OCW-UPM Estadística para Ingeniería Civil y Medioambiental

Autores: E. M. García del Toro, C. Hermoso, E. J. Huertas

PROBLEMAS Y EJERCICIOS RESUELTOS

TEMA 6 - INTERVALOS DE CONFIANZA

Ejercicios resueltos de Inferencia mediante Estimación Puntual e Intervalos de Confianza

En esta serie de ejercicios resueltos, vamos a trabajar con **todas y cada una** de las estimaciones que se incluyen en los contenidos de este Tema. Los tipos de estimaciones que pueden preguntarnos son los siguientes:

1) Estimación puntual:

   a) Estimación de parámetros poblacionales en la distribución binomial

   b) Estimación de parámetros poblacionales en la distribución normal

2) Estimación por intervalos de confianza:

   a) Intervalo de confianza para la media de una población de varianza conocida

   b) Intervalo de confianza para la media de una distribución normal de varianza desconocida

   c) Intervalo de confianza para la varianza poblacional de una distribución normal

   d) Intervalo de confianza para la proporción (distribución binomial)

Observación importante sobre notación: En lo que sigue, recuérdese que se denota por z_p, $t_{n;p}$, $\chi^2_{n;p}$, y $F_{n_1,n_2;p}$, respectivamente, el valor de la abcisa de una $N(0,1)$ normal estándar, t_n de Student, $\chi^2_{n;p}$ de Pearson, y $F_{n_1,n_2;p}$ de Fisher, que deja a su **derecha** -bajo la correspondiente función de densidad- un área de probabilidad p.
El cálculo de estas abcisas (también llamadas **valores críticos**) puede realizarse rápidamente en Statgraphics como se ha visto en las précticas de la asignatura, o, si no disponemos de computador, puede hacerse mediante las tablas de las respectivas distribuciones.

Estimación puntual y por intervalo de confianza para la media de una población normal

Varianza poblacional σ^2 conocida

Ejercicio 1 - Se desea estimar el caudal medio μ (en metros cúbicos por segundo) de cierto producto químico que circula por la tubería de una fábrica. Se supone que dicho caudal sigue una distribución normal, de media μ y de varianza (poblacional) conocida $\sigma^2 = 144\ \left(m^3/s\right)^2$. Se llevan a cabo las siguientes 25 medidas (en m^3/s) de dicho caudal, en instantes elegidos al azar:

90.8,	81.5,	83.2,	79.7,	93.3,	65.0,	88.9,	90.9,	85.0,	72.5,
77.6,	101.4,	82.5,	93.6,	105.9,	71.1,	87.2,	75.0,	93.6,	61.8,
108.9,	88.7,	93.7,	80.6,	72.6					

a) Hallar una estimación puntual del caudal medio μ de producto químico que circula por la tubería y de la varianza σ^2, mediante el estimador puntual que considere más adecuado en cada caso **(Ejemplos del caso** $1b$**) Estimación puntual de parámetros poblacionales en la distribución normal)**.
b) Se considera que la estimación puntual obtenida en el apartado anterior puede estar bastante alejada del verdadero valor poblacional del caudal medio μ que circula por la tubería. Para realizar entonces una inferencia más realista de este valor, se pide calcular tres Intervalos de Confianza distintos para μ, uno al 99%, otro al 95% y otro al 90% de confianza (**Ejemplos del caso** $2a$**) Estimación por Intervalo de confianza para la media de una población de varianza conocida)**.
c) Si forzamos al Intervalo de Confianza a tener un tamaño de 2 m^3/s, ¿qué nivel de confianza le corresponde a este intervalo?

Solución ejercicio 1:

Apartado a) En lo que sigue denotaremos por n al tamaño muestral, en este caso igual a 25. Según se ha estudiado en las clases de teoría, el estimador más adecuado para la media poblacional desconocida μ es la *media muestral*

$$\bar{x} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right).$$

Utilizando Statgraphics (o una calculadora) con la muestra proporcionada en el enunciado, obtenemos la media muestral $\bar{x} = 85\ m^3/s$.

Para estimar la varianza poblacional σ^2, el estimador puntual más adecuado es la *cuasi-varianza muestral*

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right),$$

que en Statgraphics (o con una calculadora) también se calcula directamente. Para esta muestra es $s^2 = 139.04$ $(m^3/s)^2$. La cuasivarianza tiene unidades de la media al cuadrado.

