# Laboratorio N.1

Introduccion a Los Metodos Estadisticos
Estimadores

Diana Carolina Arias Sinisterra Cod. 1528008
Kevin Steven Garcia Chica Cod. 1533173
Cesar Andres Saavedra Vanegas Cod. 1628466

**Universidad Del Valle**
Facultad De Ingenieria
Estadistica
Octubre
2017

# Índice



# Índice de figuras

# 1. Situación 1

## 1.1. Estimadores

Los estimadores propuestos son:

$$\hat{\theta}_1 = 2\bar{X} - 1$$

$$\hat{\theta}_2 = X_{Max} + X_{Min} - 1$$

$$\hat{\theta}_3 = X_{Max}$$

$$\hat{\theta}_4 = Me(X)$$

Para evaluar los estimadores anteriores obtuvimos las siguientes cuatro gráficas:

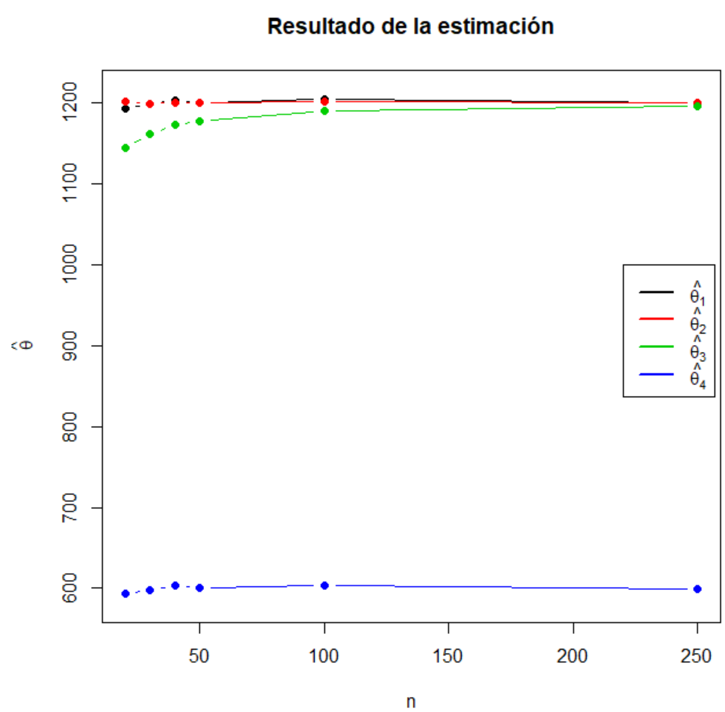


Figura 1: Gráfica del resultado de la estimación del parámetro real $\theta = 1200$

Como podemos observar en la gráfica,los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ son los que mas se acercan al parámetro real $\theta = 1200$ con cualquier tamaño de muestra, aunque el estimador $\hat{\theta}_3$ también se acerca pero después de tamaño de muestra $n = 100$, lo cual nos dice que son mejores los dos anteriores. El estimador $\hat{\theta}_4$ esta muy lejos del parámetro real, a simple vista podríamos decir que ese estimador serviría no para estimar $\theta$ en si, si no, para estimar la media de la distribución en general que es aproximadamente 600.

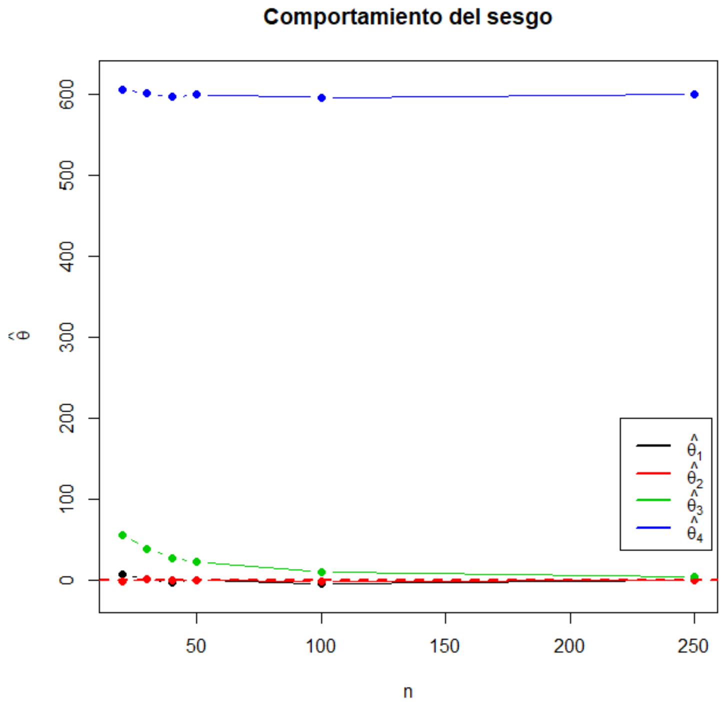


Figura 2: Gráfica del sesgo de los distintos estimadores variando los tamaños de muestra

En esta gráfica vemos que al igual que en la anterior, los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ son los mejores, ya que tienen un sesgo casi igual a 0 para todos los tamaños de muestra, aunque vemos también que el sesgo del estimador $\hat{\theta}_3$ también tiende a 0 a medida que se aumenta los tamaños de muestra. El estimador $\hat{\theta}_4$ es el menos acertado ya que su sesgo es muy elevado. En general, basándonos en esta gráfica, los mejores estimadores hasta ahora son $\hat{\theta}_1$ y $\hat{\theta}_2$

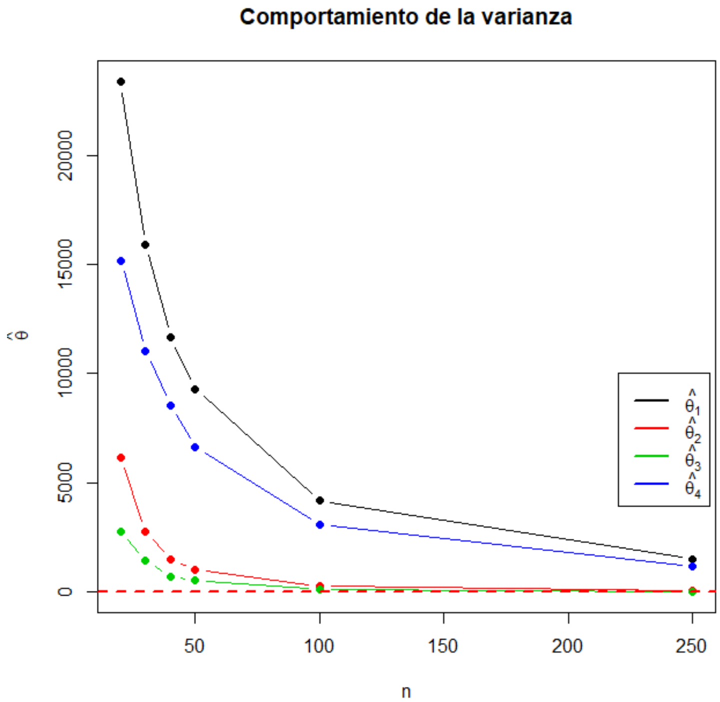


Figura 3: Gráfica de la varianza de los distintos estimadores variando los tamaños de muestra

