Cerdeira_Diiuorio_TP4

2025-05-28

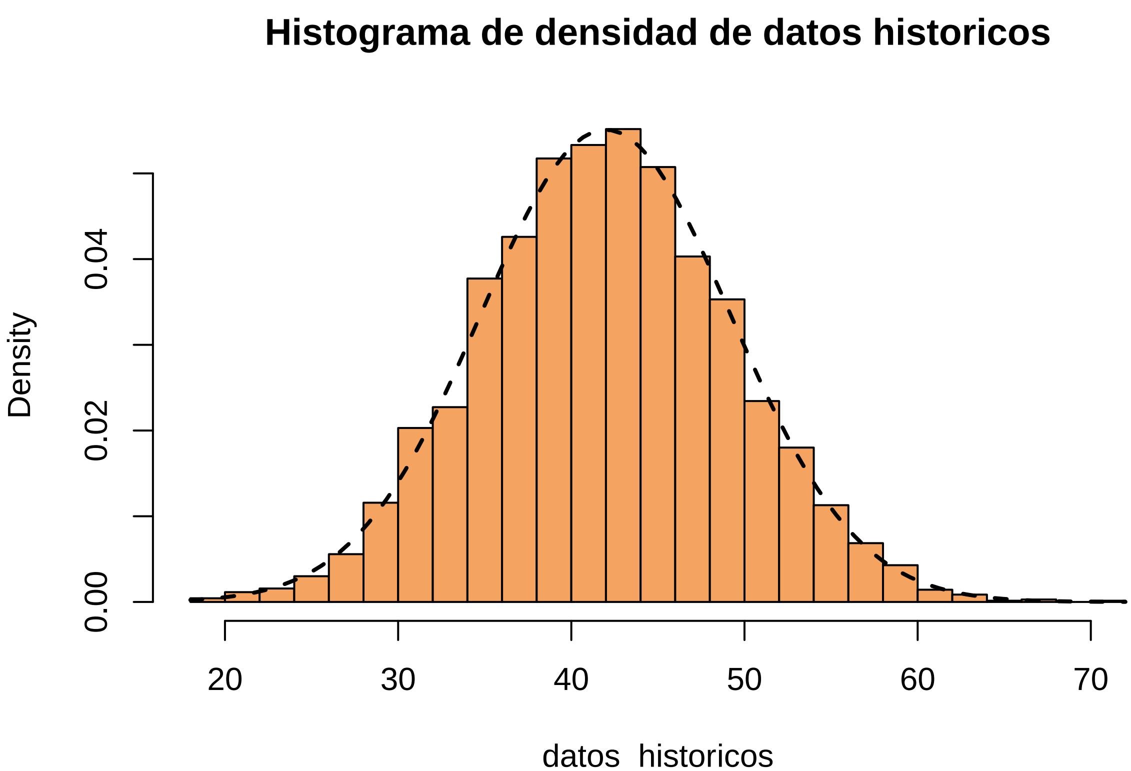
Ejercicio 1

(A)

La **media muestral** de los datos historicos es **41.98**.

La **varianza muestral** de los datos historicos es **52.43**.

(B)



Los datos parecen distribuirse de forma normal.

Ejercicio 2

Primero definimos:

- μ_0 = media del "play-delay" de la versión anterior = 41.9847
- σ_0^2 = varianza del "play-delay" de la versión anterior = 52.4325

Ahora definimos:

$$X_i = \text{El "play-delay" del i-ésimo usuario evaluado con la nueva versión. } 1 \leq i \leq 200$$

Y contamos con la siguiente distribución de nuestra muestra:

$$X_1, ..., X_{200} \sim \mathcal{N}(\mu, 52.43) \text{ iid}$$

Donde:

$$\mu = \text{media del "play-delay"}$$

Para estimar μ vamos calcular la **media muestral** de la muestra con los 200 nuevos usuarios.

