
Práctica 5. Contrastes paramétricos en una población

1. Contrastes sobre la media

El contraste de hipótesis sobre una media sirve para tomar decisiones acerca del verdadero valor poblacional de la media de una variable aleatoria.

1.1. Contraste sobre la media cuando la desviación típica poblacional es conocida

Esta técnica es válida solamente si la muestra es aleatoria y la población es Normal o el tamaño muestral, n , es grande (en la práctica, $n \geq 30$).

Para hacer este test hay que seleccionar **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1-Sample Z**. Esta opción también nos da el intervalo de confianza para la media poblacional, μ .

Abrimos el archivo de datos **Pulse.mtw**. Vamos a suponer que conocemos el valor de la desviación típica poblacional de la variable **Pulse1** (pulso antes de correr), $\sigma = 10$ pulsaciones por minuto. Comprobemos si se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que el pulso medio poblacional antes de correr es mayor que 70 pulsaciones por minuto. Si μ denota la media poblacional de la variable $X = \text{Pulso antes de correr}$, el contraste es $H_0 : \mu \leq 70$ frente a $H_1 : \mu > 70$.

En la Práctica 4 ya hemos comprobado que la muestra de resultados de la variable **Pulse1** es aleatoria. Además, el tamaño muestral es grande ($n = 92$). Por tanto, podemos utilizar este procedimiento estadístico.

Seleccionamos la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1-Sample Z**. En **Samples in columns** seleccionamos, de la lista de variables de la izquierda, la columna o columnas para las cuales se va a realizar este tipo de contraste; en nuestro caso, '**Pulse1**'. Dejamos desactivada la opción **Summarized data**. En **Standard deviation** tecleamos el valor de la desviación típica poblacional, σ , que suponemos que es **10**. Activamos **Perform hypothesis test** y en **Hypothesized mean** especificamos el valor, μ_0 , con el que se compara la media poblacional, que es **70**. Si pulsamos el botón **Options** nos aparece un nuevo cuadro de diálogo con las siguientes opciones:

Confidence level: Por defecto se muestra un intervalo de confianza al 95 % para la media poblacional μ . Se puede introducir un valor entre 1 y 99 para solicitar otro nivel de confianza. En nuestro caso, podemos dejar lo que aparece por defecto, es decir, 95.

Alternative: Aquí se especifica cuál es la hipótesis alternativa: **less than** significa que la hipótesis alternativa es $H_1 : \mu < \mu_0$, **not equal** significa que la hipótesis alternativa es $H_1 : \mu \neq \mu_0$ y

greater than significa que la hipótesis alternativa es $H_1 : \mu > \mu_0$. Tengamos en cuenta que con la opción **less than** el intervalo de confianza para la media será del tipo $(-\infty, b)$, con la opción **not equal** el intervalo de confianza para la media será del tipo (a, b) y con la opción **greater than** el intervalo de confianza para la media será del tipo $(a, +\infty)$. En nuestro caso, tenemos que seleccionar **greater than** ya que la hipótesis alternativa es $H_1 : \mu > 70$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 0'003, claramente menor que el nivel de significación, $\alpha = 0'05$. En consecuencia, rechazamos la hipótesis nula y, por tanto, aceptamos la hipótesis alternativa; es decir, aceptamos que la media poblacional de la variable **Pulse 1** es mayor que 70 pulsaciones por minuto. El intervalo de confianza al 95 % para la media poblacional, asociado a este contraste de hipótesis, es $(71'15, +\infty)$.

También se puede realizar este contraste de hipótesis si sabemos el tamaño muestral y el resultado de la media muestral. Veámoslo con un ejemplo:

En el volumen de Julio de 1992 de Economics Abstracts, la media del número de palabras por resumen es 79'56, con una varianza de 615'04. Se extrae una muestra aleatoria simple de 30 resúmenes escritos en alemán y se observa que la media del número de palabras por resumen es 67'47. Se quiere decidir si existe una diferencia significativa entre la media de palabras por resumen de los escritos en alemán y la media de palabras por resumen de todos los de este volumen.

