Taller 1 simulacion

Kevin Valencia Romero y Tatiana Mora Acosta

2022-03-11

1. Al comparar los caracteres morfométricos de una especie de langostinos de gran importancia económica en Argentina se observaron diferencias de tamaño entre macho y hembra. En la siguiente tabla se presentan los datos (Largo total mm) de las morfometrías de machos y hembras obtenidos en diferentes recolectas (Ruiz & Mendia, 2008).

Table 1: Tabla #1

Hembras			Machos		
183.2	182.5	166.8	140.9	173.9	118.9
184.1	190.0	196.3	121.7	177.4	140.0
183.0	178.1	193.3	173.8	154.8	192.7
204.3	193.2	187.3	154.5	177.5	134.4
176.5	180.4	185.8	109.2	153.4	175.0
179.0	184.3	189.3	150.7	138.7	169.8
188.3	189.2	195.5	203.3	136.7	153.9
186.8	189.1	202.4	163.0	165.3	176.7
202.2	203.1	210.8	137.7	126.7	150.0

a. Haga un histograma con cinco clases y determine la distribución de los datos para cada sexo. Explique acerca de la distribución del tamaño para cada género.

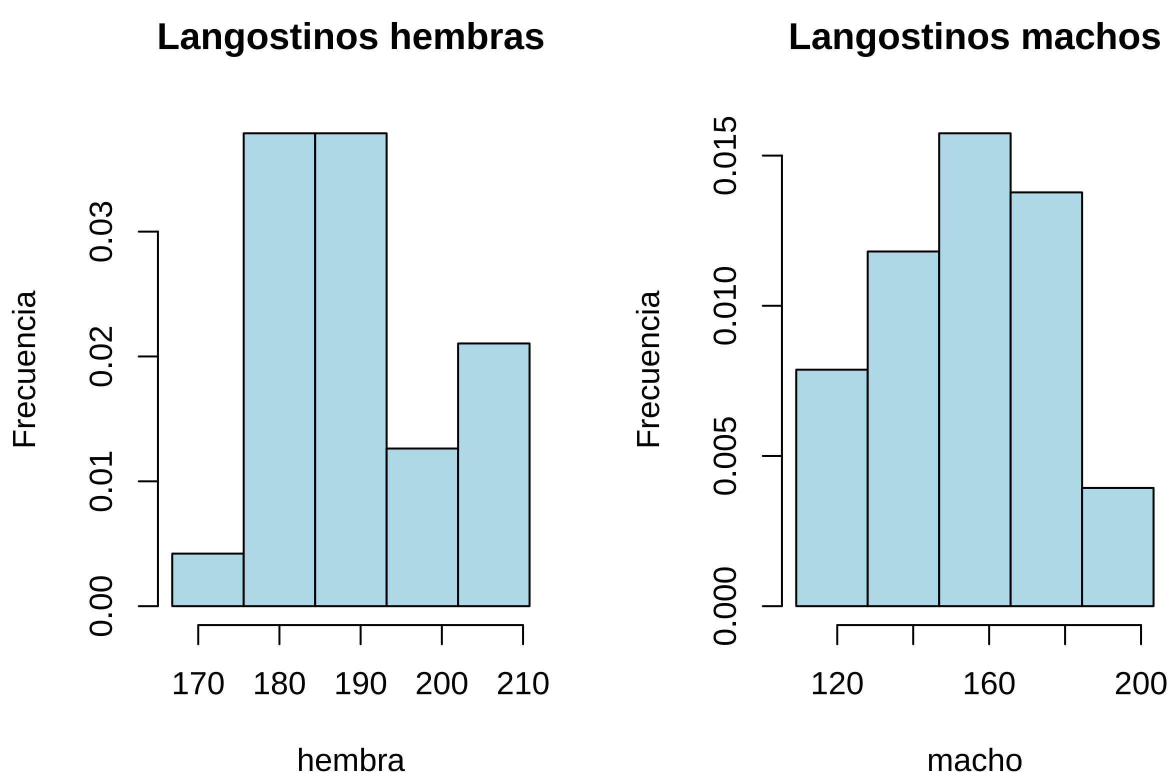
```r
hembra<- c(183.2,184.1,183.0,204.3,176.5,179.0,188.3,186.8,202.2,182.5,190.0,178.1
,193.2,180.4,184.3,189.2,189.1,203.1,166.8,196.3,193.3,187.3,185.8,189.3,195.5
,202.4,210.8)

macho<- c(140.9,121.7,173.8,154.5,109.2,150.7,203.3,163.0,137.7,173.9,177.4
,154.8,177.5,153.4,138.7,136.7,165.3,126.7,118.9,140.0,192.7,134.4,175.0
,169.8,153.9,176.7,150.0)

hembraOrdenado<- sort(hembra);hembraOrdenado
```

```
##  [1] 166.8 176.5 178.1 179.0 180.4 182.5 183.0 183.2 184.1 184.3 185.8 186.8
## [13] 187.3 188.3 189.1 189.2 189.3 190.0 193.2 193.3 195.5 196.3 202.2 202.4
## [25] 203.1 204.3 210.8
```

```r
machoOrdenado<- sort(macho);machoOrdenado
```

```
##  [1] 109.2 118.9 121.7 126.7 134.4 136.7 137.7 138.7 140.0 140.9 150.0 150.7
## [13] 153.4 153.9 154.5 154.8 163.0 165.3 169.8 173.8 173.9 175.0 176.7 177.4
## [25] 177.5 192.7 203.3
```

```
par(mfrow=c(1,2))

hist(hembra,breaks =seq(166.8,210.8, by=8.8),main = "Langostinos hembras", ylab = "Frecuencia", freq=FA
hist(macho, breaks =seq(109.2,203.3, by=18.82),main = "Langostinos machos", ylab = "Frecuencia", freq=F
```