El estimador de μ es:

$$\hat{\mu} = \overline{X}$$

Según la muestra que tenemos, la estimación de μ es:

$$\hat{\mu}_{obs} = \overline{X}_{obs} = 42.99$$

Ejercicio 3

(A)

Proponemos las siguientes Hipótesis Nula e Hipótesis Alternativa

$$H_0 : \mu \leq \mu_0 \quad vs. \quad H_1 : \mu > \mu_0$$

Es decir,

$$H_0 : \mu \leq 41.98 \quad \text{vs.} \quad H_1 : \mu > 41.98$$

Para simplificar el problema, suponemos:

$$H_0 : \mu = 41.98$$

Esto se puede hacer ya que el caso límite de H_0 es cuando $\mu = 41.98$. Por lo tanto, si puedo rechazar esa H_0 entonces puedo rechazar también todos los valores de

$$\mu \leq 41.98$$

Por ende, las hipótesis quedarían así:

$$H_0 : \mu = 41.98 \quad \text{vs.} \quad H_1 : \mu > 41.98$$

Estadístico:

Para hallar el estadístico partimos del estimador de μ:

$$\hat{\mu} = \overline{X}$$

Asumimos distribución de $\overline{X}$ bajo H_0:

$$\overline{X} \sim \mathcal{N}(41.98, \frac{52.43}{200}) \quad \text{bajo } H_0$$

Lo estandarizamos:

$$T = \frac{\overline{X} - 41.98}{\sqrt{52.43/200}} \sim \mathcal{N}(0,1) \quad \text{bajo } H_0$$

T es nuestro estadístico. Nos sirve porque es una función de la muestra aleatoria y tiene distribución conocida bajo H_0.

(B)

$$\alpha = 0.05$$

Vamos a rechazar H_0 cuando $\overline{X} > C$ donde C es un número tal que:

$$\alpha = P(EI) = P(\text{Rechazar } H_0|H_0 \text{ es V}) = P(\overline{X} > C|H_0 \text{ es V})$$
$$\implies \text{Queremos C tal que } P(\overline{X} > C|H_0 \text{ es V}) = 0.05$$

$$P(\overline{X} > C|H_0 \text{ es V}) = P(\frac{\overline{X} - 41.98}{\sqrt{52.43/200}} > \frac{C - 41.98}{\sqrt{52.43/200}}|H_0 \text{ es V })$$

$$\textit{como} \quad \frac{\overline{X} - 41.98}{\sqrt{52.43/200}} \sim \mathcal{N}(0,1) \quad \text{bajo } H_0$$

Entonces lo llamamos Z ya que

$$Z \sim \mathcal{N}(0,1)$$

Por lo tanto,

$$P(\frac{\overline{X} - 41.98}{\sqrt{52.43/200}} > \frac{C - 41.98}{\sqrt{52.43/200}}|H_0 \text{ es V }) = P(Z > q)$$