Vamos a suponer que la varianza del número de palabras por resumen de los escritos en alemán coincide con la varianza del número de palabras por resumen de todos los de este volumen. Así pues, los datos que tenemos son los siguientes:

$$\begin{aligned}\mu_0 &= 79'56, \\ \sigma^2 &= 615'04 \Rightarrow \sigma = \sqrt{615'04} = 24'8, \\ \bar{X} &= 67'47, \\ n &= 30.\end{aligned}$$

La variable observada en la población no puede ser Normal pues es discreta, pero como el tamaño muestral es 30, entonces podemos aplicar esta técnica. Así pues, consideramos el siguiente contraste de hipótesis:

$$\begin{aligned}H_0 &: \mu = 79'56, \\ H_1 &: \mu \neq 79'56.\end{aligned}$$

Seleccionamos la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1-Sample Z**. Activamos la opción **Summarized data**, con lo cual se desactiva automáticamente la opción **Samples in columns**. En **Sample size** tenemos que teclear el tamaño muestral, que es **30** y en **Mean** tenemos que teclear el resultado de la media muestral, que es **67,47**. En **Standard deviation** tecleamos el valor de la desviación típica poblacional, σ , que suponemos que es **24,8**. Activamos **Perform hypothesis test** y en **Hypothesized mean** especificamos el valor, μ_0 , con el que se compara la media poblacional, que es **79,56**. Pulsamos en **Options** y, en el cuadro de diálogo resultante, en **Alternative** seleccionamos **not equal** puesto que nuestra hipótesis alternativa es $H_1 : \mu \neq 79'56$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 0'008, claramente menor que los niveles de significación usuales ($\alpha = 0'05$ ó $\alpha = 0'01$). En consecuencia, rechazamos la hipótesis nula y, por tanto, aceptamos que existe diferencia significativa entre la media del número de palabras por resumen en alemán y la media del número de palabras por resumen de todos ellos. El intervalo de confianza al 95 % para la media poblacional, asociado a este contraste de hipótesis, es $(58'60, 76'34)$.

1.2. Contraste sobre la media cuando la desviación típica poblacional es desconocida

Igual que en el apartado anterior, esta técnica es válida solamente si la muestra es aleatoria y la población es Normal o el tamaño muestral, n , es grande (en la práctica, $n \geq 30$).

Para realizar este contraste paramétrico hay que seleccionar **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1-Sample t**. La manera de utilizar esta opción es la misma que la explicada en el apartado anterior.

Con el archivo de datos **Pulse.mtw**, veamos si se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que el pulso medio poblacional antes de correr es igual a 71 pulsaciones por minuto. Lo que queremos comprobar es si la media poblacional de la variable **Pulse1** es igual a 71 pulsaciones por minuto, suponiendo ahora desconocida la desviación típica poblacional (lo cual es cierto). Si μ denota la media poblacional de la variable **Pulse1**, el contraste es $H_0 : \mu = 71$ frente a $H_1 : \mu \neq 71$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 0'107, claramente mayor que el nivel de significación, $\alpha = 0'05$, por lo que podemos aceptar la hipótesis nula; es decir, aceptamos que la media poblacional del número de pulsaciones por minuto antes de correr es igual a 71. El intervalo de confianza al 95 % para la media poblacional de dicha variable es (70'59, 75'15).

También se puede realizar este contraste de hipótesis si sabemos el tamaño muestral, el resultado de la media muestral y el resultado de la cuasi-desviación típica muestral. Veámoslo con un ejemplo:

El número medio de libros por estante de una biblioteca es 24. Extraída una muestra de 91 estantes de libros de matemáticas se obtiene una media de 25 libros, con una cuasi-desviación típica de 1'5. Queremos decidir si existe diferencia significativa entre el número medio de libros de matemáticas por estante y el número medio de libros por estante.

La variable $X = \text{“Número de libros de matemáticas por estante”}$ no puede ser Normal porque es discreta; pero como $n = 91 \geq 30$ entonces se puede utilizar este procedimiento.

