UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Instituto de Cálculo

IECD 2C2024 - Grupo 3

Trabajo Práctico Final

Aaron Bernal Huanca
Nicolás Spisso
Marina Badaracco

Profesora: María Eugenia Szretter Noste

6 de diciembre de 2024

1 OOP en R

1.1 Breve intro opinada

1.1.1 Pregunta 1

Consideremos solamente vectores de las clases *character*, *numeric* y *logical*. Un vector que sólo contenga valores logical será claramente de esta clase (por ejemplo, `c(T, F)`). Si contiene al menos un valor numeric y n valores logical tendrá asignada la clase numeric, como `c(T, F, 1)`. Por último, un vector con al menos un valor character y cantidades arbitrarias de elementos logical y numeric tendrá asignada la clase character, por ejemplo `c(T, F, 1, "1")`.

Esto nos lleva a pensar que en `R` las clases están jerarquizadas de esta manera: logical → numeric → character, donde character es la clase más flexible de las tres, por poder englobar variables de más tipos, aunque lo haga perdiendo información. Esto último se refiere a que ya no se pueden realizar operaciones con las cantidades convertidas.

Para que las asignaciones sean consistentes, debe haber cierta transformación de los elementos que no eran originalmente de la clase asignada. En efecto, las variables logical se transforman de la siguiente manera: T → "TRUE", T → 1; mientras que las numeric: 1 → "1".

1.2 Métodos genéricos

1.2.1 Pregunta 2

Ejecutamos las dos cláusulas que pide el enunciado: `class(density)` y `class(density(1:500))`, obteniendo `function` y `density`, respectivamente. La diferencia radica en que `density` es una función en `R`, mientras que `density(1:500)` es la función `density` aplicada a los datos 1:500 (los enteros en el intervalo [1, 500]), lo cual genera una densidad.

1.3 Introspección: methods

1.3.1 Pregunta 3

El genérico `print` sabe despachar `191` clases. Por otro lado, la función `density` cuenta con otros 5 métodos además de `plot`.

1.4 Fijando ideas: mi.t.test

1.4.1 Pregunta 4

El comando `class(unclass(t_test))` devuelve `list`, esto ocurre porque la clase de `t_test` es `htest`, que esta estructurado sobre la clase implícita `list` a la que no se le puede extraer nada.

2 Test de Wilcoxon de rango signado para una muestra

2.1 Motivación: diseños experimentales apareados

2.1.1 Pregunta 5

Lema 2.1. *Bajo la hipótesis nula $H_0 : F_X = F_Y$, la distribución de $D = X - Y$ es simétrica alrededor del cero.*

Proof. Sabemos que bajo H_0 $F_X = F_Y$ vale $F_{X,Y} = F_{Y,X}$. Por lo tanto, $(X, Y) \sim (Y, X)$, lo que implica que $h(X, Y) \sim h(Y, X)$ para toda $h : \mathbb{R}^2 \to \mathbb{R}$ medible. En particular, tomando $h(X, Y) = X - Y$ tenemos que $X - Y \sim Y - X$. Entonces, para $d \in \mathbb{R}$:

$$
\begin{aligned}
F_D(d) &= P(D \leq d) && (1)\\
&= P(X - Y \leq d) && (2)\\
&= P(Y - X \leq d) && (3)\\
&= P(-D \leq d) && (4)\\
&= P(D \geq -d) && (5)\\
&= 1 - P(D \leq -d) && (6)\\
&= 1 - F_D(-d) && (7)
\end{aligned}
$$

donde (2) es igual a (3) ya que $X - Y \sim Y - X$ con la función h elegida. Luego, F_D es simétrica alrededor del cero ya que $F_D(d) = 1 - F_D(-d)\, \forall d \in \mathbb{R}$. □

2.2 Descripción del test

2.2.1 Pregunta 6

Lema 2.2. *Los siguientes estadísticos son equivalentes:*

- $T = \sum_{i=1}^{n} signo(X_i) R_i$
- $T^+ = \sum_{i=1}^{n} \mathbb{1}\{X_i > 0\} R_i$
- $T^- = \sum_{i=1}^{n} \mathbb{1}\{X_i < 0\} R_i$