donde

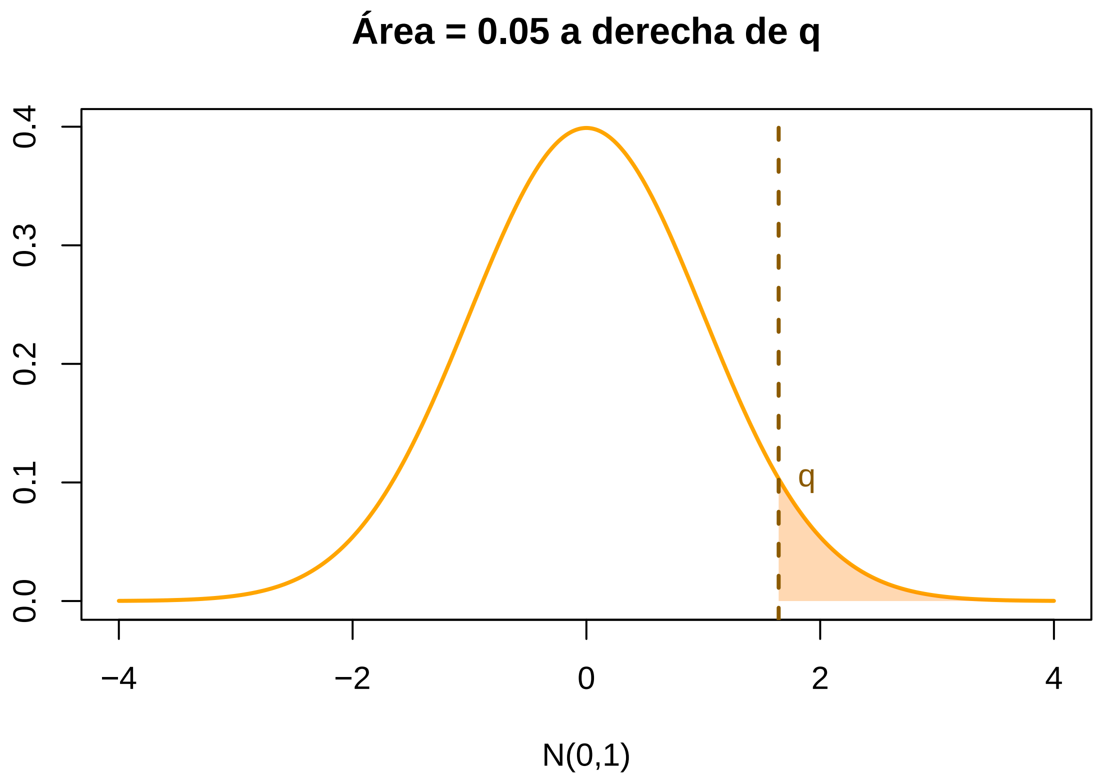
$$q = \frac{C - 41.98}{\sqrt{52.43/200}}$$

Según la Hipótesis Alternativa, decidimos rechazar H_0 cuando:

$$T = \frac{\overline{X} - 41.98}{\sqrt{52.43/200}} > q$$

$$\text{Queremos } q/P(Z > q) = \alpha = 0.05$$

```r
alpha <- 0.05
q <- qnorm(1-alpha, mean = 0, sd = 1)
```



Entonces, q es un cuantil tal que me deja área $\alpha = 0.05$ a derecha o área 1 - $\alpha = 0.95$ a izquierda .

$$\text{Por lo tanto } q = Z_{\alpha} = Z_{0.05}$$

Nuestra Región de Rechazo quedaría

$$R = \{T > Z_{0.05}\}$$

Por la tabla de una Normal Estándar, podemos ver que

$$q = Z_{0.05} = 1.64$$

Como resultado, la Región de Rechazo para este test es:

$$R = \{T > 1.64\}$$

Ejercicio 4

(A)

Para utilizar el test construido, el primer paso es calcular el promedio de los valores de la muestra "datos nueva versión". Una vez calculado, evaluamos nuestro estadístico observado. Si este pertenece en la Región de Rechazo, rechazamos H_0. En el caso de no pertenecer a la región, no confirmamos que H_0 sea verdadera sino que no tenemos la suficiente evidencia para rechazarla.

Como vimos antes:

$$\overline{X}_{obs} = 42.99$$

Ahora, buscamos al estadístico observado de esta muestra.

$$T_{obs} = \frac{\overline{X}_{obs} - 41.98}{\sqrt{52.43/200}}$$

$$T_{obs} = \frac{42.99 - 41.98}{\sqrt{52.43/200}} = 1.97$$

Mi estadístico observado es $T_{obs} = 1.97$

Como podemos ver, el estadístico observado pertenece a la región de rechazo. Tenemos suficiente evidencia para rechazar H_0 porque $1.97 > 1.64$

Con esta evidencia rechazamos la actualización ya que aumenta el "play-delay", es decir, la esperanza del "play-delay" con la versión nueva es mayor a la actual, y es por eso que enviamos el código a revisión.

