Simulacion: Tarea 01

Bernardo Mondragón Brozon, Karen Delgado Curiel, Rayan Garcia Fabian

August 28, 2018

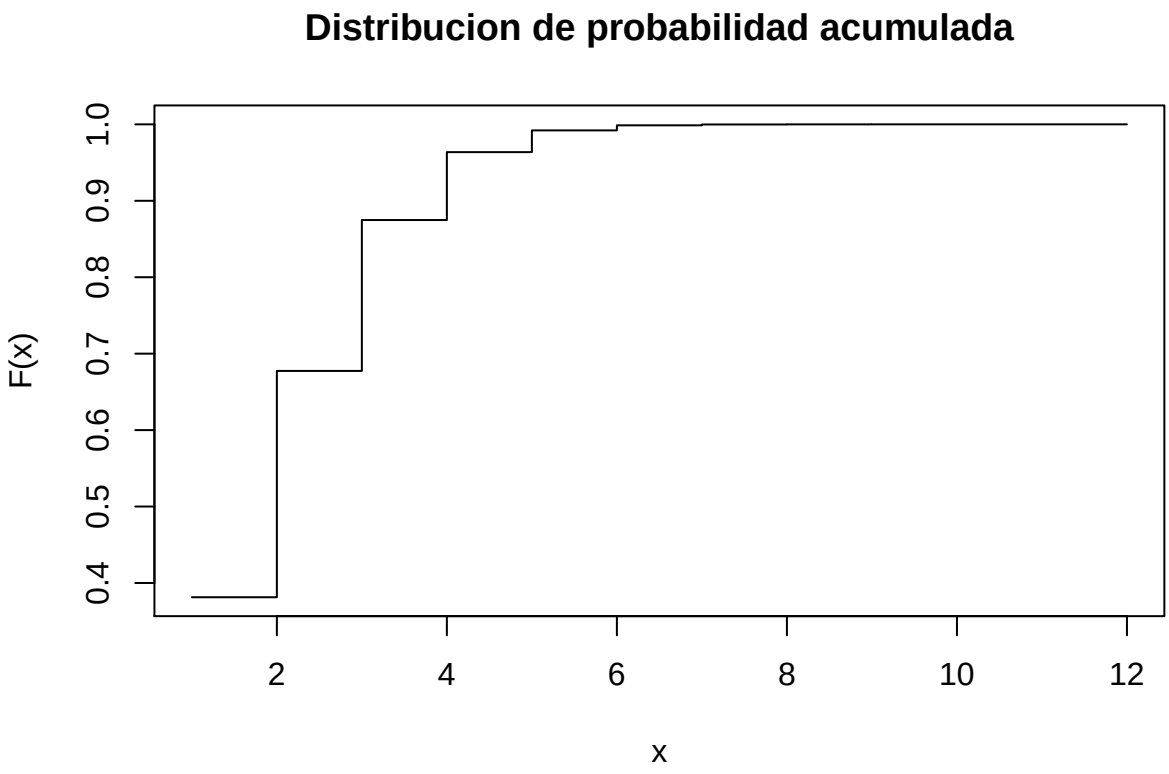
Problema 1

```
x <- c(1,2,3,4,5,6,7,8,9,10,11,12)
probs <- pbinom(q=x, size=12, prob=1/6)
table <- data.frame(x, probs)
names(table) <- c("x", "Pr{X<=x}")
kable(table, caption="Distribucion de X")
```

Table 1: Distribucion de X

x	Pr{X<=x}
1	0.3813326
2	0.6774262
3	0.8748219
4	0.9636500
5	0.9920750
6	0.9987075
7	0.9998445
8	0.9999866
9	0.9999992
10	1.0000000
11	1.0000000
12	1.0000000