Distribución de los datos para las hembra: los datos de la hembra determinan una distribución N , con un coeficiente de asimetria negativo ya que tiene una cola hacia la izquierda. Apreciando que es mas empinada a partir de los intervalos (175-193), se puede decir que es Leptocurtica

Distribución de los datos para los machos: los datos del macho determinan una distribución N , con un coeficiente de asimetria positivo ya que tiene una cola hacia la derecha. Apreciando que no es tan empinada como el anterior histograma pero tampoco es tan achatada como la platicurtica, se puede decir que es Mesocurtica.

b. Hallar el promedio y la desviación estándar para cada sexo. ¿Qué puede concluir?

```
mediaHembra<-mean(hembra);mediaHembra
```

```
## [1] 189.0667
```

```
desvHembra <-sd(hembra);desvHembra
```

```
## [1] 9.831151
```

```r
mediaMacho<-mean(macho);mediaMacho
```

```
## [1] 154.4667
```

```r
desvMacho<-sd(macho);desvMacho
```

```
## [1] 23.06462
```

En promedio, el tamaño de los langostinos hembra es de 189.0667 mm, mientras que el tamaño de los langostinos machos en promedio es de 154.4667 mm.

El tamaño de los langostinos hembra se desvía de la media aproximadamente 9.831151 mm, a diferencia del tamaño de los langostinos machos que se desvía de la media aproximadamente 23.06462 mm.