(B)

α=0.01

$$P(EI) = P(\text{Rechazar } H_0 | H_0 \text{ es V}) = 0.01$$

La Región de Rechazo con $\alpha = 0.01$ será:

$$R = \{T > Z_\alpha\} = \{T > Z_{0.01}\} = \{T > 2.33\}$$

El cuantil $Z_{0.01}$ deja menos área a la derecha que el cuantil $Z_{0.05}$. Esto implica que el primero está gráficamente a derecha del segundo ($Z_{0.05} < Z_{0.01}$).

α=0.1

$$P(EI) = P(\text{Rechazar } H_0 | H_0 \text{ es V}) = 0.1$$

La Región de Rechazo con $\alpha = 0.1$ será:

$$R = \{T > Z_\alpha\} = \{T > Z_{0.1}\} = \{T > 1.28\}$$

El cuantil $Z_{0.1}$ deja más área a la derecha que el cuantil $Z_{0.05}$. Esto implica que el primero está gráficamente a izquierda del segundo ($Z_{0.1} < Z_{0.05}$).

El α desplaza el punto de corte para la Región de Rechazo. A mayor α, menor es el punto de corte del cual arranca la Región de Rechazo. Aplicado a este test, a mayor α, mando más seguido actualizaciones a auditar, ya que el estadístico observado puede ser menor para rechazar H_0 y decir que el "play-delay" de la nueva versión es mayor al actual.

(C)

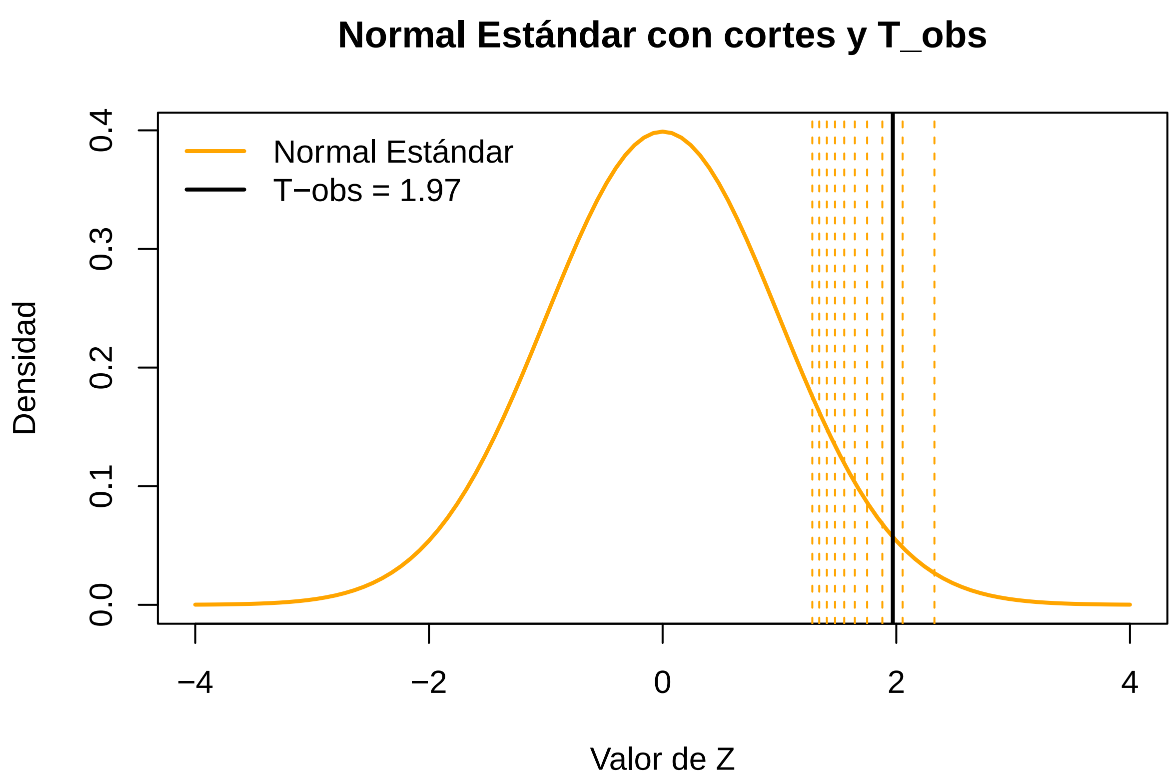
Para cada α, calculamos el Z_α: el punto que marca el comienzo de la región de rechazo.

