4.iii) Usando un generador de números aleatorios, generar 200 datos $\{x_i\}_{i=1}^{200}$, $X_i \sim$ Pareto con $\theta = 1.2$

```r
set.seed(1)
n <- 200
theta <- 1.2
x <- EnvStats::rpareto(n,
                       location = 1,
                       shape = theta)
```

4.iv) Con los datos $\{x_i\}_{i=1}^{200}$ obtenidos en iii), dar un intervalo de confianza (95%) de $\theta + \ln(\theta)$

Primero, notemos que

$$g : \mathbb{R} \to \mathbb{R} \,,$$
$$\theta + \ln(\theta) = g(\theta) \implies \frac{\partial g(\theta)}{\partial \theta} = 1 + \frac{1}{\theta}$$

Además, recordando el resultado del primer inciso, y utilizando el principio de invarianza, obtenemos

$$f : \mathbb{R} \to \mathbb{R} \,,$$
$$V = \theta^2 = f(\theta) \implies \hat{V}_n = f(\hat{\theta}_{MV}) = \hat{\theta}_{MV}^2 \quad , \text{ donde } \hat{\theta}_{MV} = \frac{n}{\sum_{i=1}^n \ln(X_i)}$$

Una vez planteado lo anterior, para calcular el intervalo de confianza de $g(\theta)$ conviene emplear el siguiente pivote:

$$\sqrt{n}\left(\frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}} \hat{V}_n \frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}}'\right)^{-1/2} \Big(g(\hat{\theta}_{MV}) - g(\theta)\Big) \xrightarrow{D} N(0,1)$$

Desarrollando el pivote:

$$\begin{aligned}
\sqrt{n}\left(\frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}} \hat{V}_n \frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}}'\right)^{-1/2} \Big(g(\hat{\theta}_{MV}) - g(\theta)\Big) &= \sqrt{n}\Big(g'(\hat{\theta}_{MV})^2 (\hat{\theta}_{MV})^2\Big)^{-1/2} \Big(g(\hat{\theta}_{MV}) - g(\theta)\Big) \\
&= \sqrt{n}\Big(g'(\hat{\theta}_{MV})\hat{\theta}_{MV}\Big)^{-1} \Big(g(\hat{\theta}_{MV}) - g(\theta)\Big) \\
&= \sqrt{n}\frac{g(\hat{\theta}_{MV}) - g(\theta)}{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}} \xrightarrow{D} N(0,1)
\end{aligned}$$

Con esto podemos construir el intervalo de confianza. Sean q_1 y q_2 dos cuantiles de la normal estándar tal que $q_1 < q_2$. Entonces:

$$\begin{aligned}
&\mathbb{P}\left[q_1 < \sqrt{n}\frac{g(\hat{\theta}_{MV}) - g(\theta)}{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}} < q_2\right] = 0.95 \\
\iff &\mathbb{P}\left[q_1 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}} < g(\hat{\theta}_{MV}) - g(\theta) < q_2 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}}\right] = 0.95 \\
\iff &\mathbb{P}\left[-g(\hat{\theta}_{MV}) + q_1 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}} < -g(\theta) < -g(\hat{\theta}_{MV}) + q_2 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}}\right] = 0.95 \\
\iff &\mathbb{P}\left[g(\hat{\theta}_{MV}) - q_1 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}} > g(\theta) > g(\hat{\theta}_{MV}) - q_2 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}}\right] = 0.95 \\
\iff &\mathbb{P}\left[g(\hat{\theta}_{MV}) - q_2 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}} < g(\theta) < g(\hat{\theta}_{MV}) - q_1 \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}}\right] = 0.95
\end{aligned}$$

El intervalo de confianza para $g(\theta) = \theta + \ln(\theta)$ es entonces:

$$\begin{aligned}
IC_{g(\theta)} &= \left(g(\hat{\theta}_{MV}) - q_{97.5\%}^{N(0,1)} \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}} \ , \ g(\hat{\theta}_{MV}) - q_{2.5\%}^{N(0,1)} \frac{g'(\hat{\theta}_{MV})\hat{\theta}_{MV}}{\sqrt{n}}\right) \\
&= \left(\hat{\theta}_{MV} + \ln(\hat{\theta}_{MV}) - q_{97.5\%}^{N(0,1)} \frac{\hat{\theta}_{MV} + 1}{\sqrt{n}} \ , \ \hat{\theta}_{MV} + \ln(\hat{\theta}_{MV}) - q_{2.5\%}^{N(0,1)} \frac{\hat{\theta}_{MV} + 1}{\sqrt{n}}\right) ,
\end{aligned}$$

donde $\hat{\theta}_{MV} = \dfrac{n}{\sum_{i=1}^n \ln(X_i)}$

Usando los datos generados en iii), computemos el intervalo. Primero, observemos que dado el valor verdadero de θ, $g(\theta)$ sería:

```r
g_theta <- theta + log(theta)
g_theta
```

```
## [1] 1.382322
```

