

# Estadística

## Práctica 8

### Intervalos de confianza para la diferencia de parámetros

## Contenido

1. Introducción .....	3
2. Ejemplo: Intervalo de confianza para la diferencia de medias de calificaciones de estudiantes del turno de mañana y del turno de tarde (muestras grandes e independientes) .....	8
3. Ejemplo: Intervalo de confianza para la diferencia de medias del tiempo que duermen los estudiantes en días laborables y en días festivos (muestras grandes y pareadas) .....	11
4. Ejemplo: Intervalo de confianza para la diferencia de proporciones de estudiantes que tienen móvil Android en el turno de mañana y en el turno de tarde (muestras grandes e independientes) .....	13
5. Ejercicios propuestos .....	15

## 1. Introducción

Con esta práctica se utiliza R y RStudio para estimar (inferir) con un nivel de confianza determinado, el intervalo en el que se encuentra la diferencia de los parámetros (media y proporción) de dos poblaciones a partir de una muestra de cada una de las poblaciones.

Como en la práctica anterior, se puede calcular el intervalo en el que se puede asegurar que se encuentra la diferencia con un determinado nivel de confianza, obteniendo lo que se conoce como Intervalo de Confianza, con un límite inferior  $L_i$  y un límite superior  $L_s$ :

$$\text{Intervalo de confianza con nivel de confianza } 1 - \alpha = (L_i, L_s)$$

En la expresión del nivel de confianza aparece un valor  $\alpha$  (alfa) que se denomina nivel de error o nivel de significación, y representa lo contrario que el nivel de confianza.

Cuando el intervalo de confianza:

- Sí contiene el valor 0, se puede afirmar que la diferencia de los parámetros de ambas poblaciones NO es significativa.
- No contiene el valor 0, se puede afirmar que la diferencia entre los parámetros de ambas poblaciones SÍ es significativa y además:
  - o Si los extremos del intervalo son negativos, se puede afirmar que el parámetro de la primera población es menor que el de la segunda.
  - o Si los extremos del intervalo son positivos, se puede afirmar que el parámetro de la primera población es mayor que el de la segunda.

En la siguiente tabla se indican las fórmulas para calcular las diferencias de los parámetros poblacionales y su posible programación utilizando el lenguaje R. En el caso de la media se distingue entre muestras independientes y muestras relacionadas o pareadas. En este último caso realmente se comparan dos variables diferentes, y se manejan parejas de valores, normalmente un valor de la variable en una situación (por ejemplo, el peso de una persona) y otro valor de la misma variable en otra situación (por ejemplo, el peso de la misma persona al finalizar una dieta).

En la tabla sólo se muestran las situaciones sobre las que se probarán ejemplos en esta práctica, por ejemplo, no aparecen las fórmulas para la diferencia de medias con varianzas poblacionales conocidas, o con varianzas poblacionales desconocidas pero iguales. En la tabla se considera el caso general de varianzas poblacionales desconocidas y diferentes.



