ga.

$$\begin{aligned}
 P(p_{(j)} < \alpha/k_0) &\leq \sum_{i \in H_0} P(p_i < \alpha/k_0) \\
 &\leq \sum_{i \in H_0} \alpha/k_0 = \alpha
 \end{aligned}$$

## Bonferroni-Holmi meetod

Võime kasutada ka Bonferroni-Holmi meetodil korrigeeritud p-väärtust:

$$p_{(j)}^* = \max_{i=1 \dots j} \{ \min[(k-i+1)p_{(i)}, 1] \}$$

## Bonferroni-Holmi meetod

P-väärtused:

0.04	0.001	0.0051	0.006	0.008					
					0.0081	0.009	0.08	0.23	0.78

Järjestatud p-väärtused

0.001	0.0051	0.006	0.008	0.0081					
					0.009	0.04	0.08	0.23	0.78

Olulisuse nivood:

0.0050	0.0056	0.0062	0.0071	0.0083					
					0.0100	0.0125	0.0167	0.0250	0.0500

## Sõltumatud testid...

Kasutame üksiktestil olulisuse nivood  $\alpha^*$

$$P(\text{Test } i \text{ teeb I-liiki vea}) \leq \alpha^*$$

$$P(\text{Test } i \text{ EI tee I-liiki viga}) \geq (1 - \alpha^*)$$

$$P(k \text{ testi ei tee I-liiki viga}) \geq (1 - \alpha^*)^k$$

$$P(k \text{ testi peale esineb I-liiki viga}) \leq 1 - (1 - \alpha^*)^k$$

Vali  $\alpha^*$  nii, et  $1 - (1 - \alpha^*)^k = \alpha$ , siis ei ületa I-liiki vea tegemise tõenäosus kõigi testide peale kokku  $\alpha$ -t, ehk vali:

$$\alpha^* = 1 - (1 - \alpha)^{1/k}$$

## Bonferroni vs Šidak

2 testi

Bonferroni:  $\alpha_{\text{Test}} = 0,025$

Šidak:  $\alpha_{\text{Test}} = 0,02532$

1000 testi

Bonferroni:  $\alpha_{\text{Test}} = 0,00005$

Šidak:  $\alpha_{\text{Test}} = 0,0000513$

Bonferroni meetodit saab kasutada ka sõltuvate testide korral...

## Sõltuvad testid...

- Igalt inimeselt küsitakse 100 küsimust ja hakatakse siis uurima, kas tegemist on riskifaktoriga. Aga küsitud tunnused on sõltuvad (inimene kes tarvitab palju leiba/saia tarvitab ka tõenäolisemalt palju võid...)
- Geenimutatsioonide mõju uurimine...;

Miks geenimutatsioonide testid on sõltuvad ?  
Aheldus...

<u>A/T</u>	<u>G/A</u>	<u>T/C</u>	<u>...</u>	<u>T/C</u>
A	G	C		C
A	G	C		T
A	G	C		T
T	A	T		C
T	A	T		C
T	A	T		C

Sõltuvus snippide vahel viib sõltuvate testideni (kus kontrollime mutatsiooni mõju fenotüübile ehk mingile meid huvitavale vaadeldavale organismi omadusele (sama p-väärtus kõigis testides...)).

## Permutatsioonitest

### Mitmene testimine sõltuvate testide korral

1. Sega Y-tunnuse (fenotüüpide) väärtused juhuslikult
2. arvuta p-väärtused kõigi X-tunnuste jaoks (kontrolli kõigi geenimutatsioonide ehk snippide mõju uuritavale fenotüübile – aga fenotüübid on juhuslikud. Tänu fenotüüpide segiajamisele teame, et kõigi testide puhul kehtib  $H_0$ )
3. Jäta meelde väikseim p-väärtus mida nägid (juhuse tahtel, kui seost tunnuste vahel pole, siis kui väikest p-väärtust võime veel näha)
4. Korda samme 1-3 palju kordi (tee analüüsid üle peale Y-tunnuse väärtuste uut segiajamist). Näiteks korda kõiki analüüse 10 000 korda.
5. Kasutades meeldejäetud pisemaid p-väärtuseid otsi välja, kui väikest p-väärtust võid juhuse tahtel näha (leia nähtud p-väärtuste 0,05-kvantil)
6. Otsustuskriteerium: kui sinu poolt tegelikus uuringus nähtud p-väärtus on väiksem kui ennist leitud 0,05-kvantil, siis otsusta  $H_1$ -kasuks. Muidu jää nullhüpoteesi juurde.

Järeldus:

Permutatsioonitesti tegemiseks pead kõvasti arvutit kulutama.