Taller Preparatorio Parcial #1

Hernan Supelano   Daniel Felipe Cendales

Estimación histograma

Antes, cargamos las librerías necesarias

```
library(sm)
```

```
## Package 'sm', version 2.2-5.7: type help(sm) for summary information
```

1. Realizamos una simulación de tamaño 100 de $X \sim \mathcal{N}(\mu = 17, \sigma^2 = 4)$

```
## Ajustamos los parámetros iniciales
n <- 100                                         # Tamaño de la muestra
mu <- 20                                         # Media teórica
sigma <- 2                                       # Desv. estándar teórica
set.seed(n)                                      # Fijamos la semilla

# Tomamos la muestra
datos <- rnorm(n = n, mean = mu, sd = sigma)     # X ~ N(20, 4)

## Estimación histograma
histograma <- hist(datos, plot = FALSE)
```

a. Notemos que:

```
# Extraemos la estimación histograma de la densidad
densidad1 <- histograma$density

# Extraemos los conteos y los extremos de un intervalo
ni <- histograma$counts                          # Conteos en cada intervalo
cortes <- histograma$breaks[1:2]                 # Los dos primeros puntos
b <- diff(cortes)/2                              # Distancia sobre 2
densidad2 <- ni / (2*n*b)                        # Calculamos la densidad

# Comparamos los vectores
all.equal(densidad1, densidad2)                  # Hacemos la comparación
```

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

b. Sea x el punto medio de un intervalo. Vamos a plantar dos casos:

- **Caso 1:** la distancia del punto medio de un intervalo a uno de los extremos es b, lo que implica que la longitud de los intervalos es de $2b$. Entonces

$$\begin{aligned} n_j &= \#\{x_i : x - b \leq x_i \leq x + b\} \\ &= \#\{x_i : -b \leq x_i - x \leq b\} \\ &= \#\left\{x_i : -1 \leq \frac{x_i - x}{b} \leq 1\right\} \\ &= \sum_{i=1}^{n} I_{[-1,1]}\left(\frac{x_i - x}{b}\right) \end{aligned}$$

Y por ende la estimación histograma toma la forma:

$$\begin{aligned} \widehat{f_H}(x) &= \frac{n_j}{2bn} \\ &= \frac{1}{2bn}\sum_{i=1}^{n} I_{[-1,1]}\left(\frac{x_i - x}{b}\right) \\ &= \frac{1}{nb}\sum_{i=1}^{n} \frac{1}{2} I_{[-1,1]}\left(\frac{x_i - x}{b}\right) \end{aligned}$$

- **Caso 2:** la longitud del intervalo es b. Es decir que la distancia del punto medio a uno de los extremos es $b/2$. Entonces

$$\begin{aligned} n_j &= \#\left\{x_i : x - \frac{b}{2} \leq x_i \leq x + \frac{b}{2}\right\} \\ &= \#\left\{x_i : -\frac{b}{2} \leq x_i - x \leq \frac{b}{2}\right\} \\ &= \#\left\{x_i : -\frac{1}{2} \leq \frac{x_i - x}{b} \leq \frac{1}{2}\right\} \\ &= \sum_{i=1}^{n} I_{\left[-\frac{1}{2},\frac{1}{2}\right]}\left(\frac{x_i - x}{b}\right) \end{aligned}$$

Y por ende la estimación histograma toma la forma:

$$\begin{aligned} \widehat{f_H}(x) &= \frac{n_j}{bn} \\ &= \frac{1}{bn}\sum_{i=1}^{n} I_{\left[-\frac{1}{2},\frac{1}{2}\right]}\left(\frac{x_i - x}{b}\right) \end{aligned}$$

c. Gráfico de curvas teóricas y sus estimaciones.

Primero debemos calcular el b óptimo. Para ello, recordemos la fórmula (asumimos normalidad sobre f)

$$b = 3.491 \cdot \hat{\sigma} \cdot n^{-1/3}$$

donde $\hat{\sigma} = \frac{\text{RIC}}{1.35}$

Pero el b óptimo necesita de σ y no de $\hat{\sigma}$. Entonces lo primero que hacemos será realizar varias simulaciones para encontrar el valor óptimo.

Calculemos entonces las cantidades necesarias:

```r
# Inicio de las simulaciones
N <- 500                                          # Cantidad de simulaciones
muestras <- matrix(0, nrow = n, ncol = N)        # Guardamos 500 muestras de tamaño 100
b0s <- NULL                                       # Almacenamos los b's

for(i in 1:N)
{
    muestras[, i] <- rnorm(n, mean = mu, sd = sigma)    # Toma de la muestra
    cuartiles <- quantile(muestras[, i],                 # Primer y tercer cuartil
                          probs = c(1, 3)/4, names = FALSE)
    RICs <- diff(cuartiles)                              # Rango intercuartílico
    sigmas_hat <- min(RICs / 1.35, sd(muestras[, i]))
    b0s[i] <- 3.491 * sigmas_hat * n^(-1/3)             # b'óptimo de cada muestra
}

# Cálculo del b óptimo como promedio de todos los b's
( b_Optimo <- mean(b0s) )
```

