Tarea 0

Jose Aznar

February 2, 2018

En el primero problema debíamos simular una distribución binomial (n=12 y p=1/6) por dos métodos y calcular la probabilidad de que hubiera más de 7 aciertos.

```r
# problema 1 - a)
# primero obtenemos una tabla donde pondremos los valores
dist1 <- table(0:12)
# iteramos simulando una binomial con p=1/6 y n=12
for(i in 0:12){
  dist1[i+1] <- dbinom(i, 12, prob=1/6)
}
# sacamos la distribución acumulada
dist1 <- cumsum(dist1)
# imprimimos el resultado
print(dist1, digits=12)
```

```
##              0              1              2              3              4
## 0.112156654785 0.381332626268 0.677426194899 0.874821907320 0.963649977909
##              5              6              7              8              9
## 0.992074960498 0.998707456435 0.999844455739 0.999986580652 0.999999213977
##             10             11             12
## 0.999999971977 0.999999999541 1.000000000000
```

```r
# problema 1 - b)
# primero obtenemos una tabla donde pondremos los valores
dist2 <- table(0:12)
# iteramos simulando una binomial con p=1/6 y n=12
for(i in 0:12){
  dist2[i+1] <- pbinom(i, 12, prob=1/6)
}
# imprimimos el resultado
print(dist2, digits=12)
```

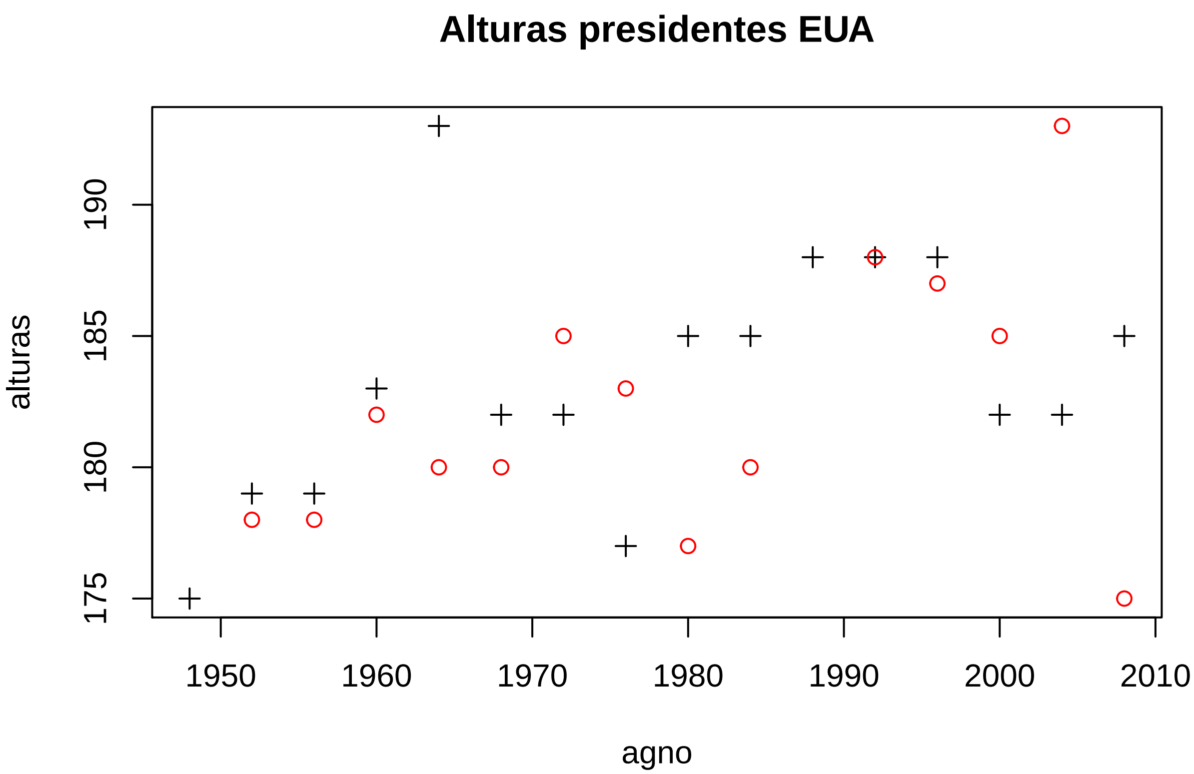
```
##
##              0              1              2              3              4
## 0.112156654785 0.381332626268 0.677426194899 0.874821907320 0.963649977909
##              5              6              7              8              9
## 0.992074960498 0.998707456435 0.999844455739 0.999986580652 0.999999213977
##             10             11             12
## 0.999999971977 0.999999999541 1.000000000000
```

```r
# problema 1 - c)
# calculamos P(X>7) como 1 - P(X<=6)
prob <- 1 - dist2[7]
# imprimimos el resultado
print(prob, digits=12)
```

```
##                6
## 0.00129254356463
```