En esta gráfica podemos observar que los estimadores $\hat{\theta}_2$ y $\hat{\theta}_3$ tienen menor varianza que los otros dos, pero notemos, que la diferencia de varianzas entre $\hat{\theta}_2$ y $\hat{\theta}_3$ es casi la mitad a favor de $\hat{\theta}_3$ en los tamaños de muestra mas pequeños, indicandonos esto, que el estimador de menor varianza es efectivamente $\hat{\theta}_3$.

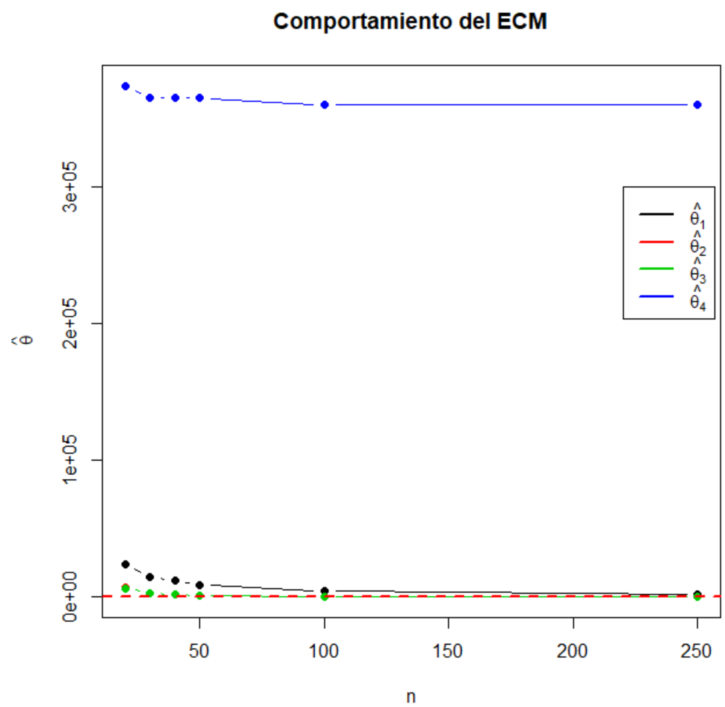


Figura 4: Gráfica del ECM de los distintos estimadores variando los tamaños de muestra

En la gráfica anterior, notamos que el menor Error cuadrático medio lo presentan los estimadores $\hat{\theta_2}$ y $\hat{\theta_3}$, con casi el mismo en todos los tamaños de muestra, por esa razón no logramos ver la linea y los puntos que representan el estimador $\hat{\theta_2}$, pero en el tamaño de muestra n=5 se alcanza a observar una pequeña porción del punto rojo que representa el estimador $\hat{\theta_2}$, obstruido por el punto verde que representa el estimador $\hat{\theta_3}$, lo que nos asegura que no hay diferencias significativas entre estos estimadores con respecto al ECM.Entonces, para decidir cual de los dos estimadores es mejor, debemos ver las 4 gráficas en conjunto, ya que el ECM no me deja claro esto, por ser tan parecido en ambos.

En conclusión, tomando en cuenta las 4 gráficas, podríamos decir que para tamaños de muestra menor a 100, $\hat{\theta_2}$ es el mejor estimador para $\theta = 1200$ ya que su sesgo es menor que el de $\hat{\theta_3}$, su estimación es mas cercana al valor real del parámetro y ademas su varianza no es muy alta.

# 2. Situación 2

## 2.1. Punto A.

**Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores insesgados tales que:**

$$\hat{\theta}_3 = a\hat{\theta}_1 + (1-a)\hat{\theta}_2$$

Aplico Esperanza a ambos lados
$E[\hat{\theta}_3] = E[a\hat{\theta}_1 + (1-a)\hat{\theta}_2]$
$E[\hat{\theta}_3] = E[a\hat{\theta}_1] + E[(1-a)\hat{\theta}_2]$
$E[\hat{\theta}_3] = aE[\hat{\theta}_1] + (1-a)E[\hat{\theta}_2]$ **Con $\hat{\theta}_1$ y $\hat{\theta}_2$ estimadores insesgados**
$E[\hat{\theta}_3] = a\theta + (1-a)\theta$
$E[\hat{\theta}_3] = a\theta + \theta - a\theta$

$$E[\hat{\theta}_3] = \theta$$

$\therefore \hat{\theta}_3$ **Es un estimador insesgado.**

## 2.2. Punto B.

**El coeficiente de variación de $\hat{\theta}_3$ es**

$$CV[\hat{\theta}_3] = \frac{\sqrt{Var[\hat{\theta}_3]}}{E[\hat{\theta}_3]}$$

Hallamos la $Var[\hat{\theta}_3]$
$Var[\hat{\theta}_3] = Var[a\hat{\theta}_1 + (1-a)\hat{\theta}_2] + 2Cov[a\hat{\theta}_1, (1-a)\hat{\theta}_2]$, ya que no sabemos si $\hat{\theta}_1$ y $\hat{\theta}_2$ son independientes
Entonces, distribuyendo la Varianza y por las propiedades de Covarianza:
$Var[\hat{\theta}_3] = Var[a\hat{\theta}_1] + Var[(1-a)\hat{\theta}_2] + 2a(1-a)Cov[\hat{\theta}_1, \hat{\theta}_2]$
$Var[\hat{\theta}_3] = a^2Var[\hat{\theta}_1] + (1-a)^2Var[\hat{\theta}_2] + (2a-2a^2)Cov[\hat{\theta}_1, \hat{\theta}_2]$