Proof. Primero calculamos $T^+ - T^-$:

$$
\begin{aligned}
T^+ - T^- &= \sum_{i=1}^{n} \mathbb{1}\{X_i > 0\} R_i - \sum_{i=1}^{n} \mathbb{1}\{X_i > 0\} R_i && (8)\\
&= \sum_{i=1}^{n} [\mathbb{1}\{X_i > 0\} - \mathbb{1}\{X_i < 0\}] R_i. && (9)
\end{aligned}
$$

Sea $g(x) = \mathbb{1}\{x > 0\} - \mathbb{1}\{x < 0\}$. Notemos que:

$$g(X_i) = \begin{cases} 1 & \text{si } X_i > 0 \\ -1 & \text{si } X_i < 0 \\ 0 & \text{si } X_i = 0. \end{cases} \tag{10}$$

entonces $g(X_i) = signo(X_i)$. Volviendo a (9):

$$\sum_i^n g(X_i)R_i = \sum_i^n signo(X_i)R_i = T. \tag{11}$$

Veamos ahora $T^+ + T^-$:

$$\begin{aligned} T^+ + T^- &= \sum_i^n \mathbb{1}\{X_i > 0\}R_i + \sum_i^n \mathbb{1}\{X_i > 0\} && (12)\\ &= \sum_i^n [\mathbb{1}\{X_i > 0\} + \mathbb{1}\{X_i < 0\}]R_i && (13)\\ &= \sum_i^n \mathbb{1}\{X_i \neq 0\}R_i && (14)\\ &= \sum_i^n R_i && (15)\\ &= \frac{n(n+1)}{2}. && (16) \end{aligned}$$

El paso (14) es posible ya que se sigue de $P(X_i = 0) = 0$ que $\mathbb{1}\{X_i \neq 0\} = 1$. Se tiene entonces el siguiente sistema de dos ecuaciones lineales con tres incógnitas:

$$T^+ - T^- = T, \tag{17}$$

$$T^+ + T^- = \frac{n(n+1)}{2}. \tag{18}$$

Conociendo a uno de los tres estadísticos (por ejemplo, T), el sistema anterior se convierte en un sistema de dos ecuaciones lineales con dos incógnitas (por ejemplo, T^+ y T^- si conocemos T). □

2.3 Distribución de T^+ bajo la hipótesis nula

2.3.1 Pregunta 7

Lema 2.3. *Sean las hipótesis $H_0 : \theta = 0$ y $H_1 : \theta > 0$, mutuamente excluyentes para $\theta \in \Theta = [0, +\infty]$. Bajo H_0, $|X_i|$ es independiente de signo(X_i).*

Proof. Como se trata de una muestra aleatoria, signo(X_i) es independiente de $|X_j|$ para $i \neq j$. Falta demostrar que signo(X_i) es independiente de $|X_i|$ $\forall i$.

Decimos que X, Y son independientes si

$$P(X \in A \cap Y \in B) = P(X \in A) \cdot P(Y \in B) \ \forall A, B \text{ borelianos.} \tag{19}$$

Equivalentemente, basta verlo para todo par de intervalos (porque su σ-álgebra genera a los borelianos).

Ademas, como signo(X_i) discreta, alcanza ver con cada valor de su rango $\{-1, 1\}$.

Tanto las variables signo(X_i) como $|X_i|$ son iid por ser funciones de variables iid, por lo cual podemos considerar el caso particular $i = 1$ sin pérdida de generalidad.

