Entrega 1 - Tests chi cuadrado

Coudet & Czarnievicz

September 2018

El código de esta entrega puede encontrarse en el siguiente repositorio de Github.

Ejercicio 2.1

Encontrar los dos primeros momentos del estadístico X_n^2 para cuando:

- La hipótesis nula es verdadera.
- La hipótesis nula es falsa.

$$X_n^2 = \sum_{i=1}^{k} \frac{(U_i - n\pi_i)^2}{n\pi_i} = \sum_{i=1}^{k} \frac{Ui^2}{n\pi_i} - n$$

Caso A: La hipótesis nula es verdadera

<u>**Primer momento**</u>

$$E_{H_0'}(X_n^2) = E_{H_0'}\left(\sum_{i=1}^{k} \frac{(U_i - n\pi_{i0})^2}{n\pi_{i0}}\right) = \sum_{i=1}^{k} \frac{1}{n\pi_{i0}} E_{H_0'}\Big((U_i - n\pi_{i0})^2\Big) =$$

Observe que el término $E((U_i - n\pi_{i0})^2)$, bajo la hipótesis nula, corresponde a $E((U_i - E(U_i))^2)$ dado que $E(U_i) = n\pi_{io}$. Con lo cual tenemos la fórmula de la varianza de U_i, la cual es conocida ya que son binomiales.

$$= \sum_{i=1}^{k} \frac{1}{n\pi_{i0}} V_{H_0'}(U_i) = \sum_{i=1}^{k} \frac{1}{n\pi_{i0}} n\pi_{i0} * (1 - \pi_{i0}) = \sum_{i=1}^{k} \frac{1}{\cancel{n\pi_{i0}}} \cancel{n\pi_{i0}} - \sum_{i=1}^{k} \frac{1}{\cancel{n\pi_{i0}}} \cancel{n\pi}_{i0}^{\cancel{2}} =$$

$$= k - \underbrace{\sum_{i=1}^{k} \pi_{i0}}_{1} = k - 1$$

Observe que es la esperanza de una VA con distribución chi-cuadrado y $k-1$ grados de libertad, lo cual es consistente con la distribución límite del estadístico.

Segundo momento

$$E_{H_0'}\Big[\Big(X_n^2\Big)^2\Big] = E_{H_0'}\left[\left(\sum_{i=1}^{k} \frac{(U_i - n\pi_{i0})^2}{n\pi_{i0}}\right)^2\right] = \frac{1}{n^2} E_{H_0'}\left[\left(\sum_{i=1}^{k} \frac{(U_i - n\pi_{i0})^2}{\pi_{i0}}\right)^2\right] =$$

$$= E_{H_0'}\Big[\Big(\underbrace{\sum_{i=1}^{k} \frac{U_i^2}{n\pi_i}}_{A} - \underbrace{n}_{B}\Big)^2\Big] = E_{H_0'}\left[\left(\sum_{i=1}^{k} \frac{U_i^2}{n\pi_i}\right)^2 - 2\sum_{i=1}^{k} \frac{U_i^2}{n\pi_i}\, n + n^2\right] =$$

Lo cual se obtiene a partir de que $(A-B)^2 = A^2 - 2AB + B^2$ y aplicando linealidad de la esperanza obtenemos que:

$$= \underbrace{E_{H_0'}\left[\left(\sum_{i=1}^{k} \frac{U_i^2}{n\pi_i}\right)^2\right]}_{\text{TÃ©rmino A}} - \underbrace{2n\, E_{H_0'}\left[\sum_{i=1}^{k} \frac{U_i^2}{n\pi_i}\right]}_{\text{TÃ©rmino B}} + n^2$$

TÃ©rmino B

$$2nE_{H_0'}\left[\sum_{i=1}^{k} \frac{Ui^2}{n\pi_{i0}}\right] = 2\cancel{n}\sum_{i=1}^{k} \frac{1}{\cancel{n}\pi_{i0}} \quad \underbrace{E_{H_0'}(U_i^2)}_{V_{H_0'}(U_i)\,+\,E^2_{H_0'}(U_i)} \quad =$$

Dado que las U_i son binomiales, conocemos sus varianzas y sus esperanzas bajo la hipótesis nula:

$$= 2\sum_{i=1}^{k} \frac{1}{\pi_{i0}}(n\pi_{i0}(1-\pi_{i0}) + n^2\pi_{i0}^2) = 2\sum_{i=1}^{k} \frac{1}{\cancel{\pi_{i0}}} n\cancel{\pi_{i0}}(1 - \pi_{i0} + n\pi_{i0}) =$$

$$= 2n\underbrace{\sum_{i=1}^{k} 1}_{k} - 2n\underbrace{\sum_{i=1}^{k} \pi_{i0}}_{1} + 2n^2\underbrace{\sum_{i=1}^{k} \pi_{i0}}_{1} = 2nk - 2n + 2n^2 = 2n(k+n-1)$$