Es decir, $g(\theta)\Big|_{\theta=1.2} = 1.3823216$. Tengamos esto en cuenta, pues el intervalo de confianza debería contener este valor dentro de su rango. Calculemos entonces este intervalo:

```r
theta_mv <- n / sum(log(x)) # Estimado de máxima verosimilitud
g_theta_mv <- theta_mv + log(theta_mv) # g(theta_mv)

alpha <- 0.05 # significancia. confianza = 1 - alpha
q1 <- qnorm(alpha/2) # cuantil 2.5%
q2 <- qnorm(1 - alpha/2) # cuantil 97.5%

# Intervalo
IC <- c(
  g_theta_mv - q2 * (theta_mv + 1) / sqrt(n),
  g_theta_mv - q1 * (theta_mv + 1) / sqrt(n)
)
IC
```

```
## [1] 1.090549 1.702497
```

Esto es, el intervalo de confianza de $\theta + \ln(\theta)$ al 95% de confianza es:

$$IC_{g(\theta)} = \Big(1.0905488\ ,\ 1.7024967\Big)$$

El cual contiene al valor verdadero de $g(\theta)$.

4.v) Con los datos $\{x_i\}_{i=1}^{200}$ obtenidos en iii), usando el estadístico de Wald probar la hipótesis

$$\mathrm{H}_0 : \theta + \ln(\theta) = 1.3 \quad \text{vs}$$
$$\mathrm{H}_1 : \theta + \ln(\theta) \neq 1.3$$

con nivel de significancia del 5%

Bajo H_0 el estadístico de Wald es

$$\mathrm{W}_n = \left(\sqrt{n}\ \ (g(\hat{\theta}_{MV}) - 1.3)\right)^2 \left(\frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}} \hat{V}_n \frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}}\right)^{-1} \xrightarrow{D} \chi_1^2$$

Desarrollando:

$$\begin{aligned}
\mathrm{W}_n &= \left(\sqrt{n}\ \ (g(\hat{\theta}_{MV}) - 1.3)\right)^2 \left(\frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}} \hat{V}_n \frac{\partial g(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}}\right)^{-1} \\
&= \left(\sqrt{n}\ \ (\hat{\theta}_{MV} + \ln(\hat{\theta}_{MV}) - 1.3)\right)^2 (g'(\hat{\theta}_{MV})^2 (\hat{\theta}_{MV})^2)^{-1} \\
&= \frac{n\ \ (\hat{\theta}_{MV} + \ln(\hat{\theta}_{MV}) - 1.3)^2}{(g'(\hat{\theta}_{MV})(\hat{\theta}_{MV}))^2} \\
&= \frac{n\ \ (\hat{\theta}_{MV} + \ln(\hat{\theta}_{MV}) - 1.3)^2}{(\hat{\theta}_{MV} + 1)^2} \\
\iff \mathrm{W}_n &= n \left(\frac{\hat{\theta}_{MV} + \ln(\hat{\theta}_{MV}) - 1.3}{\hat{\theta}_{MV} + 1}\right)^2 \xrightarrow{D} \chi_1^2
\end{aligned}$$

Con esto podemos computar el estadístico

```
wald_n <- n * ((g_theta_mv - 1.3) / (theta_mv + 1))^2
wald_n
```

```
## [1] 0.3822842
```

Es decir, $\mathrm{W}_n = 0.3822842$. El valor crítico al 5% de significancia, $\Sigma_{5\%}$ sería

```
alpha <- 0.05
q <- qchisq(alpha, df = 1, lower.tail = F)
q
```

```
## [1] 3.841459
```

