Resolución de la Guía de Estimación de Parámetros

Ejercicio 1 (estimación de la media)

X v.a.c.: Duración de los focos (hs)
n=30 (grande), $\bar{x} = 780$, $\sigma = 40$ $(\sigma^2\ conocido)$, población aproximadamente normal, $1 - \alpha = 0.96$
Como la muestra es grande utilizaremos la Normal Estándar (Z).
Lo haremos "a mano" ya que los datos están resumidos (no contamos con el vector de observaciones).

$$\bar{x} \pm z_{\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}}$$

```
# EJERCICIO 1
media=780
n=30
sd=40
alfa=0.04
media-qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim inf
media+qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim sup
```

Respuesta: $765.0015 < \mu < 794.9985$

Ejercicio 2 (estimación de la media)

X v.a.c.: Estatura de estudiantes de la carrera LSI (cm)
n=50 (grande), $\bar{x} = 174.5$, $s = 6.9$ $(\sigma^2\ desconocido)$, $1 - \alpha = 0.98$
Como la muestra es grande utilizaremos la Normal Estándar (Z).
Lo haremos "a mano" ya que los datos están resumidos (no contamos con el vector de observaciones).

$$\bar{x} \pm z_{\frac{\alpha}{2}}.\frac{s}{\sqrt{n}}$$

```
# EJERCICIO 2
media=174.5
n=50
sd=6.9
alfa=0.02
media-qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim inf
media+qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim sup
```

Respuesta: a) $172.23 < \mu < 176.77$

b) El error máximo de estimación es $E = \frac{\text{lim sup} - \text{lim inf}}{2} = 2.27$

c)

```
# Repetir el IC al 90%
alfa=0.10
media-qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim inf
media+qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim sup
# Repetir el IC al 95%
alfa=0.05
media-qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim inf
media+qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim sup
# Repetir el IC al 99%
alfa=0.01
media-qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim inf
```

```
media+qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n) #lim sup
```

IC al 90%: $172.89 < \mu < 176.11$ en el cual $E = \frac{\text{lim sup} - \text{lim inf}}{2} = 1.61$

IC al 95%: $172.59 < \mu < 176.41$ en el cual $E = \frac{\text{lim sup} - \text{lim inf}}{2} = 1.91$

IC al 99%: $171.99 < \mu < 177.01$ en el cual $E = \frac{\text{lim sup} - \text{lim inf}}{2} = 2.51$

d) Cuanto mayor es el nivel de confianza, mayor es el tamaño del error máximo de estimación.

Ejercicio 3 (estimación de la media)

X v.a.c.: Distancia recorrida por autos en la ciudad de Paraná en un año (km)
n=100 (grande), $\bar{x} = 23500$, $s = 3900$ (σ^2 $desconocido$), $1 - \alpha = 0.99$
Como la muestra es grande utilizaremos la Normal Estándar (Z).
Lo haremos "a mano" ya que los datos están resumidos (no contamos con el vector de observaciones).

$$\bar{x} \pm z_{\frac{\alpha}{2}} . \frac{s}{\sqrt{n}}$$

```
# EJERCICIO 3
media=23500
n=100
sd=3900
alfa=0.01
media-qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n)
media+qnorm(alfa/2,lower.tail=FALSE)*sd/sqrt(n)
```

Respuesta: a) $22495.43 < \mu < 24504.57$

b) El error máximo de estimación es $E = \frac{\text{lim sup} - \text{lim inf}}{2} = 1004.57$

Ejercicio 4 (estimación de la media, cálculo de tamaño de muestra)

Se utilizó en ejercicio 1 la siguiente fórmula, siendo:

$$E = z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}} \rightarrow \quad n = \left(\frac{z_{\frac{\alpha}{2}} . \sigma}{E} \right)^2$$

Despejamos n para un E=10 y al 96% de confianza.