En el segundo inciso, había que imprimir una comparacion entre las alturas de los presidentes que ganaron y su oponente en las ultimas 16 elecciones de EUA.

```r
# problema 2
library(car)
# primero ponemos la altura y el a?o en un objeto
ganador <- c(185, 182, 182, 188, 188, 188, 185, 185, 177, 182, 182, 193, 183, 179, 179, 175)
oponente <- c(175, 193, 185, 187, 188, 173, 180, 177, 183, 185, 180, 180, 182, 178, 178, 173)
agno <- c(2008, 2004, 2000, 1996, 1992, 1988, 1984, 1980, 1976, 1972, 1968, 1964, 1960, 1956, 1952, 1948
alturas <- data.frame(ganador, oponente, agno)
# ahora los graficamos
# la gr?fica principal con el ganador
plot(agno, ganador,
     xlab="agno", ylab="alturas",
     main="Alturas presidentes EUA", pch=3)
# la grafica superpuesta secundaria con el oponente
points(agno, oponente, col=2)
```



En el siguiente problema teníamos que simular una distribución Poison con 1000 y 10000 observaciones y compararla con los resultados teoricos.

```r
# problema 3
require(plotrix)
```

```
## Loading required package: plotrix
```

```r
# generamos 1000 observaciones de una poison con lambda=0.61
mil_obs <- rpois(1000, 0.61)
# desplegamos su media
mean(mil_obs)
```

```
## [1] 0.61
```

```r
# desplegamos su varianza
var(mil_obs)
```

```
## [1] 0.6225225
```