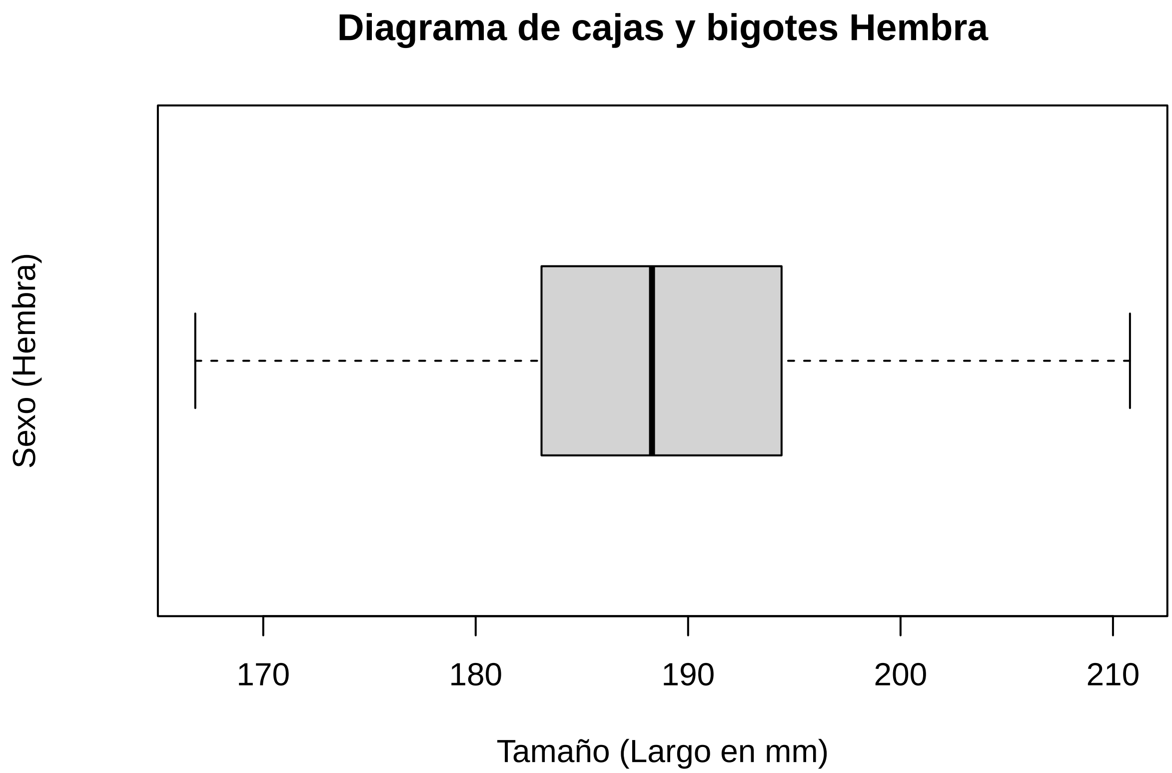
c. Halle un intervalo para la media del Largo total por sexo con un nivel de confianza del 97%. Provea la interpretación respectiva.

```r
#Datos hembras
numeroHembras <-27
EHembras <-(desvHembra/sqrt(numeroHembras))
margenE <- 2.17*EHembras
limInferiorHembras <-mediaHembra-margenE
limSuperiorHembras <-mediaHembra+margenE

#Datos machos
numeroMachos <-27
EMachos <-(desvMacho/sqrt(numeroMachos))
margenE <- 2.17*EMachos
limInferiorMachos <-mediaMacho-margenE
limSuperiorMachos <-mediaMacho+margenE
```

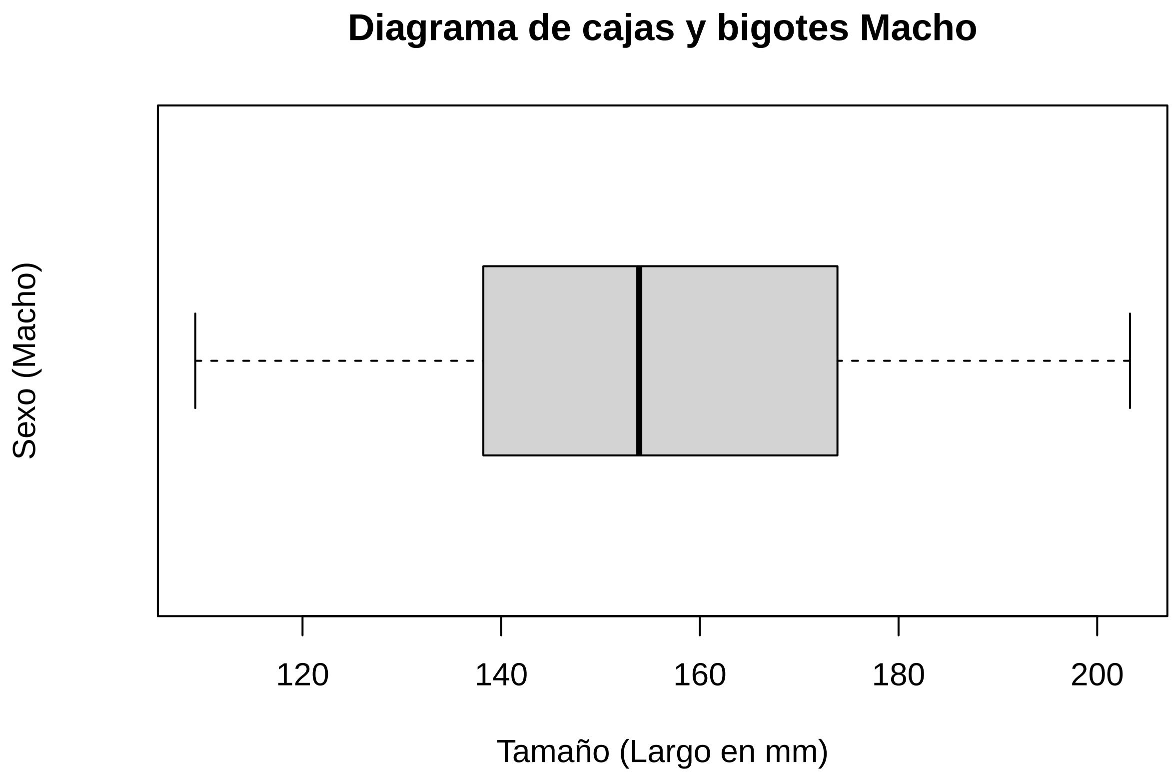
Con base a un intervalo de confianza del 0,97 de certeza se puede afirmar que el tamaño de los Langostinos hembra, fluctua entre 184.961 y 193.172, y el de los Langostinos macho varia entre 144.83 y 164.098.