$$10 = z_{0.02} . \frac{40}{\sqrt{n}} \rightarrow \quad 10 = 2.05 \, . \frac{40}{\sqrt{n}} \rightarrow \quad n = \left[\frac{(2.05). \, 40}{10} \right]^2 \rightarrow n = 68$$

El valor de z se lo puede obtener de R Studio con: qnorm(0.02, lower.tail=FALSE). O bien hacer el cálculo completo de n con la sintaxis dispuesta debajo.
Siempre el resultado se redondea al entero siguiente.

```
#EJERCICIO 4
# tamaño de muestra para E=10
error=10
alfa=0.04
sd=40
(qnorm(alfa/2,lower.tail=FALSE)*sd/error)^2
```

Respuesta: $n = 68$

Ejercicio 5 (estimación de la media, cálculo de tamaño de muestra)

X v.a.c.: Tiempo en perforar una placa metálica (segundos)
a) $1-\ \alpha = 0.95,\ E = 12, \quad \sigma = 40,\ \ n = ?$
Se parte de esta fórmula:

$$E = z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}} \ \rightarrow \ n = \left(\frac{z_{\frac{\alpha}{2}} . \sigma}{E} \right)^2$$

Despejamos n para un E=12 y al 95% de confianza.

$$12 = z_{0.025} . \frac{40}{\sqrt{n}} \rightarrow \ 12 = 1.96 \ . \frac{40}{\sqrt{n}} \rightarrow \ \ n = \left[\frac{(1.96). 40}{12} \right]^2 \rightarrow n = 43$$

El valor de z se lo puede obtener de R Studio con: qnorm(0.025, lower.tail=FALSE). O bien hacer el cálculo completo de n con la sintaxis dispuesta debajo.
Siempre el resultado se redondea al entero siguiente.

```
#EJERCICIO 5
# tamaño de muestra para E=12
error=12
alfa=0.05
sd=40
(qnorm(alfa/2,lower.tail=FALSE)*sd/error)^2
```

Respuesta: $n = 43$

b) Se repite para el procedimiento para alcanzar $\frac{1}{2}E = 6.$

```
# tamaño de muestra para E=12/2=6
error=6
alfa=0.05
sd=40
(qnorm(alfa/2,lower.tail=FALSE)*sd/error)^2
```

Respuesta: $n = 171$

Ejercicio 6 (estimación de la media)

X v.a.c.: Contenido de azúcar en barras de cereal (gr)
n=20 (chico), $\bar{x} = 11.3$, $s = 2.45$ (σ^2 *desconocido*), población normal, $1 - \alpha = 0.95$
Como la muestra es chica, la varianza poblacional desconocida y la población normal, utilizaremos la T de Student.
Lo haremos "a mano" ya que los datos están resumidos (no contamos con el vector de observaciones).

$$\bar{x} \pm t_{\frac{\alpha}{2}, \nu = n-1} . \frac{s}{\sqrt{n}}$$

```
# EJERCICIO 6
n=20
media=11.3
sd=2.45
alfa=0.05
media-qt(alfa/2,df=n-1,lower.tail = FALSE)*sd/sqrt(n)
media+qt(alfa/2,df=n-1,lower.tail = FALSE)*sd/sqrt(n)
```

Respuesta: $10.15 < \mu < 12.45$

Ejercicio 7 (estimación de la media)

X v.a.c.: Diámetros de piezas metálicas de forma cilíndrica (cm)
n=9 (chico), población normal, $1-\alpha=0.99$
Como la muestra es chica, la varianza poblacional desconocida y la población normal, utilizaremos la T de Student.
Al contar con un vector de observaciones, se puede realizar directo con la función t.test(…)$conf.int

```
# EJERCICIO 7
diametros=c(1.01,0.97,1.03,1.04,0.99,0.98,0.99,1.01,1.03)
t.test(diametros,conf.level=0.99)$conf.int
```

Respuesta: $0.978<\mu<1.033$

Ejercicio 8 (estimación de la media)