Los datos conocidos son:

$$\begin{aligned}\mu_0 &= 24, \\ S &= 1'5, \\ \bar{X} &= 25, \\ n &= 91.\end{aligned}$$

El contraste de hipótesis que vamos a hacer es el siguiente:

$$\begin{aligned}H_0 &: \mu = 24, \\ H_1 &: \mu \neq 24.\end{aligned}$$

Seleccionamos la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1-Sample t**. Activamos la opción **Summarized data**, con lo cual se desactiva automáticamente la opción **Samples in columns**. En **Sample size** tenemos que teclear el tamaño muestral, que es **91**, en **Mean** tenemos que teclear el resultado de la media muestral, que es **25**, y en **Standard deviation** tenemos que teclear el resultado de la cuasi-desviación típica muestral, que es **1,5**. Activamos **Perform hypothesis test** y en **Hypothesized mean** especificamos el valor, μ_0 , con el que se compara la media poblacional, que es **24**. Pulsamos en **Options** y, en el cuadro de diálogo resultante, en **Alternative** seleccionamos **not equal** puesto que nuestra hipótesis alternativa es $H_1 : \mu \neq 24$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 0, el mínimo posible y, por supuesto, claramente menor que los niveles de significación usuales ($\alpha = 0'05$ ó $\alpha = 0'01$). En consecuencia, rechazamos la hipótesis nula y, por tanto, aceptamos que existe diferencia significativa entre el número medio de libros de matemáticas por estante y el número medio de libros por estante. El intervalo de confianza al 95 % para la media poblacional, asociado a este contraste de hipótesis, es (24'688, 25'312).

2. Contrastes sobre la varianza

El contraste de hipótesis sobre una varianza sirve para tomar decisiones acerca del verdadero valor poblacional de la varianza de una variable aleatoria. **Minitab** realiza el contraste solamente en el caso en el que la media poblacional es desconocida.

Esta técnica es válida solamente si la muestra es aleatoria y la población es Normal.

Para hacer el contraste de hipótesis sobre una varianza poblacional hay que seleccionar **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1 Variance**. Esta opción también se utiliza para realizar un test sobre la desviación típica poblacional.

En la Práctica 4 ya hemos comprobado que la muestra de resultados de la variable **Pulse1** (del archivo de datos **Pulse.mtw**) es aleatoria, y que la variable **Pulse1** es Normal. Por tanto, podemos utilizar este procedimiento estadístico para comprobar si se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que la varianza poblacional del pulso antes de correr es menor que 130 pulsaciones al cuadrado. Si σ^2 denota la varianza poblacional de la variable $X = \text{Pulso antes de correr}$, el contraste es $H_0 : \sigma \geq 130$ frente a $H_1 : \sigma^2 < 130$.

Seleccionamos, por tanto, la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1 Variance**. En el cuadro de diálogo resultante, arriba a la derecha, seleccionamos **Enter variance** (si quisiéramos realizar un contraste sobre la desviación típica poblacional, seleccionaríamos **Enter standard deviation**); en **Samples in columns** se selecciona, de la lista de variables de la izquierda, la columna o columnas para las cuales se va a realizar este tipo de contraste; en nuestro caso se selecciona '**Pulse1**'. Dejamos desactivada la opción **Summarized data**. Activamos **Perform hypothesis test** y en **Hypothesized variance** se especifica el valor, σ_0^2 , con el que se compara la varianza poblacional, que es **130**. Si pulsamos el botón **Options** nos aparece un nuevo cuadro de diálogo con las siguientes opciones:

Confidence level: Por defecto se muestra un intervalo de confianza al 95 % para la varianza poblacional σ^2 . Se puede introducir un valor entre 1 y 99 para solicitar otro nivel de confianza. En nuestro caso, podemos dejar lo que aparece por defecto, es decir, 95.

Alternative: Aquí se especifica cuál es la hipótesis alternativa: **less than** significa que la hipótesis alternativa es $H_1 : \sigma^2 < \sigma_0^2$, **not equal** significa que la hipótesis alternativa es $H_1 : \sigma^2 \neq \sigma_0^2$ y **greater than** significa que la hipótesis alternativa es $H_1 : \sigma^2 > \sigma_0^2$. Tengamos en cuenta que con la opción **less than** el intervalo de confianza para la varianza será del tipo $(-\infty, b)$, con la opción **not equal** el intervalo de confianza para la varianza será del tipo (a, b) y con la opción **greater than** el intervalo de confianza para la varianza será del tipo $(a, +\infty)$. En nuestro caso, tenemos que seleccionar **less than** ya que la hipótesis alternativa es $H_1 : \sigma^2 < 130$.