Término A

$$E_{H_0'}\left[\left(\sum_{i=1}^{k}\frac{U_i^2}{n\pi_i}\right)^2\right] = \frac{1}{n^2}\sum_{i=1}^{k}E_{H_0'}\left[\frac{U_i^4}{\pi_i^2}\right] + \frac{2}{n^2}\sum_{i<j}E_{H_0'}\left[\frac{U_i^2}{\pi_i}\frac{U_j^2}{\pi_j}\right]$$

Para calcular dicha esperanza, usamos los momentos factoriales.

$$E(X\,(X-1)\,(X-2)\,\cdots\,(X-k+1)) = E\left[\frac{X!}{(X-k)!}\right] =$$

$$= \sum \frac{x!}{(x-k)!}\,\frac{n!}{x!(n-x)!}\,p^x(1-p)^{n-x} =$$

$$= \frac{n!}{(n-k)!}p^k \underbrace{\sum \frac{(n-k)!}{(x-k)!(n-x)!}\,p^{x-k}(1-p)^{n-x}}_{1} = \frac{n!}{(n-k)!}p^k$$

Entonces, si $X \sim \text{Binomial}(n,p)$:

$$E(X) = np$$
$$E(X(X-1)) = \frac{n!}{(n-2)!}p^2 = E(X^2) - E(X) = E(X^2) - np$$
$$\Rightarrow E(X^2) = n(n-1)p^2 + np$$

$$E(X(X-1)(X-2)) = E(X^3 - 3X^2 + 2X) =$$
$$E(X^3) - 3E(X^2) + 2E(X) =$$
$$\frac{n!}{(n-3)!}p^3 = n(n-1)(n-2)p^3$$
$$\Rightarrow E(X^3) = n(n-1)(n-2)p^3 + 3\left[n(n-1)p^2 + np\right] - 2np$$

$$E(X(X-1)(X-2)(X-3)) = E(X^4) - 6E(X^3) + 11E(X^2) - 6E(X) =$$

$$\frac{n!}{(n-4)!}p^4 = n(n-1)(n-2)(n-3)p^4$$
$$\Rightarrow E(X^4) = n(n-1)(n-2)(n-3)p^4 +$$
$$+6\left[n(n-1)(n-2)p^3 + 3\left[n(n-1)p^2 + np\right] - 2np\right] -$$

$$-11\left[n(n-1)p^2+np\right]+6np$$

Entonces, dado que las U_i bajo la hipótesis nula tienen distribución binomial de parámetros (n,π_{i0}) llegamos a que:

$$E_{H_0'}\left(U_i^4\right)=n(n-1)(n-2)(n-3)\pi_{i0}^4+6\left[n(n-1)(n-2)\pi_{i0}^3+3\left[n(n-1)\pi_{i0}^2+n\pi_{i0}\right]-2n\pi_{i0}\right]-$$

$$-11\left[n(n-1)\pi_{i0}^2+n\pi_{i0}\right]+6n\pi_{i0}$$

Ahora bien, hay que calcular también $E_{H_0'}\left[\frac{U_i^2}{\pi_i}\frac{U_j^2}{\pi_j}\right]$

$$U_i^2U_j^2=\left(\sum_{a=1}^{n}\mathrm{I}_{A_i}(X_a)\right)^2\left(\sum_{b=1}^{n}\mathrm{I}_{A_j}(X_b)\right)^2=$$

$$=\left[\sum_{a=1}^{n}\mathrm{I}_{A_i}^{\cancel{2}}(X_a)+2\sum_{a<b}\mathrm{I}_{A_i}(X_a)\mathrm{I}_{A_i}(X_b)\right]*\left[\sum_{a=1}^{n}\mathrm{I}_{A_j}^{\cancel{2}}(X_a)+2\sum_{a<b}\mathrm{I}_{A_j}(X_a)\mathrm{I}_{A_j}(X_b)\right]$$