**Como $Var[\hat{\theta}_1] = \sigma_1^2$ y $Var[\hat{\theta}_2] = \sigma_2^2$ entonces**

$Var[\hat{\theta}_3] = a^2[\sigma_1^2] + (1-a)^2[\sigma_2^2] + (2a-2a^2)Cov[\hat{\theta}_1, \hat{\theta}_2]$
Sabemos que $Cov[\hat{\theta}_1, \hat{\theta}_2] = E[\hat{\theta}_1\hat{\theta}_2] - E[\hat{\theta}_1] * E[\hat{\theta}_2]$, entonces:
Como $E[\hat{\theta}_1] = \theta$ y $E[\hat{\theta}_2] = \theta$ tenemos:
$Cov[\hat{\theta}_1, \hat{\theta}_2] = E[\hat{\theta}_1\hat{\theta}_2] - \theta^2$, por tanto:
$Var[\hat{\theta}_3] = a^2[\sigma_1^2] + (1-2a+a^2)[\sigma_2^2] + (2a-2a^2)(E[\hat{\theta}_1\hat{\theta}_2] - \theta^2)$
Como sabemos que $E[\hat{\theta}_3] = \theta$

$\therefore$ **El coeficiente de variación para $\hat{\theta}_3$ es:**

$$CV[\hat{\theta}_3] = \frac{\sqrt{a^2\sigma_1^2 + (1-a)^2\sigma_2^2 + (2a-2a^2)(E[\hat{\theta}_1\hat{\theta}_2] - \theta^2)}}{\theta}$$

## 2.3. Punto C.

Como en este punto sabemos que $\hat{\theta}_1$ y $\hat{\theta}_2$ son independientes, entonces la $Cov[\hat{\theta}_1, \hat{\theta}_2] = 0$ y por tanto:
$Var[\hat{\theta}_3] = Var[a\hat{\theta}_1 + (1-a)\hat{\theta}_2]$
Distribuyendo la varianza y aplicando sus propiedades, nos queda:
$Var[\hat{\theta}_3] = Var[a\hat{\theta}_1] + Var[(1-a)\hat{\theta}_2]$
$Var[\hat{\theta}_3] = a^2Var[\hat{\theta}_1] + (1-a)^2Var[\hat{\theta}_2]$
$Var[\hat{\theta}_3] = a^2\sigma_1^2 + (1-a)^2\sigma_2^2$
Note que $Var[\hat{\theta}_3]$ se convierte en una función que depende de a, ya que, $\sigma_1^2$ y $\sigma_2^2$ son conocidas, entonces, debemos encontrar a, que haga mínima dicha función, en otras palabras, todo se reduce a encontrar el mínimo de la función. Para ello procedemos de la siguiente manera:
1. Encontramos la primera derivada de la función con respecto a la variable a:

$$f'(a) = 2a\sigma_1^2 - 2(1-a)\sigma_2^2$$

2. Igualamos el resultado de la primera derivada a 0 y despejamos la variable que nos interesa obtener, en este caso, despejamos a:

$$2a\sigma_1^2 - 2(1-a)\sigma_2^2 = 0$$

$$2a\sigma_1^2 - (2-2a)\sigma_2^2 = 0$$

$$2a\sigma_1^2 - 2\sigma_2^2 + 2a\sigma_2^2 = 0$$

$$2a\sigma_1^2 + 2a\sigma_2^2 = 2\sigma_2^2$$

$$a(2\sigma_1^2 + 2\sigma_2^2) = 2\sigma_2^2$$

$$a = \frac{2\sigma_2^2}{2\sigma_1^2 + 2\sigma_2^2}$$

$\therefore a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$
3. Ahora, debemos encontrar la segunda derivada y ver si es positiva o negativa para saber si encontramos un mínimo o un máximo:

$$f''(a) = 2\sigma_1^2 + 2\sigma_2^2 > 0$$

Como nos dio que la segunda derivada parcial es siempre positiva, concluimos que el a hallado anteriormente es un mínimo. Por lo tanto, para hacer que la $Var[\hat{\theta}_3]$ sea mínima, debemos escoger a como $a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$.

# 3. Situación 3

## 3.1. Punto A.

Sea $X$ una distribución Poisson($\lambda$) con, $n = 30$ y $E[x] = (\lambda)$

$$M_1' = \frac{1}{30}\sum_{i=1}^{30} x_i$$

$$\mu_1' = \lambda$$

$M_1' = \frac{1}{30}\sum_{i=1}^{30} x_i = \lambda = \mu_1'$
$\therefore \hat{\lambda} = \frac{1}{30}\sum_{i=1}^{30} x_i = \overline{X}$

Ahora es necesario probar si es insesgado:

$E[\hat{\lambda}] = E[\frac{1}{30}\sum_{i=1}^{30} x_i]$
$E[\hat{\lambda}] = \frac{1}{30}E[\sum_{i=1}^{30} \lambda_i]$
$E[\hat{\lambda}] = \frac{1}{30}30\lambda$
$E[\hat{\lambda}] = \lambda$

$\therefore$ **Un estimador insesgado para $\lambda$ es $\hat{\lambda} = \frac{1}{30}\sum_{i=1}^{30} x_i = \overline{X}$**

## 3.2. Punto B.

Sea $C = 3X + X^2$

$E[C] = E[3X + X^2]$
$E[C] = E[3X] + E[X^2]$
$E[C] = 3E[X] + E[X^2]$

Sabiendo que:
$V[X] = E[X^2] - E^2[X]$

Entonces:
$E^2[X] = V[X] + E[X^2]$

$\therefore E^2[X] = \lambda + (\lambda)^2$

Reemplazando $E^2[X]$ en:
$E[C] = 3\lambda + E[X^2]$

Obtenemos que: $E[C] = 3\lambda + \lambda + (\lambda)^2$
$\therefore E[C] = 4\lambda + (\lambda)^2$

## 3.3. Punto C.

Ahora debemos sugerir un estimador insesgado para $E[C] = 4\lambda + \lambda^2$
Como $\hat{\lambda} = \frac{1}{30}\sum_{i=1}^{30} x_i = \overline{X}$ es un estimador insesgado para $\lambda$ , creemos que $4\hat{\lambda} + \hat{\lambda}^2$ es insesgado para $E[C] = 4\lambda + \lambda^2$
Para ver si eso es correcto, encontraremos la esperanza del estimador propuesto, y con respecto a ese resultado, transformaremos nuestro estimador para que efectivamente nos de insesgado para $E[C] = 4\lambda + \lambda^2$, entonces:

$$E[4\hat{\lambda} + \hat{\lambda}^2] = E[4\hat{\lambda}] + E[\hat{\lambda}^2]$$
$$= 4E[\hat{\lambda}] + E[\hat{\lambda}^2]$$

reemplazando el valor del estimador en $\hat{\lambda}$, nos queda:

$$= 4E[\frac{1}{30}\sum_{i=1}^{30} x_i] + E[\frac{1}{30}\sum_{i=1}^{30} x_i]^2$$

$$= \frac{4}{30}E[\sum_{i=1}^{30} x_i] + \frac{1}{900}E[\sum_{i=1}^{30} x_i]^2$$

Denotemos $T = \sum_{i=1}^{30} x_i$ . Sabemos que xi tiene distribución Poisson($\lambda$), por tanto T tendrá distribución Poisson($30\lambda$). Entonces, sustituyendo el T en la ecuación, tendríamos:

