Muestreo. Distribuciones muestrales

1.- El precio medio del m^2 en la venta de casas nuevas durante el último año en una determinado barrio periférico de una ciudad fue de 115000 La desviación típica de la población fue de 25000 €. Se toma una muestra aleatoria de 100 casas nuevas de esta ciudad.

a) ¿Cuál es la probabilidad de que la media muestral de los precios de venta sea menor que 110000 €?

b) ¿Cuál es la probabilidad de que la media muestral de los precios de venta esté entre 113000 €y 117000 €?

c) ¿Cuál es la probabilidad de que la media muestral de los precios de venta esté entre 114000 €y 116000 €?

d) Sin hacer cálculos, razonar en cuál de los siguientes rangos resulta más probable que se encuentre la media muestral de los precios de venta:

113000 €-	115000 €
114000 €-	116000 €
115000 €-	117000 €
116000 €-	118000 €

2.- Se ha tomado una muestra de 16 directores de oficina de corporaciones de una gran ciudad, con el fin de estimar el tiempo medio que emplean en desplazarse para ir a su trabajo. Supongamos que la distribución de dichos tiempos en la población sigue una normal con media 87 minutos y desviación típica 22.

a) ¿Cuál es el error estándar de la media muestral de los tiempos de desplazamiento?

b) ¿Cuál es la probabilidad de que la media muestral sea inferior a 100 minutos?

c) ¿Cuál es la probabilidad de que la media muestral sea superior a 80 minutos?

d) ¿Cuál es la probabilidad de que la media muestral esté entre 85 y 95 minutos?

e) Supongamos que se toma una segunda muestra de quince directores, independiente de la anterior. Sin hacer los cálculos, razonar si la probabilidades calculadas en los apartados b), c) y d) serán mayores, menores o iguales para esta segunda muestra. Utilizar gráficos para ilustrar las respuestas.

3.- Una compañía produce cereales para el desayuno. La media del peso que contienen las cajas de estos cereales es de doscientos gramos y su desviación típica de seis gramos. La distribución de los pesos de la población es normal. Se eligen cuatro cajas, que pueden considerarse como una muestra aleatoria del total de la producción.

[1] Sol.: a) 0.0228; b) 0.5762; c) 0.3108: d) el intervalo 114000 €-116000 €

[2] Sol.: a) 5.5 ; b) 0.9909; c) 0.8980; d) 0.5671; e) es menor en los tres casos.

a) ¿Cuál es el error estándar de la media muestral del peso de las cuatro cajas?

b) ¿Cuál es la probabilidad de que la media del peso de esas cuatro cajas sea inferior que 197 gramos?

c) ¿Cuál es la probabilidad , en media, el peso de estas cuatro cajas esté entre 105 y 195 gramos?

d) ¿Cuál es la probabilidad de que la suma del peso de estas cuatro cajas sea menor de 800 gr.?

e) Se eligen al azar dos de estas cuatro cajas ¿Cuál es la probabilidad de que, en media, el contenido de estas dos cajas pese entre 195 y 200 gramos?

4.- La tasa de rentabilidad de ciertos tipos de acciones sigue una distribución con desviación típica 3.8. Se extrae una muestra de tales acciones con el fin de estimar el precio medio.

a) ¿Qué tamaño ha de tener la muestra para asegurarnos que la probabilidad de que la media muestral difiera de la media poblacional en una cantidad superior a 1 sea menor que 0.1?

b) Sin realizar los cálculos razonar si será preciso un tamaño muestral mayor o menor que el requerido en el apartado a) para garantizar que la probabilidad de que la media muestral difiera de la media poblacional en más de 1 sea inferior a 0.05.

c) Sin realizar los cálculos razonar si será preciso un tamaño muestral mayor o menor que el requerido en el apartado a) para garantizar que la probabilidad de que la media muestral difiera de la media poblacional en más de 1.5 sea inferior a 0.1.