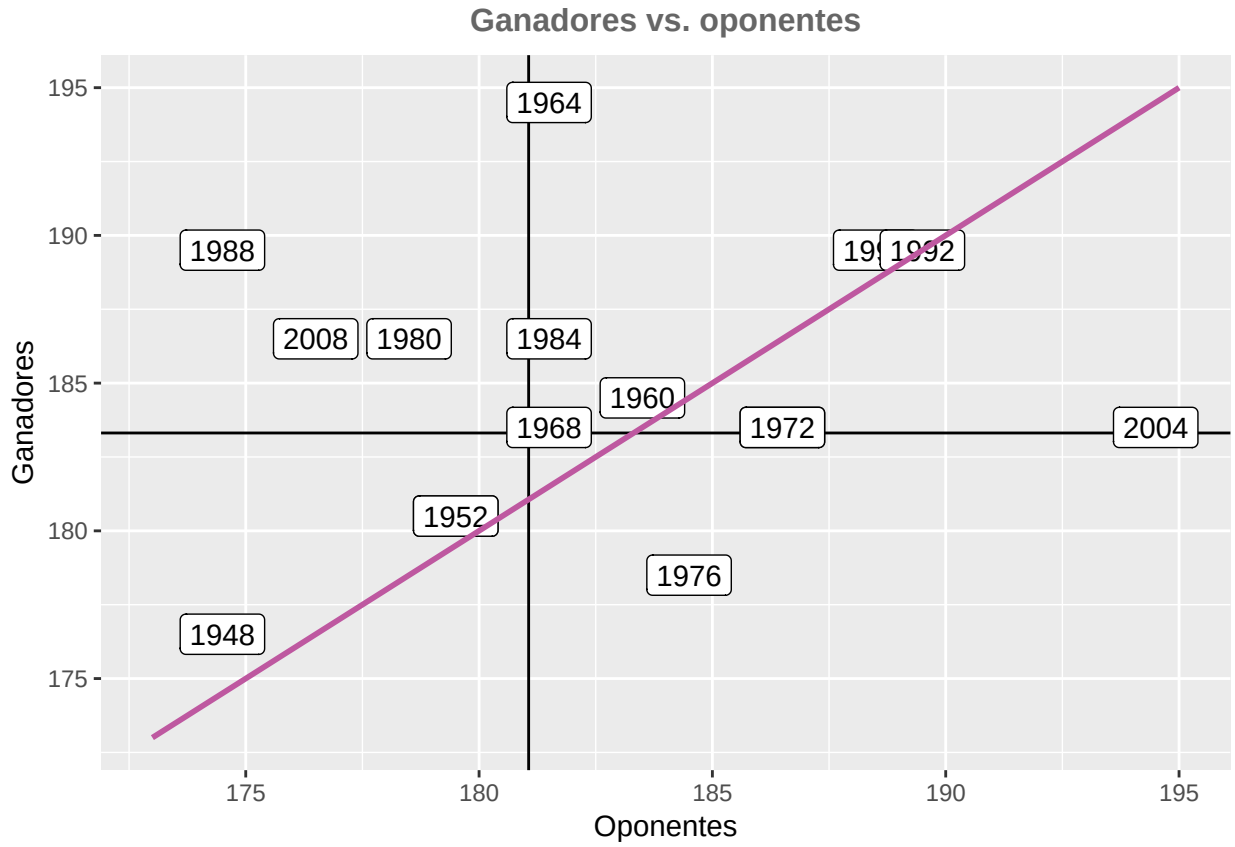
```
plot(probs, type="s", main="Distribucion de probabilidad acumulada", xlab="x", ylab="F(x)")
```



Problema 2

```
df <- read.csv("../HeightData.csv", sep=",", header = TRUE)
trim <- function (x) gsub("^\\s+|\\s+$", "", x)
getHeight <- function(Height) {
  res1 <- strsplit(x=as.character(Height), split="in")
  centimeters <- res1[1][[1]][2]
  res2 <- strsplit(x=centimeters, split="cm")
  number <- as.character(trim(res2))
  return(as.numeric(number))
}
winners_heights <- c()
losers_heights <- c()
for (i in 1:length(df$Height)) {
  height_w <- getHeight(df[i,]$Height)
  winners_heights <- c(winners_heights, height_w)
  height_l <- getHeight(df[i,]$Height.1)
  losers_heights <- c(losers_heights, height_l)
}
df <- data.frame(df$Year, df$Winner, winners_heights, df$Opponent, losers_heights)
names(df) <- c("Year", "Winner", "W.Height", "Opponent", "O.Height")
scatter_plot <- ggplot(data=df, mapping=aes(x=O.Height, y=W.Height)) +
  #geom_point(color=moradito) +
```

```
  geom_hline(aes(yintercept=mean(df$W.Height))) +
  geom_vline(aes(xintercept=mean(df$O.Height))) +
  geom_label(label=df$Year, nudge_x = 1.5, nudge_y = 1.5) +
  stat_function(fun=function(x) {x}, size=1, color=moradito) +
  ylim(c(173, 195)) +
  xlim(c(173, 195)) +
  labs(x="Oponentes", y="Ganadores") +
  ggtitle("Ganadores vs. oponentes") +
  theme(plot.title = element_text(hjust = 0.5, color = "#666666", face = "bold", size = 12))
scatter_plot
```