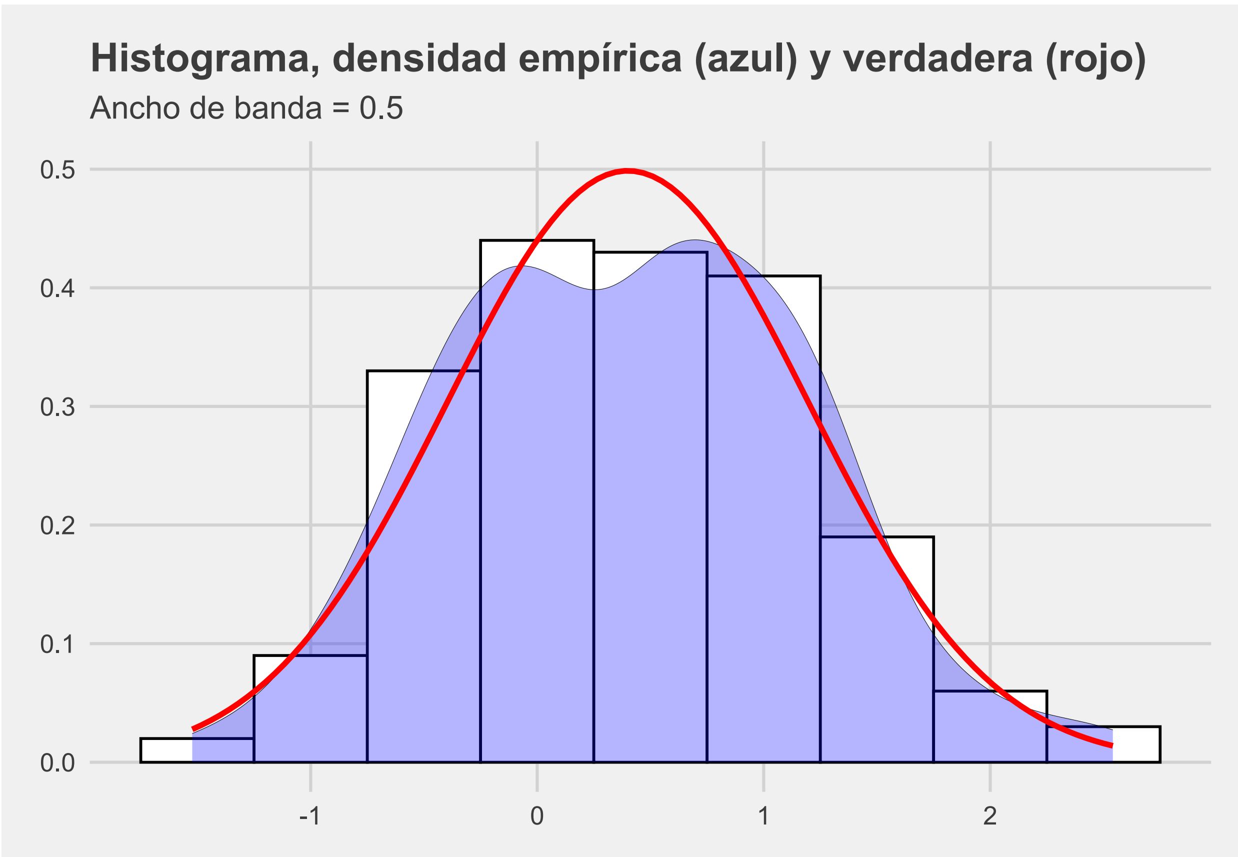
Observemos que

$$\mathrm{W}_n = 0.3822842 < 3.8414588 = \Sigma_{5\%}$$

Por lo tanto, aceptamos la hipótesis nula. Esto era de esperarse, dado que el valor 1.3 está dentro del intervalo de confianza con el mismo nivel de confianza.

5.iii) Usando un generador de números aleatorios, generar 200 datos $\{x_i\}_{i=1}^{200}$, $X_i \sim N(\mu, \sigma^2)$. con $\mu = 0.4$ y $\sigma^2 = 0.64$

```r
rm(list = ls())
set.seed(3)
n <- 200
mu <- 0.4
sigma2 <- 0.64
x <- rnorm(n = 200, mean = mu, sd = sqrt(sigma2))
```



5.iv) Con los datos $\{x_i\}_{i=1}^{200}$ obtenidos en iii), dar un intervalo de confianza (95%) de $\mu^2+e^{\sigma^2}$

Dado que $\mu^2+e^{\sigma^2}=h(\theta)$, $h:\mathbb{R}^2\to\mathbb{R}$,

Primero, notemos que

$$h:\mathbb{R}\times\mathbb{R}^+\to\mathbb{R}\ ,$$
$$\mu^2+e^{\sigma^2}=h(\theta)\implies\frac{\partial h(\theta)}{\partial\theta}=\begin{pmatrix}\frac{\partial h(\theta)}{\partial\mu} & \frac{\partial h(\theta)}{\partial\sigma^2}\end{pmatrix}=\begin{pmatrix}2\mu & e^{\sigma^2}\end{pmatrix}$$