Podemos comprobar, en la ventana de sesión, que el p-valor (para el método Standard) es 0'338, claramente mayor que el nivel de significación, $\alpha = 0'05$. En consecuencia, aceptamos la hipótesis nula y, por tanto, no podemos aceptar la hipótesis alternativa; es decir, no podemos aceptar que

la varianza poblacional del pulso antes de correr es menor que 130 pulsaciones al cuadrado. El intervalo de confianza al 95 % para la varianza poblacional, asociado a este contraste de hipótesis (con el método Standard), es $(-\infty, 158)$. El intervalo de confianza al 95 % para la desviación típica poblacional, asociado a este contraste de hipótesis (con el método Standard), es $(-\infty, 12'6)$.

También se puede realizar este contraste de hipótesis si sabemos el tamaño muestral y el resultado de la cuasi-varianza muestral. Veámoslo con un ejemplo:

Se sabe que las calificaciones en la asignatura A es una variable Normal de media y varianza desconocidas. Se extrae una muestra aleatoria simple de 81 alumnos de la asignatura A , obteniéndose una media de 6'8 puntos, con una cuasi-varianza de 1'69 puntos al cuadrado, en las calificaciones de dichos alumnos. Sabemos que la varianza de las calificaciones en otra asignatura B es de 2'6 puntos al cuadrado. Queremos saber si la verdadera varianza de las calificaciones en la asignatura A es menor que la varianza en las calificaciones en la asignatura B .

Como la cuasi-varianza muestral es $S^2 = 1'69 < 2'6$, esta evidencia debe ser compatible con la hipótesis alternativa. Así pues, vamos a realizar el siguiente contraste:

$$\begin{aligned} H_0 : \sigma^2 &\geq 2'6, \\ H_1 : \sigma^2 &< 2'6. \end{aligned}$$

Seleccionamos la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1 Variance**. En el cuadro de diálogo resultante, arriba a la derecha, seleccionamos **Enter variance**. Activamos la opción **Summarized data**, con lo cual se desactiva automáticamente la opción **Samples in columns**. En **Sample size** tenemos que teclear el tamaño muestral, que es **81**, y en **Sample variance** tenemos que teclear el resultado de la cuasi-varianza muestral, que es **1,69**. Activamos **Perform hypothesis test** y en **Hypothesized variance** se especifica el valor, σ_0^2 , con el que se compara la varianza poblacional, que es **2,6**. Pulsamos en **Options** y, en el cuadro de diálogo resultante, en **Alternative** seleccionamos **less than** puesto que nuestra hipótesis alternativa es $H_1 : \sigma^2 < 2'6$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 0'006, claramente menor que los niveles de significación usuales ($\alpha = 0'05$ ó $\alpha = 0'01$). En consecuencia, rechazamos la hipótesis nula y, por tanto, aceptamos que la varianza de las calificaciones en la asignatura A es menor que la varianza de las calificaciones en la asignatura B . El intervalo de confianza al 95 % para la varianza poblacional, asociado a este contraste de hipótesis, es $(-\infty, 2'24)$.

3. Contrastes sobre la proporción

Supongamos una población en la que observamos una característica que sólo tiene dos resultados o modalidades, que podemos denominar *éxito* y *fracaso*. Sea p la proporción poblacional de éxitos.

Para hacer el contraste de hipótesis sobre la proporción poblacional de éxitos, p , hay que seleccionar **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1 Proportion**. Esta opción también nos da el intervalo de confianza para p .

Recordemos que en la hoja de datos **Pulse.mtw** la variable **Smokes** tenía solamente dos resultados: 1=Sí fuma, 2=No fuma. Vamos a comprobar si se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que el porcentaje poblacional de fumadores es menor que 35 %; es decir, si la proporción poblacional de fumadores es menor que 0'35. El contraste es $H_0 : p \geq 0'35$ frente a $H_1 : p < 0'35$.

Seleccionamos la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1 Proportion**. En el cuadro de diálogo resultante, en **Samples in columns** seleccionamos, de la lista de variables de la izquierda, la columna '**Smokes**'; dejamos desactivada la opción **Summarized data**; activamos **Perform hypothesis test** y en **Hypothesized proportion** especificamos el valor, p_0 , con el que comparamos la proporción poblacional, que es **0,35**. Si pulsamos el botón **Options** nos aparece un nuevo cuadro de diálogo con las siguientes opciones:

Confidence level: Por defecto se muestra un intervalo de confianza al 95 % para la proporción poblacional p . Se puede introducir un valor entre 1 y 99 para solicitar otro nivel de confianza. En nuestro caso, podemos dejar lo que aparece por defecto, es decir, 95.