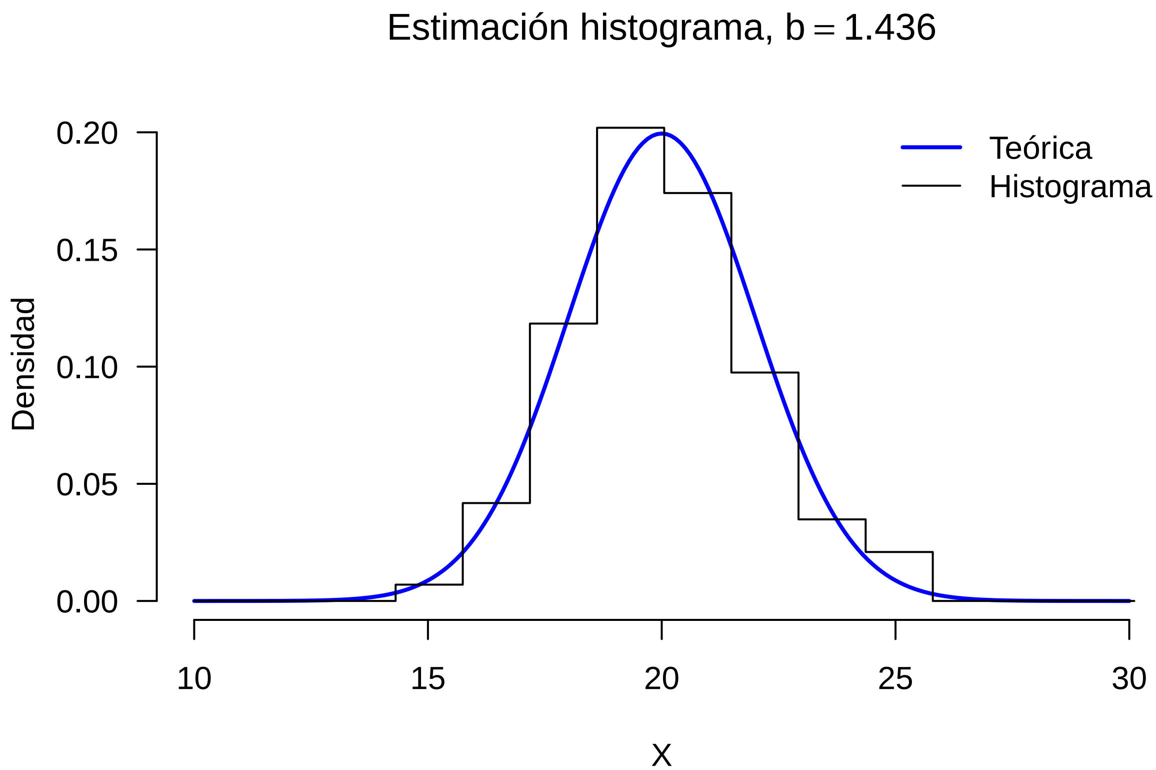
```
## [1] 1.436161
```

```r
# b óptimo obtenido de la muestra
RIC_datos <- diff(quantile(datos, probs = c(1, 3)/4, names = FALSE))
sigma_hat <- min(RIC_datos / 1.35, sd(datos))
3.491 * sigma_hat * n^(-1/3)
```

```
## [1] 1.409223
```

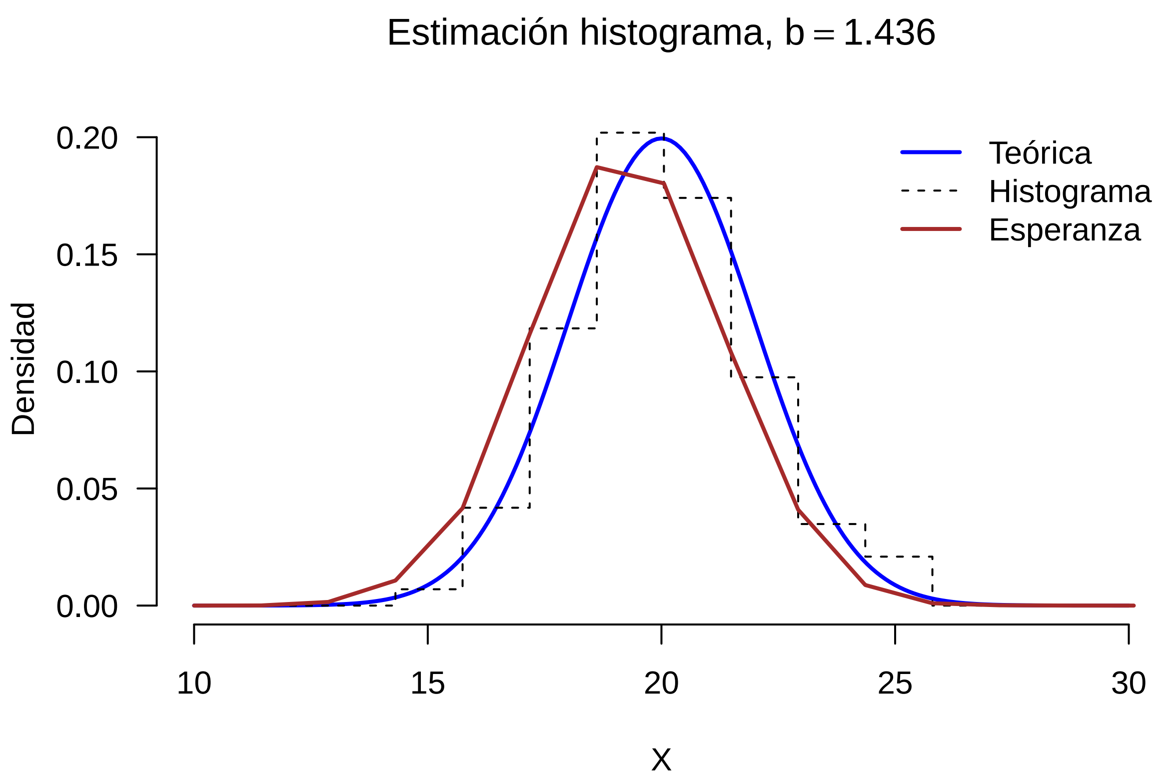
Ya calculado el b óptimo, podemos hacer la gráfica de la densidad teórica y la estimación que obtuvimos

```r
# A 5 desviaciones estándar (de la media), en una normal, estará el 99%
lImites <- mu + c(-1, 1) * 5 * sigma

# Extremos de los intervalos
puntos <- seq(from = lImites[1], to = lImites[2] + 1, by = b_Optimo)

# Histograma con el ancho de banda óptimo
Hist <- hist(datos, breaks = puntos, plot = FALSE)

# Curva teórica
rango <- seq(from = lImites[1], to = lImites[2], by = 0.1)
plot(x = rango, y = dnorm(rango, mu, sigma), frame = FALSE,
     xlab = "X", ylab = "Densidad", type = "l", las = 1,
     lwd = 2, ylim = c(0, max(Hist$density)), col = "blue",
     main = bquote("Estimación histograma, b" == .(round(b_Optimo, 3))))

# Gráfico del histograma
lines(x = puntos, y = c(Hist$density, 0), type = "s")

# Leyenda
legend(x = "topright", lwd = 2:1, col = c("blue", "black"),
       legend = c("Teórica", "Histograma"), bty = "n")
```