Y tomando esperanza:

$$E(U_i^2U_j^2)_{H_0'}=\underbrace{E_{H_0'}\left[\sum_{a=1}^{n}\mathrm{I}_{A_i}(X_a)\sum_{a=1}^{n}\mathrm{I}_{A_j}(X_a)\right]}_{n(n-1)\pi_{i0}\pi_{j0}}+$$

$$\underbrace{+2E_{H_0'}\left[\sum_{a=1}^{n}\mathrm{I}_{A_i}(X_a)\sum_{a<b}\mathrm{I}_{A_j}(X_a)\mathrm{I}_{A_j}(X_b)\right]}_{n^2(n-1)\pi_{i0}\pi_{j0}^2}+$$

$$\underbrace{+2E_{H_0'}\left[\sum_{a<b}\mathrm{I}_{A_i}(X_a)\mathrm{I}_{A_i}(X_b)\sum_{a=1}^{n}\mathrm{I}_{A_j}(X_a)\right]}_{n^2(n-1)\pi_{j0}\pi_{i0}^2}+$$

$$\underbrace{+4E_{H_0'}\left[\sum_{a<b}\mathrm{I}_{A_i}(X_a)\mathrm{I}_{A_i}(X_b)\sum_{a<b}\mathrm{I}_{A_j}(X_a)\mathrm{I}_{A_j}(X_b)\right]}_{4n^2(n-1)^2\pi_{i0}^2\pi_{j0}^2}$$

Por lo tanto, volviendo sobre el segundo momento:

$$E_{H_0'}\Big[\big(X_n^2\big)^2\Big] = \underbrace{E_{H_0'}\left[\left(\sum_{i=1}^{k}\frac{U_i^2}{n\pi_i}\right)^2\right]}_{\text{TÃ©rmino A}} \underbrace{-2n\,E_{H_0'}\left[\sum_{i=1}^{k}\frac{U_i^2}{n\pi_i}\right]}_{\text{TÃ©rmino B}} + n^2 =$$

$$= \frac{1}{n^2}\sum_{i=1}^{k}E_{H_0'}\left[\frac{U_i^4}{\pi_i^2}\right] + \frac{2}{n^2}\sum_{i<j}E_{H_0'}\left[\frac{U_i^2}{\pi_i}\frac{U_j^2}{\pi_j}\right] - 2n(k+n-1) + n^2 =$$

$$= \frac{1}{n^2}\sum_{i=1}^{k}\frac{1}{\pi_{i0}^2}E_{H_0'}\left[U_i^4\right] + \frac{2}{n^2}\sum_{i<j}\frac{1}{\pi_{i0}\pi_{j0}}E_{H_0'}\left[U_i^2U_j^2\right] - 2n(k+n-1) + n^2 =$$

$$\frac{1}{n^2}\sum_{i=1}^{k}\frac{1}{\pi_{i0}^2}\Big[n(n-1)(n-2)(n-3)\pi_{i0}^4 + 6\Big[n(n-1)(n-2)\pi_{i0}^3 + 3\Big[n(n-1)\pi_{i0}^2 + n\pi_{i0}\Big] - 2n\pi_{i0}\Big]$$

$$-11\Big[n(n-1)\pi_{i0}^2 + n\pi_{i0}\Big] + 6n\pi_{i0}\Big] +$$

$$+\frac{2}{n^2}\sum_{i<j}\frac{1}{\pi_{i0}\pi_{j0}}\Big(n(n-1)\pi_{i0}\pi_{j0} + n^2(n-1)\pi_{i0}\pi_{j0}^2 + n^2(n-1)\pi_{j0}\pi_{i0}^2 + 4n^2(n-1)^2\pi_{i0}^2\pi_{j0}^2\Big) +$$

$$+2n(k+n-1) + n^2$$

Caso B: La hipótesis nula es falsa. Es decir, los verdaderos valores de los parámetros π_i son diferentes a los postulados en la hipótesis nula π_{i0}.

Primer momento

$$E(X_n^2) = E\left(\sum_{i=1}^{k}\frac{(U_i - n\pi_{i0})^2}{n\pi_{i0}}\right) = \sum_{i=1}^{k}\frac{1}{n\pi_{i0}}E((U_i - n\pi_{i0})^2) = \sum_{i=1}^{k}\frac{1}{n\pi_{i0}}E((\underbrace{U_i - n\pi_i}_{A} + \underbrace{n\pi_i - n\pi_{i0}}_{B})^2) =$$

$$= \sum_{i=1}^{k}\frac{1}{n\pi_{i0}}E((U_i - n\pi_i)^2) - \cancel{2 * \sum_{i=1}^{k}\frac{1}{n\pi_{i0}} * \underbrace{E((U_i - n\pi_i) * (n\pi_i - n\pi_{i0}))}_{(n\pi_i - n\pi_{i0})(\underbrace{E(U_i) - n\pi_i}_{0})}} + \sum_{i=1}^{k}\frac{1}{n\pi_{i0}}E((n\pi_i - n\pi_{i0})^2) =$$