Para $t > 0$ tenemos:

$$\begin{aligned}
P(\text{signo}(X_1) = 1, |X_1| \leq t) &= P(X_1 > 0, -t \leq X_1 \leq t) && (20)\\
&= P(0 \leq X_1 \leq t) && (21)\\
&= P(X_1 \leq t) - P(X_1 \leq 0) && (22)\\
&= F_{X_1}(t) - F_{X_1}(0); && (23)
\end{aligned}$$

mientras que el producto de las probabilidades es:

$$\begin{aligned}
P(\text{signo}(X_1) = 1) \cdot P(|X_1| \leq t) &= P(X_1 > 0) \cdot P(-t \leq X_1 \leq t) && (24)\\
&= F_{X_1}(0) \cdot [F_{X_1}(t) - F_{X_1}(-t)] && (25)\\
&= \frac{1}{2}[F_{X_1}(t) - F_{X_1}(-t)] && (26)\\
&= \frac{1}{2}[F_{X_1}(t) - (1 - F_{X_1}(t))] && (27)\\
&= F_{X_1}(t) - \frac{1}{2} && (28)\\
&= F_{X_1}(t) - F_{X_1}(0); && (29)
\end{aligned}$$

Bajo H_0, la distribución de X_1 es simétrica alrededor del 0, por lo cual:

$$P(X_1 \leq 0) = F_{X_1}(0) = \frac{1}{2}. \tag{30}$$

Notando que las expresiones (23) y (29) son iguales, llegamos a:

$$P(\text{signo}(X_1) = 1, |X_1| \leq t) = P(\text{signo}(X_1) = 1) \cdot P(|X_1| \leq t), \tag{31}$$

Como

$$\mathbb{P}(D \cap C) = \mathbb{P}(D)\mathbb{P}(C) \Leftrightarrow \mathbb{P}(D^c \cap C) = \mathbb{P}(D^c)\mathbb{P}(C)$$

en el sentido de eventos independientes. Si $D = \{\text{signo}(X_1) = 1\}$, se tiene que:

$$\mathbb{P}(\text{signo}(X_1) \neq 1) = \mathbb{P}(\text{signo}(X_1) = -1)$$

porque X_1 es continua, la probabilidad de valer 0 es nula. Entonces,

$$P(\text{signo}(X_1) = -1, |X_1| \leq t) = P(\text{signo}(X_1) = -1) \cdot P(|X_1| \leq t)$$

Luego, signo(X_i) es independiente de $|X_i|$ como se quería demostrar. $\square$

2.3.2 Pregunta 8

Corolario 2.3.1. *Bajo H_0, los vectores rango $\mathbf{R} = (R_1, \ldots, R_n)$ y antirrangos $\mathbf{D} = (D_1, \ldots, D_n)$ correspondientes a $|\mathbf{X}|$ son independientes del vector de signos $\mathbf{S} = (signo(X_1), \ldots, signo(X_n))$ de la muestra original $\mathbf{X}$.*

Proof. Llamemos $\mathbf{R} = (R_1, \ldots, R_n)$ al vector de rangos y $\mathbf{D} = (D_1, \ldots, D_n)$ al vector de antirrangos correspondientes a $|\mathbf{X}|$, mientras que $\mathbf{S} = (\text{signo}(X_1), \ldots, \text{signo}(X_n))$ es el vector de signos de la muestra original $\mathbf{X}$. En la Pregunta anterior demostramos que signo(X_i) es independiente de $|X_i|$, lo cual es válido para todo i debido a la independencia de las variables. Entonces, de forma equivalente, los vectores aleatorios $\mathbf{signo}(\mathbf{X}) = \mathbf{S}$ y $|\mathbf{X}|$ son independientes. Como los vectores $\mathbf{R}$ y $\mathbf{D}$ dependen únicamente de $|\mathbf{X}|$, resultan ser vectores aleatorios independientes del vector aleatorio $\mathbf{S}$. $\square$