Alternative: Aquí se especifica cuál es la hipótesis alternativa: **less than** significa que la hipótesis alternativa es $H_1 : p < p_0$, **not equal** significa que la hipótesis alternativa es $H_1 : p \neq p_0$ y **greater than** significa que la hipótesis alternativa es $H_1 : p > p_0$. Tengamos en cuenta que con la opción **less than** el intervalo de confianza para la proporción será del tipo $(-\infty, b)$, con la opción **not equal** el intervalo de confianza para la proporción será del tipo (a, b) y con la opción **greater than** el intervalo de confianza para la proporción será del tipo $(a, +\infty)$. En nuestro caso, tenemos que seleccionar **less than** ya que la hipótesis alternativa es $H_1 : p < 0'35$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 1, el máximo posible y, por supuesto, claramente mayor que el nivel de significación, $\alpha = 0'05$. En consecuencia, aceptamos la hipótesis nula y, por tanto, no podemos aceptar que la proporción poblacional de fumadores es menor que 0'35; es decir, no podemos aceptar que el porcentaje poblacional de fumadores es menor que 35 %. El intervalo de confianza al 95 % para la proporción poblacional, asociado a este contraste de hipótesis, es $(-\infty, 0'774287)$.

También se puede realizar el contraste de hipótesis sobre una proporción poblacional si sabemos el tamaño muestral y el número de éxitos en la muestra. Veámoslo con un ejemplo:

Deseamos conocer la postura de los bibliotecarios frente a la informatización de las bibliotecas. Para ello, preguntamos a 150 de ellos (elegidos aleatoria e independientemente) sobre este tema, obligándoles a manifestarse a favor o en contra. El resultado es que 82 se manifiestan a favor y consiguientemente, 68 en contra. ¿Es compatible este resultado con que la proporción de bibliotecarios (en el colectivo total) a favor de informatizar las bibliotecas es la misma que la proporción de bibliotecarios en contra?

Sea p la proporción de bibliotecarios (en el colectivo total) a favor de informatizar las bibliotecas. El contraste que hemos de realizar es $H_0 : p = 0'5$ frente a $H_1 : p \neq 0'5$. Seleccionamos la opción **Stat** \Rightarrow **Basic Statistics** \Rightarrow **1 Proportion**. En el cuadro de diálogo resultante, activamos la opción **Summarized data**, con lo cual se desactiva automáticamente la opción **Samples in columns**. En **Number of events** tenemos que teclear el número de éxitos en la muestra, que es **82** y en **Number of trials** tenemos que teclear el tamaño muestral, que es **150**. Activamos **Perform hypothesis test** y en **Hypothesized proportion** especificamos el valor, p_0 , con el que se compara la proporción poblacional, que es **0,5**. Pulsamos en **Options** y, en el cuadro de diálogo resultante, en **Alternative** seleccionamos **not equal** puesto que nuestra hipótesis alternativa es $H_1 : p \neq 0'5$.

Podemos comprobar, en la ventana de sesión, que el p-valor es 0'288, claramente mayor que los niveles de significación usuales ($\alpha = 0'05$ ó $\alpha = 0'01$). En consecuencia, aceptamos la hipótesis nula y, por tanto, aceptamos que la proporción de bibliotecarios (en el colectivo total) a favor de informatizar las bibliotecas es la misma que la proporción de bibliotecarios en contra. El intervalo de confianza al 95 % para la proporción poblacional, asociado a este contraste de hipótesis, es $(0'463428, 0'628026)$.

4. Ejercicios propuestos

5.1.

- Crea un nuevo proyecto de *Minitab*.
- Abre la hoja de datos **Prestamos.mtw** (datos del Ejercicio 1.1).
- ¿Se puede aceptar, con un nivel de significación de 0'05, que la media poblacional del porcentaje anual de préstamos por usuario es igual a 70? ¿Por qué?
- ¿Se puede aceptar, con un nivel de significación de 0'05, que la varianza poblacional del porcentaje anual de préstamos por usuario es igual a 140? ¿Por qué?
- Graba el proyecto con el siguiente nombre: **Ejercicio5-1.mpj**

5.2.