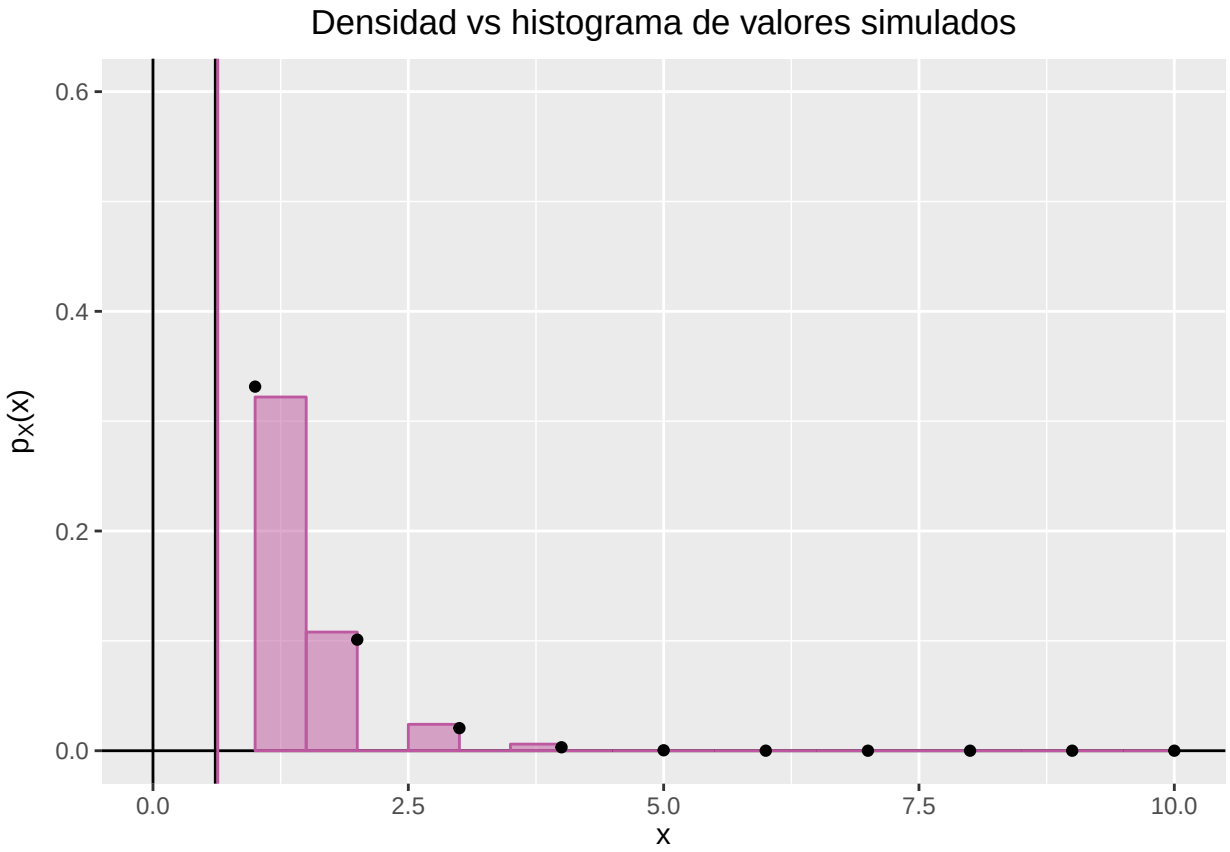
Las lineas rectas en el grafico anterior respresentan las medias.