2.3.3 Pregunta 9

Lema 2.4. *Bajo $H_0 : \theta = 0$, $F \in \Omega_s$, sean $W_j = \mathbb{1}\{X_{D_j} > 0\}$, entonces:*

$$W_1, \ldots, W_n \sim^{iid} Be(1/2).$$

Proof. Sea a_j, con $1 \leq j \leq n$, la realización de $W_j = \mathbb{1}\{X_{D_j} > 0\}$; por lo cual, $a_j \in \{0, 1\}$. Sea $\mathbf{d}$ un elemento de $S_n = \{(d_1, ..., d_n) : 1 \leq d_i \neq d_j \leq n\}$, el conjunto de todas las permutaciones posibles de índices. Utilizando probabilidad total:

$$P(W_1 = a_1, ..., W_n = a_n) = \sum_{\mathbf{d}:\mathbf{d}\in S_n} P(\mathbf{W} = \mathbf{a}|\mathbf{D} = \mathbf{d}) \cdot P(\mathbf{D} = \mathbf{d}) \tag{32}$$

$$= \sum_{\mathbf{d}:\mathbf{d}\in S_n} P[\mathbb{1}(X_{D_1} > 0) = a_1, ..., \mathbb{1}(X_{D_n} > 0) = a_n|D_1 = d_1, ..., D_n = d_n] \cdot P(\mathbf{D} = \mathbf{d}) \tag{33}$$

$$= \sum_{\mathbf{d}:\mathbf{d}\in S_n} P[\mathbb{1}(X_{d_1} > 0) = a_1, ..., \mathbb{1}(X_{d_n} > 0) = a_n] \cdot P(\mathbf{D} = \mathbf{d}) \tag{34}$$

El paso (34) es posible bajo H_0 por la independencia de las variables $\mathbf{X}$ y $\mathbf{D}$, debido a que el vector de antirrangos $\mathbf{D}$ depende del orden de las *magnitudes* de las X_i, y no de sus signos. Dado que $X_{d_1}, ..., X_{d_n}$ son independientes, y como $\mathbb{1}(X_{d_i} > 0)$ es una transformación que depende sólo del signo, se tiene:

$$P[\mathbb{1}(X_{d_i} > 0) = a_i] = \begin{cases} P(X_{d_i} > 0), \text{ si } a_i = 1, \\ P(X_{d_i} \leq 0), \text{ si } a_i = 0. \end{cases} \tag{35}$$

Bajo H_0 tenemos simetría de las X_i respecto a cero:

$$P(X_{d_i} > 0) = P(X_{d_i} \leq 0) = \frac{1}{2}; \tag{36}$$

por lo cual:

$$P[\mathbb{1}(X_{d_1} > 0) = a_1, ..., \mathbb{1}(X_{d_n} > 0) = a_n] = \prod_i^n P[\mathbb{1}(X_{d_i} > 0) = a_i] = \left(\frac{1}{2}\right)^n. \quad (37)$$

Reeemplazando (37) en (34):

$$P(W_1 = a_1, ..., W_n = a_n) = \frac{1}{2^n} \sum_{\mathbf{d}:\mathbf{d}\in S_n} P(\mathbf{D} = \mathbf{d}) \quad (38)$$

$$= \frac{1}{2^n}, \quad (39)$$

puesto que la última suma es sobre todos los valores posibles de $\mathbf{d}$, por lo cual vale 1. Con esto demostramos que $W_1, ..., W_n$ son independientes e idénticamente distribuidas como Be(1/2) bajo H_0. □

2.4 Distribución exacta de T^+

2.4.1 Pregunta 10

Bajo H_0, se tiene que para el soporte de T^+ ($[0, 1, \ldots, \frac{n(n+1)}{2}]$)

$$p_n(t) = \mathbb{P}(T^+ = t) = \frac{\#S_{n,t}}{2^n}$$

donde $S_{n,t} = \{A \subset \{1, \ldots, n\} : \sum_{a\in A} a = t\}$
Veamoslo para $n = 5$ y distintos valores t.

t	$S_{5,t}$	$\#S_{5,t}$	$p_5(t)$
0	{∅}	1	1/32
1	{{1}}	1	1/32
2	{{2}}	1	1/32
3	{{3}, {1,2}}	2	2/32
4	{{4}, {1,3}}	2	2/32
5	{{1,4}, {2,3}, {5}}	3	3/32
6	{{2,4}, {1,2,3}, {1,5}}	3	3/32
7	{{3,4}, {1,2,4}, {2,5}}	3	3/32
8	{{1,3,4}, {3.5}, {1,2,5}}	3	3/32
9	{{2,3,4}, {4,5}, {1,3,5}}	3	3/32
10	{{1,2,3,4}, {1,4,5}, {2,3,5}}	3	3/32

