Inferenza Statistica

Note in R - Simulazioni Monte Carlo

V. Gioia (and N. Torelli, and Pappadà R.)

19/11/2024

Contents

Simulazione Monte Carlo

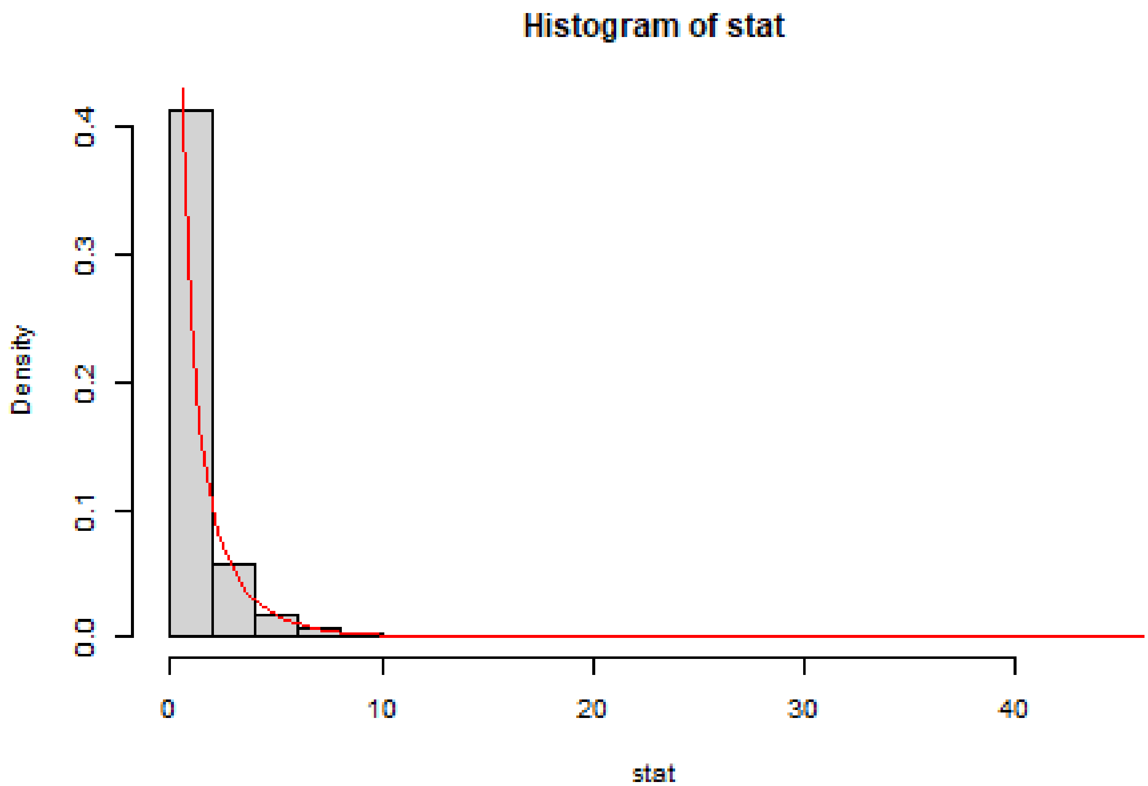
Esercizio 1

Sia $(Y_1, \ldots, Y_{20})$ un campione casuale proveniente da Y distribuita secondo una normale con media 0 e varianza σ^2. Siano $\overline{Y}$ la media campionaria e S^2 la varianza campionaria (corretta).

- Si determini la distribuzione di probabilità della v.a. $20\overline{Y}^2/S^2$

Dalla teoria abbiamo verificato che essa risulta una $F_{1,19}$. Verifichiamolo mediante simulazione

```r
set.seed(13)
# Fissiamo un valore per sigma
sigma <- 2
# Numero replicazioni
R <- 100000
n <- 20
stat <- rep(NA, R)
for(i in 1 : R){
  x <- rnorm(n, 0 ,sigma)
  s2 <- sum((x- mean(x))^2)/(n-1)
  stat[i] <- n*mean(x)^2/s2
}
hist(stat, freq = FALSE, breaks = 30)
curve(df(x, 1, 19), from = 1e-15, col = "red", add = TRUE)
```



- Si determini la costante c per la quale $P(-c < S/\overline{Y} < c) = 0.05$.

Dalla teoria abbiamo verificato che la costante c risulta essere pari

$$\sqrt{20}/c = t_{0.975;19} \implies c = \sqrt{20}/t_{0.975;19} \implies c = 2.1367$$

```
set.seed(10)
# Fissiamo un valore per sigma
sigma <- 2
# Numero replicazioni
R <- 100000
n <- 20
stat <- rep(NA, R)
count <- rep(NA, R)
for(i in 1 : R){
  x <- rnorm(n, 0 ,sigma)
  s <- sqrt(sum((x- mean(x))^2)/(n - 1))
  stat[i] <- s/mean(x)
  thresh <- sqrt(n)/qt(0.975, n - 1)
  count[i] <- (stat[i] >= -thresh & stat[i] <= thresh)
}
mean(count)
```

```
## [1] 0.05029
```

In alternativa potevamo svolgere lo stesso sfruttando la distribuzione di Fisher, secondo cui $n/c^2 = f_{0.95;1,19} = 4.38075 \implies c = \sqrt{n/f_{0.95;1,19}}$

Esercizio 2

Una macchina per l'imbottigliamento dell'olio 'e calibrata in modo tale che la quantit'a effettiva di olio in ogni bottiglia, X, 'e una variabile aleatoria normale con media $0.95l$. Si consideri un campione di 20 bottiglie e siano $\bar{X}$ la media campionaria. Si determinino i valori h, k in modo che

- $P(|\bar{X} - 0.95| > hD) = 0.05$, con $D^2 = \frac{1}{20}\sum_{i=1}^{20}(X_i - 0.95)^2$

Dalla teoria abbiamo trovato che

$$h = \frac{f_{0.95;1,19}}{19 + f_{0.95;1,19}} = 0.433$$

Verifichiamolo mediante simulazione Monte Carlo

```
set.seed(11)
sigma <- 2
R <- 100000
n <- 20
stat <- rep(NA, R)
count <- rep(NA, R)
for(i in 1 : R){
  x <- rnorm(n, 0.95, sigma)
  D2 <- sum((x - 0.95)^2)/n
  stat[i] <- abs(mean(x) - 0.95) / sqrt(D2)
  count[i] <- stat[i] > sqrt(qf(0.95, 1, n - 1)/((n - 1) + qf(0.95, 1, n - 1)))
}
mean(count)
```

```
## [1] 0.05064
```

- $P(|\bar{X} - 0.95| > kG) = 0.05$, con $G^2 = \frac{1}{20}\sum_{i=1}^{20}(X_i - \bar{X})^2$

Dalla teoria abbiamo trovato che $k = \frac{t_{0.975,n-1}}{\sqrt{n-1}} = 0.4801726$

```
set.seed(13)
sigma <- 2
R <- 100000
n <- 20
stat <- rep(NA, R)
count <- rep(NA, R)
for(i in 1 : R){
  x <- rnorm(n, 0.95, sigma)
  G2 <- sum((x - mean(x))^2)/n
  stat[i] <-abs(mean(x) - 0.95) / sqrt(G2)
  count[i] <- stat[i] > qt(0.975, n - 1)/sqrt(n - 1)
}
mean(count)
```

```
## [1] 0.0508
```