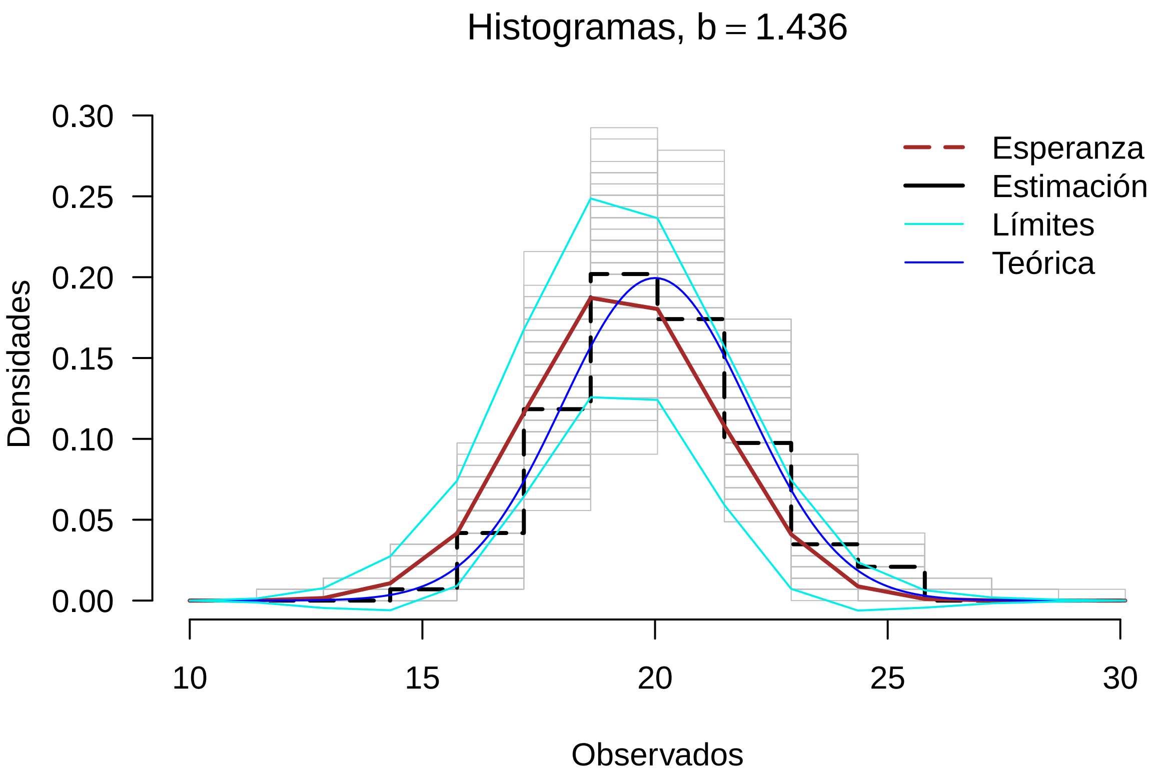
Para agregar la curva de $E\left\{\widehat{f_H}(x)\right\}$ vamos a usar la muestras generadas, calculamos el promedio y las varianzas

```
# Obtenemos los histogramas de cada muestra
histogramas <- apply(muestras, 2, function(k){
                         c(hist(k, breaks = puntos, plot = FALSE)$density, 0)
                     })
E_fh <- apply(histogramas, 1, mean)              # Valor esperado
Sd_fh <- apply(histogramas, 1, sd)               # Desviación estándar

plot(x = rango, y = dnorm(rango, mu, sigma), frame = FALSE,
     xlab = "X", ylab = "Densidad", type = "l", las = 1,
     lwd = 2, ylim = c(0, max(Hist$density)), col = "blue",
     main = bquote("Estimación histograma, b" == .(round(b_Optimo, 3))))

# Gráfico del histograma
lines(x = puntos, y = c(Hist$density, 0), type = "s", lty = 2)

# Esperanza
lines(x = puntos, y = E_fh, col = "brown", lwd = 2, type = "l")

# Leyenda
legend(x = "topright", lwd = c(2, 1, 2), col = c("blue", "black", "brown"),
       legend = c("Teórica", "Histograma", "Esperanza"), bty = "n",
       lty = c(1, 2, 1))
```



Para intentar ver el comportamiento de la varianza, podemos graficar todos los histogramas obtenidos y después agregamos unas "bandas de confianza"

```
# Todos los histogramas obtenidos
matplot(x = puntos, y = histogramas, type = "s", las = 1,
        col = "gray", lty = 1, frame = FALSE, lwd = 0.5,
        xlab = "Observados", ylab = "Densidades",
        main = bquote("Histogramas, b" == .(round(b_Optimo, 3))))

# El histograma de la muestra inicial
lines(x = puntos, y = c(Hist$density, 0), col = "black",
      lwd = 2, type = "s", lty = 5)

# Valor esperado
lines(x = puntos, y = E_fh, col = "brown", lwd = 2, type = "l")

# Estimación teórica
lines(rango, dnorm(rango, mean = mu, sd = sigma), col = "blue")

# "Bandas de confianza"
matlines(x = puntos, y = cbind(E_fh - 1.96 * Sd_fh,
                               E_fh + 1.96 * Sd_fh),
         type = "l", col = "cyan2", lty = 1, lwd = 1)

# Leyenda
legend(x = "topright", bty = "n", col = c("brown", "black", "cyan2", "blue"),
       lwd = rep(2:1, each = 2), lty = c(5, 1, 1, 1),
       legend = c("Esperanza", "Estimación", "Límites", "Teórica"))
```



d. Gráfico de la curva $V\left\{\widehat{f_H}(x)\right\}$

El gráfico anterior sugería que en los extremos la varianza disminuye y en la parte central es donde más alta es esta. Veamos gráficamente que esto es así:

```
# Gráfico de la varianza
plot(x = puntos, y = Sd_fh^2, main = "Gráfico de la varianza", frame = FALSE,
     xlab = "Observados", ylab = "Valores", type = "l", col = "orange")
```

Estimación Kernel de la densidad

a. Tomando los datos del inicio, realizamos el cálculo del h de Silverman

```
h_Silverman <- 0.9 * sigma_hat * n^(-1/5)
```

El valor de este ancho de banda es aproximadamente 0.671

b. Si asumimos un Kernel Gaussiano, la fórmula del *MISE* viene dada por:

$$MISE(\hat{f}_k) = \frac{R(K)}{nh} + \frac{\sigma_k^4 h^4}{4} R(f^{''}) + O\left(\frac{1}{n}\right) + O\left(h^5\right)$$

Como estamos usando un kernel gaussiano con $\sigma^2 = 1$, entonces $R(K) = \frac{1}{2\sqrt{\pi}\sigma_k}$ y $R(f^{''}) = \frac{3}{8\sigma^5\sqrt{\pi}}$ ya que muestreamos de una distribución normal.

$$MISE(\hat{f}_k) \quad = \quad \frac{1}{2nh\sigma_k\sqrt{\pi}} + \frac{3\sigma_k^4 h^4}{32\sigma^5\sqrt{\pi}} + O\left(\frac{1}{n}\right) + O\left(h^5\right)$$

Haciendo uso del *h de Silverman* obtenemos un valor para el *AMISE* de

```
( MISE <- 1/(2*n*h_Silverman*sqrt(pi)) + 3*h_Silverman^4/(32*sd(datos)^5*sqrt(pi)) )
```

```
## [1] 0.004505032
```

c. Primero necesitamos estimar la densidad para los datos simulados

```
# Estimamos la densidad de los datos originales
densidad_hSilver <- density(datos, bw = h_Silverman, kernel = "gaussian",
                            n = 512, from = min(datos), to = max(datos))
```

```r
f_x1 <- densidad_hSilver$y[1]
```

Entonces, la estimación en el mínimo de la muestra es $\hat{f}_k(x_{(1)}) = 0.0265689$.