Apartado b) En las clases de teoría se ha visto que la expresión para este Intervalo de Confianza es:

$$I = \left[\bar{x} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right], \tag{1}$$

donde $z_{\alpha/2}$ es el valor de la abcisa en la distribución normal estándar $N(0,1)$ que deja a su derecha un área de valor $\alpha/2$ (ver figura 1)

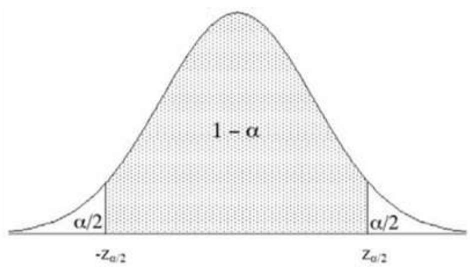

Figure 1: Valores críticos y gráfica de la pdf de la distribución Normal Estándar N(0,1)

Dado que la confianza es $(1-\alpha)$, y queremos tres intervalos $I_{0.99}$, $I_{0.95}$ e $I_{0.90}$, los respectivos valores de α son 0.01, 0.05, y 0.1. Estos valores de α nos sirven para calcular $z_{\alpha/2}$ en cada uno de los tres casos. Utilizando la herramienta `disttool`, tenemos que (esto puede hacerse también con la tabla de la N(0,1))

Para $\alpha = 0.01$, $\alpha/2 = 0.005$ y $z_{\alpha/2} = z_{0.005} = 2.5758$.
Para $\alpha = 0.05$, $\alpha/2 = 0.025$ y $z_{\alpha/2} = z_{0.025} = 1.96$.
Para $\alpha = 0.1$, $\alpha/2 = 0.05$ y $z_{\alpha/2} = z_{0.05} = 1.6449$.

Trabajo del alumno: Obtener de nuevo los valores críticos anteriores $z_{0.005}$, $z_{0.025}$, y $z_{0.05}$, pero esta vez utilizando el comando `norminv`. Obviamente, el resultado debe ser el mismo.

Los tres intervalos pedidos son entonces:

$$\begin{aligned}
I_{0.99} &= \left[\bar{x} - z_{0.005}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{0.005}\frac{\sigma}{\sqrt{n}}\right] = \left[85 - 2.5758\frac{\sqrt{144}}{\sqrt{25}},\ 85 + 2.5758\frac{\sqrt{144}}{\sqrt{25}}\right] = [78.82\,,\ 91.18]\ m^3/s,\\
I_{0.95} &= \left[\bar{x} - z_{0.025}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{0.025}\frac{\sigma}{\sqrt{n}}\right] = \left[85 - 1.96\frac{\sqrt{144}}{\sqrt{25}},\ 85 + 1.96\frac{\sqrt{144}}{\sqrt{25}}\right] = [80.30\,,\ 89.7]\ m^3/s,\\
I_{0.90} &= \left[\bar{x} - z_{0.05}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{0.05}\frac{\sigma}{\sqrt{n}}\right] = \left[85 - 1.6449\frac{\sqrt{144}}{\sqrt{25}},\ 85 + 1.6449\frac{\sqrt{144}}{\sqrt{25}}\right] = [81.05\,,\ 88.95]\ m^3/s,
\end{aligned}$$

cuyos tamaños respectivos son $12.36\ m^3/s$, $9.4\ m^3/s$ y $7.9\ m^3/s$.

Todas las probabilidades que se calculan en esta práctica, pueden calcularse también con esta herramienta.

apartado c)
En la fórmula (1) vemos que, en esta caso, el ancho del intervalo es siempre

$$L = 2 \cdot z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Si obligamos a que L tenga tamaño fijo, como conocemos σ y n, podemos despejar $z_{\alpha/2}$ y con esto averiguar α. La confianza asociada a este intervalo de longitud fija será simplemente $(1-\alpha)$. En nuestro caso:

$$\begin{aligned} 2 &= 2 \cdot z_{\alpha/2} \frac{\sqrt{144}}{\sqrt{25}} \to z_{\alpha/2} = \frac{\sqrt{25}}{\sqrt{144}} = 0.4167 \underset{MATLAB}{\to} \\ & \quad z_{\alpha/2} \text{ deja a su izquierda un área de } 0.66155, \text{ por lo que} \\ \alpha/2 &= 1 - 0.66155 = 0.33845 \to \alpha = 2 \cdot 0.33845 = 0.677 \end{aligned}$$

que es un valor de α enorme, y que corresponde a un valor de confianza realmente muy bajo, de tan solo $(1-\alpha) = 1 - 0.677 = 0.323 \to 32.3\%$ de confianza.

