

6.16 EJEMPLOS

6.16.1 EJERCICIOS RESUELTOS

Ejercicio 1 Sea una m.a.s. de tamaño n de una población $N(\mu, \sigma)$, considerando como estimador de la media la expresión

$$T = K \sum_{i=1}^n i \cdot X_i$$

determinar K para que el estadístico T sea insesgado.

Solución.-

T será un estimador insesgado de μ si verifica:

$$E[T] = \mu$$

Por tanto:

$$E \left[K \sum_{i=1}^n i \cdot X_i \right] = K \sum_{i=1}^n i \cdot E[X_i] = K \cdot \sum_{i=1}^n i \cdot \mu = K \cdot \mu \cdot \sum_{i=1}^n i$$

$$E \left[K \sum_{i=1}^n i \cdot X_i \right] = K \cdot \mu \cdot \frac{n+1}{2} \cdot n$$

Por tanto, distinguiremos dos posibilidades:

-) Si $\mu = 0$

Para cualquier valor de K , T será insesgado de μ ya que $E[T] = 0$

-) Si $\mu \neq 0$

$$T \text{ será insesgado de } \mu \Leftrightarrow K = \frac{2}{n \cdot (n+1)}$$

Ejercicio 2 Calcular el percentil $t_{0.95}$ y $t_{0.25}$ en cada uno de los siguientes casos:

1. En una distribución t -Student con 3 grados de libertad.
2. En una distribución t -Student con 30 grados de libertad.
3. En una distribución t -Student con 52 grados de libertad.
4. En una distribución t -Student con 120 grados de libertad.

Solución.-

1. Recordemos que $t_{0.95}$ es aquel número real que verifica:

$$p[T \leq t_{0.95}] = 0.95$$

Para encontrar este valor en la tabla de la distribución t-Student bastará:

-) Localizar en la primera columna los grados de libertad, en este caso: 3.
-) Localizar en la primera fila la probabilidad acumulada, en nuestro caso: 0.95.
-) Movernos horizontal y verticalmente desde las posiciones anteriores hasta cruzarnos en el punto $t_{0.95}$.

Por tanto el percentil $t_{0.95}$, en una t-Student con 3 grados de libertad será el valor:

$$t_{0.95} = \mathbf{2.3534}$$

es decir, si desde el valor 2.3534 nos movemos horizontalmente hasta la primera columna, llegaremos al valor 3 (grados de libertad), y si lo hacemos verticalmente hacia la primera fila llegaremos al valor 0.95 (probabilidad acumulada).

Como en la tabla únicamente tenemos tabulada la t-Student para colas probabilísticas que van desde 0.75 hasta 0.999, para calcular el percentil $t_{0.25}$, tendremos que realizar la siguiente consideración:

$$p[T \leq t_{0.25}] = 1 - p[T \geq t_{0.25}]$$

Como la distribución t-Student es simétrica, se verifica:

$$t_{0.25} = -t_{0.75}$$

y resulta:

$$p[T \leq t_{0.25}] = 1 - p[T \leq t_{0.75}]$$

Por tanto, buscando en la tabla con los datos:

Grados de libertad : 3
Cola de probabilidad : 0.75

tenemos:

$$t_{0.25} = -t_{0.75} = \mathbf{-0.7649}$$

2. En el caso de 30 grados de libertad actuaremos de modo similar al caso anterior, pero buscando en la fila 30 de la tabla. Resultando:

$$t_{0.95} = \mathbf{1.6973}$$

y

$$t_{0.25} = -t_{0.75} = \mathbf{-0.6828}$$

3. En el caso que nos ocupa observamos que para 52 grados de libertad no hay tabulación de la t-Student en la tabla, en tal caso podremos actuar de dos formas posibles:

- (a) Aproximar el valor del percentil con la tabulación de 50 grados de libertad, resultando:

$$t_{0.95} = \mathbf{1.6759}$$

y

$$t_{0.25} = -t_{0.75} = \mathbf{-0.6794}$$

- (b) Realizar una interpolación lineal . Es decir, en el caso $t_{0.95}$, consideraríamos la recta que pasa por los puntos:

$$(50, 1.6759) \quad \text{y} \quad (60, 1.6706)$$

y predecimos el valor cuando $x = 52$.

La recta viene dada por la expresión:

$$y - y_0 = \frac{(y_1 - y_0)}{(x_1 - x_0)} (x - x_0)$$

$$y = y_0 + \frac{(y_1 - y_0)}{(x_1 - x_0)} (x - x_0)$$

sustituyendo,

$$y = 1.6759 + \frac{(1.6706 - 1.6759)}{(60 - 50)} (x - 50)$$

simplificando resulta:

$$y = 1.7024 - 0.00053x$$

Por tanto, podemos predecir el valor $t_{0.95}$ realizando una interpolación lineal simple:

$$x = 52 \quad \implies \quad y = 1.7024 - 0.00053 * 52 = \mathbf{1.6748}$$

podemos observar que el valor obtenido es muy similar a 1.6759 .

Del mismo modo se calcularía el percentil $t_{0.25}$.

4. En este caso tenemos dos opciones para calcular, *de modo aproximado*, los percentiles $t_{0.95}$ y $t_{0.25}$.

- (a) Utilizar la última fila de la tabla, que es una aproximación para situaciones con mas de 100 grados de libertad. De este modo resulta:

$$t_{0.95} = \mathbf{1.6449}$$

y

$$t_{0.25} = -t_{0.75} = \mathbf{-0.6745}$$

- (b) O bien, tener en cuenta que:

$$t\text{-Student con } n \text{ grados de libertad} \rightarrow N(0,1) \text{ cuando } n \rightarrow \infty$$

por tanto, bastará realizar las aproximaciones siguientes:

$$t_{0.95} \simeq Z_{0.95} \quad \text{y} \quad t_{0.25} \simeq Z_{0.25}$$

resultando:

$$t_{0.95} \simeq 1.645$$

$$t_{0.25} \simeq -Z_{0.75} = -0.675$$

Ejercicio 3 Calcular el percentil $\chi_{n,0.95}^2$ y $\chi_{n,0.25}^2$ en cada uno de los siguientes casos:

1. $n=5$
2. $n=30$.

Solución.-

1. Para utilizar la tabla de la distribución χ^2 seguiremos la misma filosofía que con la t-Student, pero teniendo en cuenta que en la tabla de la χ^2 , a diferencia de la t-Student, p representa la probabilidad acumulada por la derecha y no por la izquierda como viene siendo habitual. Tenemos que hacer constar que en algunos textos, aparece la tabla de la función de distribución de la χ^2 .

