TP4 stats

Romain PEREIRA

27 Avril 2018

1) Test de Student

1) Test sur une distribution $N(\mu = 1, \sigma = 2)$ (en supposant σ inconnue)

(a) Risque de 1ère espèce

H_0 : $\mu = \mu_0$

H_1 : $\mu = \mu_1$

Soit $\boxed{\alpha = \mathbb{P}(choisir \quad H_1 \mid H_0 \quad est \quad vraie)}$ est l'erreur de 1ère espèce.

(b) Lemme de Neyman-Pearson

La région critique optimale au seuil α ('zone de rejet') est W, et peut se mettre sous la forme:

$$\boxed{W = \{\Lambda_n > K_\alpha\}} \quad avec \quad \boxed{\Lambda_n = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{S_n}} \quad et \quad \boxed{K_\alpha = \mu_0 + \frac{\sigma_0}{\sqrt{n}} F^{-1}_{T_{n-1}}(1 - \alpha)}$$

où $F_{T_{n-1}}$ est la fonction de répartition de T_{n-1} (Loi de Student à $n-1$ degrés de liberté).

(c) Programmation de la règle de décision.

```
# Règle de décision : 'S_n' suit elle une loi N('u0', *) ou N('u1', *) ?
# sous un risque de 1ère espèce 'a'
# la fonction renvoie True si l'hypothèse H0 : u = u0 est rejeté, False sinon
delta <- function(S_n, a, u0, u1) {
  n        <- length(S_n)
  X_n_bar  <- 1.0 / n * sum(S_n)
  sigma_2  <- 1.0 / (n - 1) * sum((S_n - X_n_bar) * (S_n - X_n_bar))
  lambda_n <- sqrt(n / sigma_2) * (X_n_bar - u0)
  K_a      <- qt(p=1.0 - a, df=n-1)
  return (lambda_n > K_a)
}

if (delta(rnorm(n=20, mean=1, sd=2), 0.05, 1.0, 1.5)) {
  print("H0 est rejeté: u=u1")
} else {
  print("H0 non rejeté: u=u0")
}
```

```
## [1] "H0 non rejeté: u=u0"
```

2) Test sur $N = 100$ distributions $N(\mu = 1, \sigma = 2)$ (en supposant σ inconnue)

(a)

```r
for (i in 1:100) {
  print(paste(i, " : ", delta(rnorm(n=20, mean=1, sd=2), 0.05, 1.0, 1.5), sep=""))
}
```

```
## [1] "1 : FALSE"
## [1] "2 : FALSE"
## [1] "3 : FALSE"
## [1] "4 : FALSE"
## [1] "5 : FALSE"
## [1] "6 : FALSE"
## [1] "7 : FALSE"
## [1] "8 : FALSE"
## [1] "9 : FALSE"
## [1] "10 : FALSE"
## [1] "11 : FALSE"
## [1] "12 : FALSE"
## [1] "13 : FALSE"
## [1] "14 : FALSE"
## [1] "15 : FALSE"
## [1] "16 : FALSE"
## [1] "17 : FALSE"
## [1] "18 : FALSE"
## [1] "19 : FALSE"
## [1] "20 : TRUE"
## [1] "21 : FALSE"
## [1] "22 : FALSE"
## [1] "23 : FALSE"
## [1] "24 : FALSE"
## [1] "25 : FALSE"
## [1] "26 : FALSE"
## [1] "27 : FALSE"
## [1] "28 : FALSE"
## [1] "29 : FALSE"
## [1] "30 : FALSE"
## [1] "31 : FALSE"
## [1] "32 : FALSE"
## [1] "33 : FALSE"
## [1] "34 : FALSE"
## [1] "35 : FALSE"
## [1] "36 : FALSE"
## [1] "37 : FALSE"
## [1] "38 : FALSE"
## [1] "39 : FALSE"
## [1] "40 : FALSE"
## [1] "41 : FALSE"
## [1] "42 : FALSE"
## [1] "43 : FALSE"
## [1] "44 : FALSE"
## [1] "45 : FALSE"
## [1] "46 : FALSE"
## [1] "47 : FALSE"
```

```
## [1] "48 : FALSE"
## [1] "49 : FALSE"
## [1] "50 : FALSE"
## [1] "51 : FALSE"
## [1] "52 : FALSE"
## [1] "53 : FALSE"
## [1] "54 : FALSE"
## [1] "55 : FALSE"
## [1] "56 : FALSE"
## [1] "57 : FALSE"
## [1] "58 : FALSE"
## [1] "59 : FALSE"
## [1] "60 : FALSE"
## [1] "61 : FALSE"
## [1] "62 : FALSE"
## [1] "63 : TRUE"
## [1] "64 : FALSE"
## [1] "65 : FALSE"
## [1] "66 : FALSE"
## [1] "67 : FALSE"
## [1] "68 : FALSE"
## [1] "69 : FALSE"
## [1] "70 : FALSE"
## [1] "71 : FALSE"
## [1] "72 : FALSE"
## [1] "73 : FALSE"
## [1] "74 : FALSE"
## [1] "75 : FALSE"
## [1] "76 : FALSE"
## [1] "77 : FALSE"
## [1] "78 : FALSE"
## [1] "79 : FALSE"
## [1] "80 : FALSE"
## [1] "81 : FALSE"
## [1] "82 : FALSE"
## [1] "83 : FALSE"
## [1] "84 : FALSE"
## [1] "85 : FALSE"
## [1] "86 : FALSE"
## [1] "87 : FALSE"
## [1] "88 : FALSE"
## [1] "89 : FALSE"
## [1] "90 : FALSE"
## [1] "91 : FALSE"
## [1] "92 : FALSE"
## [1] "93 : FALSE"
## [1] "94 : FALSE"
## [1] "95 : FALSE"
## [1] "96 : FALSE"
## [1] "97 : FALSE"
## [1] "98 : FALSE"
## [1] "99 : FALSE"
## [1] "100 : FALSE"
```

On observe que pour certains echantillons, le test de Student rejette l'hypothèse $H_0 : \mu = \mu_0$ (alors qu'on sait que cette hypothèse est vraie)

(b)

```
for (a in c(0.2, 0.1, 0.05, 0.01)) {
  # TODO
}
```

On pose:

$$Z_n = \frac{L(X;\mu_1)}{L(X;\mu_0)}$$

On peut montrer que

$$Z_N = exp(\frac{1}{\sigma_0^2}(\mu_1 - \mu_0)\sum_{i=1}^{n} X_i) \quad exp(-\frac{n}{2\sigma_0^2}(\mu_1^2 - \mu_0^2))$$

$$\boxed{W = \{(x_1, ..., x_n); \frac{1}{n}\sum_{i=1}^{n} x_i > K_\alpha\}} \quad et \quad \boxed{K_\alpha = \mu_0 + \frac{\sigma_0}{\sqrt{n}}\phi^{-1}(1-\alpha)}$$