Varianza poblacional σ^2 desconocida

d) En el mismo ejercicio anterior, hallar de nuevo los tres intervalos de confianza, pero en el caso en que la varianza poblacional σ^2 sea desconocida **(ejemplo del caso** $2b$**))**.

Solución:
apartado d) Si, como suele ocurrir en casos reales, no se conoce la varianza poblacional σ^2, habrá que estimarla con la cuasi-varianza muestral de las 25 observaciones. En este caso, vimos en las clases de teoría que este Intervalo de Confianza viene dado por:

$$I = \left[\bar{x} \mp t_{n-1;\alpha/2} \frac{s}{\sqrt{n}}\right] = \left[\bar{x} - t_{n-1;\alpha/2} \frac{s}{\sqrt{n}},\ \bar{x} + t_{n-1;\alpha/2} \frac{s}{\sqrt{n}}\right], \tag{2}$$

donde $t_{n-1;\alpha/2}$ es el valor de la abcisa en la distribución t de Student de $n-1$ grados de libertad, que deja a su derecha un área de valor $\alpha/2$ (ver figura 3)

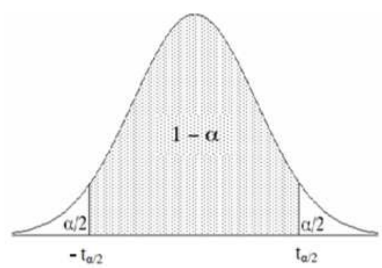

Figure 2: Valores críticos y gráfica de la pdf de la distribución t de Student

Si la estimación puntual obtenida fue $s^2 = 139.04$, utilizando por ejemplo el comando `tinv` de MATLAB para calcular $t_{24;0.005} =$ `tinv(1-0.005,24)`$= 2.7969$, $t_{24;0.025} =$ `tinv(1-0.025,24)`$= 2.0639$, $t_{24;0.05} =$ `tinv(1-0.05,24)`$=$ 1.7109, tendremos los siguientes Intervalos de Confianza:

$$\begin{aligned} \tilde{I}_{0.99} &= \left[\bar{x} - t_{24;0.005} \frac{s}{\sqrt{n}},\ \bar{x} + t_{24;0.005} \frac{s}{\sqrt{n}}\right] = \left[85 - 2.7969 \frac{\sqrt{139.04}}{\sqrt{25}},\ 85 + 2.7969 \frac{\sqrt{139.04}}{\sqrt{25}}\right] = [78.40\,,\ 91.60]\ m^3/s, \\ \tilde{I}_{0.95} &= \left[\bar{x} - t_{24;0.025} \frac{s}{\sqrt{n}},\ \bar{x} + t_{24;0.025} \frac{s}{\sqrt{n}}\right] = \left[85 - 2.0639 \frac{\sqrt{139.04}}{\sqrt{25}},\ 85 + 2.0639 \frac{\sqrt{139.04}}{\sqrt{25}}\right] = [80.13\,,\ 89.87]\ m^3/s, \\ \tilde{I}_{0.90} &= \left[\bar{x} - t_{24;0.05} \frac{s}{\sqrt{n}},\ \bar{x} + t_{24;0.05} \frac{s}{\sqrt{n}}\right] = \left[85 - 1.7109 \frac{\sqrt{139.04}}{\sqrt{25}},\ 85 + 1.7109 \frac{\sqrt{139.04}}{\sqrt{25}}\right] = [80.97\,,\ 89.03]\ m^3/s, \end{aligned}$$

cuyos tamaños respectivos son 13.2 m^3/s, 9.74 m^3/s y 8.06 m^3/s. Fijémonos que los resultados son parecidos a los del apartado **b)**.

Trabajo del alumno: Obtener de nuevo los valores críticos anteriores $t_{24;0.005}$, $t_{24;0.025}$, y $t_{24;0.05}$, pero esta vez utilizando la herramienta `disttool`. Obviamente, el resultado debe ser el mismo.

En esta práctica, seguimos trabajando los siguientes casos de *inferencia estadística*, de acuerdo con los contenidos del correspondiente tema de la asignatura. En esta ocasión hallaremos intervalos de confianza para la varianza poblacional de una distribución normal, e intervalos de confianza para la proporción.