$$= \frac{4}{30}E[T] + \frac{1}{900}E[T]^2$$

Ahora, si T se distribuye Poisson($30\lambda$), entonces $E[T] = 30\lambda$ y $V[T] = 30\lambda$. Por propiedades tenemos que $V[T] = E[T^2] - E^2[T]$, entonces, sustituyendo y resolviendo, $E[T^2] = 30\lambda - 900\lambda^2$
Sustituyendo todo en la ecuación, tenemos lo siguiente:

$$= \frac{4}{30}(30\lambda) + \frac{1}{900}(30\lambda - 900\lambda^2)$$

$$= 4\lambda + \frac{1}{30}(\lambda - 30\lambda^2)$$

$$= 4\lambda + \frac{1}{30}\lambda - \lambda^2$$

Entonces, debemos modificar nuestro estimador inicial, para que efectivamente su esperanza nos de $4\lambda + \lambda^2$
Note que para obtener esa esperanza solo debemos restar $\frac{1}{30}\lambda$ y cambiarle el signo a $-\lambda^2$ al resultado de la esperanza
Para lograr eso, propusimos el estimador $4\hat{\lambda} - \hat{\lambda}^2 + \frac{1}{30}\hat{\lambda}$
Ahora veremos si este nuevo estimador propuesto si es insesgado para $E[C] = 4\lambda + \lambda^2$

$$E[4\hat{\lambda} - \hat{\lambda}^2 + \frac{1}{30}\hat{\lambda}] = E[4\hat{\lambda}] - E[\hat{\lambda}^2] + E[\frac{1}{30}\hat{\lambda}]$$

$$= 4E[\hat{\lambda}] - E[\hat{\lambda}^2] + \frac{1}{30}E[\hat{\lambda}]$$

Haciendo la misma sustitución que antes, nos queda:

$$= \frac{4}{30}E[T] - \frac{1}{900}E[T^2] + \frac{1}{900}E[T]$$

$$= \frac{4}{30}(30\lambda) - \frac{1}{900}(30\lambda - 900\lambda^2) + \frac{1}{900}(30\lambda)$$

$$= 4\lambda - \frac{1}{30}(\lambda - 30\lambda^2) + \frac{1}{30}\lambda$$

$$= 4\lambda - \frac{1}{30}\lambda + \lambda^2 + \frac{1}{30}\lambda$$

$$= 4\lambda + \lambda^2$$

$\therefore$ **Un estimador insesgado para** $E[C] = 4\lambda + \lambda^2$ **es** $4\hat{\lambda} - \hat{\lambda}^2 + \frac{1}{30}\hat{\lambda}$ **donde:** $\hat{\lambda} = \frac{1}{30}\sum_{i=1}^{30} x_i = \overline{X}$

# 4. Situación 4

## 4.1. Punto A.

Se tiene una población que se encuentra definida por una variable aleatoria X con distribución Exponencial:

$$f(x) = 5e^{-5x}; x > 0$$

Para la cual se generaron 1000 muestras aleatorias de tamaño n, para la cual se construyeron histogramas con los valores de las medias y se les sobrepuso la curva normal de parámetros $\mu$ y $\frac{\sigma}{\sqrt{n}}$.

Esta simulación se realizo por medio del software R y para n con tamaño de muestra de 5, 10, 20, 30, 50, 100.

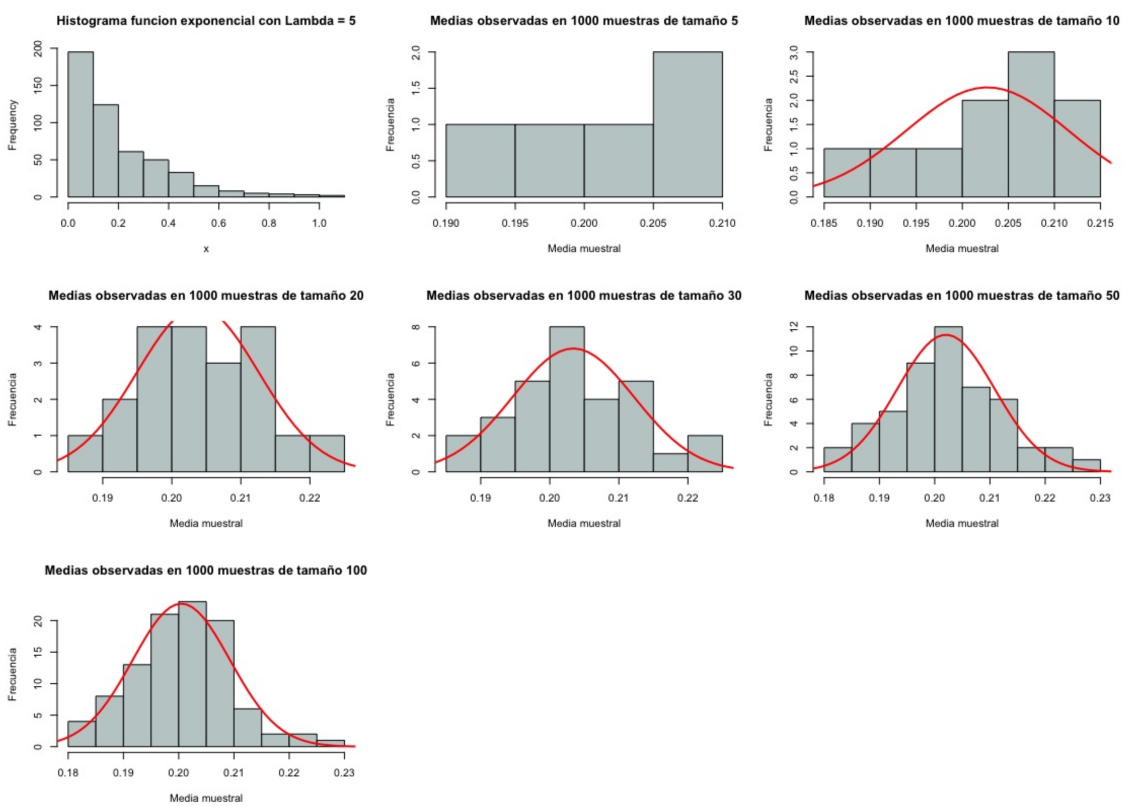


Figura 5: Histograma de frecuencias para una distribución Exponencial $\lambda = 5$

En la gráfica se puede observar claramente como el comportamiento de una función Exponencial va cambiando al aumentar el tamaño de muestra n hasta llegar a tener el comportamiento de una función normal de parámetros $\mu$ y $\frac{\sigma}{\sqrt{n}}$.