Problema 3

Con n = 1000

```
obs_1 <- rpois(n=1000, lambda=0.61)
media_1 <- mean(obs_1)
variance_1 <- var(obs_1)
df_1 <- data.frame(obs_1)
vals <- dpois(seq(1, 10, by=1), lambda=0.61)
xs <- seq(1, 10, by=1)
df_pois <- data.frame(vals, xs)
```

```r
hist_vs_den <- ggplot() +
  geom_hline(yintercept = 0, size = .5) +
  geom_vline(xintercept = 0, size = .5) +
  labs(x = TeX('$x$'), y = TeX('$p_X(x)$')) +
  ggtitle("Densidad vs histograma de valores simulados") +
  theme(plot.title = element_text(hjust = 0.5)) +
  geom_vline(aes(xintercept=0.61)) +
  geom_vline(aes(xintercept=mean(obs_1)), color=moradito) +
  xlim(0,10) +
  ylim(0,0.6) +
  geom_histogram(data=df_1, aes(x=obs_1, y=.001*..count..), breaks=seq(1, 10, by=0.5), color=moradito, f
  geom_point(data=df_pois, aes(x=xs, y=vals))
hist_vs_den
```



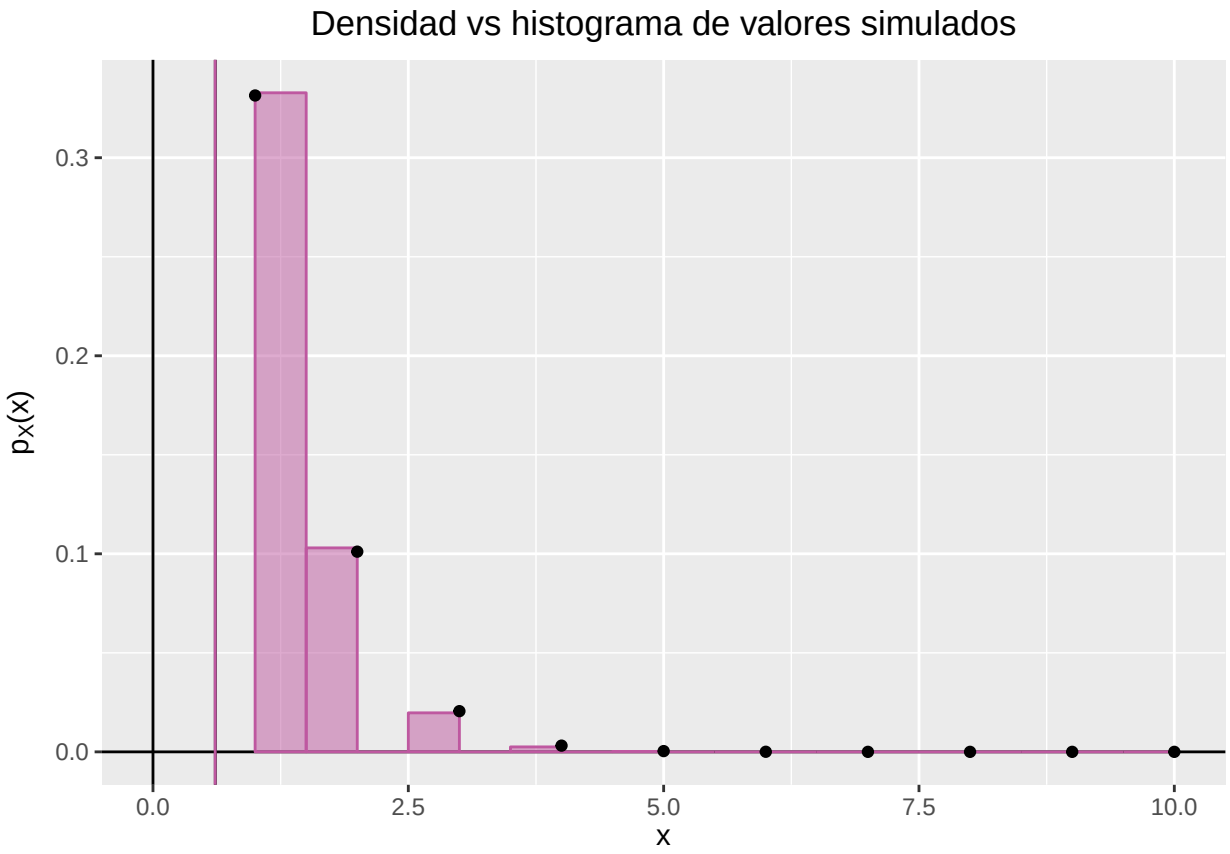

```r
tablita <- data.frame(c("Media", "Varianza"), c(mean(obs_1), var(obs_1)), c(0.61, 0.61))
names(tablita) <- c(" ", "Simulacion", "Teoricos")
kable(tablita, caption="Comparacion de los valores teoricos con los obtenidos por simulacion")
```

