

Lineaarsed mudelid
Praktikum
Testi võimsus

Vaatame loengu ülesande näitel kolme erinevat võimalust testi võimsust arvutada ning arutleme, millal eelistada ühte ja millal mõnda teist meetodit.

Näide 1

Kavatseme uuringusse kaasata 30 inimest Lõuna-Eestis, 30 inimest Kesk-Eestis, 30 inimest Põhja-Eestis. Oletame, et uuritava tunnuse keskvärtus on Lõuna-Eestis 100, Kesk-Eestis 110, Põhja-Eestis 120. Uuritava tunnuse standardhälve Lõuna-Eestis on 20, oletame et ka muudes piirkondades võiks uuritava tunnuse standardhälve olla sarnane.

Kui tõenäoliselt suudaksime tõestada, et vaadeldud kolmes piirkonnas pole uuritava tunnuse keskvärtus samasugune (kui kasutame andmete analüüsimisel dispersioonanalüüsi)?

Variant 1

Dispersioonanalüüsi võimsust saab leida R'i käsu `power.anova.test` abil (kui kõigis piirkondades tehakse sama arv vaatluseid):

```
power.anova.test(3, n=30, between.var=var(c(100,110, 120)), within.var=400)
```

Sama käsku saab kasutada ka leidmaks valimi suurust, mida kasutades saavutatakse soovitud võimsus:

```
power.anova.test(3, power=0.8,  
  between.var=var(c(100,110, 120)), within.var=400)
```

Käsu `power.anova.test` arvutusi saab ka soovi korral ise korrata:

```
gruppe=3  
between.var=var(c(100,110, 120))  
within.var=400  
n=30  
  
# Leiame mittetsentraalsuse parameetri väärtuse.  
# Selleks märkame, et samasuurte grupi suuruste korral -  
# tasakaalustatud valimi korral -  
# on  $EY^2(I-P_{X0})EY = n * (gruppide arv - 1) * between.var :$   
  
ncp = (gruppe - 1) * n * (between.var/within.var)  
kriitiline = qf(0.95, df1=gruppe - 1, df2=n*gruppe-gruppe)  
v6imsus = 1- pf(kriitiline, df1=gruppe - 1, df2=n*gruppe - gruppe,  
  ncp=ncp)  
v6imsus
```

Või praktilisem näide – graafiku koostamine, kust saab erinevate valimi mahtude korral lugeda välja saavutatavat võimsust. Graafikul on eraldi ära märgitud võimsuse 0,8 saavutamiseks vajalik valimi suurus, sest tavaks on uuringutelt nõuda vähemalt 80% võimsust – madalama võimsuse korral sageli uuringut korraldama ei hakatagi:

```
n=2:50
ncp <- (gruppe - 1) * n * (between.var/within.var)
kriitiline = qf(0.95, df1=gruppe - 1, df2=n*gruppe-gruppe)
v6imsus = 1- pf(kriitiline, df1=gruppe - 1, df2=n*gruppe - gruppe, ncp=ncp)

plot(n, v6imsus, type="l", xlab="n (igast piirkonnast)",
      ylab="Võimsus", ylim=c(0,1))

# Võimsuse 0,8 saavutamiseks vajalik valimi suurus
v6imsus80=power.anova.test(3, power=0.8,
between.var=var(c(100,110, 120)), within.var=400)$n
arrows(0, 0.8, v6imsus80, 0.8, length=0, col="gray70", lwd=2)
arrows(v6imsus80, 0.8, v6imsus80, 0, length=0.1, col="gray70", lwd=2)
```

Keerukamate katseplaanide jaoks (näiteks kui eri piirkondadest võetakse erinev arv vaatluseid) peame paratamatult arvutused tegema ise, ilma power.anova.test-käsu abita. Oletame, et tahame võtta Lõuna- ja Põhja-Eestist kaks korda suurema valimi (mõlemast 36) kui Kesk-Eestist (18 inimest, kokku $n=90$), siis milline tuleks testi võimsus?

```
# Mittetsentraalsuse parameetri leidmine
n_grupi_kohta= c(36,18,36)
grupp =rep(1:3, n_grupi_kohta)
# Keskväärtuste vektor (iga vaatluse jaoks tema keskväärtus):
EY =rep(c(100, 110, 120), n_grupi_kohta)
# Leiame lihtsama mudeliga keskväärtuste prognoosimisel tehtavad vead
vead=residuals(lm(EY~1))
ncp=sum(vead**2)/400

# Kriitilise väärtuse leidmine
kriit=qf(0.95, 2, 90-3)

# Testi võimsus
1-pf(kriit, 2, 90-3, ncp=ncp)
```

Näeme, et antud juhul on testi võimsus suurem kui siis, kui võtaksime igast maakonnast sama arvu uuritavaid. Seega meid huvitava testi puhul saavutaksime samasuure valimi korral suurema võimsuse valides Lõuna- ja Põhja-Eestist kaks korda rohkem uuritavaid kui Kesk-Eestist.

Ülesanne 1.

Kui suur on meie võime samas uuringus tõestada Lõuna-Eesti ja Kesk-Eesti keskväärtuste erinevust? Leia mõlema katseplaani (plaan A: kõigist piirkondadest sama arv vaatluseid; plaan B: Lõuna- ja Põhja-Eestist võtame kaks korda rohkem uuritavaid kui Kesk-Eestist) jaoks tõenäosus tuvastada Lõuna- ja Kesk-Eesti erinevust!