Recordando el intervalo desarrollado con el método delta, tenemos que un intervalo de confianza para $f(x)$ (considerando el sesgo despreciable, i.e. $E\left(\hat{f}_K(x_{(1)})\right) = f(x_{(1)})$)

$$IC = \left[\sqrt{\hat{f}_k(x_{(1)})} \pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{R(K)}{4nh}}\right]^2$$

Entonces, el intervalo de confianza viene dado por:

```r
# Construimos el intervalo con alpha = 0.05
alpha <- 0.05                    # Nivel de significancia
z_c <- qnorm(1 - alpha/2)        # Cuantil de la normal

R_k <- (2*sqrt(pi))^(-1)

(sqrt(f_x1) + c(-1, 1) * z_c * sqrt(R_k/(4*n*h_Silverman)))^2
```

```
## [1] 0.009895201 0.051313586
```

d. Primero calculemos las cantidades necesarias. Entonces

```r
# La desviación estándar ya la habíamos calculado con anterioridad
mu_est <- mean(datos)                         # Media estimada
h_normal <- 1.059 * sigma_hat * n^(-1/5)     # h estimado
```

Por ende tenemos que:

- El valor de $h = 0.7899$
- El valor de $\hat{\mu} = 20.006$
- El valor de $\hat{\sigma} = 2.041$

Como cambiamos el h, debemos re calcular la estimación kernel

```r
# Estimación kernel con el h de referencia Normal
densidad_hNormal <- density(datos, bw = h_normal, kernel = "gaussian",
                            n = 512, from = min(datos), to = max(datos))

f_xn <- densidad_hNormal$y[512]         # Estmación de E[f(xn)]
V_fxn <- (dnorm(0, 0, sd = sqrt(2*h_normal^2)) *
          dnorm(max(datos), mu_est,
                sd = sqrt(sd(datos)^2 + h_normal^2/2)) -
          dnorm(max(datos), mu_est,
                sd = sqrt(sd(datos)^2 + h_normal^2))^2)/n
```

Por ende tenemos que:

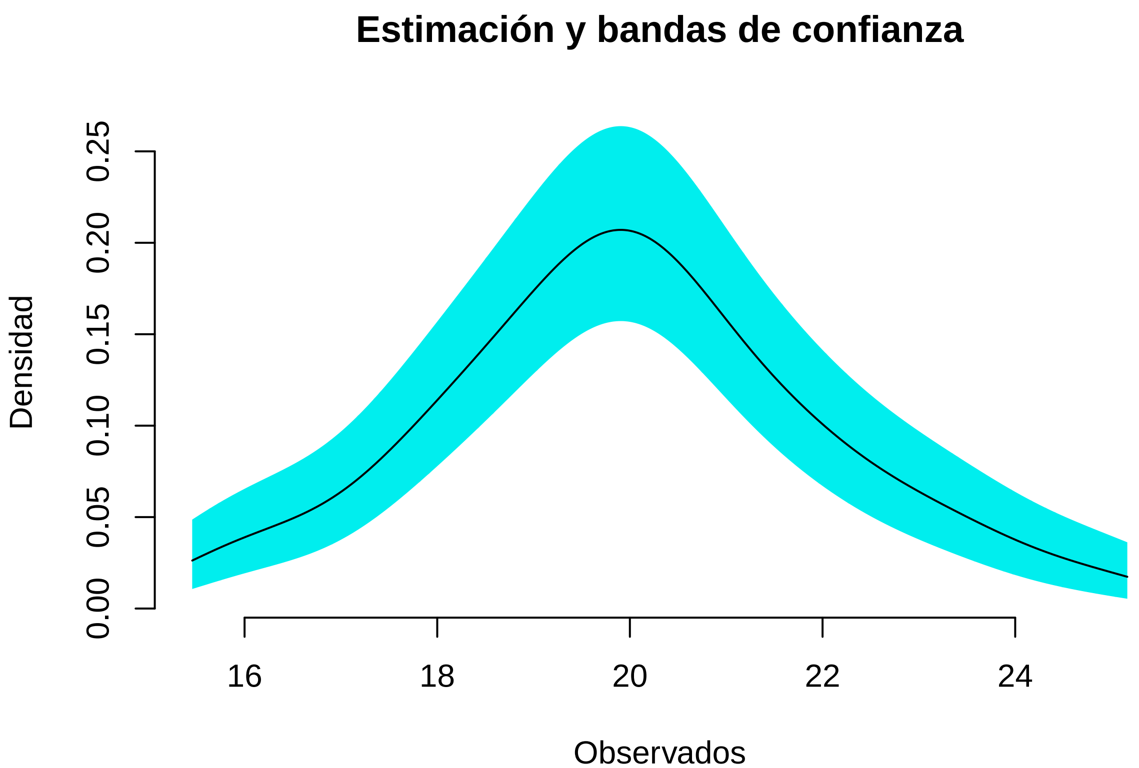
- El valor de $\hat{E}\left[\widehat{f_K}\left(x_{(n)}\right)\right] = 0.0174$
- El valor de $V\left[\widehat{f_K}\left(x_{(n)}\right)\right] = 3.3 \times 10^{-5}$
- Las bandas de referencia a la normal son

```r
f_xn + c(-1, 1) * 2.575 * sqrt(V_fxn)
```

```
## [1] 0.002509719 0.032207437
```