Diferencia a calcular	Varianzas conocidas	Intervalo de confianza (Li, Ls)	Código R
Diferencia de medias de dos poblaciones ( $\mu_1 - \mu_2$ )	Si	$L_i = \bar{x} - \bar{y} - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$ $L_s = \bar{x} - \bar{y} + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$	<pre>x=muestra de X y=muestra de Y n=length(x) m=length(y) mx=mean(x) my=mean(y) dx^2=varianza poblacional de X dy^2=varianza poblacional de Y z=qnorm(1-alfa/2,0,1)  Li=mx-my-z*sqrt((dx^2/n)+(dy^2/m))  Ls=mx-my+z*sqrt((dx^2/n)+(dy^2/m))</pre>
	No $n, m \geq 30$	$L_i = \bar{x} - \bar{y} - z_{\frac{\alpha}{2}} \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}$ $L_s = \bar{x} - \bar{y} + z_{\frac{\alpha}{2}} \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}$	<pre>x=muestra de X y=muestra de Y n=length(x) m=length(y) mx=mean(x) my=mean(y) sx=sd(x) sy=sd(y) z=qnorm(1-alfa/2,0,1)  Li=mx-my-z*sqrt((sx^2/n)+(sy^2/m))  Ls=mx-my+z*sqrt((sx^2/n)+(sy^2/m))</pre>
	No $n, m < 30$ $\sigma_X = \sigma_Y$	$L_i = \bar{x} - \bar{y} - \left(t_{\frac{\alpha}{2}, n+m-2}\right) (s_p) \sqrt{\frac{1}{n} + \frac{1}{m}}$ $L_s = \bar{x} - \bar{y} + \left(t_{\frac{\alpha}{2}, n+m-2}\right) (s_p) \sqrt{\frac{1}{n} + \frac{1}{m}}$ $s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}}$	<pre>x=muestra de X y=muestra de Y n=length(x) m=length(y) mx=mean(x) my=mean(y) sx=sd(x) sy=sd(y) sp=sqrt(((n-1)*sx^2+(m-1)*sy^2)/(n+m-2)) t=qt(1-alfa/2,n+m-2)  Li=m1-m2-t*sp*sqrt((1/n)+(1/m))  Ls=m1-m2+t*sp*sqrt((1/n)+(1/m))</pre>
	No $n, m < 30$ $\sigma_X \neq \sigma_Y$	$L_i = \bar{x} - \bar{y} - t_{\frac{\alpha}{2}, v} \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}$ $L_s = \bar{x} - \bar{y} + t_{\frac{\alpha}{2}, v} \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}$ $v = \frac{\left(\frac{s_X^2}{n} + \frac{s_Y^2}{m}\right)^2}{\frac{s_X^4}{n^2(n-1)} + \frac{s_Y^4}{m^2(m-1)}}$	<pre>x=muestra de X y=muestra de Y n=length(x) m=length(y) mx=mean(x) my=mean(y) sx=sd(x) sy=sd(y)  v=((sx^2/n+sy^2/m)^2)/((sx^4/(n^2*(n-1)))+(sy^4/(m^2*(m-1)))) t=qt(1-alfa/2, v)  Li=mx-my-t*sqrt((sx^2/n)+(sy^2/m))  Ls=mx-my+t*sqrt((sx^2/n)+(sy^2/m))</pre>



Diferencia a calcular	Varianzas conocidas	Intervalo de confianza (Li, Ls)	Código R
	Muestras pareadas $n \geq 30$	$L_i = \bar{d} - z_{\frac{\alpha}{2}} \frac{S_D}{\sqrt{n}}$ $L_s = \bar{d} + z_{\frac{\alpha}{2}} \frac{S_D}{\sqrt{n}}$ $D = X - Y$	<pre>d=muestra de X - muestra de Y n=length(d) m=mean(d) s=sd(d) z=qnorm(1-alfa/2,0,1)  Li=m-z*(s/sqrt(n)) Ls=m+z*(s/sqrt(n))</pre>
	Muestras pareadas $n < 30$	$L_i = \bar{d} - t_{\frac{\alpha}{2}, n-1} \frac{S_D}{\sqrt{n}}$ $L_s = \bar{d} + t_{\frac{\alpha}{2}, n-1} \frac{S_D}{\sqrt{n}}$ $D = X - Y$	<pre>d= muestra de X - muestra de Y n=length(d) m=mean(d) s=sd(d) t=qt(1-alfa/2,n-1)  Li=m-t*(s/sqrt(n)) Ls=m+t*(s/sqrt(n))</pre>

Diferencia a calcular	Muestras	Intervalo de confianza (Li, Ls)	Código R
Diferencia de proporciones de dos poblaciones ( $p_1 - p_2$ )	$n, m \geq 30$	$L_i = \hat{p}_1 - \hat{p}_2 - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ $L_s = \hat{p}_1 - \hat{p}_2 + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	<pre>n1,n2=tamaño de muestras p1,p2=proporciones en muestras z=qnorm(1-alfa/2,0,1)  Li=p1-p2-z*sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))  Ls=p1-p2+z*sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))</pre>

Existen en R funciones predefinidas que se pueden utilizar para los cálculos de algunos intervalos sin tener que programar las fórmulas de la tabla anterior.

Son las funciones predefinidas `t.test()` y `prop.test()`, y la función `z.test()` del paquete `BSDA`.