X v.a.c.: Tiempos de secado de cierta pintura vinílica (hs)
n=15 (chico), población normal, $1-\alpha=0.95$
Como la muestra es chica, la varianza poblacional desconocida y la población normal, utilizaremos la T de Student.
Al contar con un vector de observaciones, se puede realizar directo con la función t.test(…)$conf.int

```
# EJERCICIO 8
tiempos=c(3.4,2.5,4.8,2.9,3.6,2.8,3.3,5.6,
          4.4,4.0, 5.2,3.0,4.8,3.7,2.8)
t.test(tiempos,conf.level = 0.95)$conf.int
```

Respuesta: $3.25<\mu<4.32$

Ejercicio 9 (estimación de la media)

X v.a.d.: Número de caracteres tipeados por minuto por graduados de la carrera
n=12 (chico), $\bar{x}=79.3$, $s=7.8$ (σ^2 $desconocido$), población normal, $1-\alpha=0.95$
Como la muestra es chica, la varianza poblacional desconocida y la población normal, utilizaremos la T de Student.
Lo haremos “a mano” ya que los datos están resumidos (no contamos con el vector de observaciones).

$$\bar{x}\pm t_{\frac{\alpha}{2},\nu=n-1}\cdot\frac{s}{\sqrt{n}}$$

```
# EJERCICIO 9
79.3-qt(0.025,df=11,lower.tail=FALSE)*7.8/sqrt(12) #lim inf
79.3+qt(0.025,df=11,lower.tail=FALSE)*7.8/sqrt(12) #lim sup
```

Respuesta: $74.34<\mu<84.26$

Ejercicio 10 (estimación de la diferencia de medias)

Anulado de esta guía, sólo trabajamos con ejercicios de diferencia de medias cuando se cuenta con vectores de observaciones.

Ejercicio 11 (estimación de la diferencia de medias)

Anulado de esta guía, sólo trabajamos con ejercicios de diferencia de medias cuando se cuenta con vectores de observaciones.

Ejercicio 12 (estimación de la diferencia de medias)

Definimos las variables según el orden en el cual queremos la diferencia (NOGOPAINT-ALBA)
X_1 v.a.c.: Tiempo de secado de pintura Nogopaint (hs)
X_2 v.a.c.: Tiempo de secado de pintura Alba (hs)
$n_1 = 15$ (chico), $n_2 = 15$ (chico), $\sigma_1^2\ y\ \sigma_2^2\ \ desconocidas$ e IGUALES, poblaciones normales
$1 - \alpha = 0.95$
Como los tamaños muestrales son chicos, las varianzas poblacionales son desconocidas y las poblaciones subyacentes son normales, utilizaremos la T de Student.
Al contar con dos vectores de observaciones, se puede realizar directo con la función t.test(…)$conf.int

```
# EJERCICIO 12
nogopaint=c(4.7, 3.9, 4.5, 5.5, 4.0, 5.3, 4.3,
            6.0, 5.2, 3.7, 5.5, 6.2, 5.1, 5.4, 4.8)
alba=c(3.5, 2.7, 3.9, 4.2, 3.6, 2.7, 3.3, 5.2,
       4.2, 2.9, 4.4, 5.2, 4.0, 4.1, 3.4)
t.test(nogopaint,alba, var.equal=TRUE, conf.level = 0.95)$conf.int
```

Respuesta: $0.55 < \mu_1 - \mu_2 < 1.69$ Con un 95% de confianza podemos concluir que existe diferencia significativa entre las medias (ya que 0∉IC) y es mayor el tiempo de secado de la pintura Nogopaint, en promedio (porque los signos positivos en ambos extremos indican que la media de población 1 es mayor que la 2). Por lo tanto, la marca de látex que brinda tiempo de secado más rápido es Alba.