Para calcular el percentil $\chi_{5,0.95}^2$, tendremos en cuenta que hemos de buscar en la tabla utilizando los valores:

$$\begin{aligned}\nu &= 5 && (\text{Grados de libertad}) \\ p &= 1 - 0.95 = 0.05 && (\text{Cola de la derecha})\end{aligned}$$

resultando:

$$\chi_{5,0.95}^2 = \mathbf{11.0705}$$

Para calcular el percentil $\chi_{5,0.25}^2$, tendremos en cuenta que hemos de buscar en la tabla utilizando los valores:

$$\begin{aligned}\nu &= 5 \\ p &= 1 - 0.25 = 0.75\end{aligned}$$

resultando:

$$\chi_{5,0.25}^2 = \mathbf{2.6746}$$

2. El percentil $\chi_{30,0.95}^2$, se buscará en la tabla utilizando los valores:

$$\begin{aligned}\nu &= 30 \\ p &= 1 - 0.95 = 0.05\end{aligned}$$

resultando:

$$\chi_{30,0.95}^2 = \mathbf{43.7730}$$

El percentil $\chi_{30,0.25}^2$, se buscará en la tabla utilizando los valores:

$$\begin{aligned}\nu &= 30 \\ p &= 1 - 0.25 = 0.75\end{aligned}$$

resultando:

$$\chi_{30,0.25}^2 = \mathbf{24.4776}$$

Ejercicio 4 Calcular los percentiles $F_{8,7;0.99}$ y $F_{8,7;0.01}$

Solución.-

Para buscar en la tabla de la F-Snedecor el percentil $F_{8,7;0.99}$ hemos de tener en cuenta que:

$$\begin{aligned} df_1 &= 8 \text{ (1ª Fila de la tabla)} \\ df_2 &= 7 \text{ (1ª Columna de la tabla)} \\ 0.99 &= \text{Probabilidad acumulada (Última columna de la tabla)} \end{aligned}$$

El valor donde se cruzan todos estos datos será el percentil buscado.

Por tanto:

$$F_{9,7;0.99} = \mathbf{6.840}$$

En el caso del percentil $F_{8,7;0.01}$ nos encontramos con el problema de que en las tablas únicamente aparece la probabilidad acumulada para los valores: 0.95 , 0.975 , 0.99 y 0.999. Por ello, nos vemos obligados a utilizar la siguiente propiedad de la distribución F-Snedecor:

$$F_{n,m;\alpha} = \frac{1}{F_{m,n;1-\alpha}}$$

En el caso que nos ocupa:

$$F_{8,7;0.01} = \frac{1}{F_{7,8;1-0.01}} = \frac{1}{F_{7,8;0.99}} = \frac{1}{6.178} = \mathbf{0.1618}$$

Ejercicio 5 Una empresa fabrica bombillas, que tienen una duración distribuida aproximadamente en forma normal, con una desviación típica de 40 h.

Si una muestra de 30 bombillas tiene una vida promedio de 780 h. obtener un intervalo de confianza del 96% para la vida media de la población de todas las bombillas.

Solución.-

Se trata de obtener un intervalo de confianza para la media poblacional μ , supuesto que la varianza es conocida, para ello tendremos en cuenta que:

$$X \equiv \text{El tiempo de vida de las bombillas}$$

$$X \sim N(\mu, 40)$$

$$\alpha = 0.04$$

$$n = 30$$

En tal situación :

$$p \left[\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] = 1 - \alpha$$

$$p \left[\bar{X} - Z_{0.98} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{0.98} \frac{\sigma}{\sqrt{n}} \right] = 1 - \alpha$$

es decir $\left(\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$ es un intervalo de confianza para la media poblacional con un nivel de confianza de $(1 - \alpha) \cdot 100\%$.

Buscando $Z_{0.98}$ en la tabla de la F.d.D. de la $N(0,1)$ tenemos que:

$$Z_{0.98} = 2.06$$

y sustituyendo se obtiene el intervalo de confianza:

$$\left(780 - 2.06 \frac{40}{\sqrt{30}}, 780 + 2.06 \frac{40}{\sqrt{30}} \right)$$

$$(764.95, 795.04)$$

Como conclusión diremos que se puede tener una confianza del 96% de que una muestra de tamaño 30 proporcione una estimación de μ dentro del intervalo $(764.95, 795.04)$.

Ejercicio 6 ¿De qué tamaño debiera ser la muestra si se desea tener una confianza del 96% de que el estimador de μ difiera de éste en menos de 3 horas?

Solución.-

Teniendo en cuenta que la condición:

$$p \left[\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] = 1 - \alpha$$

es equivalente a:

$$p \left[|\bar{X} - \mu| < +Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] = 1 - \alpha$$

y siendo \bar{X} el estimador de μ y puesto que se admite un error máximo de $e = 3$ horas, tenemos que:

$$e = Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Rightarrow n = \left(\frac{Z_{1-\frac{\alpha}{2}} \cdot \sigma}{e} \right)^2$$

sustituyendo:

$$\begin{aligned} e &= 3 \\ \sigma &= 40 \\ Z_{1-\frac{\alpha}{2}} &= Z_{0.98} = 2.06 \end{aligned}$$

se tiene

$$n = \left(\frac{2.06 \cdot 40}{3} \right)^2 = 754.4$$

Es decir, el tamaño de la muestra debe ser mayor o igual a **755** para poder tener una confianza del 96% de que el tiempo de vida de las bombillas estará en el intervalo:

$$(\bar{X} - e, \bar{X} + e)$$

Ejercicio 7 ¿Cuál sería la probabilidad de que μ no difiera de 780 h. mas de 10 h.?

Solución.-

En primer lugar determinaremos el nivel de significación:

$$e = 10 \text{ h.} \Rightarrow \frac{Z_{1-\frac{\alpha}{2}} \cdot \sigma}{e} = 10$$

$$\Rightarrow Z_{1-\frac{\alpha}{2}} = \frac{10 \cdot \sqrt{n}}{\sigma}$$

Sustituyendo

$$\begin{aligned} n &= 30 \\ \sigma &= 40 \end{aligned}$$

resulta:

$$Z_{1-\frac{\alpha}{2}} = \frac{54.7}{40} \Rightarrow Z_{1-\frac{\alpha}{2}} = 1.36$$

Por tanto,

$$p[Z < 1.36] = 1 - \frac{\alpha}{2}$$

despejando α se tiene:

$$\frac{\alpha}{2} = 1 - p[Z < 1.36] \Rightarrow \frac{\alpha}{2} = 1 - 0.91308 \Rightarrow \frac{\alpha}{2} = 0.086$$

$$\alpha = 0.172$$

La solución vendrá dada por:

$$p[\bar{X} - e < \mu < \bar{X} + e] = 1 - \alpha = \mathbf{0.828}$$

Ejercicio 8 Una muestra aleatoria simple de 20 estudiantes obtuvo una calificación media de 72, con $S^2 = 16$. Bajo hipótesis de normalidad encontrar un intervalo de confianza del 98% para σ^2

Solución.-

El intervalo de confianza buscado vendrá dado por:

$$\left(\frac{(n-1) S_c^2}{\chi_{1-\frac{\alpha}{2}}^2}, \frac{(n-1) S_c^2}{\chi_{\frac{\alpha}{2}}^2} \right)$$

Donde $\chi^2_{1-\frac{\alpha}{2}}$ y $\chi^2_{\frac{\alpha}{2}}$ son los percentiles de una distribución χ^2_{19} . Buscando en la tabla de la χ^2 , obtenemos que:

$$\begin{aligned}\chi^2_{19,0.99} &= 36.1908 \\ \chi^2_{19,0.01} &= 7.6327\end{aligned}$$

Resultando,

$$\left(\frac{19 \cdot 16}{36.19}, \frac{19 \cdot 16}{7.6327} \right)$$

El intervalo de confianza al 98% para la varianza vendrá dado por:

$$(8.4001, 39.832)$$

Ejercicio 9 *Se ha tomado una muestra aleatoria simple de 25 individuos, a los que se les ha medido el nivel de hemoglobina en sangre, obteniéndose una media muestral de 110 mg/cc y una cuasidesviación típica muestral de 20 mg/cc. Obtenga un intervalo de confianza para la media poblacional al 90% de confianza.*

Solución.-

Sea

$$X \equiv \text{Nivel de Hemoglobina en Sangre}$$

Es necesario suponer normalidad en los datos, es decir:

$$X \sim N(\mu, \sigma)$$

En este caso, desconocemos el valor de σ , pero sí sabemos cuanto vale su estimador muestral, es decir S_c .