e. Bandas de variabilidad de $E\left[\widehat{f}_k(x)\right]$

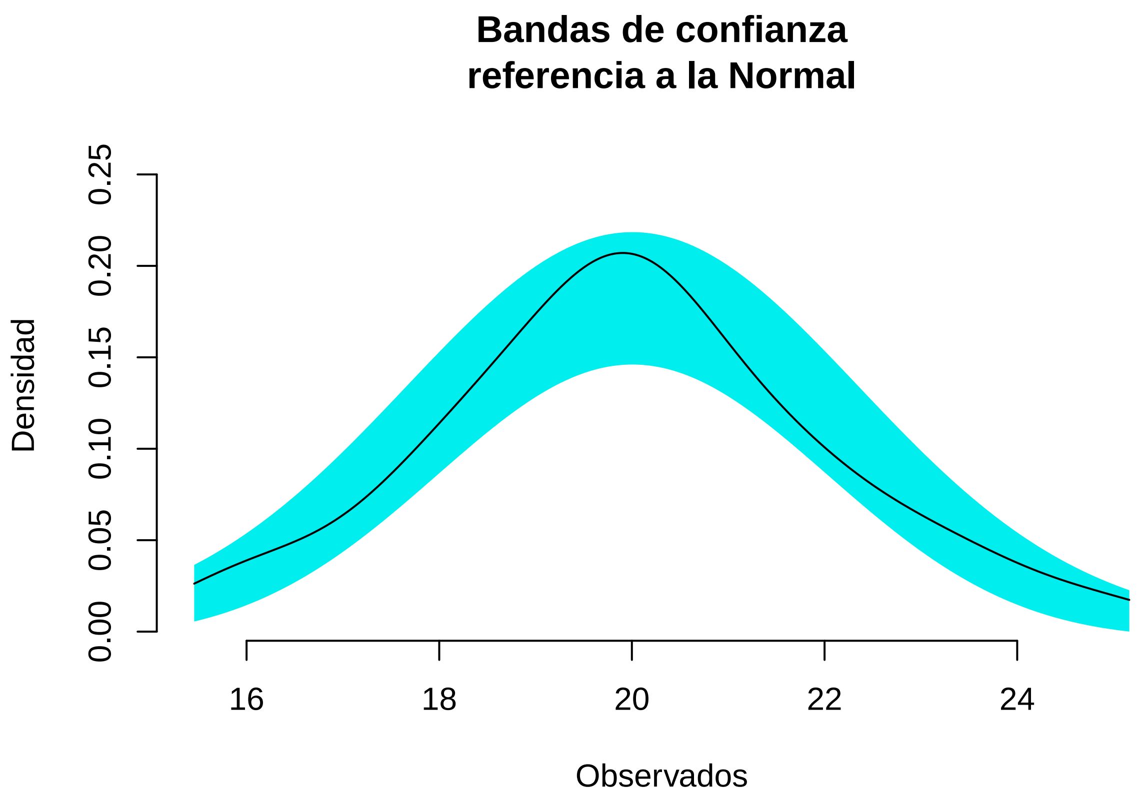
```r
# Bandas de referencia
x_obs <- densidad_hNormal$x
y_dens <- densidad_hNormal$y

band_inf <- (sqrt(y_dens) - z_c * sqrt(R_k/(4*n*h_normal)))^2
band_sup <- (sqrt(y_dens) + z_c * sqrt(R_k/(4*n*h_normal)))^2


# Gráfico de la densidad
plot(x = x_obs, y = y_dens, type = "n", frame = FALSE,
     ylim = c(min(band_inf), max(band_sup)), lwd = 2,
     xlab = "Observados", ylab = "Densidad",
     main = "Estimación y bandas de confianza")

polygon(x = c(x_obs, x_obs[512:1]), border = NA,
        y = c(band_inf, band_sup[512:1]), col = "cyan2")

lines(x = x_obs, y = y_dens)
```




```r
# Agregamos las líneas de la densidad
# matlines(x = x_obs, y = cbind(band_inf, band_sup), col = "cyan2",
#          lwd = 2, lty = 1)
```

f. Bandas de referencia a la normal

```
# Bandas de referencia a la normal
E_fx <- dnorm(x_obs, mean = mu_est,
              sd = sqrt(var(datos) + h_normal^2))
V_fx <- (dnorm(0, 0, sd = sqrt(2*h_normal^2)) *
         dnorm(x_obs, mean = mu_est,
               sd = sqrt(var(datos) + h_normal^2/2)) -
         dnorm(x_obs, mean = mu_est,
               sd = sqrt(var(datos) + h_normal^2))^2)/n

band_infN <- E_fx - qnorm(0.975)*sqrt(V_fx)
band_supN <- E_fx + qnorm(0.975)*sqrt(V_fx)

# Gráfico de la densidad
plot(x = x_obs, y = y_dens, type = "n", frame = FALSE,
     ylim = c(min(band_inf), max(band_sup)), lwd = 2,
     xlab = "Observados", ylab = "Densidad",
     main = "Bandas de confianza\nreferencia a la Normal")

polygon(x = c(x_obs, x_obs[512:1]), border = NA,
        y = c(band_infN, band_supN[512:1]), col = "cyan2")

lines(x = x_obs, y = y_dens)
```



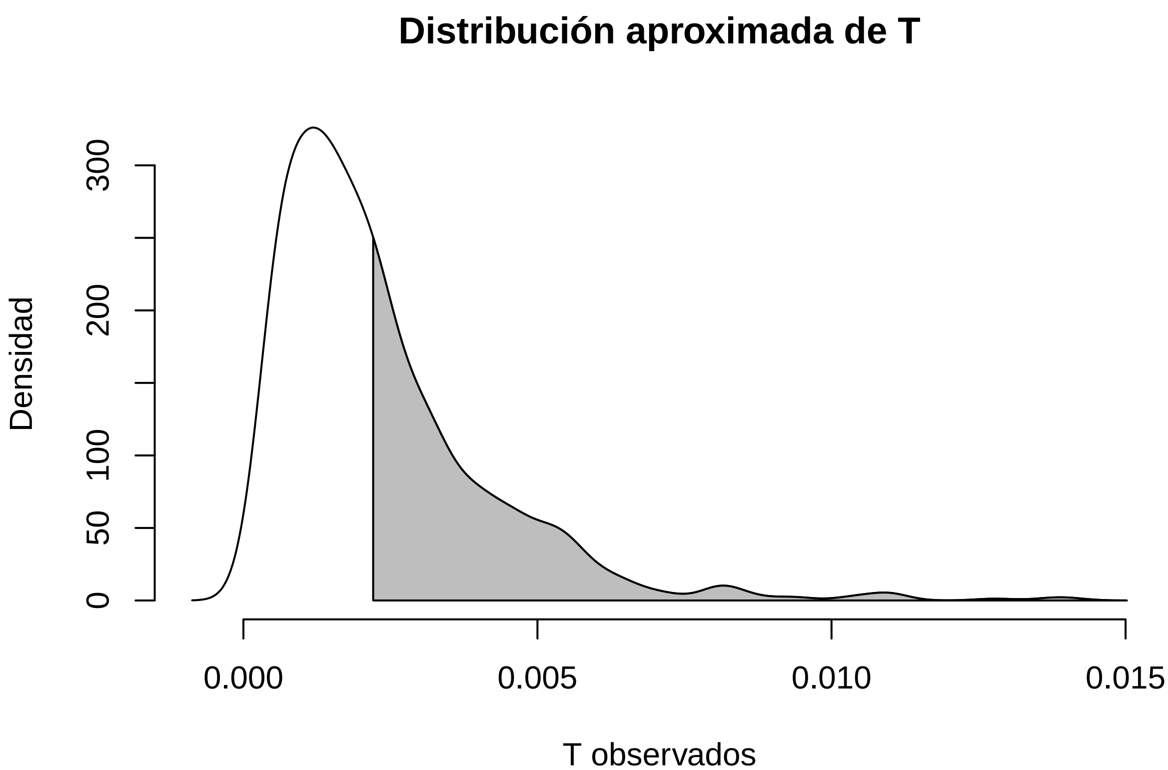
g. Test de normalidad usando *nise*

```
set.seed(7)              # Fijamos la semilla

T_obs <- nise(datos)     # T observado
num_sim <- 1000          # Número de simulaciones
```

```r
T_simul <- replicate(num_sim, expr = nise(rnorm(n)))
( pval <- sum(T_simul > T_obs)/num_sim )
```

```
## [1] 0.409
```

```r
# Graficamos la distribución asociada

densidad_t <- density(T_simul)
d_x <- densidad_t$x
d_y <- densidad_t$y

plot(d_x, d_y, frame = FALSE, type = "l",
     main = "Distribución aproximada de T",
     xlab = "T observados", ylab = "Densidad")

# Graficamos el valor-p
x_indice <- d_x[d_x > T_obs]
y_indice <- d_y[d_x > T_obs]

polygon(x = c(x_indice[1], x_indice), #border = "",
        y = c(0, y_indice), col = "gray")
```