En general, en el tema de *Estimación puntual y por intervalos de confianza*, los casos a estudio son los siguientes:

1) Estimación puntual:

   a) Estimación de parámetros poblacionales en la distribución binomial

   b) Estimación de parámetros poblacionales en la distribución normal

2) Estimación por intervalos de confianza:

   a) Intervalo de confianza para la media de una población de varianza conocida

   b) Intervalo de confianza para la media de una distribución normal de varianza desconocida

   c) Intervalo de confianza para la varianza poblacional de una distribución normal

   d) Intervalo de confianza para la proporción (distribución binomial)

Observación importante sobre notación: En lo que sigue, recuérdese que se denota por z_p, $t_{n;p}$, $\chi^2_{n;p}$, y $F_{n_1,n_2;p}$, respectivamente, el valor de la abcisa de una $N(0,1)$ normal estándar, t_n de Student, $\chi^2_{n;p}$ de Pearson, y $F_{n_1,n_2;p}$ de Fisher, que deja a su **derecha** -bajo la correspondiente función de densidad- un área de probabilidad p.

El cálculo de estas abcisas (también llamadas **valores críticos**) puede realizarse en MATLAB, respectivamente mediante los siguientes comandos:

```
norminv   %Inverse of the normal cumulative distribution function (cdf).
    X = norminv(P,MU,SIGMA) returns the inverse cdf for the normal
    distribution with mean MU and standard deviation SIGMA, evaluated at
    the values in P. Default values for MU and SIGMA are 0 and 1, respectively.

tinv      %Inverse of Student's T cumulative distribution function (cdf).
    X=tinv(P,V) returns the inverse of Student's T cdf with V degrees
    of freedom, at the values in P.

chi2inv   %Inverse of the chi-square cumulative distribution function (cdf).
    X = chi2inv(P,V)  returns the inverse of the chi-square cdf with V
    degrees of freedom at the values in P.

finv   %Inverse of the F cumulative distribution function.
    X=finv(P,V1,V2) returns the inverse of the F distribution
    function with V1 and V2 degrees of freedom, at the values in P.
```

También podremos hallar estos valores mediante la herramienta gráfica `disttool` que ya hemos visto en otras ocasiones, o en última instancia, si no disponemos de computador, mediante las tablas de las respectivas distribuciones.

Estimación por Intervalo de Confianza para la varianza σ^2 de una población normal

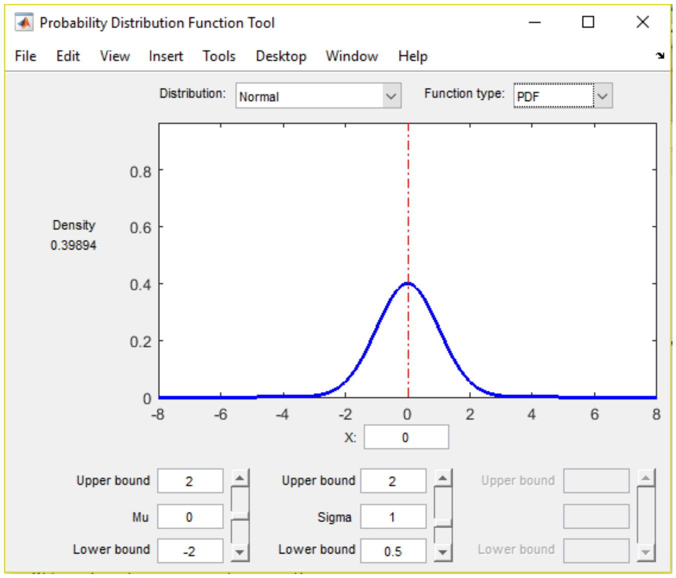

Figure 3: Ventana de la herramienta `disttool` en MATLAB

Considérese el mismo ejercicio de la práctica anterior, sobre el caudal de cierto químico que circula por la tubería de una fábrica:

Se desea estimar el caudal medio μ (en metros cúbicos por segundo) de cierto producto químico que circula por la tubería de una fábrica. Se supone que dicho caudal sigue una distribución normal, de media μ y de varianza (poblacional) conocida $\sigma^2 = 144\ \left(m^3/s\right)^2$. Se llevan a cabo las siguientes 25 medidas (en m^3/s) de dicho caudal, en instantes elegidos al azar:

90.8, 81.5, 83.2, 79.7, 93.3, 65.0, 88.9, 90.9, 85.0, 72.5, 77.6, 101.4, 82.5, 93.6, 105.9, 71.1, 87.2, 75.0, 93.6, 61.8, 108.9, 88.7, 93.7, 80.6, 72.6

Ahora se va a tratar de obtener un Intervalo de confianza para la varianza poblacional de una distribución normal (lo que corresponde a **ejemplos del caso** $2c$))

a) Hallar un Intervalo de Confianza para la varianza poblacional σ^2 de la variable en estudio, con nivel de confianza del 99%, considerando la media poblacional desconocida.

b) Hallar un Intervalo de Confianza para la varianza poblacional σ^2 de la variable en estudio, con nivel de confianza del 90%, considerando la media poblacional desconocida.