Diferencia a calcular	Varianzas conocidas	Código R	Ejemplo
Diferencia de medias de dos poblaciones ( $\mu_1 - \mu_2$ )	Si	<pre>test= z.test(muestra de X, muestra de Y, sigma.x=desviación poblacional de X, sigma.y=desviación poblacional de Y, conf.level=1-alfa) Intervalo: test\$conf.int</pre>	<pre>&gt; x=muestra de X &gt; y=muestra de Y &gt; dx=1.5 &gt; dy=2.3 &gt; alfa=0.05  &gt; test=z.test(x, y, sigma.x = dx, sigma.y = dy, conf.level = 1-alfa)  &gt; test\$conf.int [1] "Resultado"</pre>
	No $n, m \geq 30$	<pre>test= z.test(muestra de X, muestra de Y, sigma.x=desviación muestral de X, sigma.y=desviación muestral de Y, conf.level=1-alfa) Intervalo: test\$conf.int</pre>	<pre>&gt; x=rnorm(40,5,1) &gt; y=rnorm(50,6,1) &gt; alfa=0.05  &gt; test=z.test(x, y, sigma.x = sd(x), sigma.y = sd(y), conf.level = 1-alfa)  &gt; test\$conf.int [1] -1.1310858 -0.2934948</pre>
	No $n, m < 30$ $\sigma_X = \sigma_Y$	<pre>test= t.test(muestra de X, muestra de Y, conf.level=1-alfa, var.equal=TRUE) Intervalo: test\$conf.int</pre>	<pre>&gt; x=c(4.5, 5.2, 5.3, 4.7, 5, 6.3) &gt; y=c(6.7, 5.1, 6.1, 4.3) &gt; alfa=0.05  &gt; test=t.test(x, y, conf.level = 1-alfa, var.equal=TRUE)  &gt; test\$conf.int [1] -1.6043880 0.8377213</pre>
	No $n, m < 30$ $\sigma_X \neq \sigma_Y$	<pre>test= t.test(muestra de X, muestra de Y, conf.level=1-alfa, var.equal=FALSE) Intervalo: test\$conf.int</pre>	<pre>&gt; x=c(4.5, 5.2, 5.3, 4.7, 5, 6.3) &gt; y=c(6.7, 5.1, 6.1, 4.3) &gt; alfa=0.05  &gt; test=t.test(x, y, conf.level = 1-alfa, var.equal=FALSE)  &gt; test\$conf.int [1] -1.962490 1.195824</pre>
	Muestras pareadas $n \geq 30$	<pre>test= z.test(diferencia entre las muestras, sigma.x=desviación muestral de ladiferencia entre las muestras, conf.level=1-alfa) Intervalo: test\$conf.int</pre>	<pre>&gt; x=rnorm(40,5,1) &gt; y=rnorm(40,6,1) &gt; alfa=0.05 &gt; d=x-y  &gt; test=z.test(d, sigma.x = sd(d), conf.level = 1-alfa) &gt; test\$conf.int [1] -1.4781060 -0.4467718</pre>
	Muestras pareadas $n < 30$	<pre>test= t.test(muestra de X, muestra de Y, conf.level=1-alfa, paired=TRUE) Intervalo: test\$conf.int</pre>	<pre>&gt; x=c(4.5, 5.2, 5.3, 4.7) &gt; y=c(6.7, 5.1, 6.1, 4.3) &gt; alfa=0.05  &gt; test=t.test(x, y, conf.level = 1-alfa, paired=TRUE)  &gt; test\$conf.int [1] -2.482374 1.232374</pre>

Diferencia a calcular	Muestras	Código R	Ejemplo
Diferencia de proporciones de dos poblaciones ( $p_1 - p_2$ )	$n, m \geq 30$	<pre>test= prop.test(c (ng1,ng2) , c (n1,n2) , conf.level=1- alfa, correct=FALSE)</pre> <p>Intervalo: test\$conf.int</p> <p><b>NOTA:</b> n1 y n2 son los tamaños de las muestras. ng1 y ng2 son los tamaños de los grupos de datos de las muestras para los que se calculan las proporciones, es decir, las proporciones muestrales serían <math>p_1=ng1/n_1</math> y <math>p_2=ng2/n_2</math>.</p>	<pre>&gt; n1=100 &gt; n2=100 &gt; ng1=20 &gt; ng2=70 &gt; alfa=0.05  &gt; test=prop.test(c (ng1,ng2) , c (n1,n2) , conf.level = 1-alfa, correct=FALSE)  &gt; test\$conf.int [1] -0.61922 -0.38078</pre>

## 2. Ejemplo: Intervalo de confianza para la diferencia de medias de calificaciones de estudiantes del turno de mañana y del turno de tarde (muestras grandes e independientes)