2.4.2 Pregunta 11

Lema 2.5. *T^+ es simétrica alrededor de $\frac{n(n+1)}{4}$*

Proof. Sea $u_n(t) = \#S_{n,t}$. Bajo exactamente el mismo razonamiento dado en el enunciado, se puede ver que:

$$u_n(t) = \begin{cases} 1 & \text{si } n = 0 \wedge t = 0 \\ 0 & \text{si } (n = 0 \wedge t \neq 0) \vee t < 0 \vee t > \frac{n(n+1)}{2} \\ u_{n-1}(t) + u_{n-1}(t-n) & \text{sino} \end{cases}$$

Observar que vale $\forall\, t \in \mathbb{R}$ pues, si t no es entero, en ningún subproblema ocurrirá que $t' = 0$ [...]. Luego, $p_n(t) = \dfrac{u_n(t)}{2^n}$.

Sea $\theta_n = \frac{n(n+1)}{4}$, veamos que $p_n(\theta_n + t) = p_n(\theta_n - t)\ \forall t \geq 0$

$$\Rightarrow \text{ basta ver que } u_n(\theta_n + t) = u_n(\theta_n - t)$$

Para ello hacemos inducción en n:

Si $n = 1$, $\theta_n = 1/2$

$$u_1(\theta_n + t) = u_1(1/2 + t) = \begin{cases} 1 & \text{si } t = 1/2\text{.. donde } S_{1,1} = \{\{1\}\} \\ 0 & \text{sino} \end{cases}$$

Del mismo modo

$$u_1(\theta_n - t) = u_1(1/2 - t) = \begin{cases} 1 & \text{si } t = 1/2\text{.. donde } S_{1,0} = \{\emptyset\} \\ 0 & \text{sino} \end{cases}$$

$$\Rightarrow u_1(\theta_n + t) = u_1(\theta_n - t)\ \forall t \geq 0$$

Supongamos que $u_{n-1}(\theta_n + t) = u_{n-1}(\theta_n - t)\ \forall t \geq 0$, veamos que vale para u_n

Notar que $\theta_n = \theta_{n-1} + \frac{n}{2}$.

Usando la fórmula recursiva

$$\begin{aligned} u_n(\theta_n + t) &= u_{n-1}(\theta_n + t) + u_{n-1}(\theta_n + t - n) \\ &= u_{n-1}(\theta_{n-1} + \frac{n}{2} + t) + u_{n-1}(\theta_{n-1} + \frac{n}{2} + t - n) \\ &= u_{n-1}(\theta_{n-1} + (\frac{n}{2} + t)) + u_{n-1}(\theta_{n-1} - (\frac{n}{2} - t))) \end{aligned}$$

Del mismo modo

$$u_n(\theta_n - t) = u_{n-1}(\theta_{n-1} + (\frac{n}{2} - t)) + u_{n-1}(\theta_{n-1} - (\frac{n}{2} + t))$$

Por H_1, como $(\frac{n}{2} + t)) \geq 0$ y $(\frac{n}{2} - t) \geq 0$ o bien $-(\frac{n}{2} - t) \geq 0$

$$\begin{aligned} \Rightarrow & u_{n-1}(\theta_{n-1} - (\frac{n}{2} + t)) = u_{n-1}(\theta_{n-1} + (\frac{n}{2} + t)) \\ & u_{n-1}(\theta_{n-1} + (\frac{n}{2} - t)) = u_{n-1}(\theta_{n-1} - (\frac{n}{2} - t))) \end{aligned}$$

$\Rightarrow u_n(\theta_n + t) = u_n(\theta_n - t)\ \forall t \geq 0$ como se quería ver.

Luego, T^+ es simétrica alrededor de $\frac{n(n+1)}{4}$ en términos de una variable aleatoria discreta. $\square$

2.5 Fórmula recursiva para n grande

2.5.1 Pregunta 12, 13, 14

En `R`.