$$R = \{T > Z_\alpha\}$$

Si el **T observado (1.97)** es mayor al Z_α para un α dado, entonces tendremos evidencia suficiente a nivel α para rechazar H_0.

```
##    alpha z_alpha      decision
## 1   0.01   2.326 No Rechazo H0
## 2   0.02   2.054 No Rechazo H0
## 3   0.03   1.881    Rechazo H0
## 4   0.04   1.751    Rechazo H0
## 5   0.05   1.645    Rechazo H0
## 6   0.06   1.555    Rechazo H0
## 7   0.07   1.476    Rechazo H0
## 8   0.08   1.405    Rechazo H0
## 9   0.09   1.341    Rechazo H0
## 10  0.10   1.282    Rechazo H0
```

Se puede ver que para un α **igual o mayor a 0.03** ya puedo rechazar H_0. Es útil ver estos resultados de forma grafica:



Estos resultados tienen sentido: cuánto más chico es el nivel de significancia α más difícil es rechazar H_0. Esto se puede apreciar en el gráfico: los "Z críticos" (Z_α) de los niveles de significancia más chicos (0.01 y 0.02) son las líneas punteadas a la derecha del T-observado. Esto quiere decir que el valor que observamos no es "tan extremo" cómo para darnos suficiente evidencia que nos permita rechazar H_0 en niveles de significancia 0.01 o 0.02.

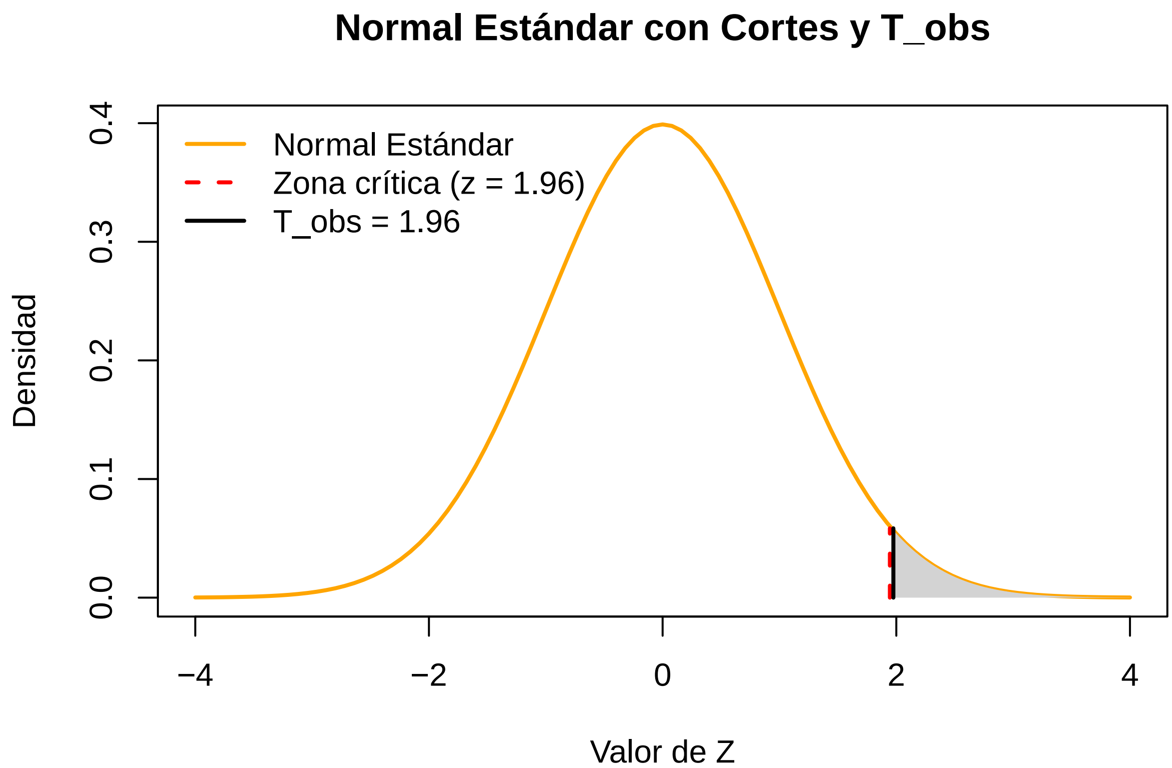
(D)

