Centro de Investigación en Matemáticas, A.C.
Inferencia Estadística

Ayudantía 2

José Miguel Saavedra Aguilar

Esto es un documento creado para ejemplificar el uso de knitr. Es recomendable consultar el libro de Yihui Xie [1] para mayor información.

```
read_chunk('Ayudantia 1.R')
```

1. Ejemplo 1

Para una v.a. $X \sim \exp$, graficamos la función de densidad para distintas tasas $\lambda = 5, 2, 0.5$.

```
# Creamos un vector de 100 puntos en el soporte.
# En nuestro caso, tomamos de 0 a 10.
x <- seq(0, 10, length = 100)
```

```
# y es la función de densidad elegida evaluada en los puntos de x
# Para obtener la función de densidad, se toma "d"+nombre de la distribución, por ejemplo dnorm
y1 <- dexp(x, rate = 5)
y2 <- dexp(x, rate = 0.5)
y3 <- dexp(x, rate = 2)
```

Ahora, graficamos las respectivas funciones de distribución de X.

```
# f es la función de distribución asociada evaluada en los puntos de x
# Para obtener la función de densidad, se toma
# "p"+nombre de la distribución, por ejemplo pnorm o pexp
f1 <- pexp(x, rate = 5 )
f2 <- pexp(x, rate = 0.5 )
f3 <- pexp(x, rate = 2 )
```

Para $\lambda = 0.5$, simulamos una muestra aleatoria de tamaño $n = 100$ datos de X.

```
# Tomamos n=100 y simulamos n datos pseudoaleatorios de la distribución elegida
# Para obtener los datos aleatorios, se toma
# "r"+nombre de la distribución, por ejemplo rnorm o rexp
n <- 100
z <- rexp(n,rate = 0.5)
```

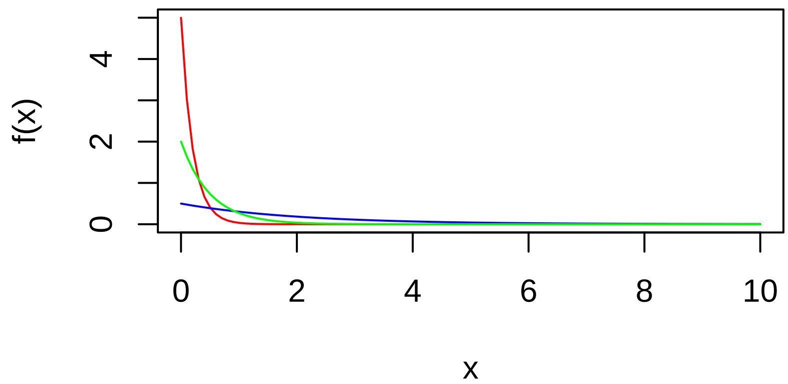

Figura 1: Densidad de una v.a. exponencial para distintas medias

Los datos que simulamos tienen media muestral 2.1625306 y varianza muestral 4.8139365. Ordenamos los puntos simulados en $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$. Les asociamos k_i definido por

$$k_i = \frac{i}{n+1} \tag{1}$$

```
#Ahora vamos a ordenar los datos como indica la tarea.
z2 <- sort(z)
# Inicializamos k =i/n+1
k <- numeric()

for (i in 1:n) {
   k[i] <- i/(n+1)
}
```

Para conocer más sobre la función de distribución empírica, pueden consultar Wikipedia.

2. Ejemplo 2

Una variable aleatoria discreta X tiene función de masa de probabilidad:

x	0	1	2	3	4
$p(x)$	0.1	0.2	0.2	0.2	0.3

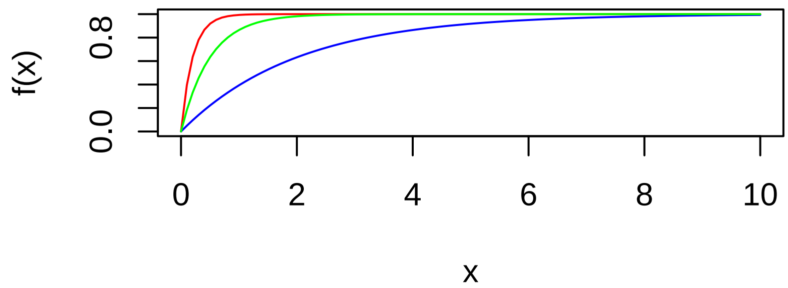

Figura 2: Distribución exponencial para distintas medias

Utilicen el teorema de la transformación inversa para generar una muestra aleatoria de tamaño 1000 de la distribución de X. Construyan una tabla de frecuencias relativas y comparen las probabilidades empíricas con las teóricas.
Repitan considerando la función de R sample.

```
#Ejemplo 2
set.seed(157017)
prob <- c(.1,.3,.5,.7,1)
```