Table 2: Comparacion de los valores teoricos con los obtenidos por simulacion

	Simulacion	Teoricos
Media	0.6340000	0.61
Varianza	0.6647087	0.61

Con n = 10000

```
obs_2 <- rpois(n=10000, lambda=0.61)
media_2 <- mean(obs_2)
variance_2 <- var(obs_2)
df_2 <- data.frame(obs_2)
vals_2 <- dpois(seq(1, 10, by=1), lambda=0.61)
xs_2 <- seq(1, 10, by=1)
df_pois_2 <- data.frame(vals_2, xs_2)
hist_vs_den_2 <- ggplot() +
  geom_hline(yintercept = 0, size = .5) +
  geom_vline(xintercept = 0, size = .5) +
  labs(x = TeX('$x$'), y = TeX('$p_X(x)$')) +
  ggtitle("Densidad vs histograma de valores simulados") +
  theme(plot.title = element_text(hjust = 0.5)) +
  geom_vline(aes(xintercept=0.61)) +
  geom_vline(aes(xintercept=mean(obs_2)), color=moradito) +
  # xlim(0,10) +
  # ylim(0,20) +
  geom_histogram(data=df_2, aes(x=obs_2, y=.0001*..count..), breaks=seq(1, 10, by=0.5), color=moradito,
  geom_point(data=df_pois_2, aes(x=xs_2, y=vals_2))
hist_vs_den_2
```

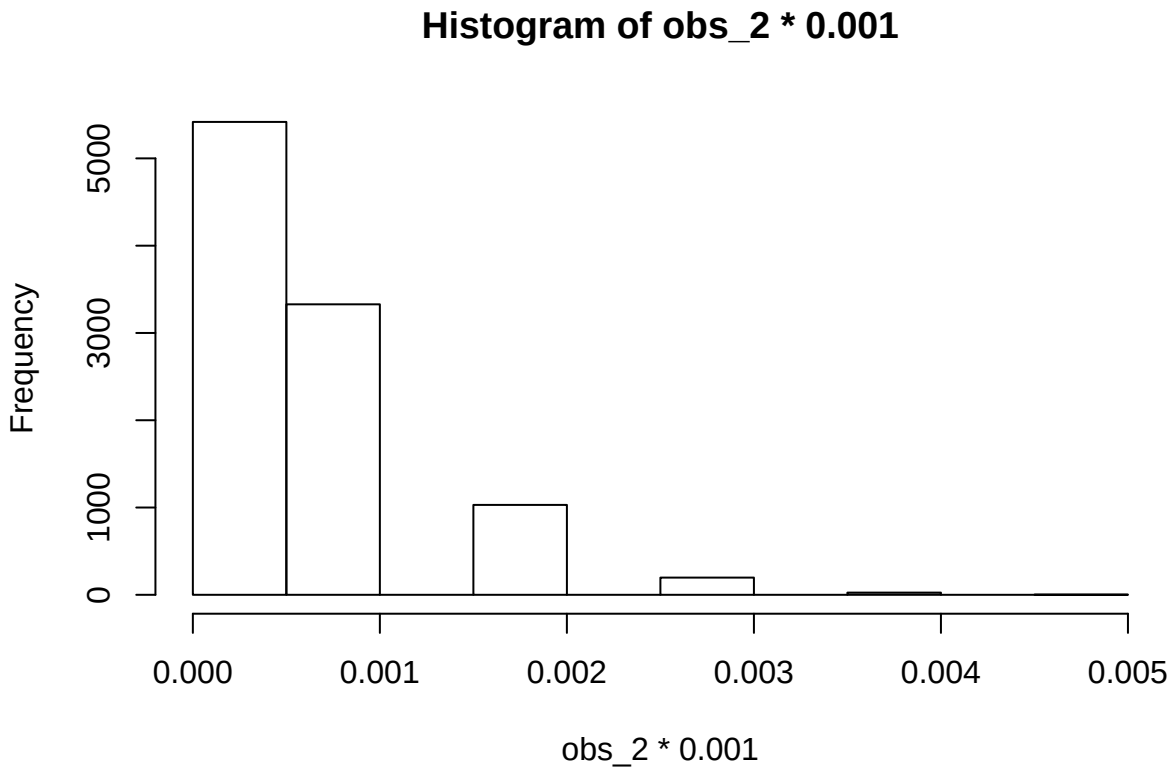



```
tablita_2 <- data.frame(c("Media", "Varianza"), c(mean(obs_2), var(obs_2)), c(0.61, 0.61))
names(tablita_2) <- c(" ", "Simulacion", "Teoricos")
kable(tablita_2, caption="Comparacion de los valores teoricos con los obtenidos por simulacion")
```