Por tanto, la estimación de intervalos de confianza para la media poblacional con desviación típica desconocida, se basa en el estadístico:

$$T = \frac{\bar{X} - \mu}{S_c/\sqrt{n}} \sim t_{n-1}$$

Un intervalo de confianza con probabilidad $(1 - \alpha)$ de contener a μ vendrá dado por:

$$\left(\bar{X} - t_{n-1, 1-\frac{\alpha}{2}} \cdot \frac{S_c}{\sqrt{n}}, \bar{X} + t_{n-1, 1-\frac{\alpha}{2}} \cdot \frac{S_c}{\sqrt{n}} \right)$$

donde $t_{n-1, 1-\frac{\alpha}{2}}$ es el percentil $1 - \frac{\alpha}{2}$ de una distribución t-Student con $n - 1$ grados de libertad.

Tenemos que:

$$\begin{aligned}\bar{X} &= 110 \\ n &= 25 \\ S_c &= 20 \\ 1 - \alpha &= 0.9\end{aligned}$$

y el percentil

$$t_{25-1, 1-\frac{0.1}{2}} = t_{24, 0.95} = 1.7109$$

Sustituyendo estos valores en la expresión del I.C. resulta:

$$\left(110 - 1.7109 \cdot \frac{20}{\sqrt{25}}, 110 + 1.7109 \cdot \frac{20}{\sqrt{25}} \right)$$

simplificando

$$(103.16, 116.84)$$

Ejercicio 10 Una compañía de TAXIS está tratando de decidir la compra de neumáticos de las marcas A o B para sus vehículos. Para estimar la diferencia de las dos marcas de neumáticos se realiza un experimento empleando 12 neumáticos de cada marca. Los neumáticos se hacen correr hasta su desgaste total, obteniéndose un recorrido medio de 33000 Km en la primera muestra y 38100 Km en la segunda.

Si sabemos que la desviación típica poblacional para la marca A es 5000 Km. y la de la marca B es 6100 Km. Obtener un intervalo de confianza para $\mu_1 - \mu_2$. Tomar $\alpha = 0.05$

Solución.-

Para realizar este tipo de contraste hemos de suponer que existe NORMALIDAD en los datos, tanto para la variable X_1 como para la variable X_2 , siendo:

$$X_1 \equiv N^{\circ} \text{ de Km. de vida de los neumáticos tipo A}$$

$$X_2 \equiv N^{\circ} \text{ de Km. de vida de los neumáticos tipo B}$$

en ambos casos es necesario suponer que:

$$X_1 \sim N(\mu_1, \sigma_1) \quad y \quad X_2 \sim N(\mu_2, \sigma_2)$$

Por tanto, al ser X_1 y X_2 independientes, se tiene que:

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

ya que:

$$\begin{aligned} X_1 - X_2 &\equiv \text{ es una v.a. normal} \\ \mu_{X_1 - X_2} &= \mu_1 - \mu_2 \\ \sigma_{X_1 - X_2}^2 &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \end{aligned}$$

Por tanto un intervalo de confianza al nivel $(1 - \alpha) \cdot 100\%$ para $\mu_{X_1 - X_2} = \mu_1 - \mu_2$, vendrá dado por:

$$\left(\overline{X_1} - \overline{X_2} - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \overline{X_1} - \overline{X_2} + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$$

Además tendremos que:

$$\begin{aligned}n_1 &= n_2 = 12 \\ \overline{X}_1 - \overline{X}_2 &= 33000 - 38100 = -5100 \\ \sigma_1 &= 5000 \\ \sigma_2 &= 6100 \\ Z_{1-\frac{\alpha}{2}} &= Z_{0.975} = 1.96\end{aligned}$$

Nuestro intervalo de confianza al 95% para la diferencia de medias será el siguiente:

$$\left((33000 - 38100) - 1.96 \cdot \sqrt{\frac{5000^2}{12} + \frac{6100^2}{12}}, (33000 - 38100) + 1.96 \cdot \sqrt{\frac{5000^2}{12} + \frac{6100^2}{12}} \right)$$

simplificando:

$$(-9562.7, -637.32)$$

Como el $0 \notin (-9562.7, -637.32)$, con un nivel de confianza del 95% podremos asegurar que las medias poblacionales pueden considerarse distintas.

Ejercicio 11 Supongamos que 100 neumáticos de cierta marca duraron en promedio 21431 kilómetros con una desviación típica poblacional de 1295 Km. Utilizando $\alpha = 0.05$, ¿podemos considerar que la duración media de los neumáticos es inferior a 22000 Km.?

Solución.-

Sea

$X \equiv$ Duración de los neumáticos.

y

$$\begin{aligned}E[X] &= \mu \\ \text{var}(X) &= \sigma^2\end{aligned}$$

El contraste de hipótesis que proponemos es el siguiente:

$$\begin{cases} H_0 : \mu = 22000 \\ H_1 : \mu < 22000 \end{cases}$$

Al ser el tamaño muestral lo suficientemente grande ($n > 30$), podemos aplicar el TCL. El estadístico que utilizaremos será:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

Al sustituir los datos: $\bar{X} = 21431$, $\mu_0 = 22000$, $\sigma = 1295$ obtenemos el siguiente valor:

$$Z_{\text{exp}} = \frac{21436 - 22000}{\frac{1295}{\sqrt{100}}} = -4.3552$$

Dado el contraste unilateral propuesto, la región crítica será la siguiente:

$$RC : Z_{\text{exp}} < Z_{\frac{\alpha}{2}}$$

$$RC : Z_{\text{exp}} < Z_{0.05} = -Z_{0.95} = -1.65$$

Al ser $Z_{\text{exp}} = -4.35 < -1.65$ rechazaremos la hipótesis nula y concluiremos diciendo que la duración media de los neumáticos puede considerarse inferior a 22000 Km.

NOTA: Si σ^2 fuese desconocida, al ser n suficientemente grande podríamos haber considerado $\sigma^2 = S^2$. Por otra parte, recordemos que cuando $n < 30$ y se desconoce σ^2 el test realizado no sería válido. En tal caso debieramos de suponer normalidad en los datos y resolver el contraste utilizando una distribución T-Student.