### ENUNCIADO

Mediante una encuesta, se saben las calificaciones de acceso a la universidad de 42 estudiantes del turno de mañana y de 32 estudiantes del turno de tarde de un curso de la asignatura Estadística del Grado en Ingeniería en Sistemas de Información de la Universidad de Alcalá, de una población de 52 matriculados en el turno de mañana y 56 en el turno de tarde. Si se supone que las calificaciones tienen una distribución Normal, responder a la siguiente pregunta:

- Dibujar un diagrama con los intervalos de confianza de las dos poblaciones por separado, con un nivel de confianza del 95%, para hacer una comparación visual.
- Calcular el intervalo de confianza para la diferencia de medias poblacionales de las calificaciones de los turnos de mañana y tarde con nivel de confianza del 95%.
- Calcular el intervalo de confianza con la función `z.test` del paquete BSDA.
- ¿Existe una diferencia significativa entre las medias de las dos poblaciones (mañana y tarde)?

### SOLUCIÓN

Primero leemos el archivo [encuesta.csv](#) y guardamos en dos vectores las muestras de notas del turno de mañana (grupos A1 y A2) y del turno de tarde (grupos B1 y B2):

```
> encuesta=read.csv2("encuesta.csv")

> encuesta.mañana=encuesta[(encuesta$GRUPO=="A1")|(encuesta$GRUPO=="A2"),]
> encuesta.tarde=encuesta[(encuesta$GRUPO=="B1")|(encuesta$GRUPO=="B2"),]
> notaM=encuesta.mañana$NOTA
> notaT=encuesta.tarde$NOTA
```

- Dibujar un diagrama con los intervalos de confianza de las dos poblaciones por separado, con un nivel de confianza del 95%, para hacer una comparación visual.

El intervalo de confianza de los alumnos de mañana es:

```
> nM=length(notaM)
> mM=mean(notaM)
> sM=sd(notaM)
> alfa=1-0.95
> z=qnorm(1-alfa/2,0,1)
> (LiM=mM-z*(sM/sqrt(nM)))
[1] 8.024987
> (LsM=mM+z*(sM/sqrt(nM)))
[1] 8.672632
```

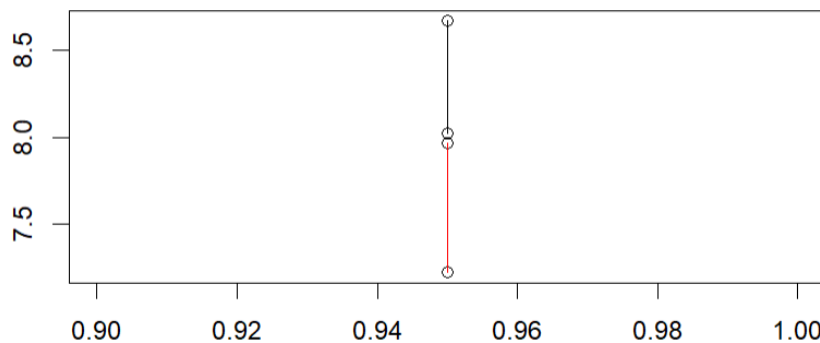


El intervalo de confianza de los alumnos de tarde es:

```
> nT=length(notat)
> mT=mean(notat)
> sT=sd(notat)
> alfa=1-0.95
> z=qnorm(1-alfa/2,0,1)
> (LiT=mT-z*(sT/sqrt(nT)))
[1] 7.220355
> (LsT=mT+z*(sT/sqrt(nT)))
[1] 7.968395
```

Podemos dibujarlos con plot y lines:

```
> plot(c(0.95,0.95,0.95,0.95),c(LiM, LsM, LiT, LsT), xlim = c(0.9,1))
> lines(c(0.95,0.95),c(LiM,LsM))
> lines(c(0.95,0.95),c(LiT,LsT),col="red")
```



Vemos que los dos intervalos no se solapan, lo cual es un indicio de que no va a haber coincidencia de medias, aunque lo tendremos que demostrar calculando el intervalo de la diferencia de medias.

**b) Calcular el intervalo de confianza para la diferencia de medias poblacionales de las calificaciones de los turnos de mañana y tarde con nivel de confianza del 95%.**

```
> alfa = 1-0.95
> z=qnorm(1-alfa/2,0,1)
> (Li=mM-mT-z*sqrt((sM^2/nM)+(sT^2/nT)))
[1] 0.2597101
> (Ls=mM-mT+ z*sqrt((sM^2/nM)+(sT^2/nT)))
[1] 1.249159
```

El intervalo de confianza para la diferencia de medias entre los grupos de mañana y tarde es (0.2597101, 1.249159).