Table 3: Comparacion de los valores teoricos con los obtenidos por simulacion

	Simulacion	Teoricos
Media	0.6089000	0.61
Varianza	0.5964004	0.61

```
hist(obs_2*.001)
```



Problema 4

```
sd.n <- function(muestra) {
  aux <- 0
  for (i in 1:length(muestra)){
    aux <- aux + muestra[i]
  }
  return(aux/length(muestra))
}

muestra <- rpois(n = 100, lambda = 0.61)
sd.n(muestra)
```

```
## [1] 0.7
```

Problema 5

```
norma <- function(vec) sqrt(sum(vec^2))

vec1 <- c(0,0,0,1)
vec2 <- c(2,5,2,4)
vec3 <- c(1,2,3,4,5,6,7,8,9,10)

norma(vec1)
```

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

```
norma(vec2)
```

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

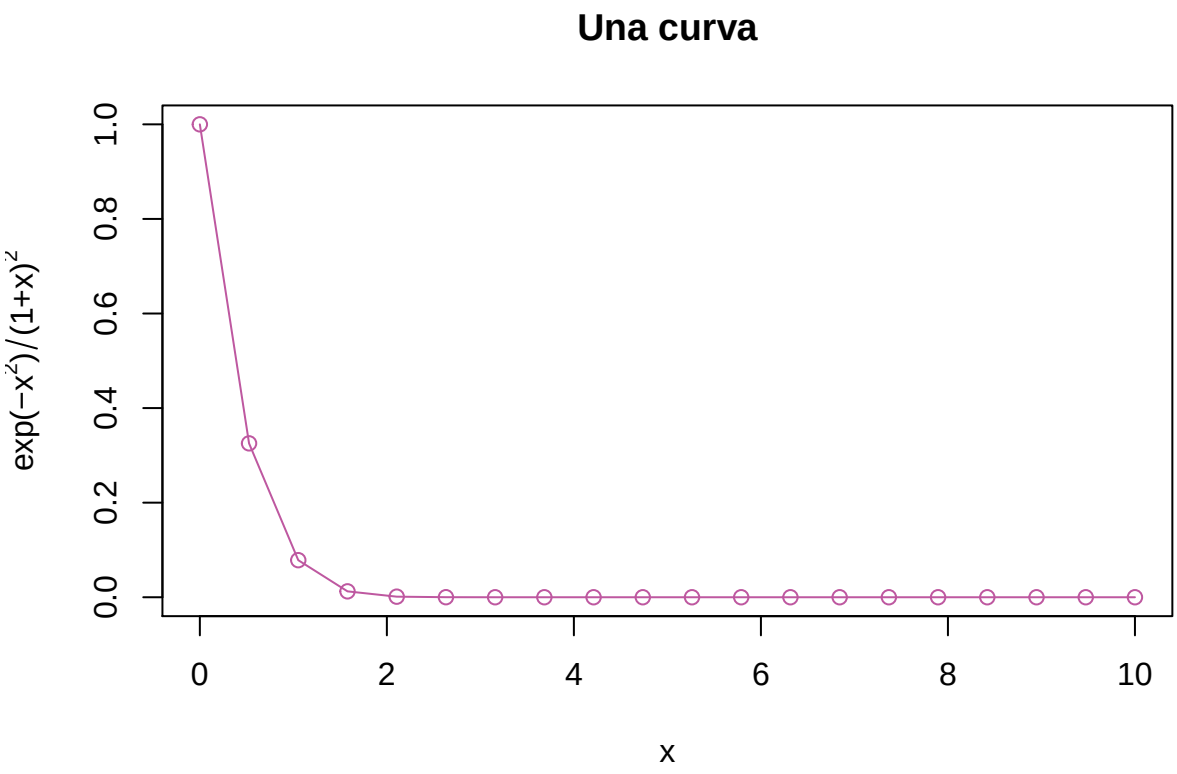
```
norma(vec3)
```

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

Problema 6