Ejercicio 12 En la comparación de dos tipos de pintura, una agencia de servicio al consumidor descubre que cuatro latas de 50 litros de una marca A cubren en promedio 512 metros cuadrados, con una cuasidesviación estandar de 31m^2 , mientras que 4 latas de 50 l. de una marca B cubren en promedio 492m^2 con una cuasidesviación estandar de 26m^2 . Considerando $\alpha = 0.05$, ¿podemos considerar ambas marcas igualmente rentables y así poder comprar la mas barata?

Solución.-

Sea

$$X_1 \equiv \text{Superficie cubierta por la marca A}$$

$$X_2 \equiv \text{Superficie cubierta por la marca B}$$

Es necesario suponer normalidad, es decir:

$$X_1 \sim N(\mu_1, \sigma_1)$$

$$X_2 \sim N(\mu_2, \sigma_2)$$

Nuestro contraste será el siguiente:

$$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$$

Pero para resolverlo, puesto que las varianzas poblacionales son desconocidas, tendremos que ver si pueden considerarse iguales o distintas; es decir, tendremos que resolver el contraste de igualdad de varianzas entre las dos marcas. Sea

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases} \iff \begin{cases} H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{cases}$$

El estadístico utilizado será:

$$F = \frac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1}$$

nuestro nivel de significación viene determinado por :

$$\alpha = 0.05$$

como el contraste es bilateral, la región crítica viene dada por:

$$F_{\text{exp}} < F_{n_1-1, n_2-1; \frac{\alpha}{2}} \quad o \quad F_{\text{exp}} > F_{n_1-1, n_2-1; 1-\frac{\alpha}{2}}$$

Operando tenemos que:

$$F_{\text{exp}} = \frac{S_1^2}{S_2^2} = \frac{31^2}{26^2} = 1.4216$$

y buscando en la tabla de la F-Snedecor , resulta:

$$\begin{aligned} F_{n_1-1, n_2-1; 1-\frac{\alpha}{2}} &= F_{3,3;0.975} = 15.44 \\ F_{n_1-1, n_2-1; \frac{\alpha}{2}} &= F_{3,3;0.025} = \frac{1}{F_{3,3;0.975}} = \frac{1}{15.44} = 0.0647 \end{aligned}$$

Como $1.4216 > 0.0647$ y $1.4216 < 15.44$ no rechazamos la hipótesis nula; por tanto podemos considerar las varianzas iguales.

Retomamos nuestro contraste de diferencia de medias:

$$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$$

Puesto que las varianzas son desconocidas pero pueden considerarse iguales, consideraremos el siguiente estadístico:

$$T = \frac{\overline{X}_1 - \overline{X}_2 - d_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

siendo

$$S_p^2 = \frac{S_1^2 \cdot (n_1 - 1) + S_2^2 (n_2 - 1)}{n_1 + n_2 - 2}$$

donde

$$S_p = \sqrt{\frac{31^2 \cdot (4 - 1) + 26^2 \cdot (4 - 1)}{4 + 4 - 2}} = 28.609$$

y

$$T_{\text{exp}} = \frac{(512 - 492)}{28.609 \sqrt{\frac{1}{4} + \frac{1}{4}}} = 0.988$$

Rechazaremos la hipótesis nula si:

$$T_{\text{exp}} \leq t_{\frac{\alpha}{2}} \quad o \quad T_{\text{exp}} \geq t_{1-\frac{\alpha}{2}}$$

$$T_{\text{exp}} \leq t_{0.025} \quad o \quad T_{\text{exp}} \geq t_{0.975}$$

Buscando en la tabla de la t-Student, tenemos:

$$T_{\text{exp}} \leq -2.44691 \quad \text{o} \quad T_{\text{exp}} \geq 2.44691$$

Como 0.99 se encuentra en $(-2.44691, 2.44691)$ no se puede rechazar la hipótesis nula. Aunque la diferencia entre las dos medias de las muestras sea muy grande, las muestras son tan pequeñas que los resultados no son concluyentes; es decir, la diferencia entre las dos medias de las muestras pueden deberse al azar.

Ejercicio 13 *Un constructor afirma que el 70% de las casas tienen calefacción. ¿Se estaría de acuerdo con tal afirmación si una inspección aleatoria muestra que 8 de 15 casas cuentan con calefacción? ($\alpha = 0.1$)*

El ejercicio nos pide la realización de un contraste de hipótesis para una proporción, donde el tamaño de la muestra n , no es superior a 30 (no podemos aplicar el TCL).

Realicemos las siguientes consideraciones teóricas para pruebas relacionadas con proporciones:

Pruebas relacionadas con proporciones

Hipótesis Nula: La proporción de éxitos en un experimento binomial es igual a cierto valor específico.

$$H_0 : p = p_0$$

Hipótesis Alternativa:

$$a) \quad H_1 : p > p_0 \quad b) \quad H_1 : p < p_0 \quad c) \quad H_1 : p \neq p_0$$

Estadístico:

$$\mathbf{X} \equiv \text{Variable aleatoria binomial con } p = p_0.$$

Región Crítica:

$$a) \quad \text{Sea } P = p[X \geq \text{número de éxitos en una muestra de tamaño } n, \text{ siendo } p = p_0] \\ \text{Si } P \geq \alpha \Rightarrow \text{Se rechaza } H_0$$

$$b) \quad \text{Sea } P = p[X \leq \text{número de éxitos en una muestra de tamaño } n, \text{ siendo } p = p_0] \\ \text{Si } P \leq \alpha \Rightarrow \text{Se rechaza } H_0$$

$$c) \quad \text{Si } x < p_0 \cdot n \\ \text{Sea } P = 2 \cdot p[X \leq \text{número de éxitos en una muestra de tamaño } n, \text{ siendo } p = p_0] \\ \text{Si } x > p_0 \cdot n \\ \text{Sea } P = 2 \cdot p[X \geq \text{número de éxitos en una muestra de tamaño } n, \text{ siendo } p = p_0]$$

$$\text{Si } P \leq \alpha \Rightarrow \text{Se rechaza } H_0$$

NOTAS:

- Cuando p_0 está próximo a 0 ó 1 se puede aproximar por la distribución de Poisson con $\lambda = np_0$.
- Si p_0 no está cerca de 0 ó 1 y $n \gg$ realizaremos la aproximación a una normal con media np_0 y $\sigma^2 = np_0q_0$.
- En el caso particular de diferencias de proporciones tendremos lo siguiente:

$$\begin{aligned} H_0 &: p_1 - p_2 = \delta \\ H_1 &: p_1 - p_2 > \delta \quad o \quad p_1 - p_2 < \delta \quad o \quad p_1 - p_2 \neq \delta \end{aligned}$$

$$\text{Estadísticos: } \hat{p}_1 - \hat{p}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2}$$

- Si $n \gg$, consideramos $\mu = p_1 - p_2$, $\sigma_{\hat{p}_1 - \hat{p}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$ y aplicamos el teorema central del límite.
- Si $p_1 = p_2 = p$ y $q_1 = q_2 = q$

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim N(0, 1)$$

Donde p y q se sustituirán por sus estimadores:

$$\begin{aligned} \hat{p} &= \frac{x_1 + x_2}{n_1 + n_2} \\ \hat{q} &= 1 - \hat{p} \end{aligned}$$

Solución.-

Nuestro contraste vendrá dado por:

$$\begin{aligned} H_0 &: p = 0.7 \\ H_1 &: p \neq 0.7 \end{aligned}$$

Siendo

$$\begin{aligned} X &\equiv \text{Nº de casas con calefacción.} \\ X &\sim B(0.7, 15) \end{aligned}$$

nuestro estadístico experimental resulta,

$$x = 8 \quad np_0 = 10.5$$

y

$$P = 2p[X \leq 8] = 2 \sum_{x=0}^8 p[X \leq x] = 0.2622 > 0.1$$

Como $0.2622 > 0.1$ no hay razones suficientes para rechazar la afirmación del constructor.