Ahora queremos encontrar el valor minimo de α para el cuál podriamos rechazar H_0 con los datos observados. Del inciso anterior, sabemos que éste valor minimo tiene que estar entre $\alpha = 0.02$ y $\alpha = 0.03$. Vamos a armar un for-loop que vaya probando todos los numeros decimales entre 0.02 y 0.03 con 3 digitos decimales y mostrar los resultados en una tabla:

```
##    alpha z_alpha      decision
## 1  0.020   2.054 No Rechazo H0
## 2  0.021   2.034 No Rechazo H0
## 3  0.022   2.014 No Rechazo H0
## 4  0.023   1.995 No Rechazo H0
## 5  0.024   1.977 No Rechazo H0
## 6  0.025   1.960    Rechazo H0
## 7  0.026   1.943    Rechazo H0
## 8  0.027   1.927    Rechazo H0
## 9  0.028   1.911    Rechazo H0
## 10 0.029   1.896    Rechazo H0
## 11 0.030   1.881    Rechazo H0
```

Vemos que el valor minimo de α para el cuál podemos rechazar H_0 es **0.025**. Esto tiene sentido si vemos que el "Z alpha" para éste nivel de significancia ($Z_{0.025}$) es igual a nuestro T-observado: **1.96**. Esto no podría ser de otra forma: el nivel mínimo de α para el cuál puedo rechazar H_0 tiene que ser igual al T-observado ya que si fuese más grande no podríamos rechazar H_0, y sí fuese más chico no sería el valor mínimo.

0.025 será lo que mida el área a la derecha del T-observado (área sombreada en el gráfico):



Es muy importante notar que el área sombreada es el p-valor. Esto tiene sentido si recordamos que el p-valor es la "probabilidad de observar valores iguales o más extremos a lo observados (en direccion a H_1), asumiendo que H_0 es verdadera". Gráficamente, el p-valor es el área a derecha del T-obs (en éste caso dónde $H_1 : \mu > \mu_0$), en otros casos podria ser el área a la izquierda, o el área en ambos extremos).

Entonces, podemos ver que existe otra manera de pensar al p-valor: "el p-valor es el nivel de significancia del α mínimo, para el cual puedo rechazar H_0 con los datos observados".

Ejercicio 5

(A)

En este ejercicio vamos a **simular** una muestra de tamaño $n = 200$ asumiendo que la hipótesis nula $H_0 : \mu = \mu_0$ es verdadera, es decir, que la nueva versión **no** modificó la media de Play-Delay con respecto a los datos históricos.

Los datos de nuestra muestra tendrán la siguiente distribución:

$$X_1, ..., X_{200} \sim N(41.98, 52.48)$$

Simulamos la muestra en R:

```
set.seed(1234)
muestra_sim <- rnorm(200, mean=media_muestral, sd= sd0)
```

Calculamos la media muestral:

```r
xbar_sim = mean(muestra_sim)
print(xbar_sim)
```

```
## [1] 41.56182
```

$$\overline{X} = 41.56$$

Calculamos el T-Observado:

```r
Tobs_sim <- (xbar_sim - media_muestral) / (sd0/sqrt(200))
print(Tobs_sim)
```

```
## [1] -0.8168396
```

$$T_{obs} = \frac{41.56 - 41.98}{\sqrt{52.43/200}} = -0.81$$

En el punto 3.b vimos que la región de rechazo para éste test con un nivel $\alpha = 0.05$ es:

$$R = \{T > 1.64\}$$

Entonces, cómo el T-observado **(-0.81)** es menor que 1.64, **NO tenemos evidencia suficiente para rechazar** H_0 con los datos observados.