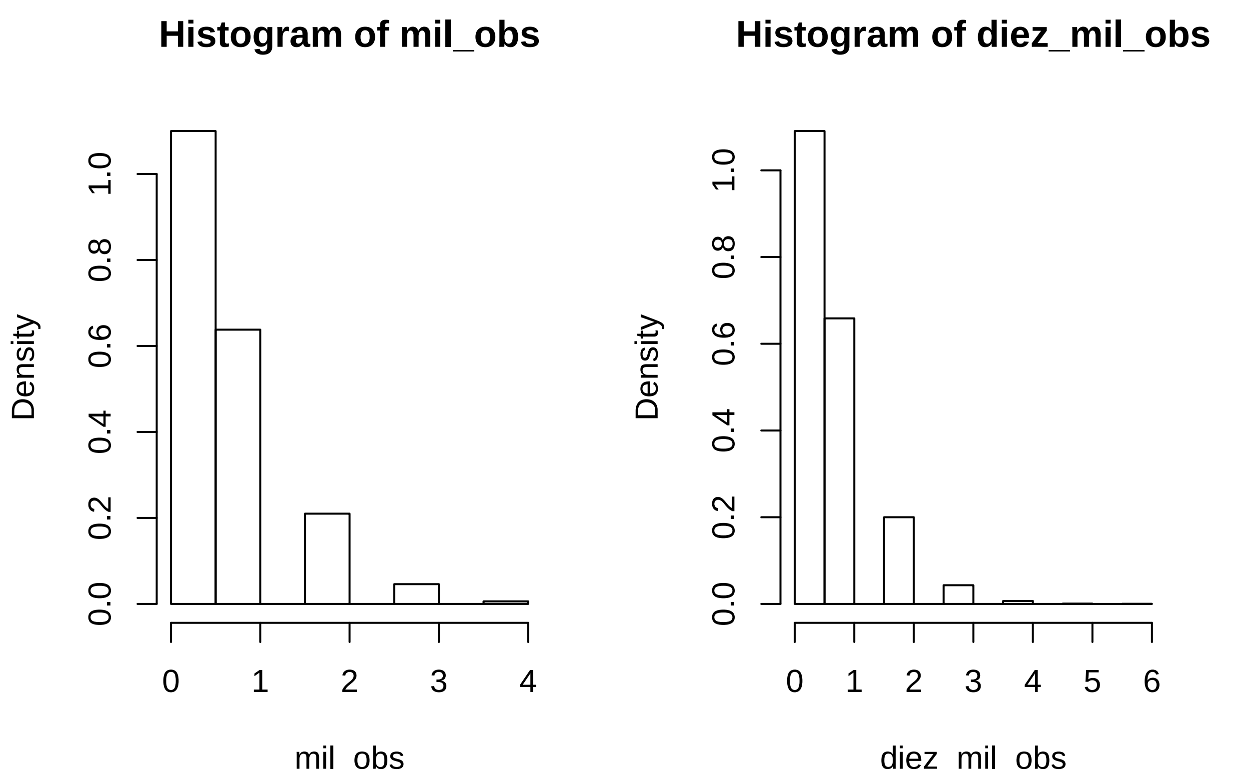
```r
# generamos 10000 observaciones de una poison con lambda=0.61
diez_mil_obs <- rpois(10000, 0.61)
# desplegamos su media
mean(diez_mil_obs)
```

```
## [1] 0.6098
```

```r
# desplegamos su varianza
var(diez_mil_obs)
```

```
## [1] 0.6174057
```

```r
# imprimimos la función de masa de probabilidad
par(mfrow=c(1,2))
hist(mil_obs, freq=FALSE)
hist(diez_mil_obs, freq=FALSE)
```



Como podemos ver, obtenemos un resultado aproximado al valor teorico de la media y la varianza (0.61), ademas, las graficas con la funcion de masa de probabilidad coinciden con la que deberia tener una Poison con lambda menor o igual a uno.

Para el problema 4, construimos una función que estima la desviacion estandar de una distribucion normal mediante el estimador de maxima verosimilitud (sesgado).

```r
sd.n <- function(muestra){
  # primero obtenemos la media y el tamaño de la muestra
  media <- mean(muestra)
  n <- length(muestra)
```

```r
  suma <- 0
  # ahora obtenemos el estimador de máxima verosimilitud para
  # la varianza de una normal que viene dado por la varianza muestral tomando
  # como estimador de miu a x barra
  for(i in 1:n){
    suma <- suma + (muestra[i] - media)**2
  }
  sigma <- (suma/n)**(1/2)

  return(sigma)
}

sd.n(rnorm(100000, mean = 10, sd = 12))
```

```
## [1] 12.0122
```

El siguiente enunciado de la tarea indicaba que hicieramos una funcion que calculara la norma euclidiana de un vector (base 2) y la probabaramos con 3 vectores muestra.

```r
norma <- function(vector){
  # calculamos la norma euclidiana sumando todos los elementos al cuadrado
  # sacando la raíz cuadrada de la suma
  suma <- 0
  for(elem in vector){
    suma <- suma + elem**2
  }
  res <- suma**(1/2)
  return(res)
}

# obtenemos la norma con la función antes especificada de vectores
v1 <- c(0,0,0,1)
norma(v1)
```

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

```r
v2 <- c(2, 5, 2, 4)
norma(v2)
```

```
## [1] 7
```