Ejercicio 14 Observar, de cara a una simulación, si el tiempo de vida de cierto microorganismo puede considerarse $N(3.5, 0.7)$

Límites	0.9-1.95	1.95-2.45	2.45-2.95	2.95-3.45	3.45-3.95	3.95-4.45	4.45-4.62	TOTAL
o_i	2	1	4	15	10	5	3	40

Solución.-

En esta situación la hipótesis nula será la siguiente:

$$H_0 : \text{El tiempo de vida del microorganismo se puede considerar } N(3.5, 0.7)$$

En primer lugar, para poder aplicar el test χ^2 de bondad de ajuste, deberemos determinar las frecuencias esperadas:

$$\bullet p[X < 1.95] = p\left[\frac{X-3.5}{0.7} < \frac{1.95-3.5}{0.7}\right] = p[Z < -2.21] = 1 - p[Z < 2.21] = 0.0135 \Rightarrow e_1 = 40 \cdot 0.0135$$

$$e_1 = 0.542$$

$$\bullet p[1.95 < X < 2.45] = p\left[\frac{1.95-3.5}{0.7} < Z < \frac{2.45-3.5}{0.7}\right] = 0.05326 \Rightarrow e_2 = 40 \cdot 0.05326$$

$$e_2 = 2.1304$$

•

$$\bullet p[X > 4.45] = p\left[\frac{X-3.5}{0.7} > \frac{4.45-3.5}{0.7}\right] = p[Z > 1.3571] = 0.0886 \Rightarrow e_7 = 40 \cdot 0.0886 = 3.54$$

Finalmente obtenemos la tabla de frecuencias esperadas:

límites	1.45-1.95	1.95-2.45	2.45-2.95	2.95-3.45	3.45-3.95	3.95-4.45	4.45-4.95
e_i	0.5	2.1	5.9	10.3	10.7	7	3.5

podemos observar que se tienen intervalos adyacentes donde las frecuencias esperadas son menores que 5. En esta situación hemos de aplicar los criterios de corrección; reduciremos el número de intervalos de 7 a 4 con el fin de que ninguna frecuencia esperada sea inferior a 5.

o_i	7	15	10	8
e_i	8.5	10.3	10.7	10.5

nos encontramos en condiciones de poder aplicar el Test Chi-cuadrado de bondad de ajuste:

$$\begin{aligned}\chi^2 &= \sum_i \frac{(o_i - e_i)^2}{e_i} \sim \chi_{n-1}^2 \\ \chi_{\text{exp}}^2 &= \frac{(7 - 8.5)^2}{8.5} + \frac{(15 - 10.3)^2}{10.3} + \frac{(10 - 10.7)^2}{10.7} + \frac{(8 - 10.5)^2}{10.5} \\ \chi_{\text{exp}}^2 &= 3.05\end{aligned}$$

Región Crítica:

$$\chi^2 > \chi_{0.95}^2(3) = 7.81472$$

como $7.8147 > 3.05$ no se puede rechazar H_0 . Por ello consideraremos que $N(3.5, 0.7)$ proporciona un buen ajuste al “tiempo de vida de los microorganismos”.

Ejercicio 15 Tenemos una m.a.s. de 200 hombres casados, con su nivel de educación y el número de hijos del matrimonio:

o_{ij}	0-1	2-3	>3
Primaria	14	37	32
Secundaria	19	42	17
Bachiller	12	17	10

Verificar la hipótesis de que el número de hijos de una familia es independiente del nivel de estudios alcanzado por sus padres.

Solución.-

Tendremos que realizar un contraste no paramétrico de independencia. Nuestra hipótesis nula será:

H_0 : Las características nivel de estudios y número de hijos son independientes entre sí.

Utilizaremos el Test χ^2 de independencia.

$$\chi^2 = \sum_i \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim \chi_{n-1}^2$$

Definamos los siguientes sucesos:

L_1 = Un individuo posee 0-1 hijos
 L_2 = Un individuo posee 2-3 hijos
 L_3 = Un individuo posee más de 3 hijos
 M_1 = Un individuo posee nivel de primaria
 M_2 = Un individuo posee nivel de secundaria
 M_3 = Un individuo posee nivel de bachiller

Sea la tabla de contingencia:

Fr. Observadas	L_1	L_2	L_3	$L_1 + L_2 + L_3$
M_1	14	37	32	83
M_2	19	42	17	78
M_3	12	17	10	39
$M_1 + M_2 + M_3$	45	96	59	200

Calculemos las frecuencias esperadas bajo hipótesis de independencia:

$$p(L_1 \cap M_1) = p(L_1) \cdot p(M_1) = \frac{45}{200} \cdot \frac{83}{200} \Rightarrow e_{11} = \frac{45}{200} \cdot \frac{83}{200} \cdot 200 = 18.67$$

$$p(L_1 \cap M_2) = p(L_1) \cdot p(M_2) = \frac{45}{200} \cdot \frac{78}{200} \Rightarrow e_{11} = \frac{45}{200} \cdot \frac{78}{200} \cdot 200 = 17.6$$

$$p(L_1 \cap M_3) = p(L_1) \cdot p(M_3) = \frac{45}{200} \cdot \frac{39}{200} \Rightarrow e_{11} = \frac{45}{200} \cdot \frac{39}{200} \cdot 200 = 8.7$$

$$p(L_2 \cap M_1) = p(L_2) \cdot p(M_1) = \frac{96}{200} \cdot \frac{83}{200} \Rightarrow e_{11} = \frac{96}{200} \cdot \frac{83}{200} \cdot 200 = 39.84$$

.....

$$p(L_3 \cap M_3) = p(L_3) \cdot p(M_3) = \frac{59}{200} \cdot \frac{39}{200} \Rightarrow e_{11} = \frac{59}{200} \cdot \frac{39}{200} \cdot 200 = 11.5$$

Resultando la siguiente tabla de frecuencias esperadas:

e_{ij}	0-1	2-3	>3	
Primaria	18.6	39.8	24.4	83
Secundaria	17.6	37.4	23.01	78
Bachiller	8.7	18.7	11.5	39
	45	96	59	200

La Región Crítica viene dada por:

$$\chi_{\text{exp}}^2 > \chi_{4,0.95}^2$$

y

$$\chi_{\text{exp}}^2 = \frac{(14 - 18.6)^2}{18.6} + \frac{(37 - 39.8)^2}{39.8} + \dots + \frac{(10 - 11.5)^2}{11.5} = 7.54$$

Como $\chi_{\text{exp}}^2 = 7.54 < \chi_{0.95}^2 = 9.48773$ no rechazamos H_0 , es decir, no hay suficiente información para poder afirmar que el n° de hijos y el nivel de estudios no son independientes, y por tanto aceptamos la hipótesis de independencia con un nivel de confianza del 95%.