Esta es la decisión correcta ya que estamos asumiendo que H_0 es verdadera. Si, por casualidad, el T-Observado hubiera sido mayor que 1.64, habríamos cometido un **Error de Tipo 1** al rechazar H_0.

(B)

Ahora vamos a repetir el procedimiento anterior **10000 veces**. Vamos a calcular la proporcion de veces que no rechazo H_0 y la proporcion de veces que sí lo rechazo (proporcion de veces que cometo un Error de Tipo 1).

Como el test que estamos usando tiene un nivel $\alpha = 0.05$, esperamos equivocarnos (rechazar H_0) sólo el **5%** de las veces y acertar el **95%** de las veces.

Armamos la simulacion con R e imprimimos los resultados:

```
## De 10000 simulaciones:

##   • Porcentaje de veces que RECHAZO H0: 4.74 %

##   • Porcentaje de veces que NO RECHAZO H0: 95.26 %
```

Vemos que los resultados que obtuvimos son lo que esperabamos: nos equivocamos casi un 5% porciento de las veces y acertamos el 95% de las veces.

(C)

El nivel de significancia α es la probabilidad de cometer un Error de Tipo 1. Es decir, α se define como:

$$\alpha = P(\text{Rechazar } H_0 | H_0 \text{ Verdadera})$$

Al construir un test de hipótesis, uno de los pasos fundamentales es establecer un nivel de significancia α, que debe elegirse en función del contexto del problema. Este nivel representa la probabilidad máxima tolerada de cometer un Error de Tipo 1, es decir, de rechazar la hipótesis nula H_0 cuando en realidad es verdadera. Un valor de α más bajo implica que somos más conservadores: disminuye la chance de falsos positivos, pero al mismo tiempo aumenta la dificultad para rechazar H_0, lo cual puede reducir el poder del test.

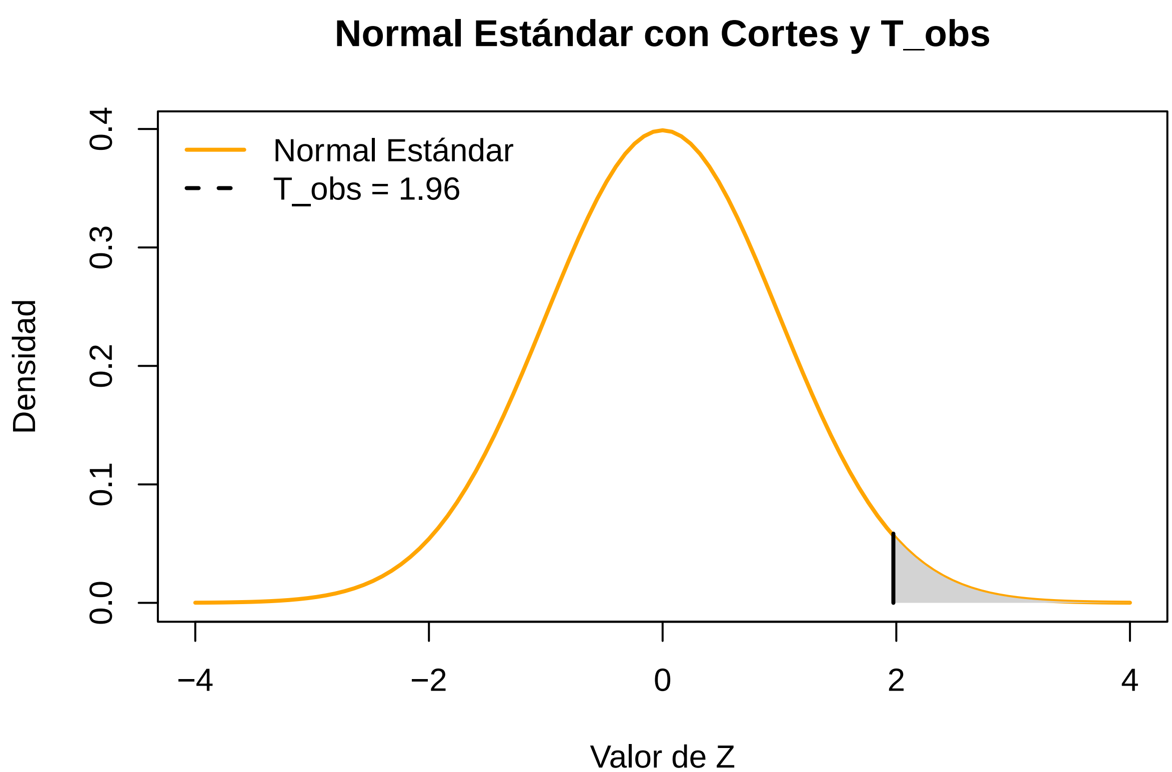
En nuestro caso particular, estamos asumiendo que H_0 es verdadera: es decir, la nueva versión de la plataforma no modificó el valor promedio del Play-Delay. Por lo tanto, si el test rechaza H_0 bajo esta suposición, estaríamos cometiendo un Error de Tipo 1.