2.6 Distribución asintótica

2.6.1 Pregunta 15

Bajo $H_0 : \theta = 0$, calculamos $\mathbb{E}(T^+)$ y $\mathrm{Var}(T^+)$

Para ello usamos que

- $T^+ = \sum_{j=1}^n j \times W_j$
- $W_j = \mathbb{1}\{X_{D_j} > 0\} \sim^{iid} Be(1/2)$ para cada j

$$\mathbb{E}(T^+) = \mathbb{E}(\sum_{j=1}^n j \times W_j) \tag{40}$$

$$= \sum_{j=1}^n j \times \mathbb{E}(W_j) \text{ (Linealidad de la esperanza)} \tag{41}$$

$$= \sum_{j=1}^n j\frac{1}{2} \tag{42}$$

$$= \frac{n(n+1)}{2}\frac{1}{2} \tag{43}$$

$$= \frac{n(n+1)}{4} \tag{44}$$

$$\begin{aligned}
\mathrm{Var}(T^+) &= \mathrm{Var}(\sum_{j=1}^{n} j \times W_j) && (45)\\
&= \sum_{j=1}^{n} \mathrm{Var}(j \times W_j) + 2\sum_{i\neq j} \mathrm{Cov}(j \times W_j, i \times W_i) && (46)\\
&= \sum_{j=1}^{n} j^2 \mathrm{Var}(W_j) + 2\sum_{i\neq j} ji\mathrm{Cov}(W_j, W_i) \text{ (Prop de la covarianza)} && (47)\\
&= \sum_{j=1}^{n} j^2 \frac{1}{4} + 2\sum_{i\neq j} ji0 \text{ (Independencia implica covarianza 0)} && (48)\\
&= \sum_{j=1}^{n} j^2 \frac{1}{4} && (49)\\
&= \frac{n(n+1)(2n+1)}{6}\frac{1}{4} && (50)\\
&= \frac{n(n+1)(2n+1)}{24} && (51)
\end{aligned}$$

2.6.2 Pregunta 16

Sea $V_j = W_j - \frac{1}{2}$ para cada j, entonces:

- $\mathbb{E}(V_j) = 0$ para cada j
- $\sigma^2 = \mathrm{Var}(V_j) = \mathrm{Var}(W_j) = \frac{1}{4}$

Sean $S = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} j \times V_j$ y $a_j = j$, claramente:

$$\begin{aligned}
\frac{\max_j |a_j|}{\sqrt{\sum_j^n a_j^2}} &= \frac{\max_j j}{\sqrt{\sum_j^n j^2}} && (52)\\
&= \frac{n}{\sqrt{\frac{n(n+1)(2n+1)}{6}}} && (53)\\
&= \sqrt{\frac{6n}{(n+1)(2n+1)}} \to_{n\to\infty} 0 && (54)
\end{aligned}$$

Entonces, por TCL de Lindeberg:

$$\frac{S}{\sqrt{\mathrm{Var}(S)}} \to^{\mathcal{D}} \mathcal{N}(0,1)$$

donde

$$\begin{aligned}
\mathrm{Var}(S) &= \frac{\sigma^2}{n}\sum_{j}^{n} a_j^2 && (55)\\
&= \frac{\sigma^2}{n}\frac{4}{4}\sum_{j}^{n} j^2 && (56)\\
&= \frac{\sigma^2}{n} 4\mathrm{Var}_n(T^+) && (57)\\
&= \frac{1/4}{n} 4\mathrm{Var}_n(T^+) && (58)\\
&= \frac{\mathrm{Var}_n(T^+)}{n}. && (59)
\end{aligned}$$

Despejamos T^+:

$$\begin{aligned}
S &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} j \times (W_j - \frac{1}{2}) && (60)\\
&= \frac{1}{\sqrt{n}}(T^+ - \sum_{j=1}^{n} j\frac{1}{2}) && (61)\\
&= \frac{1}{\sqrt{n}}(T^+ - \mathbb{E}_n(T^+)). && (62)
\end{aligned}$$