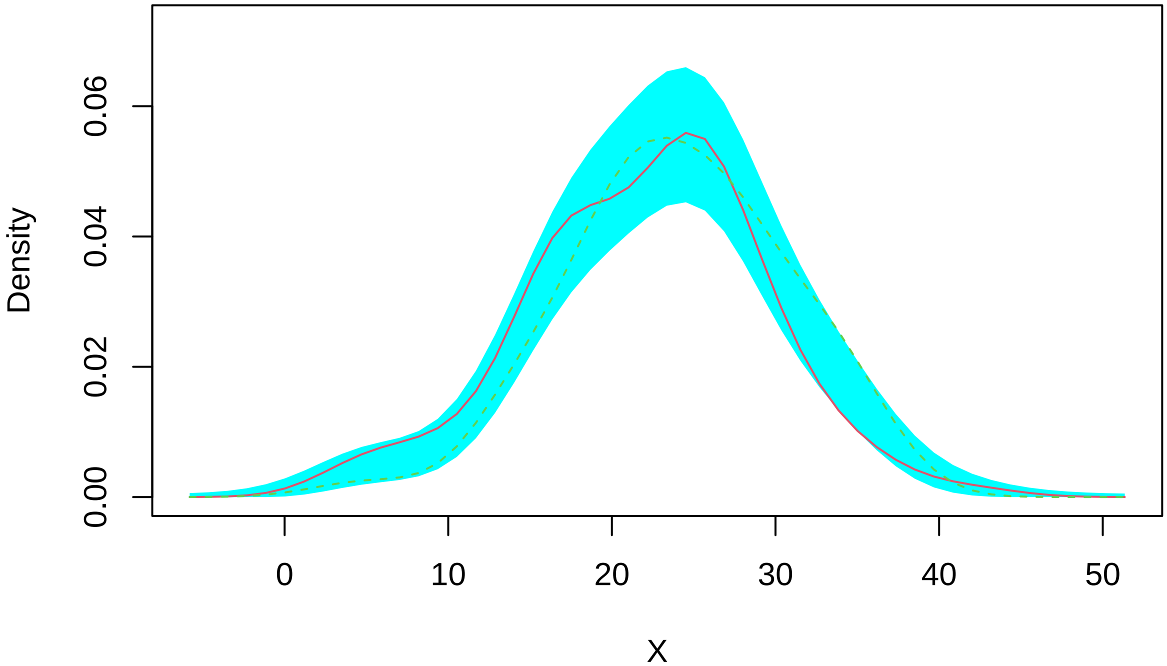
h. Comparación de distribuciones:

- **Bajo igualdad de distribuciones**: vamos a usar variables de una variable aleatoria normal

```r
# Fijamos la semilla
set.seed(31416)

X <- rnorm(200, mean = 23, sd = 7)       # Hacemos la toma de las muestras
```

```r
grupo <- rep(paste0("g", 1:2),          # Creamos los respectivos grupos
             each = 100)
sm.density.compare(x = X, group = grupo, model = "equal")
```

```
## Test of equal densities:  p-value =  0.24
```

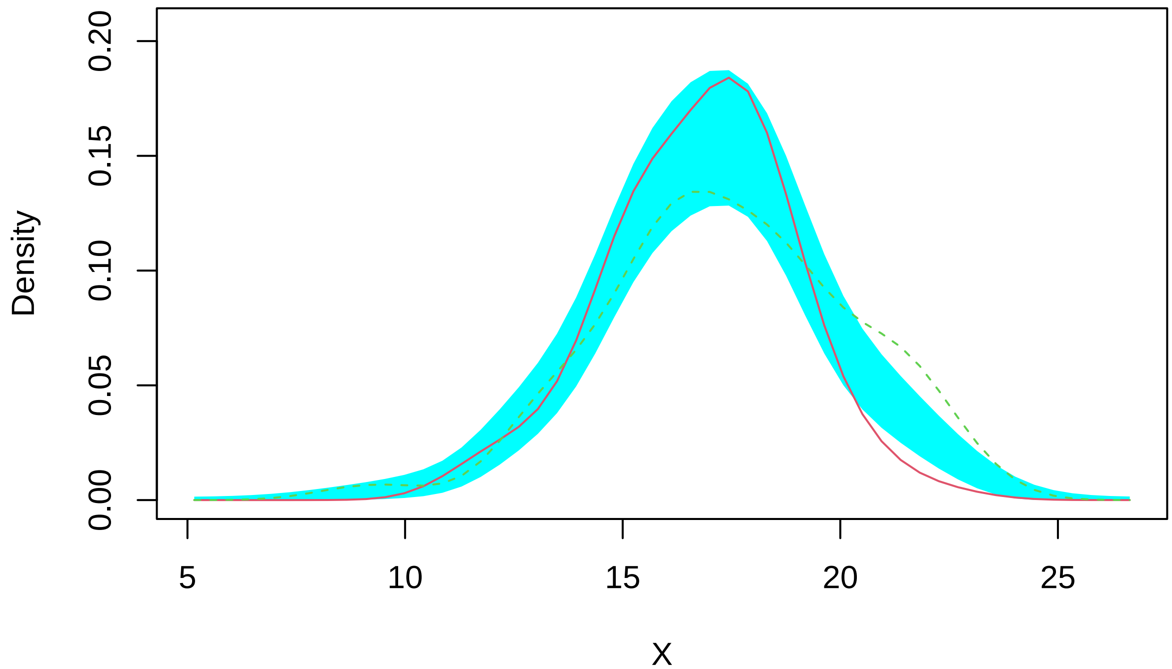


Vemos que no se rechaza la hipótesis nula de igualdad de distribuciones.