Usar la funcion curve para graficar la funcion f(x) en el intervalo 0 <= x <= 10. Luego usar la funcion integrate para caulcular el valor de la integral.

```
myFunction <- function(x) exp(-x^2)/(1+x)^2
curve(myFunction, from=0, to=10, n=20, type="o", col=moradito, xlab="x", ylab=TeX('exp(-x^2)/(1+x)^2'),
```




```
integrate(myFunction,lower=0, upper=Inf)
```

```
## 0.4378135 with absolute error < 3e-05
```

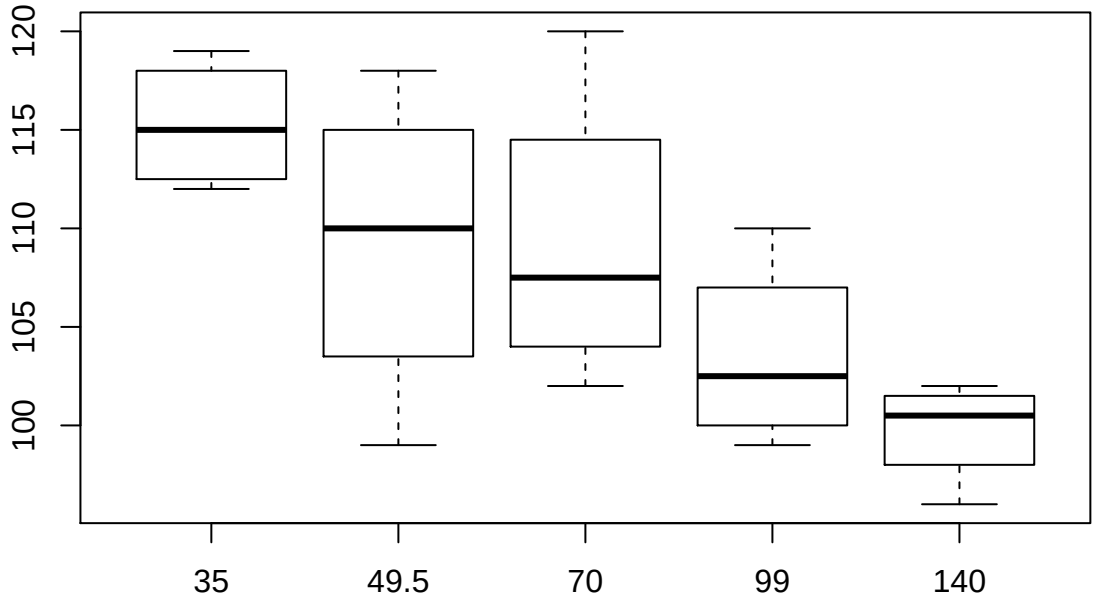
Problema 7

```
x <- matrix(data=rnorm(n=20), nrow = 10, ncol = 2)
apply(x, 1, norma)
```

```
##  [1] 0.7982793 1.6919034 0.1917173 0.4054778 1.3060986 0.6873092 1.5900296
##  [8] 1.1276743 1.5869412 0.8306302
```

Problema 8

```
presion<-c(rep(35,4),rep(49.5,4),rep(70,4),rep(99,4),rep(140,4))
factorRes<-c(112,119,117,113,108,99,112,118,120,106,102,109,110,101,99,104,100,102,96,101)
tabHojas<-as.data.frame(cbind(presion,factorRes))
boxplot(factorRes~presion)
```



Problema 9