Hay que tener en cuenta que la distribución Normal es la mas utilizada ya que sus propiedades son el fundamento de los procedimientos de Inferencia ya que la mayor parte de las variables aleatorias de los fenómenos naturales presentan un comportamiento semejante a esta, y de tal forma fue demostrado con la función exponencial.
Al tomar muestras de tamaños diferentes cada una de estas arroja un resultado distinto lo que conlleva a que la función este menos dispersa, es decir, su error típico disminuye al aumentar el tamaño de muestra ya que la normalidad de las estimaciones mejora.

## 4.2. Punto B.

Para este punto se desea realizar el mismo procedimiento anterior y observar el comportamiento para una distribución t-Student con 7 grados de libertad.

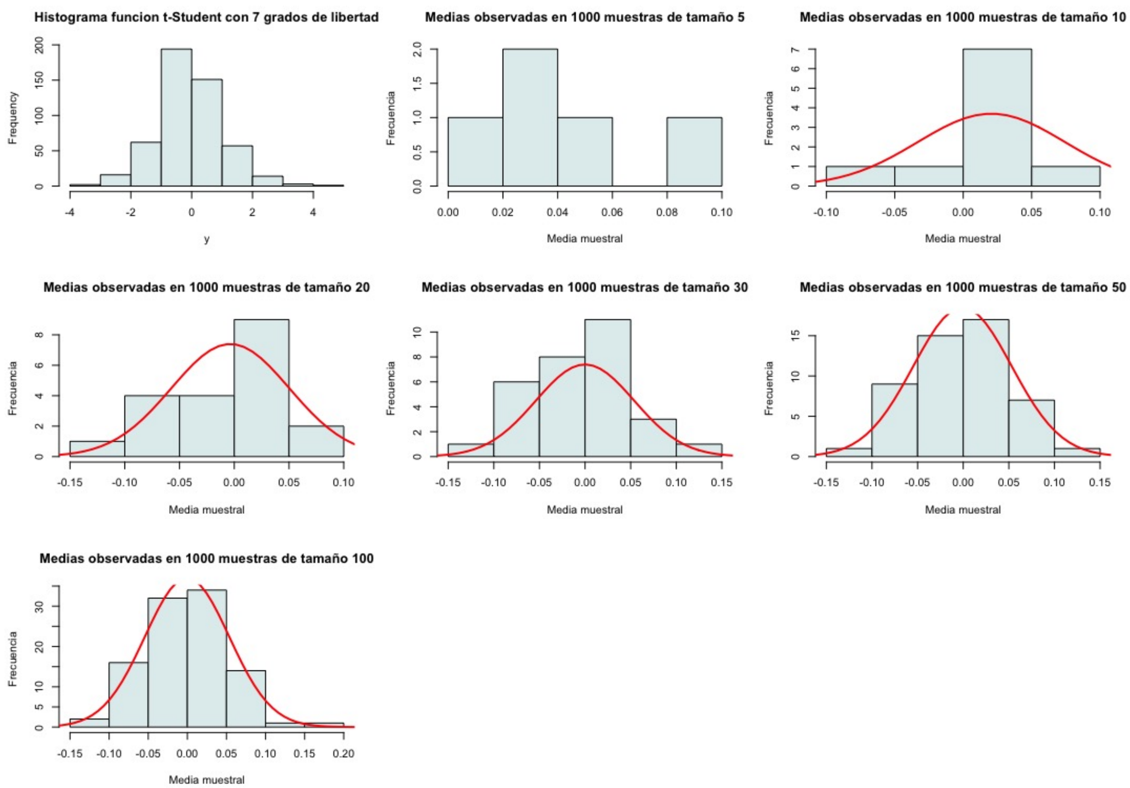


Figura 6: Histograma de frecuencias para una distribución t-Student con 7 grados de libertad

Al igual que en la distribución exponencial al aumentar el tamaño de muestra la distribución t-Student converge a un comportamiento parecido al de la normal de parámetros $\mu$ y $\frac{\sigma}{\sqrt{n}}$.

Para lo cual se comprueba el teorema del limite central que dice que entre mas grande sea el numero de la muestra, la distribución de las medias de la población original tiende a ser NORMAL, así la población original tenga una distribución diferente a la normal, ya sea exponencial, T Student, etc.

# 5. Situación 5

## 5.1. Punto A.

Tenemos la siguiente información del problema:
$\mu = 3800$ $\sigma = 400psi$ $n = 4$ y $\bar{X} = 3334$ el cual es el valor de la media sobre la cual se toma la decisión.
Para responder a la pregunta sobre el riesgo que corre el proveedor bajo este criterio de decisión, decidimos calcular la probabilidad de que $\bar{X} \leq 3334$, entonces:
Sabemos que $\bar{X} \sim N(\mu_x, \frac{\sigma}{\sqrt{n}})$, en este caso $\bar{X} \sim N(\mu_x = 3800, \frac{\sigma=400}{\sqrt{4}}) = (\mu_x = 3800, \sigma = 200)$
Entonces:
$P(\bar{X} \leq 3334) = P(Z \leq \frac{3334-3800}{200}) = P(Z \leq \frac{-466}{200}) = P(Z \leq -2.33) = 0.0099$ (Utilizando la tabla de la normal estándar)
Según lo anterior, logramos ver que la probabilidad de que la media sea menor que 3344 es tan solo de 0.99 %, entonces concluimos que el riesgo que corre el proveedor de que el cliente les devuelva un lote bajo esta regla de decisión es muy bajo, de hecho, según los resultados, de cada 100 lotes, solo le devolverá uno aproximadamente. El riesgo lo corre mas bien el comprador, ya que este, con la regla de decisión establecida no va a lograr concluir casi nunca que el proceso esta fuera de control cuando en realidad lo esta. Por ultimo, viéndolo en forma de errores tipo 1 y 2, concluimos que esta regla de decisión cuenta con una probabilidad de cometer error tipo 2 muy alta, es decir, la mayoría de veces va a concluir que el proceso esta bajo control, cuando en realidad esta fuera de control.

## 5.2. Punto B.

Si se cambia la media a $\mu = 3700$, la probabilidad que existe de que un lote sea detectado bajo estas condiciones es la siguiente:
$P(\bar{X} \leq 3334) = P(Z \leq \frac{3334-3700}{200}) = P(Z \leq \frac{-366}{200}) = P(Z \leq -1.83) = 0.0336$ (Utilizando la tabla de la normal estándar)
Con lo anterior, podemos observar que la probabilidad de que un lote sea detectado bajo esas condiciones es de 3.36 %, un poco mas alto que en el punto anterior, esto nos dice que entre mas cerca este $\mu$ del valor de la media muestral $\bar{X}$, mas alta va a ser esta probabilidad, y por consiguiente, las posibilidades de que el cliente detecte el lote van a ser mayores.