- **Bajo diferencia de distribuciones**: vamos a usar variables de una variable aleatoria normal

```r
# Fijamos la semilla
set.seed(31416)

X <- c(rnorm(102, 17, 2), rnorm(98, 17, 3))
grupo <- rep(paste0("g", 1:2), c(102, 98))
sm.density.compare(x = X, group = grupo, model = "equal")
```

```
## Test of equal densities:  p-value =  0
```



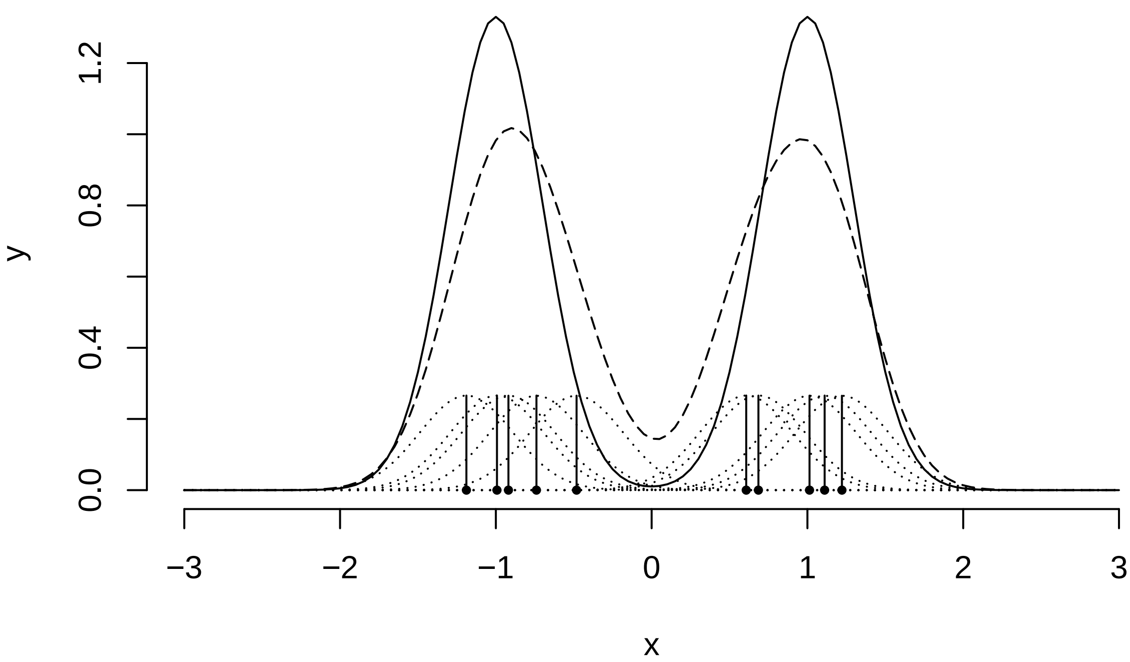
Y

podemos ver que rechazamos la hipótesis nula de igualdad de distribuciones.

i. Réplica de la función

```
# Gráfico de la distribución teórica
mu1 <- -1
mu2 <- 1
sd <- 0.3

x <- seq(-3, 3, by = 0.05)          # Rango de valores
y <- dnorm(x, mu1, sd = sd) +      # Parecen suma de normales :v
    dnorm(x, mu2, sd = sd)

set.seed(6)
x1 <- rnorm(5, mu1, sd = sd)
x2 <- rnorm(5, mu2, sd = sd)

plot(x, y, type = "l", main = "Figura 1", frame = FALSE)

X <- c(x1, x2)
# Agregamos los puntos de las simulaciones
points(x = X, y = rep(0, length(X)), pch = 19, cex = 0.5)

# Ancho de banda
bw <- 0.3
d_k <- sapply(X, function(k) dnorm(k-x, 0, sd = bw))

alturas <- apply(d_k, 2, max)
matlines(x = x, y = d_k/5, lty = 3, col = "gray0")

segments(x0 = X, y0 = rep(0, 10), x1 = X, y1 = alturas/5)
apply(d_k, 1, sum) / 5 -> estimaciOn

lines(x, estimaciOn, lty = 5, col = "gray0")
```

Figura 1



2. **Demostración**

$$
\begin{aligned}
E\left\{\widehat{f_H}(x)\right\} &= E\left\{\frac{1}{nb}\sum_{i=1}^{n} K_h\left(\frac{x-X_i}{b}\right)\right\} \\
&= \frac{1}{nb}E\left\{\sum_{i=1}^{n} K_H\left(\frac{x-X_i}{b}\right)\right\} \\
&\overset{\text{linealidad}}{=} \frac{1}{nb}\sum_{i=1}^{n} E\left\{K_H\left(\frac{x-X_i}{b}\right)\right\} \\
&\overset{\text{iid}}{=} \frac{1}{b\cancel{n}}\cancel{n}E\left\{K_H\left(\frac{x-X}{b}\right)\right\} \overset{\text{simetría}}{=} \frac{1}{b}E\left\{K_H\left(\frac{X-x}{b}\right)\right\} \\
&= \frac{1}{b}\int_{-\infty}^{\infty} K_H\left(\frac{s-x}{b}\right) f_X(s)ds \\
&\overset{t=\frac{s-x}{b}}{=} \frac{1}{\cancel{b}}\int_{-\infty}^{\infty} K_H(t) f_X(x+tb)\cancel{b}dt \\
&\overset{\text{E. Taylor}}{=} \int_{-\infty}^{\infty} K_H(t)\big[f_X(x)+O(b)\big]dt \\
&= \big[f_X(x)+O(b)\big]\cancelto{1}{\int_{-\infty}^{\infty} K_H(t)dt} \\
&= f_X(x)+O(b)
\end{aligned}
$$

3. **Demostración:**

$$
\begin{aligned}
V\left\{\widehat{f_H}(x)\right\} &= V\left\{\frac{1}{nb}\sum_{i=1}^{n} K_H\left(\frac{X_i-x}{b}\right)\right\} \\
&\overset{\text{reescalado}}{=} \frac{1}{n^2} V\left\{\sum_{i=1}^{n} K_{H_b}\left(X_i-x\right)\right\} \\
&\overset{iid}{=} \frac{1}{n^{\cancel{2}}} \cancel{n} V\left\{K_{H_b}(X-x)\right\} \\
&= \frac{1}{n} E\left\{K_{H_b}(X-x)^2\right\} - \frac{1}{n}\left[E\left\{K_{H_b}(X-x)\right\}\right]^2
\end{aligned}
$$

Notemos que, como $E\left\{K_{H_b}(X-x)\right\}$ existe, implica que existe un $c \in \mathbf{R}^+$ tal que $\left|E\left\{K_{H_b}(X-x)\right\}\right| \leq c \implies \left|E\left\{K_{H_b}(X-x)\right\}\right|^2 \leq c^2$.