Ejercicio 16 En dos ciudades se llevó a cabo una encuesta sobre el costo de la vida para obtener el gasto promedio en alimentación en familias constituidas por cuatro personas. De cada ciudad se seleccionó aleatoriamente una muestra de 20 familias y observaron sus gastos semanales en alimentación. Obteniéndose los siguientes datos:

$$\begin{aligned} \bar{X}_1 &= 135 & \bar{X}_2 &= 122 \\ S_1 &= 15 & S_2 &= 10 \end{aligned}$$

Si se supone que se muestrearon dos poblaciones independientes con distribución normal cada una y varianzas iguales, obtener los intervalos de confianza estimados del 95% y del 99% para $\mu_1 - \mu_2$. ¿Podríamos concluir que existe una diferencia real entre μ_1 y μ_2 ?

Solución.-

Queremos construir un intervalo de confianza para la diferencia de medias en poblaciones normales con varianzas desconocidas pero que pueden ser consideradas iguales.

Nuestro intervalo de confianza al nivel de confianza $(1 - \alpha) \cdot 100\%$ para la diferencia de medias será el siguiente:

$$\left(\overline{X}_1 - \overline{X}_2 - t_{1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad , \quad \overline{X}_1 - \overline{X}_2 + t_{1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

siendo

$$S_p = \sqrt{\frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}}$$

donde $t_{1-\frac{\alpha}{2}}$ es el percentil de una distribución t-Student con $n_1 + n_2 - 2$ grados de libertad. En nuestro caso $n_1 = n_2$. Por tanto el intervalo quedaría:

$$\left(\overline{X}_1 - \overline{X}_2 - t_{1-\frac{\alpha}{2}} S_p \sqrt{\frac{2}{n_1}} \quad , \quad \overline{X}_1 - \overline{X}_2 + t_{1-\frac{\alpha}{2}} S_p \sqrt{\frac{2}{n_2}} \right)$$

y

$$S_p = \sqrt{\frac{(n_1 - 1) (S_1^2 + S_2^2)}{2 \cdot n_1 - 2}}$$

El valor de S_p será el siguiente:

$$S_p = \sqrt{\frac{19 \cdot (15^2 + 10^2)}{2 \cdot 20 - 2}} = 12.748$$

El valor de los percentiles, para un t-Student con 38 grados de libertad, será:

$$\begin{aligned} t_{1-\frac{0.05}{2}} &= t_{0.975} = 2.025 \\ t_{1-\frac{0.01}{2}} &= t_{0.995} = 2.716 \end{aligned}$$

Si $\alpha = 0.05 \rightarrow$ Intervalo al 95%

$$\begin{aligned} &\left(135 - 122 - t_{0.975} \cdot 12.748 \sqrt{\frac{2}{20}} \quad , \quad 135 - 122 + t_{0.975} \cdot 12.748 \sqrt{\frac{2}{20}} \right) \\ &\left(135 - 122 - 2.025 \cdot 12.748 \sqrt{\frac{2}{20}} \quad , \quad 135 - 122 + 2.025 \cdot 12.748 \sqrt{\frac{2}{20}} \right) \\ &(4.8367, 21.163) \end{aligned}$$

Al nivel de confianza del 95% se puede considerar $\mu_1 \neq \mu_2$ porque $0 \notin (4.8367, 21.163)$.

Si $\alpha = 0.01 \rightarrow$ Intervalo al 99%

$$\begin{aligned} &\left(135 - 122 - t_{0.995} \cdot 12.748 \sqrt{\frac{2}{20}} \quad , \quad 135 - 122 + t_{0.995} \cdot 12.748 \sqrt{\frac{2}{20}} \right) \\ &\left(135 - 122 - 2.716 \cdot 12.748 \sqrt{\frac{2}{20}} \quad , \quad 135 - 122 + 2.716 \cdot 12.748 \sqrt{\frac{2}{20}} \right) \\ &(2.0511, 23.949) \end{aligned}$$

Efectivamente, podemos concluir que existen diferencias significativas para la diferencia de medias poblacionales, ya que el cero no pertenece a ninguno de los intervalos.

Ejercicio 17 Una compañía textil afirma que a lo sumo el 20% del público compra ropa de lana. Verifica esta afirmación para $\alpha = 0.01$, si una encuesta aleatoria indica que 46 de 200 clientes compran ropa de lana.

Solución.-

Sea

$X \equiv N^\circ$ de personas que utilizan una prenda de lana.

donde

$$X \sim B(200, 0.2)$$

Como n es suficientemente grande por el Teorema Central del Límite podemos aproximar esta distribución a una

$$N(np, \sqrt{npq})$$

Nuestro contraste será el siguiente:

$$\begin{cases} H_0 : p = 0.2 \\ H_1 : p \neq 0.2 \end{cases}$$

En esta situación la región de aceptación, a un nivel de significación $1 - \alpha$, de la hipótesis nula es:

$$\left(np - Z_{1-\frac{\alpha}{2}} \sqrt{npq}, np + Z_{1-\frac{\alpha}{2}} \sqrt{npq} \right)$$

resultando:

$$\left(200 * 0.2 - 2.58 * \sqrt{200 * 0.2 * 0.8}, 200 * 0.2 + 2.58 * \sqrt{200 * 0.2 * 0.8} \right)$$

$$(25.405, 54.595)$$

Como 46 está en la región de aceptación, concluiremos diciendo que no podemos rechazar la hipótesis nula, es decir no hay datos suficientes como para afirmar que la compañía textil se equivoca en sus afirmaciones.

Ejercicio 18 Se llevó a cabo un estudio para determinar el grado en el cual el alcohol entorpece la habilidad de pensamiento para llevar a cabo determinada tarea. Se seleccionaron al azar 10 personas de distintas características y se les pidió que participaran en el experimento. Después de proporcionarles la información pertinente, cada persona llevo a cabo la tarea sin nada de alcohol en su organismo. Entonces la tarea volvió a llevarse a cabo, después de que cada persona hubiese consumido una cantidad de alcohol para tener un contenido en su organismo de 0,1%. Supóngase que los tiempos antes y después (en minutos) son los siguientes:

Participante	1	2	3	4	5	6	7	8	9	10
Antes	27	24	52	45	32	35	40	30	37	20
Despues	39	45	71	61	46	58	51	34	48	30

¿Influye el consumo de alcohol en la realización de esta tarea? ($\alpha = 0.05$)

Solución.-

Estamos ante una situación de muestras pareadas, es decir, las muestras están formadas por pares de observaciones. Sea X_1 la variable que nos da el tiempo de reacción de un individuo antes de consumir alcohol, y sea X_2 la v.a. que nos da el tiempo de reacción de un individuo que si ha consumido alcohol. Hemos de suponer en ambos casos que:

$$X_1 \sim N(\mu_1, \sigma_1) \quad y \quad X_2 \sim N(\mu_2, \sigma_2)$$

Las hipótesis a contrastar son:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases} \Leftrightarrow \begin{cases} H_0 : \mu_d = \mu_1 - \mu_2 = 0 \\ H_1 : \mu_d = \mu_1 - \mu_2 < 0 \end{cases}$$

Sea

$$X_d = X_1 - X_2$$

la v.a. de la diferencia,

<i>Antes</i>	27	24	52	45	32	35	40	30	37	20
<i>Despues</i>	39	45	71	61	46	58	51	34	48	30
<i>Diferencia</i>	-12	-21	-19	-16	-14	-25	-11	-14	-11	-10

en tal caso los valores obtenidos como diferencia de las parejas de observaciones muestrales anteriores, representan una muestra aleatoria de X_d .