d. Construya un Boxplot por sexo e interprételo.

```r
#Hembras
boxplot(hembra, horizontal = TRUE, main = "Diagrama de cajas y bigotes Hembra",
     xlab = "Tamaño (Largo en mm)", ylab = "Sexo (Hembra)")
```




```
#Machos
boxplot(macho, horizontal = TRUE, main = "Diagrama de cajas y bigotes Macho",
     xlab = "Tamaño (Largo en mm)", ylab = "Sexo (Macho)")
```



El tamaño del largo de los langostinos hembra,los datos se encuentran mas dispersos a partir del cuartil 2, que hace referencia al 50% hasta el 75%, ademas se puede observar que no hay datos atipicos. En el diagrama de los machos podemos diferenciar que los datos se encuentras mas dispersos comparando al diagrama de las hembras, tomando como referencia 50% hasta el 75%, tambien se observa la inexistencia de datos atipicos.

e. comprobar la normalidad de los datos

```
ks.test(hembra,pnorm,mediaHembra, desvHembra)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  hembra
## D = 0.12885, p-value = 0.7136
## alternative hypothesis: two-sided
```

2.En un restaurante de la ciudad se sabe que la probabilidad de que se reciba un billete de $50.000 falso es de 0.015. Si se sabe que en una semana se reciben pagos con 900 billetes de $50.000, halle la probabilidad de que:

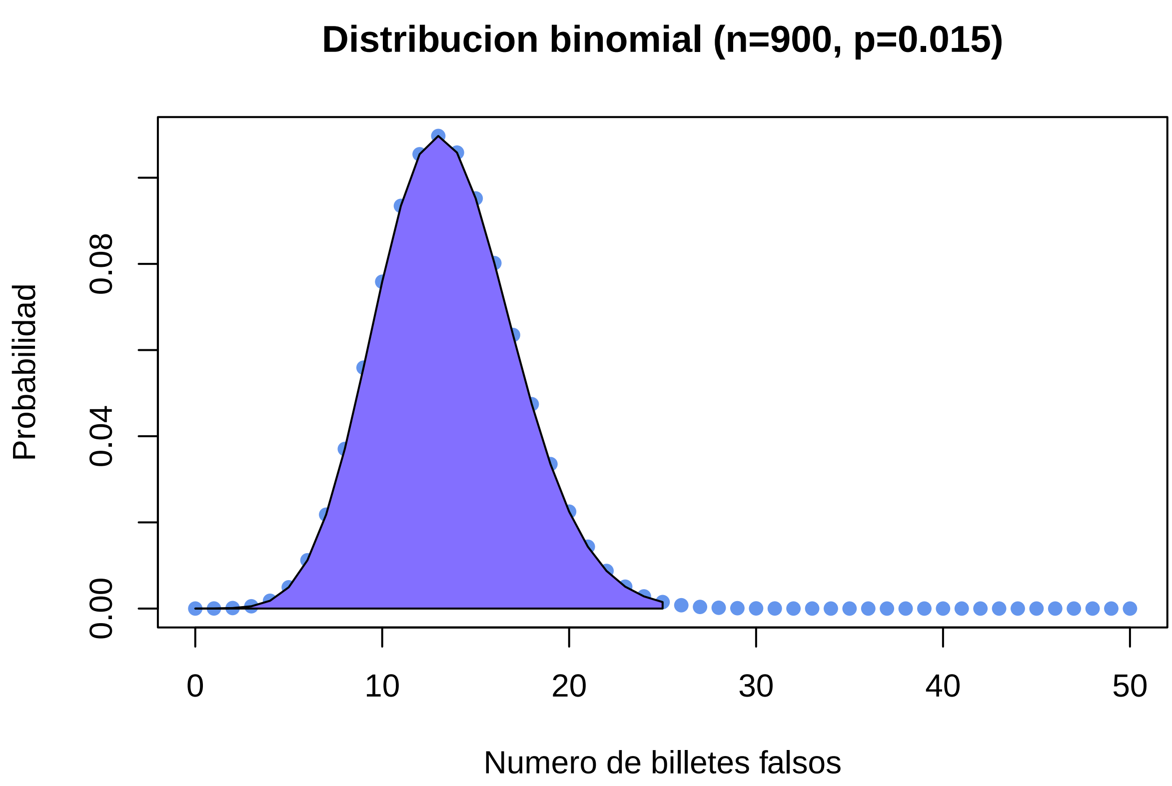
a. A lo sumo 25 billetes sean falsos.

```
## x= Numero de billetesde 50.000 falsos
n<-900
```

```r
p <- 0.015
q <- 1-p

probabilidadA<- pbinom(25,n,p);probabilidadA
```

```
## [1] 0.9985199
```

```r
longitudX<-c(0:50)
probabilidadAgrafica<-dbinom(0:50,900,0.015)

#grafica
plot(longitudX,probabilidadAgrafica,
xlab="Numero de billetes falsos",
ylab="Probabilidad",
title("Distribucion binomial (n=900, p=0.015)"),
pch=16,
col = "cornflowerblue")

#colorear el area bajo la curva
polygon(c(longitudX[longitudX <= 25 ], 25),c(probabilidadAgrafica[longitudX <= 25 ], 0),
        col = "slateblue1",
        border = 1)
```