**c) Calcular el intervalo de confianza con la función z.test del paquete BSDA.**

```
> install.packages("BSDA")
> library(BSDA)

> z.test(notaM,notaT,sigma.x=sM,sigma.y=sT,conf.level=0.95)
95 percent confidence interval:
 0.2597101 1.2491589
```

Podemos comprobar que se obtiene el mismo intervalo que aplicando las fórmulas.

**d) ¿Existe una diferencia significativa entre las medias de las dos poblaciones (mañana y tarde)?**

Como el 0 no está incluido en el intervalo (0.2597101, 1.2491589), se puede afirmar con un nivel de confianza del 95% que ***Sí hay diferencia significativa entre la media de las calificaciones de los estudiantes de los turnos de mañana y tarde, y que la media de las calificaciones de los estudiantes del turno de mañana es mayor que la media de las calificaciones de los estudiantes del turno de tarde***, porque los extremos de los intervalos son positivos.

### 3. Ejemplo: Intervalo de confianza para la diferencia de medias del tiempo que duermen los estudiantes en días laborables y en días festivos (muestras grandes y pareadas)

#### ENUNCIADO

Mediante la encuesta del ejemplo anterior se sabe el tiempo que duermen los días laborables 74 estudiantes de una población de 108 matriculados. Se ha averiguado que los días festivos los estudiantes duermen un 10% más de tiempo que en días laborables. Si se supone que el tiempo que duermen tanto en días laborables como en festivos tiene una distribución Normal, responder a la siguiente pregunta:

- Calcular el intervalo de confianza para la diferencia de medias poblacionales del tiempo que duermen los estudiantes en días laborables y festivos, con nivel de confianza del 95%.
- Calcular el intervalo de confianza con la función `z.test` del paquete BSDA
- ¿Existe una diferencia significativa entre las medias de las dos poblaciones?

#### SOLUCIÓN

- Calcular el intervalo de confianza para la diferencia de medias poblacionales del tiempo que duermen los estudiantes en días laborables y festivos, con nivel de confianza del 95%.

Primero guardamos en dos vectores las muestras de las horas que duermen los estudiantes en días laborables y en festivos:

```
> dormirL=encuesta$DORMIR
> dormirF=1.1*encuesta$DORMIR #Un 10% más que los laborables

> alfa=1-0.95
> d=dormirL-dormirF
> m=mean(d)
> s=sd(d)
> n=length(d)
> z=qnorm(1-alfa/2,0,1)
> (Li=m-z*(s/sqrt(n)))
[1] -0.6978706
> (Ls=m+z*(s/sqrt(n)))
[1] -0.6561835
```

El intervalo de confianza de la diferencia de media de las horas que duermen en laborables y festivos es (-0.6978706, -0.6561835).

- Calcular el intervalo de confianza con la función `z.test` del paquete BSDA.

```
> z.test(d, sigma.x=s ,conf.level=0.95)
95 percent confidence interval:
-0.6978706 -0.6561835
```

Podemos comprobar que se obtiene el mismo intervalo que aplicando las fórmulas.

**c) ¿Existe una diferencia significativa entre las medias de las dos poblaciones?**

Como el 0 no está incluido en el intervalo, se puede afirmar con un nivel de confianza del 95% que ***Sí hay diferencia significativa entre la media del tiempo que duermen los estudiantes los días laborables y los días festivos, y que la media del tiempo que duermen los días laborables es menor que la media del tiempo que duermen los días festivos***, porque los extremos de los intervalos son negativos.

## 4. Ejemplo: Intervalo de confianza para la diferencia de proporciones de estudiantes que tienen móvil Android en el turno de mañana y en el turno de tarde (muestras grandes e independientes)

### ENUNCIADO

Mediante una encuesta se sabe que en una muestra de 42 estudiantes del turno de mañana de la asignatura Estadística, 27 tienen teléfonos móviles con sistema operativo Android, y que en una muestra de 32 estudiantes del turno de tarde, 23 tienen también teléfono Android. La población es de 52 matriculados por la mañana y 56 por la tarde. Responder a las siguientes preguntas:

- Calcular el intervalo de confianza para la diferencia de las proporciones poblacionales de estudiantes que tienen móvil Android en los turnos de mañana y tarde, con un nivel de confianza del 95%.
- Calcular el intervalo de confianza con la función `prop.test`
- ¿Existe una diferencia significativa entre las proporciones de las dos poblaciones?