Ülesanne 2.

Kavatseme ikkagi võtta igast piirkonnast (Lõuna-Eesti; Kesk-Eesti; Põhja-Eesti) sama arvu uuritavaid (30 uuritavat igast piirkonnast). Aga seekord kavatseme kodeerida maakonna pideva tunnuseks (1: Lõuna-Eesti; 2: Kesk-Eesti; 3: Põhja-Eesti) ja dispersioonanalüüsi asemel kavatseme hinnata regressioonanalüüsi mudeli. Kui testime sirge tõusu olulisust siis milline tuleks testi võimsus? Leia lähenemist 2 kasutades õige vastus!

Vihjeks: Õige vastus on suurem kui 0,94 ja väiksem kui 0,98!

Simulatsioon

Mõnikord on kõige mugavam leida testi võimsust simulatsiooni abil. Näiteks kui võtaksime juhusliku valimi Eestis elavatest karuküttidest (valimi suurus $n=90$) ja teaksime, et 1/3 karuküttidest elab Põhja-Eestis, 1/3 Kesk-Eestis ja 1/3 Lõuna-Eestis, siis milline tuleks testi võimsus (kui karuküttidel uuritava tunnuse keskväärtused piirkonniti oleks 120; 110 ja 100)? Eelnenud arvutused eeldasid, et võtame igast piirkonnast kindla arvu uuritavaid. Juhusliku valimi korral me aga ei tea ette, mitu inimest Põhja-Eestis, mitu Kesk-Eestist ja mitu Lõuna-Eestist valimisse sattub. Siiski saame testi võimsust ka sellises situatsioonis mugavalt leida simulatsiooni abil (tõsi küll, ligikaudselt):

```
# NB! Simulatsioon võtab aega!
set.seed(1)
korduseid=20000
pvalue=rep(NA, korduseid)
for (i in 1:korduseid){
  grupp=sample(1:3, 90, replace=TRUE)
  y=c(100, 110, 120)[grupp]+rnorm(90, 0, sqrt(400))
  m=lm(y~factor(grupp))
  pvalue[i] = anova(m)[1,5]
}
# Testi võimsus (koos arvutustäpsust näitava usaldusintervalliga):
t.test(pvalue<0.05)
```

Alternatiivina simulatsioonile võiksime proovida testi võimsust leida täistõenäosuse valemit kasutades:

$$P(\text{võimsus}) = \sum_{\{n_1+n_2+n_3=90\}} P(\text{võimsus}|n_1;n_2;n_3)P(n_1;n_2;n_3),$$

ehk leiame testi tingliku võimsuse mistahes võimalike valimi suuruste korral ja seejärel leiaksime täistõenäosuse valemi abil lihtsalt testi võimsuse. See lähenemine võimaldab ka juhuslike grupisuuruste puhul leida anova-testi võimsust täpselt, aga vajalik arvutus muutub valimimahu kasvades kiiresti ebapraktiliselt keeruliseks (kuigi väikeste valimi suuruste korral võib toodud täpne arvutus osutuda kiiremaks kui mõistliku täpsusega simulatsioon!).

Selle variandi rakendusnäidis:

```
n1=rep(NA, 91*46)
n2=rep(NA, 91*46)
n3=rep(NA, 91*46)
p_komb=rep(NA, 91*46)
power_tinglik=rep(NA, 91*46)
nr=1
for (i in 0:90){
  for (j in 0:(90-i)){
    n1[nr]=i
    n2[nr]=j; n3[nr]=90-n1[nr]-n2[nr]
    # Antud valimisuuruste kombinatsiooni esinemistõenäosus:
    p_komb[nr]=dbinom(n1[nr], 90, 1/3)*dbinom(n2[nr], 90-n1[nr], 1/2)

    n_grupi_kohta= c(n1[nr], n2[nr], n3[nr])
    grupp =rep(1:3, n_grupi_kohta)
    # Keskväärtuste vektor (iga vaatluse jaoks tema keskväärtsus):
    EY=rep(c(100, 110, 120), n_grupi_kohta)
    # Leiame lihtsama mudeliga keskväärtuste prognoosimisel tehtavad vead
    vead=residuals(lm(EY~1))
    ncp=sum(vead**2)/400

    kriit=qf(1-0.05, 2, 90-3)
    power_tinglik[nr]= 1-pf(kriit, 2, 90-3, ncp=ncp)
    nr=nr+1
  }
}
power=sum(power_tinglik*p_komb)
power
```

Ülesanne 3

Vanarahvas teadis: „need lapsed, kes hommikuti putru söövad kasvavad suureks“. Kavatsame teha uuringu selle väite tõesuse kontrolliks. Jagame juhuslikult lastekodusse sattunud lapsed kaheks grupiks. Üks gruppidest saab hommikuti putru, teine teistsugust hommikusööki. Kui katses osalenud lapsed on ükskord täiskasvanuks saanud (näiteks saanud 20 aastat vanaks) siis mõeldame nende pikkused ära. Kui palju katsealuseid vajaksime tõestamaks pudrusöömise mõju pikkusele (näiteks kui suurt valimit vajame võimsuse 0,8 saavutamiseks)? Leia vajalik valimi suurus!

Kas pudru mõju testimisel peaksime mudelisse lisama ka katsealuse soo? Põhjenda oma otsust!