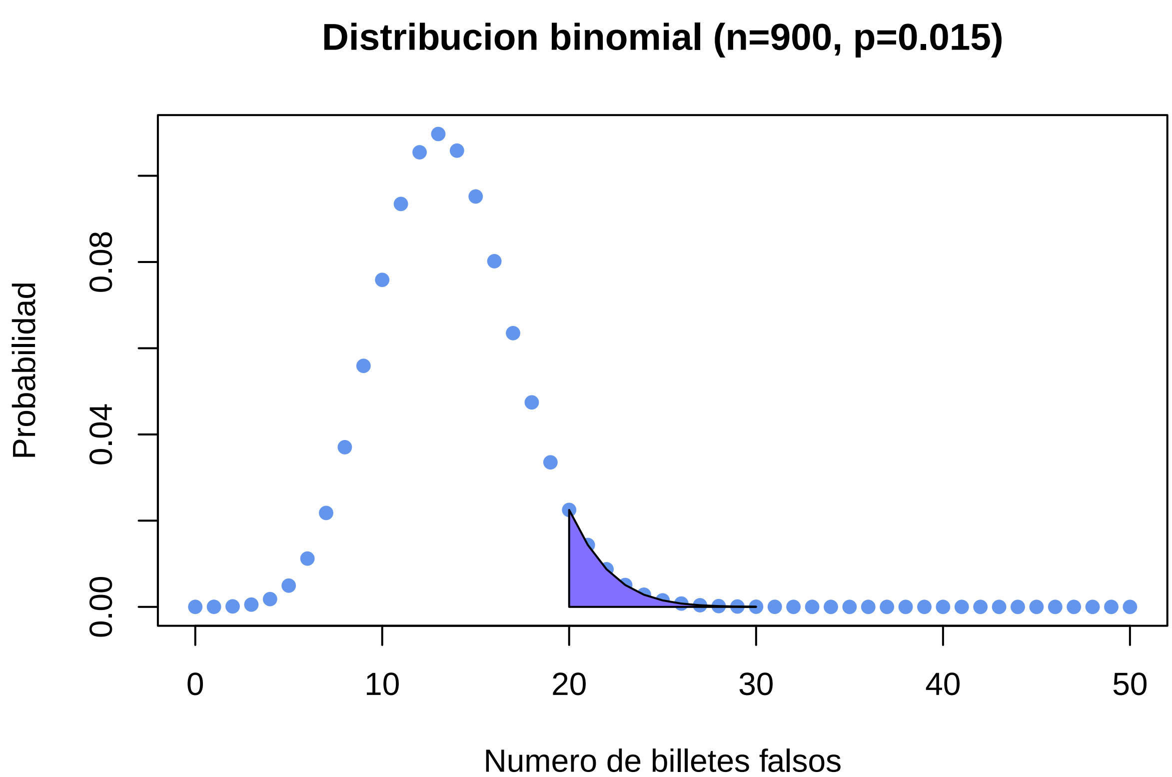


b. La cantidad de billetes falsos esté entre 20 y 30.

```
probabilidadB<- pbinom(30,n,p) - pbinom(19,n,p);probabilidadB
```

```
## [1] 0.05644795
```

```
longitudBX<-c(0:50)
probabilidadBgrafica<-dbinom(0:50,900,0.015)

#grafica
plot(longitudBX,probabilidadBgrafica,
xlab="Numero de billetes falsos",
ylab="Probabilidad",
title("Distribucion binomial (n=900, p=0.015)"),
pch=16,
col = "cornflowerblue")

#colorear el area bajo la curva

i <- longitudBX >= 20 & longitudBX <= 30
polygon(c(20,longitudBX[i],30), c(0,probabilidadBgrafica[i],0), col="slateblue1")
```