Ejercicio 13 (estimación de la diferencia de medias)

X_1 v.a.c.: Duración de películas disponibles en sitio web A (minutos)
X_2 v.a.c.: Duración de películas disponibles en sitio web B (minutos)
$n_1 = 5$ (chico), $n_2 = 7$ (chico), $\sigma_1^2\ y\ \sigma_2^2\ \ desconocidas$ y DIFERENTES, poblaciones aprox. normales
$1 - \alpha = 0.90$
Como los tamaños muestrales son chicos, las varianzas poblacionales son desconocidas y las poblaciones subyacentes son aprox. normales, utilizaremos la T de Student.
Al contar con dos vectores de observaciones, se puede realizar directo con la función t.test(…)$conf.int

```
# EJERCICIO 13
a=c(103, 94, 110, 87, 98)
b=c(97, 82, 123, 92, 175, 88, 118)
t.test(a,b, var.equal=FALSE, conf.level = 0.90)$conf.int
```

Respuesta: $-36.427 < \mu_1 - \mu_2 < 11.798$ Con un 90% de confianza podemos concluir que no existe diferencia significativa entre las medias (ya que 0ϵIC) de los tiempos de duración de las películas en sitios web A y B.

Ejercicio 14 (estimación de la diferencia de medias)

X_1 v.a.c.: Ganancia de peso del ganado porcino con DIETA TRADICIONAL (kg)
X_2 v.a.c.: Ganancia de peso del ganado porcino con DIETA NUEVA (kg)
$n_1 = 12$ (chico), $n_2 = 12$ (chico), $\sigma_1^2\ y\ \sigma_2^2\ \ desconocidas$, poblaciones normales
$1 - \alpha = 0.94$
Como los tamaños muestrales son chicos, las varianzas poblacionales son desconocidas y las poblaciones subyacentes son aprox. normales, utilizaremos la T de Student.
Al contar con dos vectores de observaciones, se puede realizar directo con la función t.test(…)$conf.int

PASO 1: ¿Las varianzas se asumen iguales o diferentes? Armamos el IC para cociente de varianzas.

```
# EJERCICIO 14
trad=c(16,24,18,19,22,12,15,16,11,14,25,31)
nueva=c(31,34,29,25,38,34,29,31,32,29,32,28)
mean(trad)
mean(nueva)
#IC para el cociente de variancias ¿es cierto que son iguales?
var.test(trad, nueva, conf.level = 0.94)$conf.int
```

El resultado es: $0.927 < \frac{\sigma_1^2}{\sigma_2^2} < 10.063$, como el $1\ \epsilon$IC podemos concluir que las varianzas poblacionales son IGUALES.

PASO 2: Realizamos el IC para diferencia de medias, asumiendo varianzas iguales.

```
# IC para diferencia de Medias
t.test(trad, nueva, var.equal=TRUE, conf.level = 0.94)$conf.int
```

Respuesta: $-16.318 < \mu_1 - \mu_2 < -8.516$ Con un 94% de confianza podemos concluir que existe diferencia significativa entre las medias (ya que 0∉IC) y es mayor la ganancia de peso con la dieta nueva, en promedio (porque los signos negativos en ambos extremos indican que la media de población 2 es mayor que la 1). Por lo tanto, la dieta es efectiva para lograr ganancia de peso.

Ejercicio 15 (estimación de la diferencia de medias)

X_1 v.a.c.: Resistencia del alambre A (ohms)
X_2 v.a.c.: Resistencia del alambre B (ohms)
$n_1 = 6$ (chico), $n_2 = 6$ (chico), σ_1^2 y σ_2^2 $desconocidas$
$1 - \alpha = 0.95$
Como los tamaños muestrales son chicos, las varianzas poblacionales son desconocidas y las poblaciones subyacentes las asumimos normales, utilizaremos la T de Student.
Al contar con dos vectores de observaciones, se puede realizar directo con la función t.test(…)$conf.int

PASO 1: ¿Las varianzas se asumen iguales o diferentes? Armamos el IC para cociente de varianzas.

```
# EJERCICIO 15
a=c(0.140,0.138,0.143,0.142,0.144,0.137)
b=c(0.135,0.140,0.136,0.142,0.138,0.140)
#IC para el cociente de variancias ¿es cierto que son iguales?
var.test(a, b, conf.level = 0.95)$conf.int
```

El resultado es: $0.155 < \frac{\sigma_1^2}{\sigma_2^2} < 7.918$, como el 1 ϵIC podemos concluir que las varianzas poblacionales son IGUALES.