$$= \sum_{i=1}^{k}\frac{1}{n\pi_{i0}}\underbrace{E((U_i - n\pi_i)^2)}_{V(U_i) = n\pi_i(1-\pi_i)} + \sum_{i=1}^{k}\frac{1}{n\pi_{i0}}\underbrace{E((n\pi_i - n\pi_{i0})^2)}_{(n\pi_i - n\pi_{i0})^2} = \sum_{i=1}^{k}\frac{\pi_i(1-\pi_i)}{\pi_{i0}} + \sum_{i=1}^{k}\frac{n^{\cancel{2}}(\pi_i - \pi_{i0})^2}{\cancel{n}\pi_{i0}} =$$

$$= \sum_{i=1}^{k}\frac{\pi_i(1-\pi_i)}{\pi_{i0}} + \sum_{i=1}^{k}\frac{n(\pi_i - \pi_{i0})^2}{\pi_{i0}}$$

Ejercicio 2.2

Ejercicio 2.5

```
U <- c(74, 92, 83, 79, 80, 73, 77, 75, 76, 91)
pi_i0 <- rep(0.1, 10)
n <- sum(U)
k <- length(U)
X2 <- sum(((U-n*pi_i0)^2)/(n*pi_i0))
valor.critico <- qchisq(0.95, df=k-1)
if(X2 > valor.critico){
      print("Rechazo H0")
} else {
      print("No rechazo H0")
}
```

```
[1] "No rechazo H0"
```

Ejercicio 2.6

```
pot <- seq(0, 1, 0.051)
pot <- pot[-1]
pot <- pot[-length(pot)]
n.beta <- NULL
for (i in 1:length(pot)) {
      n.beta[i] <- pwr.chisq.test(w=0.1, N=NULL, df=9, sig.level=0.05,
                                  power=pot[i])$N
}
```

```
as.tibble(cbind(pot, n.beta)) %>%
      ggplot(aes(x=pot, y=n.beta)) +
      geom_line(colour='navy') +
      labs(x='Potencia', y='NÃºmero de observaciones') +
      ggthemes::theme_economist() +
      theme(axis.title=element_text(face='bold', size=12))
```

```
## Warning: `as.tibble()` is deprecated, use `as_tibble()` (but mind the new semantics).
## This warning is displayed once per session.
```

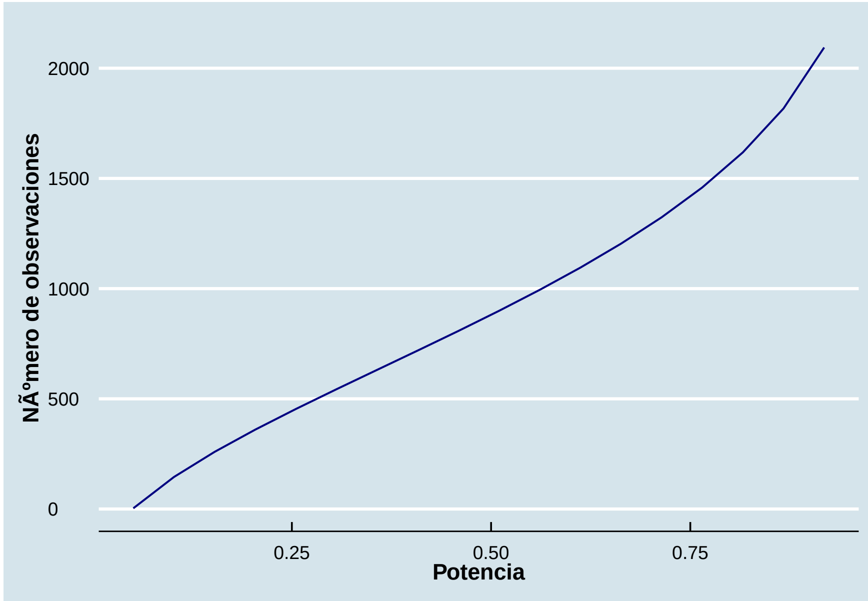

Figure 1: Gráfico del número de observaciones en función de la potencia e la prueba