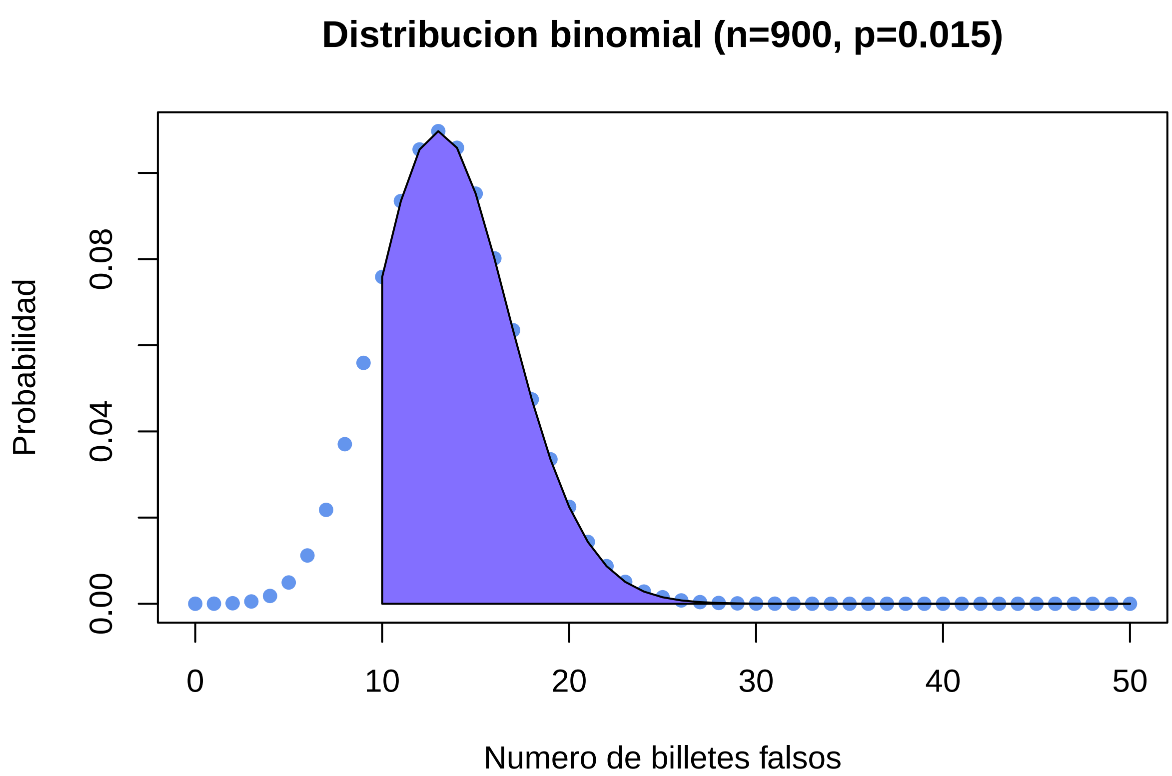


c. Más de 10 sean falsos.

```
probabilidadC<- 1-(pbinom(9,n,p));probabilidadC
```

```
## [1] 0.8666443
```

```r
##probabilidadC<- pbinom(10,n,p,lower.tail = FALSE);probabilidadC

longitudCX<-c(0:50)
probabilidadCgrafica<-dbinom(0:50,900,0.015)

#grafica
plot(longitudCX,probabilidadCgrafica,
xlab="Numero de billetes falsos",
ylab="Probabilidad",
title("Distribucion binomial (n=900, p=0.015)"),
pch=16,
col = "cornflowerblue")

#colorear el area bajo la curva
polygon(c(longitudCX[longitudCX >= 10 ], 10),c(probabilidadCgrafica[longitudCX >= 10], 0),
        col = "slateblue1",
        border = 1)
```