### SOLUCIÓN

a) Calcular el intervalo de confianza para la diferencia de las proporciones poblacionales de estudiantes que tienen móvil Android en los turnos de mañana y tarde, con un nivel de confianza del 95%.

Primero calculamos el tamaño de las muestras y de los estudiantes que usan Android en cada muestra.

```
> soM=encuesta.mañana$SO
> soT=encuesta.tarde$SO

> nM=length(soM)
> nT=length(soT)

> nMA=length(soM[soM=="Android"])
> nTA=length(soT[soT=="Android"])
```

Y se aplican las fórmulas:

```
> alfa=1-0.95

> pM=nMA/nM
> pT=nTA/nT

> z=qnorm(1-alfa/2,0,1)

> (Li=pM-pT-z*sqrt((pM*(1-pM)/nM)+(pT*(1-pT)/nT)))
[1] -0.2886515
> (Ls=pM-pT+z*sqrt((pM*(1-pM)/nM)+(pT*(1-pT)/nT)))
```

```
[1] 0.1368658
```

El intervalo de confianza de la diferencia de proporciones de estudiantes que tienen móvil Android en los turnos de mañana y tarde es (-0.2886515, 0.1368658).

**b) Calcular el intervalo de confianza con la función prop.test**

```
> prop.test(c(nMA,nTA), c(nM,nT), conf.level = 1-alfa, correct=FALSE)
95 percent confidence interval:
-0.2886515  0.1368658
```

Podemos comprobar que se obtiene el mismo intervalo que aplicando las fórmulas.

**c) ¿Existe una diferencia significativa entre las proporciones de las dos poblaciones?**

Como el 0 está incluido en el intervalo (-0.2886515, 0.1368658), se puede afirmar con un nivel de confianza del 95% que **NO hay diferencia significativa entre las proporciones de móvil Android de los estudiantes de los turnos de mañana y tarde.**

## 5. Ejercicios propuestos

- 1) Mediante una encuesta, se sabe el tiempo en minutos del viaje a la Escuela Politécnica de 42 estudiantes del turno de mañana y de 32 estudiantes del turno de tarde de un curso de la asignatura Estadística del Grado en Ingeniería en Sistemas de Información de la Universidad de Alcalá, de una población de 52 matriculados en el turno de mañana y 56 en el turno de tarde. Si se supone que las calificaciones tienen una distribución Normal, responder a la siguiente pregunta:
  - a. Calcular el intervalo de confianza para la diferencia de medias poblacionales del tiempo del viaje a la Escuela de los turnos de mañana y tarde con nivel de confianza del 95%. ¿Existe una diferencia significativa entre las medias de las dos poblaciones?
  
- 2) Mediante la encuesta se sabe el tiempo en minutos del viaje de ida a la Escuela Politécnica de 74 estudiantes de una población de 108 matriculados. Se ha averiguado que el viaje de regreso dura un 10% menos que el de ida. Si se supone que el tiempo del viaje, tanto de ida como de vuelta, tiene una distribución Normal, responder a la siguiente pregunta:
  - a. Calcular el intervalo de confianza para la diferencia de medias poblacionales del tiempo de viaje de ida y del tiempo de viaje de vuelta de los estudiantes, con nivel de confianza del 95%. ¿Existe una diferencia significativa entre las medias de las dos poblaciones?
  
- 3) Mediante la encuesta se sabe que en una muestra de 42 estudiantes del turno de mañana, 13 tienen una línea telefónica contratada con las compañías Jazzte/Orange/Simyo, y que en una muestra de 32 estudiantes del turno de tarde, 13 tienen también una línea telefónica contratada con las compañías Jazzte/Orange/Simyo. La población es de 52 matriculados por la mañana y 56 por la tarde. Responder a las siguientes preguntas:
  - a. Calcular el intervalo de confianza para la diferencia de las proporciones poblacionales de estudiantes que tienen una línea telefónica contratada con las compañías Jazzte/Orange/Simyo en los turnos de mañana y tarde, con un nivel de confianza del 95%. ¿Existe una diferencia significativa entre las proporciones de las dos poblaciones?