cv4.R

Marek

2023-03-09

```r
library(latexpdf)
```

Intervaly spolahlivosti pre parametre normalneho rozdelenia

pre strednu hodnotu μ a disperziu σ^2. Bodove odhady pozname, maju dobre statisticke vlastnosti (nevychylene, efektivne) Pri nahodnom pokuse je vzdy nejaka chyba, preto si urcujeme, akej chyby sa mozem dopustit. IS konstruujeme s vopred danou, zvolenou chybou, ktorej sa mozem dopustit. Ak robim 95% IS, tak sa dopustam 5% (nepokryje), $\alpha = 0.05$ Nahodny vyber musi byt normalne rozdeleny, da otestovat, my to predokladame. # IS pre parameter alternativneho rozdelenia, p. Zo 100 opytanych respondentov, 15 vyjadrilo ochotu dat sa zaockovat prave vakcinou ABC. Bodovy odhad najdeme ako pomer $\hat{p} = \frac{15}{100}$. (1-)100%$ IS pre tento parameter

$$\hat{p} \pm u(1-\alpha/2)\sqrt{\frac{p(1-p)}{n}}$$

Najdite 95% IS pre p

```r
alfa <- 0.05
p <- 0.15
n <- 100
IS <- p + c(-1, 1) * qnorm(1 - alfa / 2, 0, 1) * sqrt(p * (1 - p) / n)
IS
```

```
## [1] 0.08001529 0.21998471
```

Teraz vstavana funkcia

```r
binom.test(15, 100, p = 0.15)
```

```
##
##  Exact binomial test
##
## data:  15 and 100
## number of successes = 15, number of trials = 100, p-value = 1
## alternative hypothesis: true probability of success is not equal to 0.15
## 95 percent confidence interval:
##  0.08645439 0.23530750
## sample estimates:
## probability of success
##                   0.15
```

```r
binom.test(15, 100, p = 0.15)$conf.int
```

```
## [1] 0.08645439 0.23530750
## attr(,"conf.level")
## [1] 0.95
```

Zostrojte 90% IS

```r
binom.test(15, 100, p = 0.15, conf.level = 0.9)$conf.int
```

```
## [1] 0.09479401 0.22153691
## attr(,"conf.level")
## [1] 0.9
```

Lavostranny, pravostranny

```r
binom.test(15, 100, p = 0.15, alternative = 'greater')$conf.int
```

```
## [1] 0.09479401 1.00000000
## attr(,"conf.level")
## [1] 0.95
```

```r
binom.test(15, 100, p = 0.15, alternative = 'less')$conf.int
```

```
## [1] 0.0000000 0.2215369
## attr(,"conf.level")
## [1] 0.95
```

```r
######################################################################
```

IS pre μ, ak σ pozname Priklad prednaska - kopirka Najdite 95% IS, 90% IS obojstranne, 95% jednostranne podla vzorca

```r
x <- c(2445, 2450, 2453, 2462, 2463, 2463, 2466, 2471, 2474, 2475, 2475,
       2484, 2485, 2486, 2487, 2490, 2491, 2493, 2499, 2501, 2501, 2503,
       2504, 2505, 2505, 2506, 2506, 2507, 2509, 2511, 2511, 2513, 2514,
       2515, 2518, 2523, 2523, 2524, 2525, 2527, 2529, 2530, 2530, 2533,
       2535, 2536, 2537, 2539, 2560, 2571)
alfa <- 0.05
sigma <- 30
n <- 50
IS <- mean(x) + c(-1, 1) * qnorm(1- alfa / 2) * sigma / sqrt(n)
IS
```

```
## [1] 2494.945 2511.575
```

```r
alfa <- 0.1
IS <- mean(x) + c(-1, 1) * qnorm(1- alfa / 2) * sigma / sqrt(n)
IS
```

```
## [1] 2496.281 2510.239
```

```r
alfa <- 0.05
ISl <- c(mean(x) - qnorm(1 - alfa) * sigma/ sqrt(n), Inf)
ISl
```

```
## [1] 2496.281      Inf
```

```r
ISp <- c(-Inf, mean(x) + qnorm(1 - alfa) * sigma/ sqrt(n))
ISp
```

```
## [1]     -Inf 2510.239
```

IS pre strednu hodnotu, ak sigma nie je zname. Najprv vzorcom, a potom vstavanymi funkciami. Rieste rovnake ulohy, ale za predokladu, ze sigma nepozname σ nahradzame odhadom, kvantil je kvanitol studentovho t rozdelenia s parametrom stupne volnosti tu je to n - 1

```r
alfa <- 0.05
IS <- mean(x) + c(-1, 1) * qt(1 - alfa / 2, n - 1) * sd(x) / sqrt(n)
IS
```

```
## [1] 2495.285 2511.235
```

```r
alfa <- 0.1
IS <- mean(x) + c(-1, 1) * qt(1 - alfa / 2, n - 1) * sd(x) / sqrt(n)
IS
```

```
## [1] 2496.606 2509.914
```

```r
alfa <- 0.05
ISl <- c(mean(x) - qt(1 - alfa, n - 1) * sd(x) / sqrt(n), Inf)
ISl
```

```
## [1] 2496.606      Inf
```

```r
ISp <- c(-Inf, mean(x) + qt(1 - alfa, n - 1) * sd(x) / sqrt(n))
ISp
```

```
## [1]     -Inf 2509.914
```