Ejercicio 2.7

```
U <- c(315, 101, 108, 32)
pi_i0 <- c(9/16, 3/16, 3/16, 1/16)
n <- sum(U)
k <- length(U)
X2 <- sum(((U-n*pi_i0)^2)/(n*pi_i0))
valor.critico <- qchisq(0.95, df=k-1)
if(X2 > valor.critico){
      print("Rechazo H0")
} else {
      print("No rechazo H0")
}
```

```
[1] "No rechazo H0"
```

Ejercicio 2.8

```
U <- c(773, 231, 238, 59)
pi_i0 <- c(9/16, 3/16, 3/16, 1/16)
n <- sum(U)
k <- length(U)
X2 <- sum(((U-n*pi_i0)^2)/(n*pi_i0))
confianza <- pchisq(X2, df=k-1)
nivel.sig <- 1 - confianza
nivel.sig
```

```
[1] 0.02589168
```

Ejemplo 2.4

```
x = c(5.017, 0.146, 6.474, 13.291, 5.126, 8.934, 10.971, 7.863, 5.492,
      13.930, 12.708, 7.329, 5.408, 6.808, 0.923, 4.679, 2.242, 4.120,
      12.080, 2.502, 16.182, 6.592, 2.653, 4.252, 8.609, 10.419, 2.173,
      3.321, 4.086,  11.667, 19.474, 11.067, 11.503, 2.284, 0.926, 2.065,
      4.703, 3.744, 5.286, 5.497, 4.881, 0.529, 10.397, 30.621, 5.193,
      7.901, 10.220, 16.806, 10.672, 4.209, 5.699, 20.952, 12.542, 7.316,
      0.272, 4.380, 9.699, 9.466, 7.928, 13.086, 8.871, 13.000, 16.132,
      9.950, 8.449, 8.301, 16.127, 22.698, 4.335, 2.992)
n = length(x)
xbar = mean(x)
k = 6
p = 1/k
P_i = c(1:k*p)
z_i = c(0, -log(1 - P_i))
a_i = xbar * z_i
b_i = NULL
for(i in 1:(length(z_i)-1)){
      if(i == k){
            b_i[i] = - z_i[i] * exp(-z_i[i])
      } else {
            b_i[i] = z_i[i+1] * exp(-z_i[i+1]) - z_i[i] * exp(-z_i[i])
      }
}
U = rep(0, k)
```

```
for(i in 1:length(U)){
      for(j in 1:length(x)){
            if ((x[j] > a_i[i]) & (x[j] <= a_i[i+1])) {
                  U[i] = U[i] + 1
            } else {next}
      }
}
v = sum(b_i * U / p)
lambda = 1 - sum(b_i^2 / p)
X2 = sum(U^2 / (n*p)) - n
Q = v^2 / (n*lambda)
Y2 = X2 + Q
valor.critico = qchisq(p=0.95, df=k-1)
if(Y2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "Rechazo H0"
```

Ejercicio 2.15

```
U <- c(53, 41, 30, 22, 16, 12, 9, 7, 5, 5)
a_i <- c(0, 300, 600, 900, 1200, 1500, 1800, 2100, 2400, 2700, 3600)
n <- sum(U)
k <- length(U)
p <- 1/k
tita <- seq(850, 900, 0.0001)
tita <- tita[-1]
log.vero <- matrix(NA, nrow=length(tita), ncol=k)
for (i in 1:length(tita)) {
      for (j in 2:length(a_i)) {
            log.vero[i,j-1] <- U[j-1] * log(exp(-a_i[j-1]/tita[i]) -
                                              exp(-a_i[j]/tita[i]))
      }
}
log.vero <- rowSums(log.vero)
tita_ml <- tita[which.max(log.vero)]
pi_i <- NULL
```

```r
for (j in 2:length(a_i)) {
      pi_i[j-1] <- exp(-a_i[j-1]/tita_ml) - exp(-a_i[j]/tita_ml)
}
X2 <- sum( (U - n * pi_i)^2 / (n * pi_i) )
valor.critico = qchisq(p=0.95, df=k-1)
if(Y2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "No rechazo H0"
```