Además, recordando el resultado del primer inciso, y utilizando el principio de invarianza, obtenemos

$$j:\mathbb{R}\times\mathbb{R}^+\to\mathbb{R}^4\ ,$$
$$V=\begin{pmatrix}\sigma^2 & 0\\ 0 & 2\sigma^4\end{pmatrix}=j(\theta)\implies\hat{V_n}=j(\hat{\theta}_{MV})=\begin{pmatrix}\hat{\sigma}^2_{MV} & 0\\ 0 & 2\hat{\sigma}^4_{MV}\end{pmatrix}\quad ,$$

donde, por el inciso ii) sabemos que $\hat{\sigma}^2_{MV}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$

Una vez planteado lo anterior, para calcular el intervalo de confianza de $g(\theta)$ conviene emplear el siguiente pivote:

$$\sqrt{n}\left(\frac{\partial h(\hat{\theta}_{MV})}{\partial\hat{\theta}_{MV}}\hat{V_n}\frac{\partial h(\hat{\theta}_{MV})}{\partial\hat{\theta}_{MV}}'\right)^{-1/2}\Big(h(\hat{\theta}_{MV})-h(\theta)\Big)\xrightarrow{D}N(0,1)$$

Desarrollando el pivote:

$$\begin{aligned}
&\sqrt{n}\left(\frac{\partial h(\hat{\theta}_{MV})}{\partial\hat{\theta}_{MV}}\hat{V_n}\frac{\partial h(\hat{\theta}_{MV})}{\partial\hat{\theta}_{MV}}'\right)^{-1/2}\Big(h(\hat{\theta}_{MV})-h(\theta)\Big)\\
&=\sqrt{n}\left(\begin{pmatrix}2\hat{\mu}_{MV} & e^{\hat{\sigma}^2_{MV}}\end{pmatrix}\begin{pmatrix}\hat{\sigma}^2_{MV} & 0\\ 0 & 2\hat{\sigma}^4_{MV}\end{pmatrix}\begin{pmatrix}2\hat{\mu}_{MV}\\ e^{\hat{\sigma}^2_{MV}}\end{pmatrix}\right)^{-1/2}\Big(h(\hat{\theta}_{MV})-h(\theta)\Big)\\
&=\sqrt{n}\left(2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV}+\hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})\right)^{-1/2}\Big(h(\hat{\theta}_{MV})-h(\theta)\Big)\\
&=\sqrt{n}\frac{h(\hat{\theta}_{MV})-h(\theta)}{\sqrt{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV}+\hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})}}\xrightarrow{D}N(0,1)
\end{aligned}$$

Con esto podemos construir el intervalo de confianza. Sean q_1 y q_2 dos cuantiles de la normal estándar tal que $q_1 < q_2$. Entonces:

$$\mathbb{P}\left[q_1 < \sqrt{n}\frac{h(\hat{\theta}_{MV}) - h(\theta)}{\sqrt{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV} + \hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})}} < q_2\right] = 0.95$$

$$\iff \mathbb{P}\left[q_1\sqrt{\frac{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV} + \hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})}{n}} < h(\hat{\theta}_{MV}) - h(\theta) < q_2\sqrt{\frac{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV} + \hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})}{n}}\right] = 0.95$$

$$\iff \mathbb{P}\left[-h(\hat{\theta}_{MV}) + q_1\gamma < -h(\theta) < -h(\hat{\theta}_{MV}) + q_2\gamma\right] = 0.95$$

$$\iff \mathbb{P}\left[h(\hat{\theta}_{MV}) - q_1\gamma > h(\theta) > h(\hat{\theta}_{MV}) - q_2\gamma\right] = 0.95$$

$$\iff \mathbb{P}\left[h(\hat{\theta}_{MV}) - q_2\gamma < h(\theta) < h(\hat{\theta}_{MV}) - q_1\gamma\right] = 0.95 \quad ,$$

$$\text{donde } \gamma = \sqrt{\frac{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV} + \hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})}{n}}$$

El intervalo de confianza para $h(\theta) = \mu^2 + e^{\sigma^2}$ es entonces:

$$\begin{aligned} IC_{h(\theta)} &= \left(h(\hat{\theta}_{MV}) - q^{N(0,1)}_{97.5\%}\gamma\ ,\ h(\hat{\theta}_{MV}) - q^{N(0,1)}_{2.5\%}\gamma\right) \\ &= \left(\hat{\mu}^2_{MV} + e^{\hat{\sigma}^2_{MV}} - q^{N(0,1)}_{97.5\%}\gamma\ ,\ \hat{\mu}^2_{MV} + e^{\hat{\sigma}^2_{MV}} - q^{N(0,1)}_{2.5\%}\gamma\right) \ , \end{aligned}$$

$$\begin{aligned} \text{donde } \hat{\sigma}^2_{MV} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \\ \hat{\mu}_{MV} &= \bar{x}_n \\ \gamma &= \sqrt{\frac{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV} + \hat{\sigma}^2_{MV}e^{2\hat{\sigma}^2_{MV}})}{n}} \end{aligned}$$