PASO 2: Realizamos el IC para diferencia de medias, asumiendo varianzas iguales.

```
# IC para diferencia de Medias
t.test(a,b, var.equal=TRUE, conf.level = 0.95)$conf.int
```

Respuesta: $-0.0014 < \mu_1 - \mu_2 < 0.0057$ Con un 90% de confianza podemos concluir que no existe diferencia significativa entre las medias (ya que 0ϵIC) de las resistencias de los alambres A y B. Por lo tanto, no puede afirmarse que A sea más resistente que B en promedio, a partir de lo obtenido.

Ejercicio 16 (estimación de cociente de varianzas) – refiere al ejercicio 13

X_1 v.a.c.: Duración de películas disponibles en sitio web A (minutos)
X_2 v.a.c.: Duración de películas disponibles en sitio web B (minutos)
$1-\alpha = 0.90$
Al contar con vectores de observaciones, se puede realizar directo con la función var.test(…)$conf.int

```
# EJERCICIO 16
a=c(103, 94, 110, 87, 98)
b=c(97, 82, 123, 92, 175, 88, 118)
#IC para el cociente de variancias ¿es cierto que son DISTINTAS?
var.test(a, b, conf.level = 0.90)$conf.int
```

Respuesta: $0.016 < \frac{\sigma_1^2}{\sigma_2^2} < 0.454$ Como el 1 $\notin$ IC podemos concluir que las varianzas poblacionales son DIFERENTES. Por lo tanto, estuvo bien suponerlas distintas en el IC para diferencia de medias del ejercicio 13.

Ejercicio 17 (estimación de cociente de varianzas) – refiere al ejercicio 12

X_1 v.a.c.: Tiempo de secado de pintura Nogopaint (hs)
X_2 v.a.c.: Tiempo de secado de pintura Alba (hs)
$1-\alpha = 0.95$
Al contar con vectores de observaciones, se puede realizar directo con la función var.test(…)$conf.int

```
# EJERCICIO 17
nogopaint=c(4.7, 3.9, 4.5, 5.5, 4.0, 5.3, 4.3,
            6.0, 5.2, 3.7, 5.5, 6.2, 5.1, 5.4, 4.8)
alba=c(3.5, 2.7, 3.9, 4.2, 3.6, 2.7, 3.3, 5.2,
       4.2, 2.9, 4.4, 5.2, 4.0, 4.1, 3.4)
var.test(nogopaint, alba, conf.level = 0.95)$conf.int
```

Respuesta: $0.314 < \frac{\sigma_1^2}{\sigma_2^2} < 2.787$ Como el 1 ϵIC podemos concluir que las varianzas poblacionales son IGUALES. Por lo tanto, estuvo bien suponerlas iguales en el IC para diferencia de medias del ejercicio 12.

Ejercicio 18 (estimación de una varianza)

Anulado de la guía para el cursado de este año.

Ejercicio 19 (estimación de una varianza)

Anulado de la guía para el cursado de este año.

Ejercicio 20 (estimación de la proporción de éxitos)

X v.a.d.: Número de hombres africanos que padecen cierto trastorno sanguíneo

n=100, $x = 24$, $\hat{p} = \frac{24}{100}$, $1 - \alpha = 0.99$

Lo hacemos directamente con la función prop.test(…)$conf.int de R Studio.

```
# EJERCICIO 20
prop.test(x=24, n=100, conf.level=0.99)$conf.int
```

Respuesta: $0.144 < p < 0.369$. Con un 99% de confianza podemos afirmar que el real porcentaje de hombres africanos que sufren este trastorno sanguíneo está entre 14.4 y 36.9% aproximadamente.