Por ende podemos deducir que

$$
\begin{aligned}
\frac{1}{n}\left|E\left\{K_{H_b}(X-x)\right\}\right|^2 &\leq \frac{c^2}{n} \\
&\implies \frac{1}{n}\left|E\left\{K_{H_b}(X-x)\right\}\right|^2 = O\left(\frac{1}{n}\right)
\end{aligned}
$$

Nos falta analizar $\frac{1}{n}E\left\{K_{H_b}(X-x)^2\right\}$. Notemos que

$$
\begin{aligned}
\frac{1}{n}E\left\{K_{H_b}(X-x)^2\right\} &= \frac{1}{n}E\left\{\frac{1}{b^2}K_H\left(\frac{X-x}{b}\right)^2\right\} = \frac{1}{nb^2}E\left\{K_H\left(\frac{X-x}{b}\right)^2\right\} \\
&= \frac{1}{nb^2}\int_{-\infty}^{\infty} K_H\left(\frac{s-x}{b}\right)^2 f_X(s)ds \\
&\overset{t=\frac{s-x}{b}}{=} \frac{1}{nb^{\cancel{2}}}\int_{-\infty}^{\infty} K_H(t)^2 f_X(x+tb)\cancel{b}dt \\
&= \frac{1}{nb}\int_{-\infty}^{\infty} K_H(t)^2\left[f_X(x)+O(b)\right]dt \\
&= \left[f_X(x)+O(b)\right]\frac{1}{nb}\int_{-\infty}^{\infty} K_H(t)^2 dt \\
&= \frac{1}{nb}\left[f_X(x)+O(b)\right]R(K) \\
&= \frac{f_X(x)R(K)}{nb} + \frac{R(K)}{nb}O(b) = \frac{f_X(x)R(K)}{nb} + \frac{1}{nb}O(b)
\end{aligned}
$$

El término $O(b)$ expresa que la función original está acotada por la función kb para algún $k \in R$. Luego $\frac{O(b)}{b}$ estará acotado por una constante. Lo que nos lleva a deducir que $\frac{O(b)}{nb}$ está acotado por una múltiplo de $\frac{1}{n}$. Es decir, que $\frac{O(b)}{nb} = O\left(\frac{1}{n}\right)$. Resumiendo lo anterior, tenemos que

$$
\frac{1}{n}E\left\{K_{H_b}(X-x)^2\right\} = \frac{f_X(x)R(K)}{nb} + O\left(\frac{1}{n}\right)
$$

Que a su vez implica que

$$V\left\{\widehat{f_H}(x)\right\} \quad = \quad \frac{f_X(x)R(K)}{nb} + O\left(\frac{1}{n}\right) - O\left(\frac{1}{n}\right)$$

La expresión $O\left(\frac{1}{n}\right) - O(\frac{1}{n})$ es una combinación lineal de múltiplos de funciones de la forma $\frac{1}{n}$ y por ende dicha combinación será acotada por una función de la forma $\frac{1}{n}$. Es decir que $O(\frac{1}{n}) - O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right)$. Lo que nos permite completar nuestra demostración, i.e.

$$V\left\{\widehat{f_H}(x)\right\} = \frac{f_X(x)R(K)}{nb} + O\left(\frac{1}{n}\right)$$

4. **Demostración:** sea $X \sim \mathcal{N}(\mu, \sigma^2)$. Tenemos que la función de densidad de X viene dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

La derivada de esta función viene dada por:

$$\begin{aligned}
\frac{d}{dx} f_X(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\left(-\frac{\cancel{2}(x-\mu)}{\cancel{2}\sigma^2}\right) \\
&= -\frac{(x-\mu)}{\frac{\sqrt{2}}{\sqrt{2}}\sigma\cdot\sigma} f_X(x) = -\frac{1}{\sqrt{2}\sigma}\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}} f_X(x)
\end{aligned}$$

Y al elevar el cuadrado la expresión anterior se sigue que

$$\begin{aligned}
\left[f_X^{'}(x)\right]^2 &= \frac{1}{2\sigma^2}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 f_X(x)^2 \\
&= \frac{1}{2\sigma^2}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 \frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{2(x-\mu)^2}{2\sigma^2}\right\} \\
&= \frac{1}{2\sigma^2\sqrt{2\pi\sigma^2}}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 \frac{1}{\sqrt{2\pi 2\frac{\sigma^2}{2}}} \exp\left\{-\frac{(x-\mu)^2}{2\frac{\sigma^2}{2}}\right\} \\
&= \frac{1}{2\sigma^2\sqrt{2\pi\sigma^2}\sqrt{2}}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 f_{X^*}(x) \\
&= \frac{1}{4\sigma^2\sqrt{\pi\sigma^2}}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 f_{X^*}(x)
\end{aligned}$$

La función $f_{X^*}(x)$ se puede entender como la función de densidad de una variable aleatoria normal con media $\mu^* = \mu$ y varianza ${\sigma^*}^2 = \frac{\sigma^2}{2}$. Integrando sobre x la expresión anterior se tiene que

$$\begin{aligned}
R(f^{'}) &= \int_{-\infty}^{\infty} \frac{1}{4\sigma^3\sqrt{\pi}}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 f_{X^*}(x)dx \\
&= \frac{1}{4\sigma^3\sqrt{\pi}}\int_{-\infty}^{\infty}\left(\frac{x-\mu}{\frac{\sigma}{\sqrt{2}}}\right)^2 f_{X^*}(x)dx
\end{aligned}$$

La integral anterior vemos que expresa el cálculo de un valor esperado. Para ser más exactos, estamos calculando $E\left[\left(\frac{X^*-\mu^*}{\sigma^*}\right)^2\right]$. Pero esa transformación corresponde al valor esperado de una variable aleatoria *chi* cuadrado, con un grado de libertad. Como el valor esperado de esta corresponde a sus grados de libertad, se tiene que

$$R(f^{'}) = \frac{1}{4\sigma^3\sqrt{\pi}}\cancelto{1}{E\left[\left(\frac{X^*-\mu^*}{\sigma^*}\right)^2\right]} = \frac{1}{4\sigma^3\sqrt{\pi}}$$