Solución:
apartado a)

Si, como suele ocurrir en casos reales, no se conoce la media poblacional μ, la expresión para este Intervalo de Confianza es siempre:

$$I = \left[\frac{(n-1)s^2}{\chi^2_{n-1;\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{n-1;1-\alpha/2}}\right], \tag{3}$$

donde $\chi^2_{n-1;\alpha/2}$ es el valor de la abcisa en la distribución Chi cuadrado de Pearson, de $n-1$ grados de libertad, que deja a su derecha un área de valor $\alpha/2$. Por otro lado, $\chi^2_{n-1;1-\alpha/2}$ será el valor de la abcisa en la distribución Chi cuadrado de Pearson, de $n-1$ grados de libertad, que deja a su derecha un área de valor $1-\alpha/2$.

Dado que en este caso la confianza es $1-\alpha = 0.99$, el valor de significación es $\alpha = 0.01$, por lo que $\alpha/2 = 0.005$. Como siempre, podemos calcular estos dos valores críticos en MATLAB, bien con la herramienta `disttool`, o bien escribiendo en la linea de comandos `chi2inv(1-0.005,24)` y `chi2inv(0.005,24)`, obteniendo:

$$\begin{aligned}\chi^2_{n-1;0.005} &= \texttt{chi2inv}(1-0.005, 24) = 45.56,\\ \chi^2_{n-1;1-0.005} &= \texttt{chi2inv}(0.005, 24) = 9.886.\end{aligned}$$

Reemplazando estos valores en (3), obtenemos:

$$I_{0.99} = \left[\frac{24\cdot 139.04}{45.56}, \frac{24\cdot 139.04}{9.886}\right] = [73.24\,,\,337.5]\ \left(m^3/s\right)^2. \tag{4}$$

Nótese que, con un tamaño muestral tan pequeño como $n = 25$ muestras solamente, el intervalo de confianza al 99% es tan grande que carece de utilidad práctica y resulta muy poco informativo. Además, debido a que la varianza se mide en unidades de la media **al cuadrado**, hay un efecto de distorsión en el tamaño del intervalo. Los resultados son más fáciles de asimilar y comprender si se expresa en términos de la desviación típica σ. Así, por ejemplo, el intervalo anterior (4), en términos de la desviación típica (que es simplemente la raíz cuadrada de la varianza) queda

$$\sigma^2 \in [73.24\,,\,337.5]\ \left(m^3/s\right)^2 \to \sigma \in [\sqrt{73.24}\,,\,\sqrt{337.5}] = [8.56\,,\,18.37]\ m^3/s$$

que, además de parecer de un tamaño más aceptable, es más lógico pues se expresa en las mismas unidades que la media.

apartado b)

Los cálculos aquí son prácticamente iguales que en el apartado anterior, solamente que tendremos que calcular otros valores críticos, ya que ha cambiado la confianza. Como en este apartado la confianza es $1-\alpha = 0.90$, el valor de significación es $\alpha = 0.1$, por lo que $\alpha/2 = 0.05$. Así

$$\begin{aligned}\chi^2_{n-1;0.05} &= \texttt{chi2inv}(1-0.05, 24) = 36.415,\\ \chi^2_{n-1;1-0.05} &= \texttt{chi2inv}(0.05, 24) = 13.85.\end{aligned}$$

Sustituyendo estos valores en (3), obtenemos el nuevo intervalo de confianza:

$$I_{0.90} = \left[\frac{24\cdot 139.04}{36.42}, \frac{24\cdot 139.04}{13.85}\right] = [91.6\,,\,240.9]\ \left(m^3/s\right)^2. \tag{5}$$

Obsérvese que, ni siquiera reduciendo el grado de confianza al 90%, el intervalo correspondiente anterior (5) gana mucho en precisión. De nuevo, la razón es que la varianza se mide en unidades de la media **al cuadrado**, hay un efecto de distorsión en el tamaño del intervalo. Como en el caso anterior, es más fácil interpretar los resultados si se expresa (5) en términos de la desviación típica σ. Así,

$$\sigma^2 \in [91.6\,,\,240.9]\ \left(m^3/s\right)^2 \to \sigma \in [\sqrt{91.6}\,,\,\sqrt{240.9}] = [9.57\,,\,15.52]\ m^3/s$$

que, además de parecer más aceptable, es más lógico pues se expresa en las mismas unidades que la media.