Ejercicio 21 (estimación de la proporción de éxitos)

X v.a.d.: Número de viviendas que utilizan petróleo como combustible para calefacción

n=1000, $x = 228$, $\hat{p} = \frac{228}{1000}$, $1 - \alpha = 0.99$

Lo hacemos directamente con la función prop.test(…)$conf.int de R Studio.

```
# EJERCICIO 21
prop.test(x=228, n=1000, conf.level=0.99)$conf.int
```

Respuesta: $0.195 < p < 0.264$. Con un 99% de confianza podemos afirmar que el real porcentaje de vivienda que utilizan petróleo para calefaccionarse en dicha ciudad varía entre 19.5 y 26.4% aproximadamente.

Ejercicio 22 (estimación de diferencia de proporciones)

X_1 v.a.d.: Cantidad de hombres que padecen cierto trastorno sanguíneo

X_2 v.a.d.: Cantidad de mujeres que padecen cierto trastorno sanguíneo

Hombres: $n_1 = 1000, x_1 = 250, \quad \widehat{p_1} = \frac{250}{1000}$

Mujeres: $n_2 = 1000, x_2 = 275, \quad \widehat{p_2} = \frac{275}{1000}$

$1 - \alpha = 0.95$

Lo hacemos directamente con la función prop.test(…)$conf.int de R Studio.

```
# EJERCICIO 22
prop.test(x=c(250,275),n=c(1000,1000),conf.level=0.95)$conf.int
```

Respuesta: $-0.065 < p_1 - p_2 < 0.015$. Con un 95% de confianza podemos afirmar que no existe diferencia significativa entre las proporciones (ya que 0ϵIC) de hombres y mujeres que padecen este trastorno sanguíneo.

Ejercicio 23 (estimación de diferencia de proporciones)

X_1 v.a.d.: Cantidad de ingenieros eléctricos que son mujeres
X_2 v.a.d.: Cantidad de ingenieros químicos que son mujeres
Ingeniería eléctrica: $n_1 = 250$, $x_1 = 80$, $\widehat{p_1} = \frac{80}{250}$
Ingeniería química: $n_2 = 175$, $x_2 = 40$, $\widehat{p_2} = \frac{40}{175}$
$1 - \alpha = 0.90$
Lo hacemos directamente con la función prop.test(…)$conf.int de R Studio.

```
# EJERCICIO 23
prop.test(x=c(80,40),n=c(250,175),conf.level=0.90)$conf.int
```

Respuesta: $0.015 < p_1 - p_2 < 0.168$. Con un 90% de confianza podemos afirmar que existe diferencia significativa entre las proporciones (ya que 0∉IC) de mujeres en estas ingenierías, siendo mayor en la Ingeniería Eléctrica (como ambos extremos son positivos, concluimos que p_1 es mayor que p_2).

Ejercicio 24 (estimación de diferencia de proporciones)

X_1 v.a.d.: Cantidad de estudiantes que ACTUALMENTE eligen al equipo A como su favorito
X_2 v.a.d.: Cantidad de estudiantes que EN 2012 eligen al equipo A como su favorito
Encuesta 2022: $n_1 = 1000$, $x_1 = 274$, $\widehat{p_1} = \frac{274}{1000}$
Encuesta 2012: $n_2 = 760$, $x_2 = 240$, $\widehat{p_2} = \frac{240}{760}$
$1 - \alpha = 0.95$

```
# EJERCICIO 24
prop.test(x=c(274,240),n=c(1000,760),conf.level=0.95)$conf.int
```

Respuesta: $-0.086 < p_1 - p_2 < 0.002$. Con un 95% de confianza podemos afirmar que no existe diferencia significativa entre las proporciones (ya que 0ϵIC) de estudiantes que favorecen a este equipo actualmente y en 2012. Por lo tanto, la proporción de adeptos a este equipo se mantiene constante.