## 5.3. Punto C.

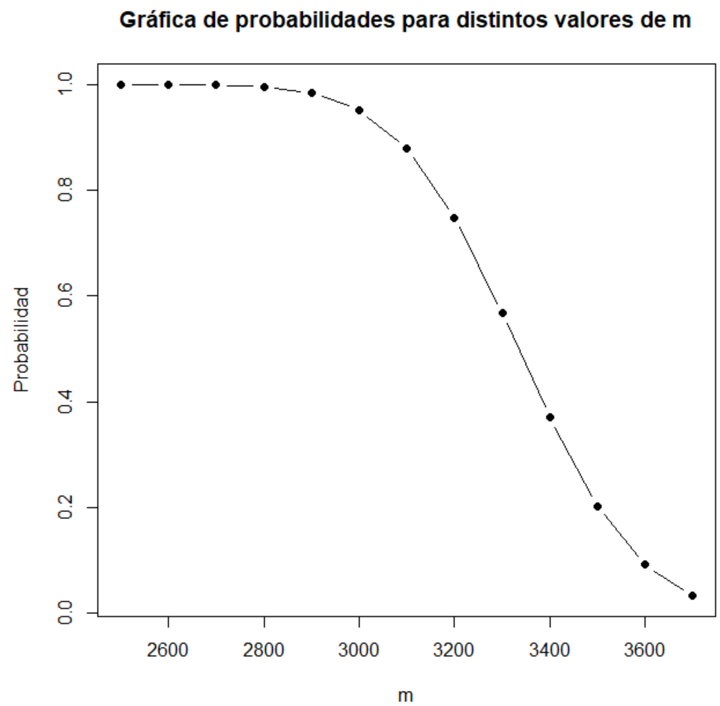


Figura 7: Gráfica de probabilidades de detectar el lote según la media $\mu$

En esta gráfica podemos ver que mientras el valor de la media $\mu$ aumenta, la probabilidad de detectar el lote va disminuyendo, esto pasa, porque la resistencia media con la cual se dice que el proceso de producción esta bajo control es de 3800 psi, entonces a medida que la media muestral aumenta, nos vamos acercando mas a la media teórica (3800 en este caso), lo que hace menos probable que se devuelva un lote de varillas metálicas, ya que mientras mas cercana sea la media muestral a la poblacional, tenemos mas certeza para concluir que las varillas se están haciendo con el proceso adecuado o bajo control.

## 5.4. Punto D.

En la gráfica que se muestra a continuación es posible apreciar la probabilidad de detectar el lote para k=5 y k=6, para lo cual se muestra la misma tendencia de la gráfica anterior y es que a medida que se incrementa la media $\mu$ la probabilidad disminuye.