3. Según un estudio del Departamento Nacional de Estadística -DANE-, la vida media para el quinquenio de 2010 a 2015 de los habitantes de Colombia es 76 años, con una varianza de 25. Se pretende hacer un estudio con el objetivo de extrapolar los resultados anteriores a una pequeña ciudad de 100.000 habitantes, considerando que el tiempo de sobrevida es normal.

a. ¿Cuántos de los habitantes de la pequeña ciudad superarán previsiblemente los 92 años?

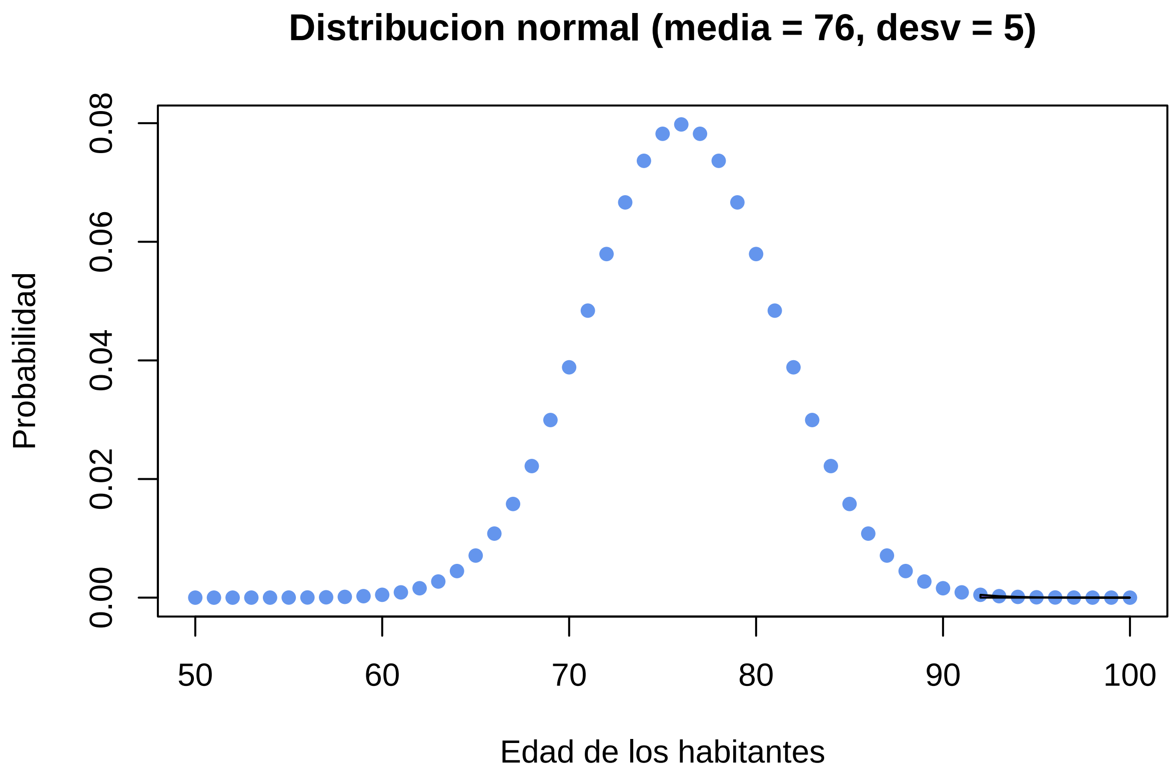
```
media<- 76
desv<- 5
n<- 100.000
longitudX<-c(50:100)
probabilidadPA<- 1-( pnorm (91, media, desv));probabilidadPA
```

```
## [1] 0.001349898
```

```
#Habitantes que superan previsiblemente los 92 años
respuesta<-(probabilidadPA*n);respuesta
```

```
## [1] 0.1349898
```

```
probabilidadDN<- dnorm(longitudX,media, desv)

#grafica
plot(longitudX,probabilidadDN,
xlab="Edad de los habitantes",
ylab="Probabilidad",
title("Distribucion normal (media = 76, desv = 5)"),
pch=16,
col = "cornflowerblue")

#colorear el area bajo la curva
polygon(c(longitudX[longitudX >= 92 ], 92),c(probabilidadDN[longitudX >= 92], 0),
        col = "slateblue1",
        border = 1)
```