pomocou vstavanych funkcii a roznych kniznic

```r
t.test(x)$conf.int #default je 95% obojstranny
```

```
## [1] 2495.285 2511.235
## attr(,"conf.level")
## [1] 0.95
```

```r
t.test(x, conf.level = 0.9)$conf.int # 90%, ale stale obojstranny
```

```
## [1] 2496.606 2509.914
## attr(,"conf.level")
## [1] 0.9
```

```r
t.test(x, alternative = 'greater')$conf.int # 95%, lavostranny
```

```
## [1] 2496.606      Inf
## attr(,"conf.level")
## [1] 0.95
```

```r
t.test(x, alternative = 'less')$conf.int # 95%, pravostranny
```

```
## [1]     -Inf 2509.914
## attr(,"conf.level")
## [1] 0.95
```

```r
#install.packages('DescTools')
library('DescTools')
MeanCI(x) # 95% obojstranny
```

```
##     mean   lwr.ci   upr.ci
## 2503.260 2495.285 2511.235
```

```r
MeanCI(x, conf.level = 0.9) # 90%, ale stale obojstranny
```

```
##     mean   lwr.ci   upr.ci
## 2503.260 2496.606 2509.914
```

```r
MeanCI(x, sides = 'left') # 95% lavostranny
```

```
##     mean   lwr.ci   upr.ci
## 2503.260 2496.606      Inf
```

```r
MeanCI(x, sides = 'right') # 95% pravostranny
```

```
##     mean   lwr.ci   upr.ci
## 2503.260     -Inf 2509.914
```

```r
library(Rmisc)
```

```
## Loading required package: lattice
```

```
## Loading required package: plyr
```

```r
#install.packages('Rmisc')
CI(x) #95%
```

```
##    upper     mean    lower
## 2511.235 2503.260 2495.285
```

```r
CI(x, ci = 0.9) #90%
```

```
##    upper     mean    lower
## 2509.914 2503.260 2496.606
```

IS pre strednu hodnotu a podmnoziny dat, Rmisc

najdeme IS pre priemerny prijem vzhladom na pohlavie, vzhladom na vzdelanie

```r
library(readxl)
data <- read_xlsx('C:/R/r_workspace/statistika/dataE.xlsx')
group.CI(data$mprij ~ data$pohlavie, data = data)
```

```
##   data$pohlavie data$mprij.upper data$mprij.mean data$mprij.lower
## 1             m         1183.562            1065         946.4381
## 2             z         1309.026            1170        1030.9738
```

```r
group.CI(mprij ~ vzdelanie, data = data)
```

```
##   vzdelanie mprij.upper mprij.mean mprij.lower
## 1         1    1042.106   923.0769    804.0478
## 2         2    1074.737   994.5652    914.3930
## 3         3    1636.938  1500.0000   1363.0616
```

IS pre disperziu

Vypocitajte 95% obojstranny, lavostranny a pravostranny pre nas priklad hore podla vzorca

```r
alfa <- 0.05
c((n - 1) * var(x) / qchisq(1 - alfa / 2, n - 1), (n - 1) * var(x) / qchisq(alfa / 2, n - 1))
```

```
## [1]  549.5342 1222.9353
```

```r
c((n - 1) * var(x) / qchisq(1 - alfa, n - 1), Inf)
```

```
## [1] 581.7065      Inf
```

```r
c(0, (n - 1) * var(x) / qchisq(alfa, n - 1))
```

```
## [1]    0.00 1137.32
```

Kniznica EnvStats

```r
#install.packages('EnvStats')
library('EnvStats')
```

```
##
## Attaching package: 'EnvStats'
```

```
## The following objects are masked from 'package:stats':
##
##     predict, predict.lm
```

```
## The following object is masked from 'package:base':
##
##     print.default
```

```r
varTest(x)$conf.int
```

```
##       LCL       UCL
##  549.5342 1222.9353
## attr(,"conf.level")
## [1] 0.95
```

```r
varTest(x, conf.level = 0.9)$conf.int
```

```
##       LCL       UCL
##  581.7065 1137.3201
## attr(,"conf.level")
## [1] 0.9
```

```r
varTest(x, alternative = 'greater')$conf.int
```

```
##      LCL      UCL
## 581.7065      Inf
## attr(,"conf.level")
## [1] 0.95
```

```r
varTest(x, alternative = 'less')$conf.int
```

```
##     LCL     UCL
##    0.00 1137.32
## attr(,"conf.level")
## [1] 0.95
```

```r
VarCI(x)
```

```
##       var    lwr.ci    upr.ci
##  787.5433  549.5342 1222.9353
```

```r
##########################################################
```

Simulujeme 100 krat nahodny vyber z normalneho rozdelenia dlzky 30, $\mu = 3$ a $\sigma = 1$, zratame 95% IS pre kazdu simulaciu, kolko je takych intervalov, ktore nepokryju μ

```r
k <- 0
for(i in 1:100){
  a <- rnorm(30, mean = 3, sd = 1);
  is <- t.test(a)$conf.int;
  if(3 > is[2])
    k <- k + 1;
  if(3 < is[1])
    k <- k + 1;
}
k
```

```
## [1] 6
```

Nakreslime hustotu N(0, 1) a zopar hustot studentovho t rozdelenia s roznymi stupnami volnosti, porovname tvary kriviek

```r
plot(function(x) dnorm(x, 0, 1), xlim = c(-5, 5))
plot(function(x) dt(x, df = 2), col = 'green', xlim = c(-5, 5), add = T)
plot(function(x) dt(x, df = 4), col = 'red', xlim = c(-5, 5), add = T)
plot(function(x) dt(x, df = 10), col = 'purple', xlim = c(-5, 5), add = T)
```