- Crea un nuevo proyecto de *Minitab*.
- Abre la hoja de datos **Transacciones.mtw** (datos del Ejercicio 1.2).
- ¿Se puede aceptar, con un nivel de significación de 0'05, que la media poblacional del porcentaje de transacciones de referencia finalizadas es menor que 86? ¿Por qué?
- ¿Se puede aceptar, con un nivel de significación de 0'05, que la desviación típica poblacional del porcentaje de transacciones de referencia finalizadas es mayor que 5? ¿Por qué?
- Graba el proyecto con el siguiente nombre: **Ejercicio5-2.mpj**

- 5.3. En una muestra aleatoria simple de 15 individuos que consultan bases de datos, el tiempo (en minutos) que están utilizando el ordenador para realizar esta tarea es:

22	13	17	14	15	18	19	14	17	20	21	13	15	18	17
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

- Crea un nuevo proyecto de *Minitab*.
 - Guarda los datos en el archivo **Minutos.mtw**
 - ¿Se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que la muestra es aleatoria? ¿Por qué?
 - ¿Se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que la variable aleatoria “tiempo empleado en consultar bases de datos por ordenador” es Normal? ¿Por qué?
 - ¿Se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que la media poblacional del tiempo empleado en consultar bases de datos por ordenador es mayor que 15 minutos? ¿Por qué?
 - ¿Se puede aceptar, con un nivel de significación de $\alpha = 0'05$, que la desviación típica poblacional del tiempo empleado en consultar bases de datos por ordenador es menor que 2 minutos? ¿Por qué?
 - Graba el proyecto con el siguiente nombre: **Ejercicio5-3.mpj**
- 5.4. El número medio de libros por estante en una biblioteca es de 24'4, con una desviación típica de 1'6. Una muestra aleatoria simple de 36 estantes de dicha biblioteca tiene una media de 25'2 libros por estante. ¿La información proporcionada por la muestra es representativa de toda la población?

- 5.5.** El número medio recomendado de usuarios servidos semanalmente por cada miembro del personal de una biblioteca es de 100. En una muestra aleatoria simple de 81 miembros del personal de las bibliotecas de una determinada región se obtiene una media de 132'88 usuarios servidos semanalmente, con una cuasi-desviación típica de 55'19. ¿Las bibliotecas de dicha región siguen la recomendación mencionada?
- 5.6.** El precio medio de los libros en rústica es de 63'4 euros, con una desviación típica de 14'8 euros. Una muestra aleatoria simple de 61 libros en rústica con ilustraciones en color tiene un precio medio de 69'5 euros, con una cuasi-desviación típica de 16'6 euros.
- ¿Permiten los datos afirmar que los libros en rústica con ilustraciones en color son más caros que el resto de libros en rústica?
 - ¿La varianza del precio de los libros en rústica con ilustraciones en color es mayor que la del precio de los libros en rústica?
- 5.7.** Se sabe que el número medio de veces que un artículo científico es citado durante los 5 siguientes años a su publicación es de 6'5. Se eligen aleatoria e independientemente 71 artículos de medicina, obteniéndose una media de 7'8 citas durante los 5 siguientes años a su publicación, con una cuasi-desviación típica de 2'3. ¿Se puede afirmar que durante los 5 siguientes años a su publicación se citan más los artículos de medicina que el resto de artículos científicos?
- 5.8.** En una biblioteca desconocemos la proporción de libros escritos en español. De una muestra aleatoria simple de 125 libros, 80 de ellos están escritos en español, y el resto en otros idiomas. Según estos datos, ¿se puede afirmar que la proporción de libros escritos en español en dicha biblioteca es mayor que 0'6?
- 5.9.** En una biblioteca escolar hay una proporción de libros prestados que se devuelven con retraso. De una muestra aleatoria simple de 250 libros, 50 de ellos se han devuelto con retraso. ¿Permiten los datos afirmar que la proporción de libros prestados que se devuelven con retraso a dicha biblioteca escolar es mayor que 0'15?