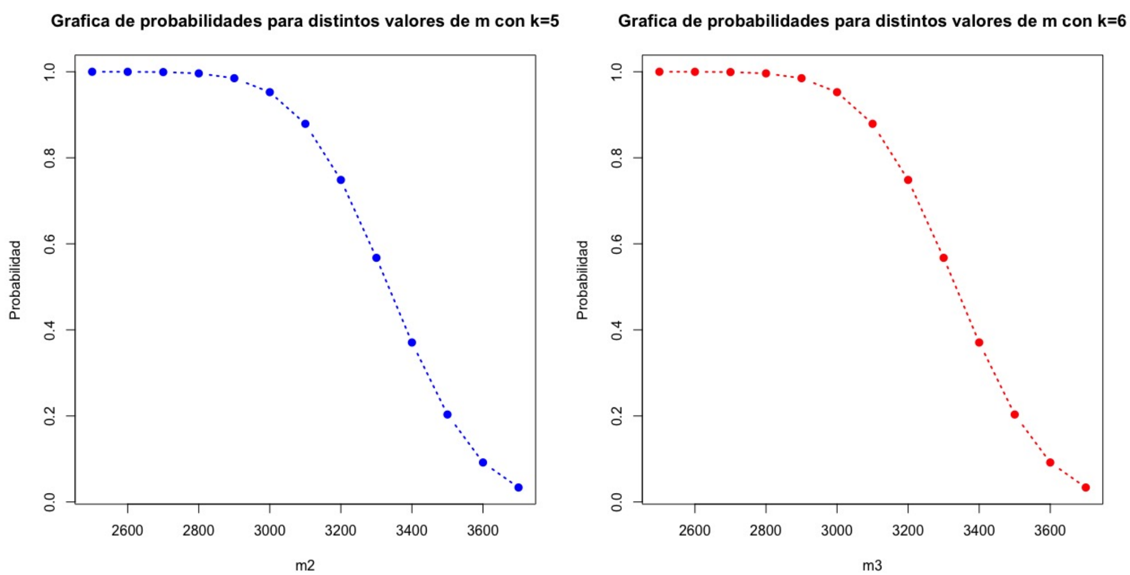


Figura 8: Gráfica de probabilidades de detectar el lote para k=5 y k=6

# 6. Scripts

```
#-------------------------------------------------------------------------------#
#Punto 1
#Estimadores planteados:
#Estimador 1:
Est1 <- function(datos) {
  Resul <- (2 * mean(datos) - 1)
  return(Resul)
}
# Estimador 2
Est2 <- function(datos) {
  Resul <- max(datos) + min(datos) - 1
  return(Resul)
}
# Estimador 3
Est3 <- function(datos) {
  Resul <- max(datos)
  return(Resul)
}
# Estimador 4
Est4 <- function(datos) {
  Resul <- median(datos)
  return(Resul)
}
n <- c(20, 30, 40, 50,100, 250)
theta <- 1200
Nsim <- 1000
NsimECM <- 1000
Pob <- seq(1, theta, 1)
Res.Es1 <- numeric(Nsim)
Res.Es2 <- numeric(Nsim)
Res.Es3 <- numeric(Nsim)
Res.Es4 <- numeric(Nsim)
ECM.Es1 <- numeric(length(n))
ECM.Es2 <- numeric(length(n))
ECM.Es3 <- numeric(length(n))
ECM.Es4 <- numeric(length(n))
Sesgo.Es1 <- numeric(length(n))
Sesgo.Es2 <- numeric(length(n))
Sesgo.Es3 <- numeric(length(n))
Sesgo.Es4 <- numeric(length(n))
theta_est <- matrix(NA, nrow = 6, ncol = 4)
var_theta_est <- matrix(NA, nrow = 6, ncol = 4)
ECM_theta_est <- matrix(NA, nrow = 6, ncol = 4)
Sesgo_theta_est <- matrix(NA, nrow = 6, ncol = 4)
```

```
# Estimacion del Sesgo, Varianza y Error Cuadratico Medio
for (j in 1:length(n)) {

  muestra <- sample(Pob, n[j])

  for (i in 1:Nsim) {
    muestra <- sample(Pob, n[j])
    Res.Es1[i] <- Est1(muestra)
    Res.Es2[i] <- Est2(muestra)
    Res.Es3[i] <- Est3(muestra)
    Res.Es4[i] <- Est4(muestra)
  }

  ECM.Es1[j] <- sum((theta - Res.Es1)^2)/Nsim
  ECM.Es2[j] <- sum((theta - Res.Es2)^2)/Nsim
  ECM.Es3[j] <- sum((theta - Res.Es3)^2)/Nsim
  ECM.Es4[j] <- sum((theta - Res.Es4)^2)/Nsim

  Sesgo.Es1[j] <- (theta - mean(Res.Es1))
  Sesgo.Es2[j] <- (theta - mean(Res.Es2))
  Sesgo.Es3[j] <- (theta - mean(Res.Es3))
  Sesgo.Es4[j] <- (theta - mean(Res.Es4))

  theta_est[j, ] <- c(mean(Res.Es1), mean(Res.Es2),mean(Res.Es3),mean(Res.Es4))
  var_theta_est[j, ] <- c(var(Res.Es1), var(Res.Es2),var(Res.Es3),var(Res.Es4))
  ECM_theta_est[j, ] <- c(ECM.Es1[j], ECM.Es2[j], ECM.Es3[j], ECM.Es4[j])
  Sesgo_theta_est[j, ] <- c(Sesgo.Es1[j], Sesgo.Es2[j],Sesgo.Es3[j],Sesgo.Es4[j])
}

theta_est <- as.data.frame(theta_est)
var_theta_est <- as.data.frame(var_theta_est)
ECM_theta_est <- as.data.frame(ECM_theta_est)
Sesgo_theta_est <- as.data.frame(Sesgo_theta_est)

#Grafica del comportamiento del sesgo:
x11()
plot(n, Sesgo_theta_est[, 1], type = "b", ylim = c(min(Sesgo_theta_est) - 10, max(Sesgo_the
                                                10), pch = 19, ylab = expression(hat(theta))
lines(n, Sesgo_theta_est[,2], type = "b", col = 2, pch = 19)
lines(n, Sesgo_theta_est[,3], type = "b", col = 3, pch = 19)
lines(n, Sesgo_theta_est[,4], type = "b", col = 4, pch = 19)
legend(x=220,y=200, c(expression(hat(theta[1])), expression(hat(theta[2])),expression(hat(t
       col = c(1, 2, 3,4), lwd = c(2, 2))
abline(h = 0, lty = 2, lwd = 2, col = 2)
```

```
#Grafica de la varianza de los estimadores
x11()
plot(n, var_theta_est[, 1], type = "b", ylim = c(min(var_theta_est) - 10, max(var_theta_est
                                                     10), pch = 19, ylab = expression(hat(t
lines(n, var_theta_est[,2], type = "b", col = 2, pch = 19)
lines(n, var_theta_est[,3], type = "b", col = 3, pch = 19)
lines(n, var_theta_est[,4], type = "b", col = 4, pch = 19)
legend(x=220,y=10000, c(expression(hat(theta[1])), expression(hat(theta[2])),expression(hat
       col = c(1, 2, 3,4), lwd = c(2, 2))
abline(h = 0, lty = 2, lwd = 2, col = 2)

#Gr?fica del resultado de la estimaci?n
x11()
plot(n, theta_est[, 1], type = "b", ylim = c(min(theta_est) - 10, max(theta_est) +
                                                   10), pch = 19, ylab = expression(hat(the
lines(n, theta_est[,2], type = "b", col = 2, pch = 19)
lines(n, theta_est[,3], type = "b", col = 3, pch = 19)
lines(n, theta_est[,4], type = "b", col = 4, pch = 19)
legend(x=220,y=1000, c(expression(hat(theta[1])), expression(hat(theta[2])),expression(hat(
       col = c(1, 2, 3,4), lwd = c(2, 2))
abline(h = 0, lty = 2, lwd = 2, col = 2)

#Comportamiento del Error Cuadratico medio
x11()
plot(n, ECM_theta_est[, 1], type = "b", ylim = c(min(ECM_theta_est) - 10, max(ECM_theta_est
                                               10), pch = 19, ylab = expression(hat(theta))
lines(n, ECM_theta_est[,2], type = "b", col = 2, pch = 19)
lines(n, ECM_theta_est[,3], type = "b", col = 3, pch = 19)
lines(n, ECM_theta_est[,4], type = "b", col = 4, pch = 19)
legend(x=220,y=300000, c(expression(hat(theta[1])), expression(hat(theta[2])),expression(ha
       col = c(1, 2, 3,4), lwd = c(2, 2))
abline(h = 0, lty = 2, lwd = 2, col = 2)
#-------------------------------------------------------------------------------#
Punto 4.a.
X~Exponencial(Lambda=5)
n=500
x <- rexp(n,5)
media1 <- mean(x)
sd1 <- sd(x)

mediaMuestral=function(n){
  muestra=rexp(n,5)
  media=mean(muestra)
  return(media)
}
```

```
m=1000
muchasMedias=replicate(m,mediaMuestral(n))
mean(muchasMedias)
sd(muchasMedias)

para5 <- muchasMedias[1:5]
para10 <- muchasMedias[1:10]
para20 <- muchasMedias[1:20]
para30 <- muchasMedias[1:30]
para50 <- muchasMedias[1:50]
para100 <- muchasMedias[1:100]

par(mfrow=c(3,3))
hist(x, main="Histograma funcion exponencial con Lambda = 5", col="azure3")