Estimacion Puntual y por Intervalo de Confianza para la proporción poblacional

Un ayuntamiento desea conocer si en su ciudad se cumple la ley sobre la obligatoriedad del uso del casco reglamentario en motoristas. Para ello, encarga un estudio en el que se observa que, de 464 motoristas, 131 de ellos no llevan casco.

a) Hallar una estimación puntual de la proporción p de motoristas que circulan sin casco **(ejemplo del caso** $1a$**))**.

b) Hallar un intervalo de confianza al 90 % para la proporción poblacional p de motoristas que circulan sin casco **(ejemplo del caso** $2d$**))**. Indique qué condición se cumple para que el intervalo calculado sea válido.

c) Un dirigente de la oposición critica que la política de concienciación del uso del casco en la ciudad es pésima, ya que el 40 % de motoristas circulan actualmente sin este elemento imprescindible de seguridad ¿Se puede dar por cierta dicha afirmación al mismo nivel de confianza?

Solución:

apartado a) Según lo estudiado en las clases de teoría, el estimador más adecuado para la proporción poblacional desconocida p es la proporción muestral $\hat{p}$, donde

$$\hat{p} = \frac{\text{número de veces que aparece el suceso A}}{\text{tamaño muestral } n}.$$

Así, en nuestro caso

$$\hat{p} = \frac{131}{464} = 0.2823 \to 28.23\% \text{ de motoristas circulan sin casco.}$$

apartado b) Hay que hallar un intervalo de confianza al 90 % para la proporción poblacional, cuya expresión vimos en las clases teóricas que es siempre:

$$I = \left[\hat{p} \mp z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\right] = \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\,,\, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\,,\right]. \tag{6}$$

El valor de la proporción muestral $\hat{p} = 0.2823$ lo hemos calculado en el apartado anterior. La confianza del 90 % indica que el nivel de significación es $\alpha = 0.1$ (ya que $90 = (1-\alpha)\cdot 100$). El valor crítico que se necesita en (6) es $z_{\alpha/2} = z_{0.1/2} = z_{0.05}$. Calculando este valor en MATLAB, por ejemplo mediante el comando `norminv`, obtenemos que $z_{0.05} =$ `norminv(1-0.05,0,1)` $= 1.645$. Así pues,

$$\begin{aligned}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} &= \sqrt{\frac{0.28233\,(1-0.28233)}{464}} = \sqrt{\frac{0.28233 \cdot 0.717\,67}{464}} \\ &= \sqrt{4.366\,8 \times 10^{-4}} = 0.020897.\end{aligned}$$

Reemplazando los cálculos en (6) tenemos

$$\begin{aligned} I_{0.90} &= \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\,;\, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\right] \\ &= [0.28233 - 1.645 \cdot 0.020897\,;\, 0.28233 + 1.645 \cdot 0.020897] \\ &= [0.248\,;\, 0.317]\,. \end{aligned}$$

Para que este intervalo de confianza sea válido, el tamaño muestral debe ser mayor que 30. En este caso es 464, por lo que esta condición se cumple.

apartado c) Como el valor del 40 % (que corresponde a una proporción del 0.4) **no** se encuentra en el interior del intervalo de confianza calculado en el apartado a), **no puede darse por cierta** la afirmación a un nivel de confianza del 90 %.

Trabajo para el alumno: Resolver el siguiente ejercicio:

En una encuesta para estudiar el grado de preocupación de la población por el medio ambiente, se ha preguntado a 460 personas de las que 130 han contestado que reciclan los envases de plástico que utilizan.

a) Hallar una estimación puntual de la proporción p de personas que reciclan.

b) Hallar un intervalo de confianza al 90 % para la proporción poblacional de personas que reciclan. Indique qué condición se cumple para que el intervalo calculado sea válido.

c) En las noticias se publica un titular que afirma que el 35 % de personas de la población reciclan los envases de plástico. ¿Se puede dar por cierta dicha afirmación al mismo nivel de confianza?