Respuesta: El numero de habitantes que superan previsiblemente los 92 años es de: 0.1349898

b. ¿Cuántos vivirán menos de 55 años o más de 75 años?

```
probabilidadPB<-( pnorm (55, media, desv))+(1-(pnorm (74, media, desv)));probabilidadPB
```

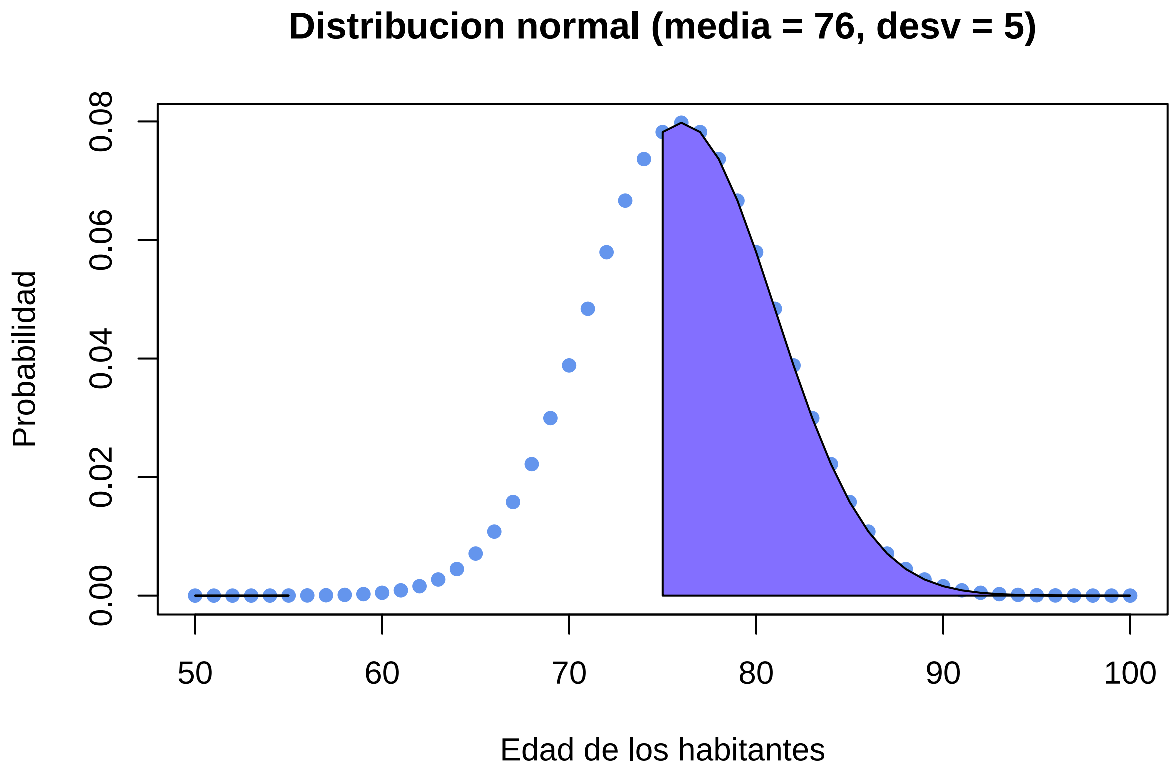
```
## [1] 0.6554351
```

```
respuesta<- probabilidadPB*n;respuesta
```

```
## [1] 65.54351
```

```
longitud<-c(50:100)
probabilidad<- dnorm(longitud,media, desv)

#grafica
plot(longitud,probabilidad,
xlab="Edad de los habitantes",
ylab="Probabilidad",
title("Distribucion normal (media = 76, desv = 5)"),
pch=16,
col = "cornflowerblue")

polygon(c(longitud[longitud >= 75 ], 75),c(probabilidad[longitud >= 75], 0),
        col = "slateblue1",
```

```
        border = 1)

polygon(c(longitud[longitud <= 55 ], 55),c(probabilidad[longitud <= 55], 0),
        col = "slateblue1",
        border = 1)
```



Respuesta: El numero de habitantes que viviran menos de 55 años o más de 75 años es de: 65.54351