El test que usamos fue diseñado con un nivel de significancia $\alpha = 0.05$, que significa que esperamos que aproximadamente el 5% de los tests conduzcan a un rechazo incorrecto de H_0 cuando esta es verdadera (que es lo que observamos en el punto anterior).

Ejercicio 6

(A)

Si asumimos que H_0 es verdadera, podemos calcular la probabilidad de observar un resultado cómo el observado o más extremo midiendo el área a la derecha del T-observado.



```r
p_valor <- 1 - pnorm(1.96)
print(p_valor)
```

```
## [1] 0.0249979
```

$$\text{p-valor} = 0.25$$

(B)

Vemos que el p-valor que calculamos es el mismo que habiamos obtenido en el punto 4.d. Como explicamos en ése punto, esto se debe a que el p-valor es el minimo nivel de significancia α para el que rechazamos H_0 con los datos observados.

Por lo tanto, con conocer el p-valor de un test nos alcanza para tomar la decision del test:

- Si p-valor $\leq \alpha$: rechazo H_0
- Si p-valor $> \alpha$: NO rechazo H_0

Ejercicio 7

(A)

H_0: "Los datos provienen de una distribución normal".

H_1: "Los datos *no* provienen de una distribución normal".

$$\begin{gathered} H_0 : X_i \sim \mathcal{N}(\mu, \sigma^2) \quad \forall i : 1 \leq i \leq 200 \\ vs. \\ H_1 : X_i \nsim \mathcal{N}(\mu, \sigma^2) \quad \text{para al menos un } i \end{gathered}$$

(B)

```r
print(shapiro.test(datos_historicos))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  datos_historicos
## W = 0.99958, p-value = 0.6934
```

```r
print(shapiro.test(datos_nueva_version))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  datos_nueva_version
## W = 0.99587, p-value = 0.867
```

Como no se especifica el valor de α para estos test, se asume que $\alpha = 0.05$.

- Si el p-valor $< 0.05 \implies$ Rechazo $H_0 \implies$ los datos no son normales.
- Si el p-valor $\geq 0.05 \implies$ No rechazo $H_0 \implies$ los datos son compatibles con una normal.

El p-valor en los datos históricos es 0.6934.

El p-valor en los datos de la nueva versión es 0.867.

En ambos casos el p-valor $\geq \alpha = 0.05$. Por lo tanto no rechazo H_0 ni para los datos históricos ni para los datos de la nueva versión. No tenemos evidencia para sostener que los datos NO provienen de una Normal.