a) Hay que hallar una estimación puntual del parámetro poblacional p. Según lo estudiado en las clases de teoría, el estimador más adecuado para la proporción poblacional desconocida p es la proporción muestral $\hat{p}$, donde

$$\hat{p} = \frac{\text{número de veces que aparece el suceso A}}{\text{tamaño muestral } n}.$$

Así, en nuestro caso

$$\hat{p} = \frac{130}{460} = 0.2826 \to 28.26\% \text{ de personas que reciclan.}$$

b) Hay que hallar un intervalo de confianza al 90 % para la proporción poblacional, cuya expresión vimos en las clases teóricas que es siempre:

$$I = \left[\hat{p} \mp z_{\alpha/2} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} \right]. \tag{7}$$

La confianza del 90 % indica que el nivel de significación es $\alpha = 0.1$ (ya que $90 = (1-\alpha) \cdot 100$). El valor crítico que se necesita en (7) es $z_{\alpha/2} = z_{0.1/2} = z_{0.05}$. Mirando este valor en la tabla de la normal estándar, obtenemos que $z_{0.05} = 1.645$. El valor de la raíz es:

$$\begin{aligned}
\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} &= \sqrt{\frac{0.28261\,(1-0.28261)}{460}} \\
&= \sqrt{\frac{0.28261 \cdot 0.71739}{464}} = \sqrt{4.3694 \times 10^{-4}} = 0.020903.
\end{aligned}$$

Reemplazando los cálculos en (7) tenemos

$$\begin{aligned}
I_{0.90} &= \left[\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\,;\, \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} \right] \\
&= [0.28261 - 1.645 \cdot 0.020903\,;\, 0.28261 + 1.645 \cdot 0.020903] \\
&= [0.28261 - 0.034385\,;\, 0.28261 + 0.034385] \\
&= [0.24822\,;\, 0.31700]\,.
\end{aligned}$$

Para que este intervalo de confianza sea válido, el tamaño muestral debe ser mayor que 30. En este caso tenemos $n = 460 > 30$, por lo que este valor se cumple.

c) Como el valor del 35 % (que corresponde a una proporción del 0.35) **no** se encuentra en el interior del intervalo de confianza calculado en el apartado a), **no puede darse por cierta** la afirmación a un nivel de confianza del 90 %.

1.- Para estudiar la evolución de la longitud del tallo de plantas de pimientos, un botánico cultivó 8 plantas en un banco de cultivo en invernadero. Midió la longitud total, en centímetros, del tallo de cada planta a las 10 semanas y a las 13 semanas, obteniendo los siguientes valores:

A las 10 semanas	10.1	10.3	10.6	10.8	10.9	11.1	11.2	11.2
A las 13 semanas	12.4	13.5	12.3	11.9	12.7	12.8	13.0	12.5

a) Calcule un intervalo de confianza al 90 % para la diferencia de alturas medias a las 13 y a las 10 semanas, suponiendo normalidad en las alturas.

b) ¿Se podría afirmar a ese nivel de confianza que el crecimiento medio ha sido de 2.8 cm? Justifique la respuesta.

2.- En los inviernos rigurosos, se utiliza sal para quitar el hielo de las carreteras de una determinada región. Para hallar la cantidad aproximada de sal que se está introduciendo en el medio ambiente por esta causa, se realizó un estudio en dicha región. Se obtuvieron las siguientes observaciones sobre la variable aleatoria X: número de toneladas métricas de sal utilizadas sobre las carreteras por semana, en zonas aleatoriamente seleccionadas, a lo largo de la región:

3900 3875 3820 3860 3840 3850 3800 3825 3790

(a) Establece una estimación puntual de μ y otra de σ.

(b) Suponiendo que la variable X está normalmente distribuida, halla un intervalo de confianza del 90% para μ y otro para σ.

3.- El porcentaje de calcio observado en dientes sanos de 10 individuos de una especie animal es:

36,6	35,9	35,6	35,4	34,9	36,5	35,6	35,2	35,6	35,4

Se pide:
a) Intervalo de confianza del 95 % para el porcentaje medio de calcio.
b) ¿Se podría aceptar que el porcentaje medio de calcio es igual a 36?
c) Intervalo de confianza del 95 % para la varianza de dicho porcentaje.
d) ¿Se podría aceptar que la varianza de dicho porcentaje es igual a 1,5?