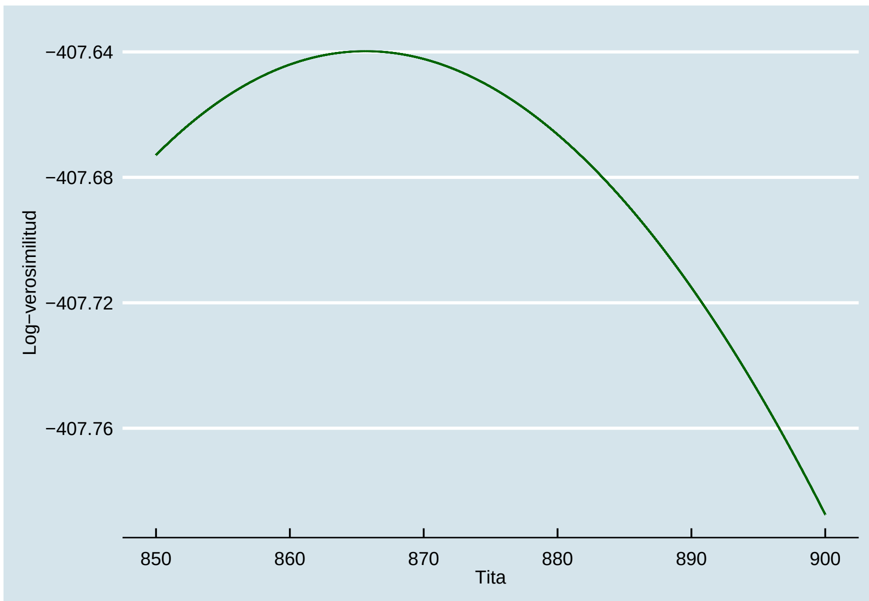

Figure 2: Gráfico del logaritmo de la verosimilitud de los datos agrupados

Ejemplo 2.5

```r
# CASO A
x <- c(8.7, 6.6, 10.0, 24.3, 7.9, 1.3, 26.2, 8.3, 0.9, 7.1, 5.9, 16.8, 6.0,
       13.4, 31.7, 8.3, 28.3, 17.1, 16.7, 19.7, 5.2, 18.9, 1.0, 3.5, 2.7,
       12.0, 8.3, 14.8, 6.3, 39.3, 4.3, 19.4, 6.5, 7.4, 3.4, 7.6, 8.3, 1.9,
       10.3, 3.2, 0.7, 19.0, 26.2, 10.0, 17.7, 14.1, 44.8, 3.4, 3.5)
n <- length(x)
```

```
k <- 6
p <- 1/k
xbar <- mean(x)
sigma2 <- (1/n)*sum((x-xbar)^2)
sigma <- sqrt(sigma2)
P_i <- c(1:k*p)
z_i <- qnorm(P_i, mean=xbar, sd=sigma)
# a_i <- xbar + z_i * sigma   # Â¿Â¿los z_i son los a_i del ejemplo??
a_i <-c(-Inf, z_i)
U = rep(0, k)
for(i in 1:length(U)){
      for(j in 1:length(x)){
            if ((x[j] > a_i[i]) & (x[j] <= a_i[i+1])) {
                  U[i] = U[i] + 1
            } else {next}
      }
}
X2 = sum(U^2 / (n*p)) - n
valor.critico <- qchisq(0.95, df=k-1)
if(X2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "Rechazo H0"
```

```
# CASO B
x <- c(8.7, 6.6, 10.0, 24.3, 7.9, 1.3, 26.2, 8.3, 0.9, 7.1, 5.9, 16.8, 6.0,
       13.4, 31.7, 8.3, 28.3, 17.1, 16.7, 19.7, 5.2, 18.9, 1.0, 3.5, 2.7,
       12.0, 8.3, 14.8, 6.3, 39.3, 4.3, 19.4, 6.5, 7.4, 3.4, 7.6, 8.3, 1.9,
       10.3, 3.2, 0.7, 19.0, 26.2, 10.0, 17.7, 14.1, 44.8, 3.4, 3.5)
x <- log(x)
n <- length(x)
k <- 6
p <- 1/k
xbar <- mean(x)
sigma2 <- (1/n)*sum((x-xbar)^2)
sigma <- sqrt(sigma2)
P_i <- c(1:k*p)
z_i <- qnorm(P_i, mean=xbar, sd=sigma)
# a_i <- xbar + z_i * sigma   # Â¿Â¿los z_i son los a_i del ejemplo??
a_i <-c(-Inf, z_i)
```

```
U = rep(0, k)
for(i in 1:length(U)){
      for(j in 1:length(x)){
            if ((x[j] > a_i[i]) & (x[j] <= a_i[i+1])) {
                  U[i] = U[i] + 1
            } else {next}
      }
}
X2 = sum(U^2 / (n*p)) - n
valor.critico <- qchisq(0.95, df=k-1)
if(X2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "No rechazo H0"
```

```
# CASO C
x <- c(8.7, 6.6, 10.0, 24.3, 7.9, 1.3, 26.2, 8.3, 0.9, 7.1, 5.9, 16.8, 6.0,
       13.4, 31.7, 8.3, 28.3, 17.1, 16.7, 19.7, 5.2, 18.9, 1.0, 3.5, 2.7,
       12.0, 8.3, 14.8, 6.3, 39.3, 4.3, 19.4, 6.5, 7.4, 3.4, 7.6, 8.3, 1.9,
       10.3, 3.2, 0.7, 19.0, 26.2, 10.0, 17.7, 14.1, 44.8, 3.4, 3.5)
x <- x^(1/4)
n <- length(x)
k <- 6
p <- 1/k
xbar <- mean(x)
sigma2 <- (1/n)*sum((x-xbar)^2)
sigma <- sqrt(sigma2)
P_i <- c(1:k*p)
z_i <- qnorm(P_i, mean=xbar, sd=sigma)
# a_i <- xbar + z_i * sigma   # Â¿Â¿los z_i son los a_i del ejemplo??
a_i <-c(-Inf, z_i)
U = rep(0, k)
for(i in 1:length(U)){
      for(j in 1:length(x)){
            if ((x[j] > a_i[i]) & (x[j] <= a_i[i+1])) {
                  U[i] = U[i] + 1
            } else {next}
      }
}
X2 = sum(U^2 / (n*p)) - n
```

```
valor.critico <- qchisq(0.95, df=k-1)
if(X2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "No rechazo H0"
```