hist(para5, main="Medias observadas en 1000 muestras de tamano 5", xlab="Media muestral", y
hist(para10, main="Medias observadas en 1000 muestras de tamano 10", xlab="Media muestral",
xfit <- seq(min(x), max(x), length = 1000)
yfit <- dnorm(xfit, mean(para10), sd = (media1)/(sqrt(500)))
yfit <- yfit * diff(h$mids[1:2]) * length(para10)
lines (xfit, yfit, col = "red", lwd = 2)
hist(para20, main="Medias observadas en 1000 muestras de tamano 20", xlab="Media muestral",
xfit <- seq(min(x), max(x), length = 1000)
yfit <- dnorm(xfit, mean(para20), sd = (media1)/(sqrt(500)))
yfit <- yfit * diff(h$mids[1:2]) * length(para20)
lines (xfit, yfit, col = "red", lwd = 2)
hist(para30, main="Medias observadas en 1000 muestras de tamano 30", xlab="Media muestral",
xfit <- seq(min(x), max(x), length = 1000)
yfit <- dnorm(xfit, mean(para30), sd = (media1)/(sqrt(500)))
yfit <- yfit * diff(h$mids[1:2]) * length(para30)
lines (xfit, yfit, col = "red", lwd = 2)
hist(para50, main="Medias observadas en 1000 muestras de tamano 50", xlab="Media muestral",
xfit <- seq(min(x), max(x), length = 1000)
yfit <- dnorm(xfit, mean(para50), sd = (media1)/(sqrt(500)))
yfit <- yfit * diff(h$mids[1:2]) * length(para50)
lines (xfit, yfit, col = "red", lwd = 2)
hist(para100, main="Medias observadas en 1000 muestras de tamano 100", xlab="Media muestral
xfit <- seq(min(x), max(x), length = 1000)
yfit <- dnorm(xfit, mean(para100), sd = (media1)/(sqrt(500)))
yfit <- yfit * diff(h$mids[1:2]) * length(para100)
lines (xfit, yfit, col = "red", lwd = 2)
#----------------------------------------------------------------------------#
Punto 4.b.
t-Student, grados de libertad = 7
n2=500
y <- rt(n2,7)
media2 <- mean(y)
```

```
sd2 <- sd(y)
hist(y)

mediaMuestral2=function(n2){
  muestra=rt(n2,7)
  media=mean(muestra)
  return(media)
}

m2=1000
muchasMedias2=replicate(m2,mediaMuestral2(n2))
mean(muchasMedias2)
sd(muchasMedias2)

tpara5 <- muchasMedias2[1:5]
tpara10 <- muchasMedias2[1:10]
tpara20 <- muchasMedias2[1:20]
tpara30 <- muchasMedias2[1:30]
tpara50 <- muchasMedias2[1:50]
tpara100 <- muchasMedias2[1:100]

t <- hist(muchasMedias2,xlab="Media muestral", ylab="Frecuencia", col="slategray1",freq=T,
          main="Histograma de las medias muestrales observadas en 1000 muestras de tamano 5
xfit <- seq(min(y), max(y), length = 1000)
yfit <- dnorm(xfit, mean=mean(muchasMedias2), sd(muchasMedias2))
yfit <- yfit * diff(t$mids[1:2]) * length(muchasMedias2)
lines (xfit, yfit, col = "red", lwd = 2)

par(mfrow=c(3,3))
hist(y, main="Histograma funcion t-Student con 7 grados de libertad",col="azure2")
hist(tpara5, main="Medias observadas en 1000 muestras de tamano 5",xlab="Media muestral", y
hist(tpara10, main="Medias observadas en 1000 muestras de tamano 10",xlab="Media muestral",
xfit <- seq(min(y), max(y), length = 1000)
yfit <- dnorm(xfit, mean(tpara10), sd(muchasMedias2))
yfit <- yfit * diff(t$mids[1:2]) * length(tpara10)
lines (xfit, yfit, col = "red", lwd = 2)
hist(tpara20, main="Medias observadas en 1000 muestras de tamano 20",xlab="Media muestral",
xfit <- seq(min(y), max(y), length = 1000)
yfit <- dnorm(xfit, mean(tpara20), sd(muchasMedias2))
yfit <- yfit * diff(t$mids[1:2]) * length(tpara20)
lines (xfit, yfit, col = "red", lwd = 2)
hist(tpara30, main="Medias observadas en 1000 muestras de tamano 30",xlab="Media muestral"
xfit <- seq(min(y), max(y), length = 1000)
yfit <- dnorm(xfit, mean(tpara30), sd(muchasMedias2))
yfit <- yfit * diff(h$mids[1:2]) * length(tpara30)
lines (xfit, yfit, col = "red", lwd = 2)
```

```
hist(tpara50, main="Medias observadas en 1000 muestras de tamano 50",xlab="Media muestral",
xfit <- seq(min(y), max(y), length = 1000)
yfit <- dnorm(xfit, mean(tpara50), sd(muchasMedias2))
yfit <- yfit * diff(t$mids[1:2]) * length(tpara50)
lines (xfit, yfit, col = "red", lwd = 2)
hist(tpara100, main="Medias observadas en 1000 muestras de tamano 100",xlab="Media muestral
xfit <- seq(min(y), max(y), length = 1000)
yfit <- dnorm(xfit, mean(tpara100), sd(muchasMedias2))
yfit <- yfit * diff(t$mids[1:2]) * length(tpara100)
lines (xfit, yfit, col = "red", lwd = 2)
#------------------------------------------------------------------------------#
Punto 5.C.
m <-c(3700, 3600, 3500, 3400, 3300, 3200, 3100, 3000, 2900,2800,2700,2600,2500)
xbarra=3334  #Media de decision
sigma=400  #Desviacion de x
k=4   #tamano de muestra
Proba=numeric(length(m))
for (i in 1:length(m)){
  prob=pnorm((xbarra-m[i])/(sigma/sqrt(k)))
  Proba[i]=prob
}

x11()
plot(m, Proba, type = "b",pch=19, ylab="Probabilidad", main="Grafica de probabilidades para
#------------------------------------------------------------------------------#
Punto 5.D.
m2 <-c(3700, 3600, 3500, 3400, 3300, 3200, 3100, 3000, 2900,2800,2700,2600,2500)
xbarra=3334  #Media de decision
sigma=400  #Desviacion de x
k2=5   #tamano de muestra
Proba2=numeric(length(m2))
for (i in 1:length(m2)){
  prob2=pnorm((xbarra-m2[i])/(sigma/sqrt(k2)))
  Proba2[i]=prob2
}

m3 <-c(3700, 3600, 3500, 3400, 3300, 3200, 3100, 3000, 2900,2800,2700,2600,2500)
xbarra=3334  #Media de decision
sigma=400  #Desviacion de x
k3=6   #tamano de muestra
Proba3=numeric(length(m3))
for (i in 1:length(m3)){
  prob3=pnorm((xbarra-m3[i])/(sigma/sqrt(k3)))
  Proba3[i]=prob3
}
```

```
par(mfrow=c(1,2))
plot(m2, Proba, type = "b",pch=19, ylab="Probabilidad", main="Grafica de probabilidades par
plot(m3, Proba, type = "b",pch=19, ylab="Probabilidad", main="Grafica de probabilidades par
```