ND6 Hipotezių tikrinimas

Matas Amšiejus

4/27/2021

1)

Buvo atsitiktinai parinkta 30 žmonių ir tirtas jų reakcijos laikas (sekundėmis) į tam tikrą signalą:

```
reakcija<-c(0.80,0.75,0.98,0.82,0.67,0.88,1.00,0.47,0.87,1.02,0.94,0.92,0.66,0.85,1.18,0.77,0.58,1.27,1
```

Tarę, kad buvo stebimas normalusis atsitiktinis dydis, esant reikšmingumo lygmeniui 0,05, patikrinkite hipotezę, kad vidutinis reakcijos laikas yra 1 sekundė.
Sprendimas
Turime $\alpha = 0.05$, $\mu_0 = 1$.

```
t.test(reakcija, mu = 1)
```

```
##
##  One Sample t-test
##
## data:  reakcija
## t = -3.7762, df = 29, p-value = 0.0007319
## alternative hypothesis: true mean is not equal to 1
## 95 percent confidence interval:
##  0.8083272 0.9430061
## sample estimates:
## mean of x
## 0.8756667
```

Matome, kad p-reikšmė yra tik ~0.0007, o tam, kad hipotezė H_0 nebūtų atmesta, reikėjo $p \geq 0.05$.

2)

Pateikti duomenys apie MgO kiekį (procentais) 15 mėginių:

```
meginiai<-c(6.2,6.8,9.8,5.0,1.4,2.6,4.0,5.0,3.2,5.9,6.9,3.4,4.8,4.4,4.0)
```

Tarę, kad buvo stebimas normalusis atsitiktinis dydis, atlikite užduotis:
a) esant 95 procentų patikimumui, įvertinkite MgO kiekio vidurkį ir dispersiją;
b) esant reikšmingumo lygmeniu 0,05, patikrinkite hipotezę, kad vidutinis MgO kiekio mėginiuose standartinis nuokrypis nedidesnis už 1.5;
c) esant reikšmingumo lygmeniui 0,05 patikrinkite hipotezę, kad vidutinis MgO kiekis mėginiuose yra 5.0.
Sprendimas
a)
Ieškosime normaliojo dydžio vidurkio pasikliautinio intervalo kai dispersija nežinoma:
I)Randame vidurkio įvertinį

```
vid<-mean(meginiai);vid
```

```
## [1] 4.893333
```

II)Randame nepaslinktaįį standartinio nuokrypio įvertinį

```
stnuok<-sd(meginiai);stnuok
```

```
## [1] 2.050946
```

III)Nustatome $\alpha/2$. $\alpha = 1 - Q$, $\alpha = 0.05$, $\alpha/2 = 0.025$.
IV)Ieškome pasikliautinių intervalų:

```
kaire<-vid-qt(p=0.025, 14, lower.tail = F)*stnuok/sqrt(15)
desine<-vid+qt(p=0.025, 14, lower.tail = F)*stnuok/sqrt(15)
c(kaire,desine)
```

```
## [1] 3.757557 6.029110
```

Ieškosime normaliojo dydžio dispersijos pasikliautinio intervalo kai vidurkis nežinomas:

```
kaire_d<-14*stnuok^2/qchisq(0.025, 14, lower.tail = F)
desine_d<-14*stnuok^2/qchisq(1-0.025, 14, lower.tail = F)
c(kaire_d,desine_d)
```

```
## [1]  2.254659 10.462284
```

Kaip vėliau matysime, šis intervalas nesutampa su testo duotu intervalu.
b)
Naudosime biblioteką *EnvStats*:

```
library(EnvStats)
```

Kadangi tirsime standartinį nuokrypį, o (pagal mano žinias) *varTest* skaičiuoja dispersiją, pakeliame 1.5 kvadratu. Gauname $\sigma^2 = 2.25$. Tada naudodami *varTest* tikriname hipotezę ir įvertiname 0.95% patikimumui dispersiją:

```
varTest(meginiai, sigma.squared= 2.25, alternative = "greater", conf.level = 0.95)
```

```
##
##  Chi-Squared Test on Variance
##
## data:  meginiai
## Chi-Squared = 26.173, df = 14, p-value = 0.02461
## alternative hypothesis: true variance is greater than 2.25
## 95 percent confidence interval:
##  2.486378      Inf
## sample estimates:
## variance
## 4.206381
```

Deja ir vėl nelygybė $0.025 \geq 0.05$ nėra tenkinama, todėl hipotezę atmesime (statistiškai reikšmingai skiriasi nuo 2.25). Taip pat keista, bet gavome, kad pasikliovimo lygmens viršutinis rėžis yra begalybė (kodėl?). **c)**

```
t.test(meginiai, mu=5)
```

```
##
##  One Sample t-test
##
## data:  meginiai
## t = -0.20143, df = 14, p-value = 0.8433
## alternative hypothesis: true mean is not equal to 5
```

```
## 95 percent confidence interval:
##  3.757557 6.029110
## sample estimates:
## mean of x
##  4.893333
```

Gavome, kad p-vertė yra $0.843 \geq 0.05$, todėl galime teigti, kad su reikšmingumo lygmeniu 0.05 vidutinis MgO kiekis statistiškai reikšmingai nesiskiria nuo 5.