El estadístico que utilizaremos para resolver el contraste será:

$$T = \frac{\bar{X}_d - \mu_d}{S_d / \sqrt{n}} \sim_{H_0} t_{n-1}$$

Siendo:

$$\begin{aligned} \bar{X}_d &= -15.3 \\ S_d^2 &= 27.17 \\ n &= 10 \\ \alpha &= 0.05 \end{aligned}$$

La Región Crítica viene dada por:

$$T_{\text{exp}} < t_{9,0.05} = -t_{9,0.95}$$

Realizando los cálculos oportunos tenemos:

$$T_{\text{exp}} = \frac{-15.3}{5.21 / \sqrt{10}} = -9.28$$

$$t_{9,0.05} = -t_{9,0.95} = -1.83$$

Como -9.28 es menor que -1.83, rechazamos la hipótesis nula; es decir el tiempo medio de reacción de una persona aumenta si ha consumido alcohol.

6.16.2 EJERCICIOS PROPUESTOS

Ejercicio 19 Se consideran los siguientes tiempos de reacción de un producto químico, en segundos:

3.4, 2.2, 3.3, 3.4, 3.5, 3.3, 4.1, 2.4, 3.3

Se pide:

1. Obtener un intervalo de confianza del 95% para el tiempo medio de reacción suponiendo la variable normal, con desviación típica $\sigma = 0.4$.
2. Determinar al tamaño muestral necesario para estimar el tiempo medio de reacción con un error menor que 0,1 con probabilidad 0,9.

Ejercicio 20 A diez trabajadores de un departamento se les aplicó un control de velocidad en el trabajo en dos momentos punta del día, cronometrando los minutos empleados en pulir una pieza a medio terminar, se obtiene:

Trabajador	1	2	3	4	5	6	7	8	9	10
1ª hora (min)	6	5	7.5	5.6	9	2.8	8.7	7.1	7.4	4.6
2ª hora (min)	8.2	6	7.4	5.7	8.5	3.2	7.6	7.9	9.1	5.3

¿Podría afirmarse, con un 6% de significación que la velocidad en el trabajo no cambia sustancialmente entre los dos momentos del día?

Ejercicio 21 Se está estudiando la distribución del color de pelo de una población: Moreno, castaño, rubio y pelirrojo. Extraída una muestra aleatoria, de ella se obtuvieron los siguientes datos:

Color de Pelo	Frecuencia
Moreno	141
Castaño	140
Rubio	99
Pelirrojo	23

Un modelo teórico asigna las siguientes probabilidades a cada uno de los grupos:

Color pelo	Frecuencia
Moreno	r^2
Castaño	$p^2 + 2pr$
Rubio	$q^2 + 2qr$
Pelirrojo	$2pq$

A partir de los datos de la muestra se obtuvieron las siguientes estimaciones de los parámetros:

$$\hat{p} = 0.2465$$

$$\hat{q} = 0.1732$$

Contrastar la hipótesis de que los datos se ajustan al modelo teórico.

Ejercicio 22 El fabricante de un determinado producto afirma que el porcentaje de unidades defectuosas que produce es de $p = 0.206$. En una muestra de 125 unidades se observaron 14 defectuosas. ¿A qué nivel de significación puede tener razón el fabricante?. ¿A qué nivel de significación no tendría razón el fabricante?

Ejercicio 23 Se sabe que la longitud en cm. de una determinada especie de coleópteros sigue una distribución normal de varianza 0.25. Capturados 6 ejemplares de dicha especie, sus longitudes fueron:

2.75 1.72 2.91 2.6 2.64 3.34

¿Se puede aceptar la hipótesis de que la población tiene una longitud media de 2.656 cm?

Ejercicio 24 Supongamos que usted es un ingeniero de una planta eléctrica manufacturera. En la cual uno de los productos es un fusible eléctrico, cuya característica más importante es el intervalo de tiempo antes de fundirse, cuando se le somete a una carga específica. Tras llevar a cabo un programa de pruebas se obtuvieron los siguientes datos (en segundos):

Día 1	Día 2
42	69
45	109
68	113
72	118
90	153

1. Obtener un intervalo del 90 por ciento de confianza, de la verdadera diferencia de las medias de las producciones de cada día. Establece todas las hipótesis necesarias e interpreta el resultado.
2. Obtener los límites de confianza del cociente de varianzas teóricas .

Ejercicio 25 Se sabe que la ganancia, en miles de pesetas, que se obtiene en una inversión a un año de 6000 euros, sigue una distribución normal con ganancia media conocida $\mu = 450$ aunque con una varianza desconocida, lo que la hace insegura. Con objeto de estimar mediante un intervalo de confianza dicha varianza, se eligieron al azar diez inversiones del tipo considerado, en donde se obtuvieron los siguientes resultados (en euros):

510, 390, 620, 400, 430, 520, 500, 600, 520, 350.

Determinar un intervalo de confianza del 90% para la desviación típica de la variable ganancia. Obtener otro intervalo del 99% para la misma variable e interpretar los resultados obtenidos.

Ejercicio 26 En un estudio sobre las caries dental de niños de cuatro zonas geográficas con distintos niveles de fluor en el agua, se tomó una muestra de 190 niños de cada zona, obteniéndose los siguientes datos:

Zona	Niños sin caries
A	55
B	21
C	50
D	64

¿Se pueden aceptar como equivalentes las cuatro zonas geográficas respecto a la presencia de caries?

Ejercicio 27 Se tomaron dos muestras de familias, una de familias rurales y otra de familias urbanas, con el objeto de estudiar las diferencias en compras de café de los dos grupos. A continuación se listan los datos obtenidos, en Kg. , comprados por familias en un periodo de 11 meses.

Familias	1	2	3	4	5	6	7	8	9	10	11
Rural	5.49	3.08	4.13	5.03	5.17	6.03	4.45	5.13	4.26	4.62	3.1
Urbana	3.76	4.22	4.17	5.03	4.85	2.09	4.45	3.58	3.86	4.13	4.40

¿Atribuiría usted la diferencia en consumos de café observada en esta muestra a fluctuaciones normales de muestreo o existe una diferencia real entre los consumos de café rural y urbano? . Seleccione su propio nivel de control del error TIPO1 y saque la conclusión correspondiente.