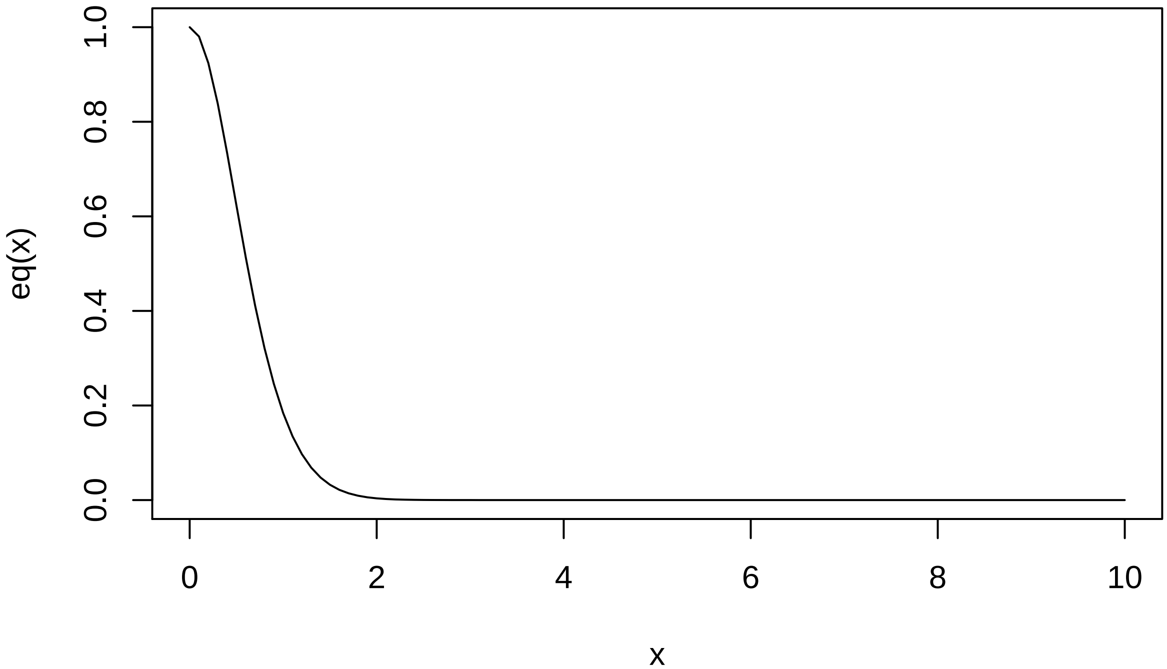
```r
v3 <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)
norma(v3)
```

```
## [1] 19.62142
```

En el problema 6 graficamos una funcion continua y luego calculamos su area bajo la curva entre cero e infinito.

```r
# problema 6
# primero definimos la ecuación
eq = function(x){exp(-x*x)/(1+x*x)}

# graficamos la función entre 0 y 10
curve(eq, from=0, to=10)
```




```r
# ahora integramos entre 0 e infinito
integrate(eq, lower=0, upper=Inf)
```

```
## 0.6716467 with absolute error < 8.3e-05
```

Luego, ocupamos la funcion de la norma euclidiana antes creada para aplicarsela a cada renglon de un vector que simula 20 observaciones de una normal.

```r
# problema 7
# primero definimos la matriz
x <- matrix(rnorm(20),10,2)

# ahora aplicamos la norma a cada uno de los elementos
apply(x, 1, norma)
```

```
##  [1] 2.2738508 1.7358988 0.2895146 2.1001187 1.2671961 2.1553222 0.5320119
##  [8] 1.6049193 0.7048932 0.6921326
```