Usando los datos generados en iii), computemos el intervalo. Primero, observemos que dado los valores verdaderos de θ, $h(\theta)$ sería:

```
h_theta <- mu^2 + exp(sigma2)
h_theta
```

```
## [1] 2.056481
```

Es decir, $h(\theta)\Big|_{\substack{\mu=0.4 \\ \sigma_2=0.64}} = 2.0564809$. Tengamos esto en cuenta, pues el intervalo de confianza debería contener este valor dentro de su rango. Calculemos entonces este intervalo:

```
theta_mv <- matrix(  # Estimadores de MV
  c(
    mu_mv = mean(x),
    sigma2_mv = var(x)*(n-1)/n
    )
  )
h_theta_mv <- theta_mv[1]^2 + exp(theta_mv[2]) # h(theta_mv)
```

```r
gamma_mv <- sqrt(
  (2*theta_mv[2] * (2*theta_mv[1]^2 + theta_mv[2]*exp(2*theta_mv[2]))
   ) / n
  ) # el elemento gamma que definimos para simplificar

alpha <- 0.05 # significancia. confianza = 1 - alpha
q1 <- qnorm(alpha/2) # cuantil 2.5%
q2 <- qnorm(1 - alpha/2) # cuantil 97.5%

# Intervalo
IC <- c(
  h_theta - q2 * gamma_mv,
  h_theta - q1 * gamma_mv
)
IC
```

```
## [1] 1.816484 2.296478
```

Esto es, el intervalo de confianza de $\mu^2 + e^{\sigma^2}$ al 95% de confianza es:

$$IC_{h(\theta)} = \Big(1.8164838\ ,\ 2.2964779\Big)$$

El cual contiene al valor verdadero de $h(\theta)$.

5.v) Con los datos $\{x_i\}_{i=1}^{200}$ obtenidos en iii), usando el estadístico de Wald probar la hipótesis

$$\text{H}_0 : \mu^2 + e^{\sigma^2} = 2 \quad \text{vs}$$
$$\text{H}_1 : \mu^2 + e^{\sigma^2} \neq 2$$

con nivel de significancia del 5%

Bajo H_0 el estadístico de Wald es

$$\text{W}_n = \left(\sqrt{n} \ \ (h(\hat{\theta}_{MV}) - 2)\right)^2 \left(\frac{\partial h(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}} \hat{V}_n \frac{\partial h(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}}^{'}\right)^{-1} \xrightarrow{D} \chi_1^2$$

Desarrollando:

$$\text{W}_n = \left(\sqrt{n} \ \ (h(\hat{\theta}_{MV}) - 2)\right)^2 \left(\frac{\partial h(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}} \hat{V}_n \frac{\partial h(\hat{\theta}_{MV})}{\partial \hat{\theta}_{MV}}^{'}\right)^{-1}$$

$$\iff \text{W}_n = n \frac{\left(h(\hat{\theta}_{MV}) - 2\right)^2}{2\hat{\sigma}^2_{MV}(2\hat{\mu}^2_{MV} + \hat{\sigma}^2_{MV} e^{2\hat{\sigma}^2_{MV}})} \xrightarrow{D} \chi_1^2$$

$$\iff \text{W}_n = \left(\frac{h(\hat{\theta}_{MV}) - 2}{\gamma}\right)^2 \xrightarrow{D} \chi_1^2$$

Calculemos el estadístico

```r
wald_n <- ((h_theta_mv - 2) / gamma_mv)^2
wald_n
```

```
## [1] 0.0228647
```

Es decir, $\text{W}_n = 0.0228647$. El valor crítico al 5% de significancia, $\Sigma_{5\%}$ sería

```r
alpha <- 0.05
q <- qchisq(alpha, df = 1, lower.tail = F)
q
```

```
## [1] 3.841459
```

Observemos que

$$\text{W}_n = 0.0228647 < 3.8414588 = \Sigma_{5\%}$$

Por lo tanto, aceptamos la hipótesis nula. Esto era de esperarse, dado que el valor de 2 está dentro del intervalo de confianza con el mismo nivel de confianza.