

Intervalos de Confianza

Francisco Paz

23/8/2019

```
library(pander) #Paquetería para sacar tablas más bonitas
library(tidyverse)
```

Intervalos de Confianza

Función de distribución acumulada empírica

$$\hat{F}_n = \frac{\sum_{i=1}^n I(X_i \leq x)}{n}$$

Del teorema 7.5 encontramos que para la creación de intervalos de confianza (desde el enfoque no paramétrico) podemos utilizar la siguiente función

$$\mathbb{P}(L(x) \leq F(x) \leq U(x)) \geq 1 - \alpha$$

Donde

$$L(x) = \max\{\hat{F}_n - \epsilon_n, 0\}$$

$$U(x) = \min\{\hat{F}_n + \epsilon_n, 1\}$$

$$\epsilon_n = \sqrt{\frac{1}{2n} \log\left(\frac{2}{\alpha}\right)}$$

Ahora el chiste es programar estas funciones

```
f_empirica <- function(x, y){
  n <- length(x)
  sum <- 0
  for (i in 1:length(x)) {
    if(x[i] <= y){
      sum = sum + 1
    }
  }
  sum/n
}

Int_inf <- function(x,y,a){
  res <- -100
  n <- length(x)
  e <- sqrt((1/(2*n))*log(2/a))
  val <- f_empirica(x,y)-e
  if(val < 0) res <- 0
  if(val >= 0) res <- val
  return(res)
}
```

```

}

Int_sup <- function(x,y,a){
  res <- -100
  n <- length(x)
  e <- sqrt((1/(2*n))*log(2/a))
  val <- f_empirica(x,y)+e
  if(val < 1) res <- val
  if(val >= 1) res <- 1
  return(res)
}

set.seed(2485)
x <- rnorm(601,0,1)
f_empirica(x,1)

## [1] 0.843594
Int_inf(x,1,0.05)

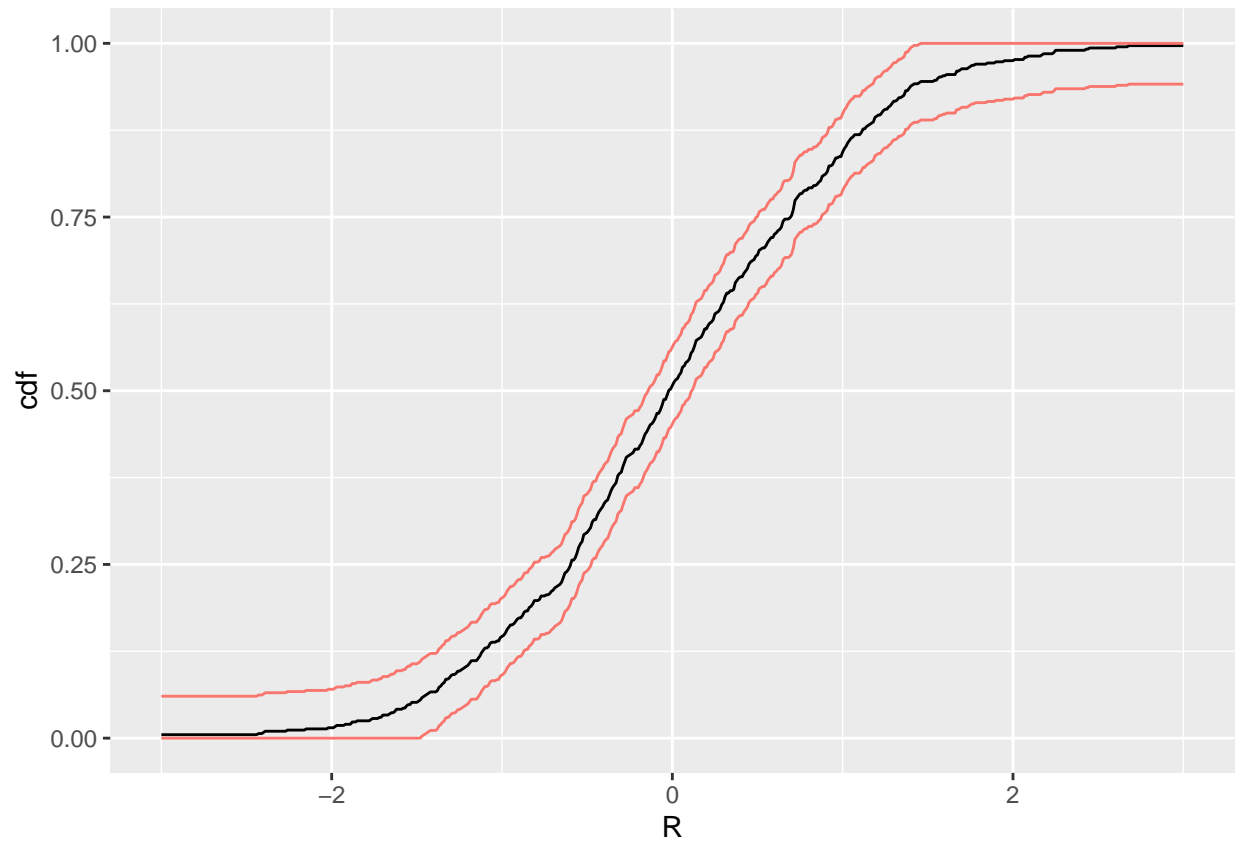
## [1] 0.7881959
Int_sup(x,1,0.05)

## [1] 0.8989921
alpha <- 0.05
emp <- c(rep(0,601))
inf <- c(rep(0,601))
sup <- c(rep(0,601))
cont <- 1
for (i in seq(-3,3,0.01)) {
  emp[cont] <- f_empirica(x,i)
  inf[cont] <- Int_inf(x,i,alpha)
  sup[cont] <- Int_sup(x,i,alpha)
  cont = cont + 1
}

datos <- as.data.frame(cbind(emp,inf,sup))

ggplot(datos, aes(x = seq(-3,3,0.01))) +
  geom_line( y = emp, show.legend = FALSE) +
  geom_line(aes(y = inf, col = 'red'), show.legend = FALSE) +
  geom_line(aes(y = sup, col = 'red'), show.legend = FALSE) +
  labs( x = 'R', y = 'cdf')

```



Podríamos experimentar con otras distribuciones u otros rangos ¿Qué deberíamos de cambiar en el código?