Ejercicio 2.16

```
x <- c(338, 336, 312, 322, 381, 302, 296, 360, 342, 334, 348, 304, 323, 310,
       368, 341, 298, 312, 322, 350, 304, 302, 336, 334, 304, 292, 324, 331,
       324, 334, 314, 338, 324, 292, 298, 342, 338, 331, 325, 324, 326, 314,
       312, 362, 368, 321, 352, 304, 302, 332, 314, 304, 312, 381, 290, 322,
       326, 316, 328, 340, 324, 320, 364, 304, 340, 290, 318, 332, 354, 324,
       304, 321, 356, 366, 328, 332, 304, 282, 330, 314, 342, 322, 362, 298,
       316, 298, 332, 342, 316, 326, 308, 321, 302, 304, 322, 296, 322, 338,
       324, 323)
n <- length(x)
k <- 8
p <- 1/k
xbar <- mean(x)
sigma2 <- (1/n)*sum((x-xbar)^2)
sigma <- sqrt(sigma2)
P_i <- c(1:k*p)
z_i <- qnorm(P_i)
z_i <-c(-Inf, z_i)
a_i <- xbar + z_i * sigma
U <- rep(0, k)
for(i in 1:length(U)){
      for(j in 1:length(x)){
            if ((x[j] > a_i[i]) & (x[j] <= a_i[i+1])) {
                  U[i] <- U[i] + 1
            } else {next}
      }
}
X2 <- sum(U^2 / (n*p)) - n
b_1 <- diff(dnorm(z_i))
b_2 <- NULL
```

```
for(i in 2:(length(z_i))){
      if (i == 2) {
            b_2[i-1] <- z_i[i] * dnorm(z_i[i])
      } else if (i == length(z_i)) {
            b_2[i-1] <- -z_i[i-1] * dnorm(z_i[i-1])
      } else {
            b_2[i-1] <- z_i[i] * dnorm(z_i[i]) - z_i[i-1] * dnorm(z_i[i-1])
      }
}
j_02 <- 1
j_12 <- 0
lambda_1 <- j_02 - sum(b_1^2/p)
lambda_2 <- 2 - sum(b_2^2/p)
lambda_3 <- j_12 - sum((b_1*b_2)/p)
alpha <- sum(b_1*U/p)
beta <- sum(b_2*U/p)
Q <- (lambda_1 * alpha^2 - 2 * lambda_3 * alpha * beta + lambda_2 * beta^2) /
      (n * (lambda_1 * lambda_2 - lambda_3^2))
Y2 <- X2 + Q
valor.critico <- qchisq(0.95, df=k-1)
if(Y2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "No rechazo H0"
```