Ejercicio 28 Se quiere estimar la diferencia de uso que se hace de dos impresoras de una determinada empresa, mediante un intervalo de confianza para la diferencia de tiempos medios de utilización de ambas máquinas. Para ello, se controló el tiempo de utilización de cada impresora durante una serie de días elegidos al azar, obteniéndose los siguientes resultados:

Impresora 1	2.4	3.1	2.5	2.6	4.1	2.7	3	2.3	3.2	3
Impresora 2	2.7	1.9	2.4	2	1.8	2.7	2.2	2.4	2.5	1.8

Suponiendo que el tiempo de utilización de cada impresora sigue una distribución normal, determinar un intervalo de confianza al 95% para la diferencia de utilización media de las impresoras suponiendo,

1. Que las varianzas de utilización de ambas máquinas son desconocidas .
2. Determinar un intervalo de confianza del 95% para el cociente de varianzas del tiempo de servicio de las dos impresoras.

Ejercicio 29 En un cierto día se obtuvieron los precios por caja seleccionando al azar diez mercados. Se desea comparar los precios de las manzanas Golden y Arenisca, sabiendo que los precios están normalmente distribuidos y que el precio en una variedad en un mercado está influido por el precio de la otra en el mismo mercado.

Mercado	Golden	Arenisca
1	26.88	29.00
2	27.00	29.25
3	26.63	28.75
4	28.13	30.00
5	27.5	29.25
6	27.25	29.25
7	28.38	30.25
8	27.63	29.50
9	27.88	29.50
10	27.00	28.75

Ejercicio 30 Si $X \sim N(\mu, \sigma)$. ¿Cómo de grande debe ser una muestra si queremos tener un 95% de seguridad de que \bar{X} no se aleja de μ más de $\frac{\sigma}{2}$?

Ejercicio 31 En dos ciudades se llevó a cabo una encuesta sobre el costo de la vida para obtener el gasto promedio en alimentación en familias constituidas por cuatro personas. De cada ciudad se seleccionó aleatoriamente una muestra de 20 familias y observaron sus gastos semanales en alimentación. Obteniéndose los siguientes datos:

$$\begin{aligned}\bar{X}_1 &= 225 & \bar{X}_2 &= 195 \\ S_1 &= 21 & S_2 &= 12\end{aligned}$$

Si se supone que se muestrearon dos poblaciones independientes con distribución normal cada una y varianzas iguales, obtener los intervalos de confianza estimados del 90% y del 99% para $\mu_1 - \mu_2$. ¿Podríamos concluir que existe una diferencia real entre μ_1 y μ_2 ?

Ejercicio 32 Las siguientes son las calificaciones obtenidas por muestras en un examen de personalidad con nueve mujeres casadas y nueve mujeres solteras:

Solteras	88	68	77	82	63	80	78	71	72
Casadas	73	77	67	74	74	64	71	71	72

Además sabemos que estos datos pueden considerarse como muestras aleatorias independientes tomadas de dos poblaciones normales.

1. Contrastar la igualdad de varianzas contra la alternativa $\sigma_1^2 \neq \sigma_2^2$ usando un nivel de significación de 0.05 (σ_1^2 es la varianza de las calificaciones obtenidas por las mujeres solteras y σ_2^2 la de las mujeres casadas)
2. ¿Podemos considerar que las mujeres casadas presenta un nivel de personalidad superior al de las solteras? ($\alpha = 0.05$)

Ejercicio 33 Supóngase que se miden los contenidos de nicotina de dos marcas de cigarrillos. En un experimento de 50 cigarrillos de la marca A y 40 de la marca B se obtienen los siguientes datos:

Marca A	Marca B
$X_A = 2.61 \text{ mg.}$	$X_B = 2.38 \text{ mg.}$
$S_A = 0.12 \text{ mg}$	$S_B = 0.14 \text{ mg}$

Realice el contraste bilateral adecuado para probar la hipótesis de que la marca A posee 0.2 mg. más de nicotina que la marca B. ($\alpha = 0.05$)

Ejercicio 34 Sea $X \sim B(n, p)$. Probar que:

1. $\hat{p} = \frac{X}{n}$ es insesgado de p .
2. $p' = \frac{X + \sqrt{\frac{n}{2}}}{n + \sqrt{n}}$ es sesgado de p .
3. p' es insesgado si $n \rightarrow \infty$

Ejercicio 35 Un estudio realizado sobre 100 usuarios revela que un automóvil recorre anualmente un promedio de 15200 Km con una desviación típica de 2250 Km.

1. (1 punto) Determine un intervalo de confianza, al 99 %, para la cantidad promedio de kilómetros recorridos.
2. (1 punto) ¿Cuál debe ser el tamaño mínimo de la muestra para que el error cometido no sea superior a 500 Km, con igual confianza?

Ejercicio 36 Los resultados de un test de sensibilidad musical realizado a los alumnos de un Conservatorio se distribuyen según una ley Normal de media 65 y desviación típica 18.

1. ¿Cuál es la distribución de la media muestral para muestras de tamaño 16?
2. Para muestras aleatorias de tamaño 100, halle la probabilidad de que su puntuación media esté comprendida entre 63.5 y 67 puntos.

Ejercicio 37 En los individuos de una población, la cantidad de colesterol en sangre se distribuye según una ley normal de media desconocida y varianza de 0.25. Hemos tomado una muestra de 10 individuos, y se ha obtenido una media muestral de 1.7 g/l.

1. Obtenga un intervalo de confianza, al 95 %, para la cantidad media de colesterol en sangre de la población.
2. ¿Qué nivel de confianza tendría un intervalo para la media cuyos límites fuesen 1.2930 y 2.107?
3. ¿Cuál tendría que ser el tamaño muestral para tener una confianza del 98% de que la media muestral no se distancie de la media poblacional mas de 0.2 g/l?

Ejercicio 38 Mendel sembró 529 plantas de guisantes usando semillas del mismo tipo y los frutos resultantes los clasificó atendiendo al color en: verde, verde amarillento y amarillo y atendiendo a la forma en: redondo, levemente rugoso y rugoso. Obtuvo los siguientes datos:

	Verde	Verde-Amarillo	Amarillo
Redondo	35	68	38
Levemente Rugoso	67	138	60
Rugoso	30	68	28

Contrastar la hipótesis de que las clasificaciones por forma y color son independientes

Ejercicio 39 Un físico toma 25 medidas independientes de la gravedad específica de cierto cuerpo. Sabe que las limitaciones de su equipo son tales que la desviación típica de cada medición es s unidades, si bien el valor medio teórico será la gravedad desconocida m .

1. Utilizando la desigualdad de Tchebychev, encuentra una cota inferior para la probabilidad de que el promedio de sus mediciones difiera de la verdadera gravedad del cuerpo m en menos de $s/4$ unidades.
2. Utilizando el Teorema Central del Límite, encuentra un valor aproximado para la probabilidad del apartado anterior.
3. Utiliza de nuevo la desigualdad de Tchebychev para encontrar el número de mediciones n para que el promedio diste de m menos de $s/4$ con probabilidad mayor de 0.99.