4.- Se tomaron aleatoriamente nueve hojas de tabaco totalmente atacadas por el pulgón verde, *myzus persicae,* y se trató una mitad de cada hoja con el insecticida dimetoato y la otra mitad con triazamato. A los quince días de tratamiento se anotó la cantidad de superficie (cm^2) de cada mitad que quedó libre del pulgón verde, con los resultados siguientes:

Hoja nº	1	2	3	4	5	6	7	8	9
Dimetoato	350	273	245	119	77	98	168	259	245
Triazamato	366	265	249	133	69	112	159	255	249

Se sospecha que el pulgón verde se está haciendo más resistente al dimetoato que al triazamato.
a) Calcule un intervalo de confianza del 95% para la diferencia entre la superficie media libre de pulgón en las zonas tratadas con dimetoato y las tratadas con triazamato.
b) ¿Es cierta la sospecha de que es mayor la resistencia del pulgón verde al dimetoato que al triazamato? Justifique la respuesta.

5.- Un partido político desea conocer su intención de voto de cara a las próximas elecciones. Para ello, encarga un sondeo sobre un total de 230 personas, de las que 69 contestan que le votarían.
a) Halle un intervalo de confianza al 90% para la verdadera proporción poblacional.
b) Indique qué condición se cumple para que el intervalo calculado sea válido.
c) Un dirigente de dicho partido asegura en sus mítines que conseguirán el 33% de los votos. ¿Se puede dar por cierta dicha afirmacion al mismo nivel de confianza?

6.- Una central de transformación de productos lácteos recibe diariamente la leche de dos granjas. Deseando estudiar la calidad de los productos recibidos se extraen dos muestras al azar y se analiza el contenido en materia grasa, obteniéndose los siguientes resultados expresados en tanto por ciento:

Granja A	$\overline{x}_A = 8,7\%$	$s_A^2 = 1,02(\%)^2$	$n_A = 33$
Granja B	$\overline{x}_B = 10,9\%$	$s_B^2 = 1,73(\%)^2$	$n_B = 27$

Construir un intervalo de confianza del 95% para la diferencia del contenido medio en grasa de leche de ambas granjas. Interpretar el resultado.

7.- En un estudio de angina de pecho en ratas, se dividió aleatoriamente a 18 animales afectados, en dos grupos de 9 individuos cada uno. A un grupo se le suministró un placebo y al otro el fármaco experimental FL113. Después de un ejercicio controlado sobre una rueda de andar, se determinó el tiempo de recuperación de cada rata. Se piensa que el FL113 reducirá el tiempo medio de recuperación. Se dispone de la siguiente información:

Placebo	$\overline{x}_1 = 329$ segundos	$s_1 = 45$ segundos	$n_1 = 9$
FL113	$\overline{x}_2 = 238$ segundos	$s_2 = 43$ segundos	$n_2 = 9$

a) Construir un intervalo de confianza del 95% para la diferencia de los tiempos medios de recuperación de las ratas que recibieron placebo y las que recibieron el FL113.
b) ¿Es cierta la suposición de que el FL113 reducirá el tiempo medio de recuperación? Justificarlo.

8.- Se inoculan dos organismos aislados durante dos epidemias distintas, a dos muestras diferentes obtenidas de la misma población. A las dos semanas enferma el 68.5% de las 200 pruebas realizadas con el primer organismo, y el 65.3% de las 150 pruebas efectuadas para el segundo caso.
Halle un intervalo de confianza del 95% de esta diferencia de proporciones encontradas.

Soluciones Estimación puntual y por intervalos de confianza
Ejercicio 1.- a) [1.4283 ; 2.2967].

Ejercicio 2.- a) $\mu \leftarrow \bar{x} = 3840$, $\sigma \leftarrow s = 35.4436$; **b)** Para la media poblacional $\mu \in [3818.025\,;\,3861.975]$, y para la desviación típica poblacional $\sigma \in [25.45\,;\,60.64]$.

Ejercicio 3.- a) $[35.287\,;\,36.053]$; **c)** $[0.1357\,;0.9592]$.

Ejercicio 4.- a) $[-5.35\,;\,10.46]$ en cm^2.

Ejercicio 5.- a) $[0.2503\,;\,0.3497]$; **b)** condición $n > 30$; **c)** Como $0.33 \in [0.2503\,;\,0.3497]$ puede darse por válida la afirmación.

Ejercicio 6.- $[-2.8041\,;\,-1.5960]$.

Ejercicio 7.- a) $[47.0179\,;\,134,982]$, hay que hallar primero un intervalo de confianza para el cociente de varianzas y decidir si se consideran las varianzas poblacionales iguales o distintas.

Ejercicio 8.- $[-0.06774\,;\,0.13174]$