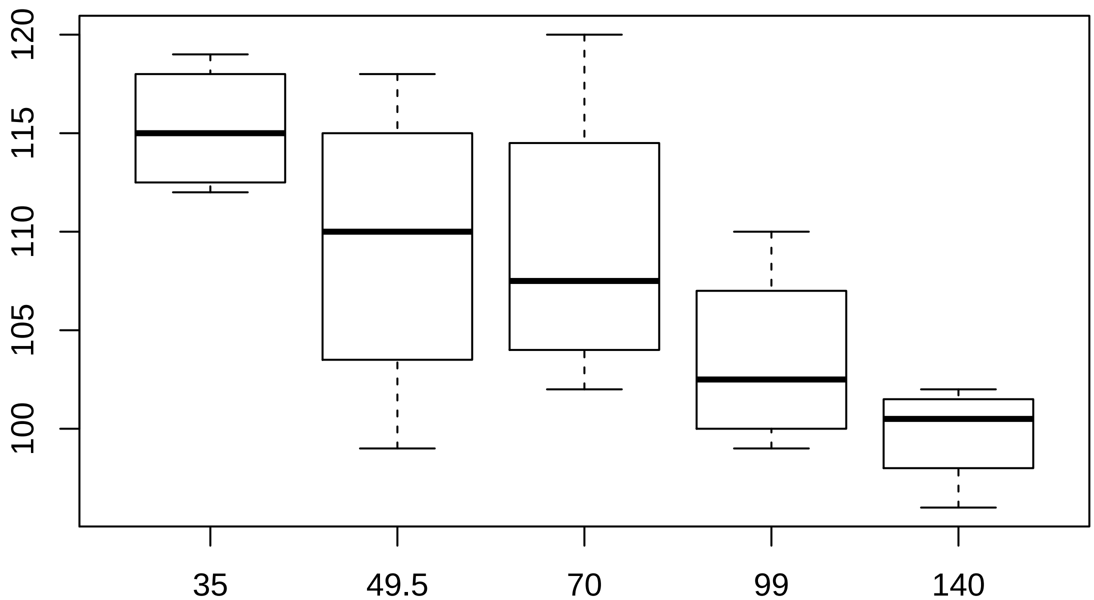
En el ulitmo problema con codigo, teniamos que ocupar la funcion boxplot para comparar diferentes valores de resistencias bajo una misma presion.

```r
# problema 8
# primero ponemos los datos en un objeto
presion <- c(35, 35, 35, 35,
              49.5, 49.5, 49.5, 49.5,
              70, 70, 70, 70,
              99, 99, 99, 99,
              140, 140, 140, 140)

resistencia <- c(112, 119, 117, 113,
                 108, 99, 112, 118,
                 120, 106, 102, 109,
                 110, 101, 99, 104,
                 100, 102, 96, 101)
```

```r
datos <- data.frame(presion, resistencia)

# ahora los graficamos
boxplot(resistencia~presion, data=datos)
```



Finalmente, en el ultimo problema (9) teniamos que clasificar sistemas por su tipo, identificar sus componentes y justificarlo.

A continuacion los resultados de los sistemas.

En una pequeña seccion de una fabrica existente se tiene un sistema un sistema estocastico, pues se tiene que estimar que tanto va a pasar un evento que depende totalmente de algun evento aleatorio; en el caso de la seccion que verifica la calidad en una fabrica de billetes, se tienen el tiempo que tarda en realizar un prueba y la cantidad de billetes por lote con errores como variables a estimar.

Por otro lado, en la congestion en una interseccion tenemos un sistema determinista, ya que, asumiendo que no hay accidentes y saben manejar los conductores, siempre que llegue una cierta cantidad de coches va a suceder lo mismo, esto esta determinado por el funcionamiento (por las reglas) de la interseccion.

Al igual que en la fabrica, en la sala de urgencias de un hospital se tiene un sistema estocastico, cuyos componentes vienen siendo el tiempo en tratar un paciente y el numero de pacientes que van llegando.

De la misma forma, en la pizzeria se tiene un sistema estocastico determinado por la cantidad de ordenes que llegan, el tiempo que se tardan en prepararlas y el tiempo que se tardan en repartirlas.

Contrariamente, la operacion de un shuttle de una agencia de autos de un aereopuerto es un sistema determinista, en el sentido de que hace viajes de ida y venida periodicamente, sin ser afectado por otros factores. Sin embargo, es cierto que los tiempos de llegada y de salida pueden ser aleatorios determinados por factores externos, como el trafico.