Ejercicio 2.17

```
x <- c(10, 51, 8, 47, 8, 5, 56, 12, 4, 5, 4, 4, 7, 6, 9, 30, 25, 12, 3, 22,
       5, 15, 4, 4, 29, 15, 4, 2, 18, 41, 3, 5, 54, 110, 24, 16, 2, 37, 20,
       2, 6, 7, 16, 2, 14, 68, 10, 16, 11, 78, 6, 17, 7, 11, 21, 15, 24, 6,
       32, 8, 11, 4, 14, 45, 17, 10, 15, 20, 4, 65, 10, 3, 5, 11, 13, 35,
       11, 34, 3, 4, 12, 7, 6, 62, 13, 36, 26, 6, 11, 6, 13, 1, 4, 36, 18,
       10, 37, 28, 4, 12, 31, 14, 3, 11, 6, 4, 10, 38, 6, 11, 24, 9, 4, 5,
       8, 135, 22, 6, 18, 49, 17, 9, 32, 27, 2, 12, 8, 93, 3, 9, 10, 3, 14,
       33, 72, 14, 4, 9, 10, 19, 2, 5, 21, 8, 25, 30, 20, 12, 19, 16)
x <- log(x)
n <- length(x)
```

```
k <- 10
p <- 1/k
xbar <- mean(x)
sigma2 <- (1/n)*sum((x-xbar)^2)
sigma <- sqrt(sigma2)
P_i <- c(1:k*p)
z_i <- qnorm(P_i)
z_i <-c(-Inf, z_i)
a_i <- xbar + z_i * sigma
U <- rep(0, k)
for(i in 1:length(U)){
      for(j in 1:length(x)){
            if ((x[j] > a_i[i]) & (x[j] <= a_i[i+1])) {
                  U[i] <- U[i] + 1
            } else {next}
      }
}
X2 <- sum(U^2 / (n*p)) - n
b_1 <- diff(dnorm(z_i))
b_2 <- NULL
for(i in 2:(length(z_i))){
      if (i == 2) {
            b_2[i-1] <- z_i[i] * dnorm(z_i[i])
      } else if (i == length(z_i)) {
            b_2[i-1] <- -z_i[i-1] * dnorm(z_i[i-1])
      } else {
            b_2[i-1] <- z_i[i] * dnorm(z_i[i]) - z_i[i-1] * dnorm(z_i[i-1])
      }
}
j_02 <- 1
j_12 <- 0
lambda_1 <- j_02 - sum(b_1^2/p)
lambda_2 <- 2 - sum(b_2^2/p)
lambda_3 <- j_12 - sum((b_1*b_2)/p)
alpha <- sum(b_1*U/p)
beta <- sum(b_2*U/p)
Q <- (lambda_1 * alpha^2 - 2 * lambda_3 * alpha * beta + lambda_2 * beta^2) /
      (n * (lambda_1 * lambda_2 - lambda_3^2))
Y2 <- X2 + Q
valor.critico <- qchisq(0.95, df=k-1)
if(Y2 < valor.critico){
      print("No rechazo H0")
} else {
      print("Rechazo H0")
```

```
}
```

```
[1] "No rechazo H0"
```

Ejercicio 2.22

```
n <- sum(U)
pi_i <- c(21, 21, 45)
pi_j <- c(20, 35, 32)
k <- (length(pi_i) - 1)*(length(pi_j) - 1)
p_ij <- pi_i %*% t(pi_j)
U <- matrix(c(5,9,7,7,5,9,8,21,16), nrow=3, byrow=T)
x2 <- 87*(sum(U^2/p_ij)-1)
valor.critico <- qchisq(0.95, df=k)
if (x2 < valor.critico) {
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "No rechazo H0"
```

Ejercicio 2.23

```
U <- matrix(c(4, 12, 35, 61, 52, 23, 7, 4, 2, 1, 0, 0, 6, 10, 12, 13, 12,
              15, 12, 11, 7, 4), nrow=2, byrow=T)
U_j <- colSums(U)
n_i <- rowSums(U)
n <- sum(U)
x2 <- n * (sum(diag(1/n_i)%*%(U^2)%*%diag(1/U_j)) - 1)
k <- (dim(U)[1] - 1)*(dim(U)[2] - 1)
valor.critico <- qchisq(0.95, df=k)
if (x2 < valor.critico) {
      print("No rechazo H0")
} else {
      print("Rechazo H0")
}
```

```
[1] "Rechazo H0"
```

Actividad de programación

```r
x2_observado <- 5.125 # tomado del ejercicio 2.5
pi <- rep(0.1, 10)
N <- 10000
n <- 800
U <- rmultinom(n = N, size = n, prob = pi)
x2 <- apply(U, 2, function(x) sum(((x-n*pi)^2)/(n*pi)) )
p <- (as.numeric(table(x2 > x2_observado)[2]))/N
alpha <- 0.05
if(p < alpha){
      print("Rechazo H0")
} else {
      print("No rechazo H0")
}
```

```
[1] "No rechazo H0"
```