3)

Firma siūlo vaistus karpoms gydyti su žinomu išgijimo procentu 40%. Buvo 25 pacientai gydyti šiais vaistais, taip pat kartu buvo duota vitamino C. 14 pacientų pagijo. Ar tai neprieštarauja prielaidai, kad išgijimo procentas yra 40%. Duomenys: (t – pagijo, n - ne)
Sprendimas
Susikuriame naują pagalbinį vektorių iš 1 ir 0, kur t yra 1, n yra 0:

```
vaistai<-c('t','n','t','n','t','t','n','t','n','n','t','n','t','n','t','n','n','t','t','n','t','t','n',
vaistai_b<-c()
for(i in c(1:(length(vaistai)))){
  if(vaistai[i]=='t'){
    vaistai_b[i]=1
  }
  else vaistai_b[i]=0
}
```

Ieškome binominio skirstinio p-value:

```
isgijo<-sum(vaistai_b)
binom.test(isgijo, length(vaistai), 0.4, alternative = "less")
```

```
##
##  Exact binomial test
##
## data:  isgijo and length(vaistai)
## number of successes = 14, number of trials = 25, p-value = 0.9656
## alternative hypothesis: true probability of success is less than 0.4
## 95 percent confidence interval:
##  0.0000000 0.7301469
## sample estimates:
## probability of success
##                   0.56
```

Taigi gauname, kad p-reikšmė yra 0.9656 (netgi labai aukšta), todėl vidutiniškas vaistų veikimo lygis (56%), su patikimumo kriterijumi 0.05 statistiškai nesiskiria nuo 40%.

4)

) Lentelėje pateiktas įvažiuojančių į Vilniaus senamiestį automobilių skaičius. Duomenys pateikti pagal 30 dienų stebėjimus.

```
masinos<-c(428,371,397,360,467,410,375,429,472,336,440,451,280,353,404,368,409,459,373,415,446,385,433,
```

Tarkime, kad buvo stebimas normalusis atsitiktinis dydis. Ar galime teigti, kad vidutinis įvažiuojančių į Vilniaus senamiestį automobilių skaičius viršija 400 automobilių per dieną?
Sprendimas

```
t.test(masinos, mu=400, alternative = "less") # tikiu, kad cia tikrai galima kazka redaguoti (gal 401?)
```

```
##
##  One Sample t-test
##
## data:  masinos
## t = -0.37884, df = 29, p-value = 0.3538
## alternative hypothesis: true mean is less than 400
## 95 percent confidence interval:
##      -Inf 411.1521
## sample estimates:
## mean of x
##     396.8
```

Gauname, kad p-reikšmė yra lygi 0.354, vadinasi vidutinis mašinų skaičius (396.8) statistiškai nesiskiria nuo 400(?) su patikimumo kriterijumi 0.05.

5)

Kuriamas naujas kompiuterių tinklas. Reikalaujama, kad jis būtų daugiau negu 99% suderintas su jau naudojama įranga. Atlikite užduotis:
a) suformuluokite hipotezę ir alternatyvą;
b) buvo atrinkta 300 programų imtis ir patikrintas jų suderinamumas su kuriamu tinklu. Gauta, kad 298 programos suderintos su tinklu. Ar remiantis šiais duomenimis galime atmesti hipotezę (pagrįskite atsakymą)
Sprendimas
a)
H_0 :Suderinamumo lygis turi būti lygus arba didesnis nei 99%;
H_1 :Suderinamumo lygis mažiau nei 99%.
b)

```
binom.test(298,300,p=0.99,alternative = "less")
```

```
##
##  Exact binomial test
##
## data:  298 and 300
## number of successes = 298, number of trials = 300, p-value = 0.8024
## alternative hypothesis: true probability of success is less than 0.99
## 95 percent confidence interval:
##  0.0000000 0.9988142
## sample estimates:
## probability of success
##              0.9933333
```

Taigi galime teigti, kad hipotezės galime neatmesti (statistiškai vienoda)

6)

Ampulėje turi būti po 300 mg tam tikro preparato. Leistinas nukrypimas nuo normos toks: standartinis nuokrypis ne didesnis už 10 mg. Patikrinus 15 naujos siuntos ampulių, jose preparato atitinkamai rasta

```
ampule<-c(310,312,298,270,280,300,305,311,290,288,302,330,320,295,289)
```