Finalmente, se obtiene la distribución asintótica de T^+

$$\begin{aligned}
\frac{S}{\sqrt{\mathrm{Var}_n(S)}} &= \frac{\frac{1}{\sqrt{n}}(T^+ - \mathbb{E}_n(T^+))}{\sqrt{\frac{\mathrm{Var}_n(T^+)}{n}}} && (63)\\
&= \frac{T^+ - \mathbb{E}_n(T^+)}{\sqrt{\mathrm{Var}_n(T^+)}} \to^{\mathcal{D}} \mathcal{N}(0,1). && (64)
\end{aligned}$$

Por lo cual, para n lo suficientemente grande:

$$T^+ \simeq \mathcal{N}(\mathbb{E}_n(T^+), \mathrm{Var}_n(T^+)). \tag{65}$$

2.6.3 Pregunta 17

En la Figura 1 se muestran la densidad de probabilidad puntual exacta de T^+ (barras grises) obtenida con `dTmas`, junto con la densidad asintótica (65) (línea roja), para $n \in \{4, 10, 20\}$, donde los valores para $\mathbb{E}_n(T^+)$ y $\mathbb{V}_n(T^+)$ se obtuvieron a través de (44) y (51), respectivamente. Puede observarse que ambas densidades coinciden razonablemente en toda la distribución para todo n, aunque a medida que n decrece la densidad asintótica parece subestimar la densidad exacta hacia las colas de la distribución. Una posible solución sería utilizar una distribución t de Student, ya que tiene colas más pesadas que la normal.

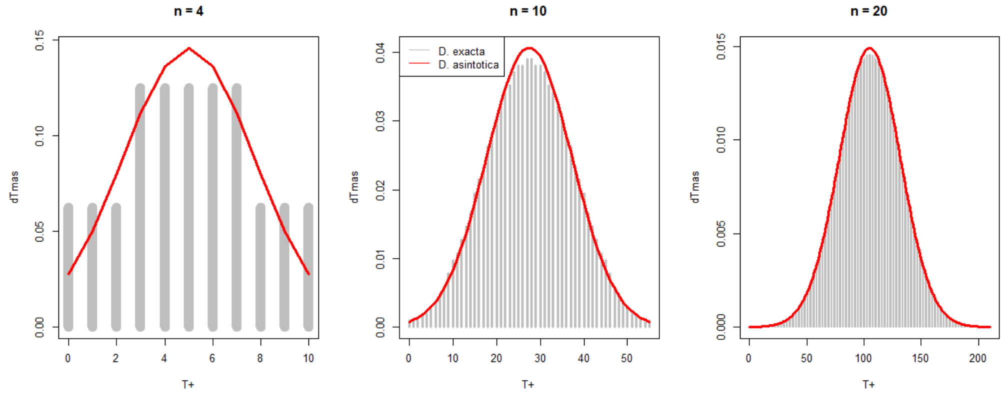

Figure 1: Densidad de probabilidad puntual exacta (barras grises) y asintótica (línea roja) para T^+.

2.7 Distribución bajo la alternativa vía bootstrap

2.7.1 Pregunta 18

En R.

2.7.2 Pregunta 19

Sean $X_1, \ldots, X_n \sim \mathcal{N}(\mu, 1)$. Se quiere testear:

$$H_0 : \mu = 0 \text{ vs. } H_1 : \mu > 0.$$

El test UMP de nivel $\alpha = 0.05$ para estas hipótesis es:

$$\phi_n = \begin{cases} 1 & \text{si } \sqrt{n}\,\overline{X}_n \geq 1.645 \\ 0 & \text{si } \sqrt{n}\,\overline{X}_n < 1.645. \end{cases}$$

La función de potencia para el test ϕ_n evaluada en $\theta_1 = 1$ es:

$$\pi_{\phi_n}(\theta_1) = \pi_{\phi_n}(1) = 1 - \Phi(1.645 - \sqrt{n}). \tag{66}$$

El resto del ejercicio se encuentra en R.