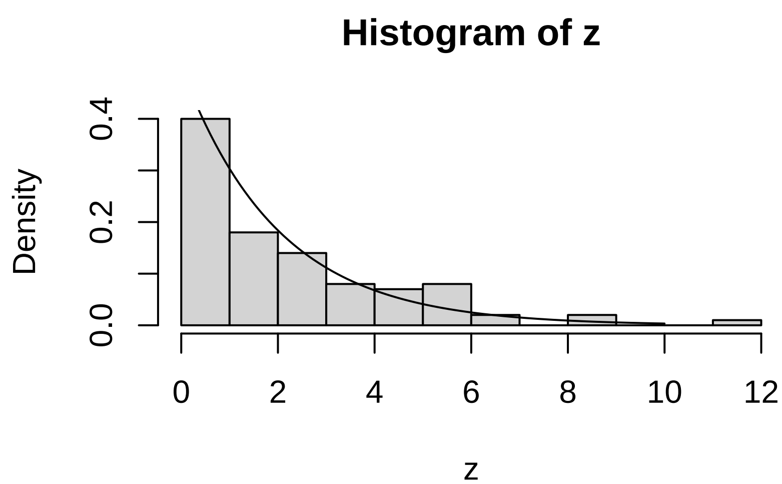

Figura 3: Histograma de una m.a. Exponencial con $\lambda = 0.5$

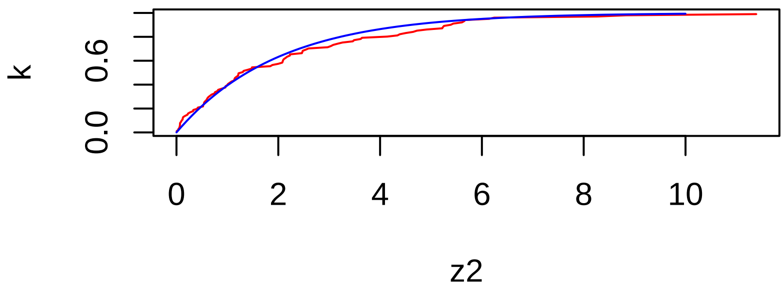

Figura 4: Comparación entre la función de densidad empírica y la teórica a partir de la muestra

```
frec <- findInterval(runif(1000),prob)
table(frec)/1000

## frec
##     0     1     2     3     4
## 0.125 0.168 0.224 0.185 0.298

set.seed(157017)
frecSample<-sample(x=0:4,size=1000,replace = TRUE,prob = c(0.1,0.2,0.2,0.2,0.3))
table(frecSample)/1000

## frecSample
##     0     1     2     3     4
## 0.103 0.195 0.224 0.185 0.293
```

3. Ejemplo 3

Obtengan una muestra de $10,000$ números de la siguiente distribución discreta:

$$p(x) = \frac{2x}{k(k+1)}, x = 1, 2, \ldots, k$$

para $k = 100$

```
#Ejemplo 3
p<-function(n){
  values=sample(1:100,n,replace=T)
  prob=(2*values)/(100*101)
  return(prob)
}
head(p(10000),n=30)
```

```
##  [1] 0.0091089109 0.0172277228 0.0106930693 0.0011881188 0.0112871287
##  [6] 0.0164356436 0.0001980198 0.0065346535 0.0196039604 0.0061386139
## [11] 0.0116831683 0.0108910891 0.0106930693 0.0194059406 0.0077227723
## [16] 0.0108910891 0.0186138614 0.0186138614 0.0182178218 0.0003960396
## [21] 0.0091089109 0.0142574257 0.0166336634 0.0172277228 0.0081188119
## [26] 0.0037623762 0.0011881188 0.0196039604 0.0061386139 0.0150495050
```

4. Ejemplo 4

Una compañía de seguros tiene 1000 asegurados, cada uno de los cuales presentará de manera independiente una reclamación en el siguiente mes con probabilidad $p = 0.09245$. Suponiendo que las cantidades de los reclamos hechos son variables aleatorias Gamma(7000,1), hagan simulación para estimar la probabilidad de que la suma de los reclamos exceda $\$500,000$.

```
#Ejemplo 4
mayor <- 0
for(i in 1:10000){
  reclammaciones<-sum(rbinom(1000, 1, 0.09245))
  montos <- sum(rgamma(reclammaciones, 7000, 1))
  if(montos > 500000){
    mayor= mayor+1
  }
}
mayor/10000
```

```
## [1] 0.9897
```

Referencias

[1] Y. Xie, *Dynamic Documents with R and knitr, Second Edition*, 2nd ed., ser. Chapman & Hall/CRC The R Series. Philadelphia, PA: Chapman & Hall/CRC, Jun. 2015.