5.- De acuerdo con los datos del ministerio de Economía y Hacienda, el 15% de las declaraciones del IRPF del último año darán lugar a una devolución. Se toma una muestra aleatoria de 10 declaraciones.

a) ¿Cuál es la media de la distribución en el muestreo de la proporción muestral de declaraciones que darán lugar a una devolución?

b) ¿Cuál es la varianza de la proporción muestral?

c) ¿Cuál es el error estándar de la proporción muestral?

d) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.8?

[3]Sol.: a) 3; b) 0.1587; c) 0.0475; d) 0.5; e) 0.3810

[4]Sol.: a) $n \geq 40$; b) mayor; c) menor

6.- El dueño de una portal de ventas de discos por Internet ha comprobado que el 20% de los clientes que acceden a su portal realizan una compra. Cierta mañana entraron en el portal 180 personas, que pueden ser consideradas como una muestra aleatoria de todos sus clientes.

a) ¿Cuál será la media de la proporción muestral de clientes que realizaron alguna compra?

b) ¿Cuál es la varianza de la proporción muestral?

c) ¿Cuál es el error estándar de la proporción muestral?

d) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.15?

7.- El administrador de una gran cadena de hospitales opina que, de entre todos sus pacientes, el 30% generará facturas que se pagarán con más de dos meses de retraso. Se toma una muestra aleatoria de 200 pacientes.

a) ¿Cuál es el error estándar de la proporción muestral de pacientes con facturas cuyo pago se retrasará dos meses?

b) ¿Cuál es la probabilidad de que la proporción muestral sea inferior a 0.25?

c) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.33?

d) ¿Cuál es la probabilidad de que la proporción muestral esté entre 0.27 y 0.33?

e) Sin realizar los cálculos, razonar en cuál de los siguientes intervalos es más probable que se encuentre la proporción muestral: 0.29-0.31; 0.30-0.32; 0.31-0.33; 0.32-0.34.

f) Supongamos que se toma al azar una muestra de 500 pacientes. Sin realizar los cálculos razonar si las probabilidades de los apartados b), c) y d) resultarán en este caso mayores, menores o iguales que las calculadas para la muestra anterior.

8.- Se toma una muestra aleatoria de 100 votantes con el fin de estimar la proporción de los mismos que están a favor de un aumento en los impuestos sobre la gasolina para contar así con un ingreso adicional para reparaciones de las autopistas. ¿Cuál es el mayor valor que puede tomar el error estándar de la proporción muestral de esta medida?

9.- Continuando en la situación del problema anterior, se decide que una muestra de 100 votantes es muy pequeña para obtener una estimación de la proporción poblacional que

[5] Sol.: a) 0.15; b) 0.01275; c) 0.1129; d) casi nula.

[6] Sol.: a) 0.2; b) ≈ 0.0009; c) 0.03; d) 0.9525

[7] Sol.: a) 0.0324; b) 0.0618; c) 0.1762; d) 0.6476; e) 0.29-0.31; f) menor, menor, mayor

[8] Sol.: 0.05

resulte suficientemente creíble. Se decide exigir que la probabilidad de que la proporción muestral difiera de la proporción poblacional (cualquiera que sea su valor) en más de 0.03 no debe ser superior a 0.05. ¿Qué tamaño ha de tener la muestra para poder garantizar que se cumple este requisito?

10.- Una compañía quiere estimar la proporción de personas que son posibles compradores de máquinas de afeitar eléctricas que ven retransmisiones partidos de La Liga de Campeones. Se toma una muestra de 120 individuos que se identificaron como posibles compradores de afeitadoras eléctricas. Supongamos que la proporción de posibles compradores de afeitadoras eléctricas en la población que ven estas retransmisiones es 0.25.

a) 0.10 es la probabilidad de que la proporción muestral exceda a la proporción poblacional ¿en qué valor?

b) 0.05 es la probabilidad de que la proporción muestral esté por debajo de la proporción poblacional ¿en qué cantidad?

c) 0.30 es la probabilidad de que la proporción muestral difiera de la proporción poblacional ¿en menos de qué cantidad?

11.- Supongamos que el 50% de los españoles adultos opina que es necesaria una revisión del sistema nacional público de hospitales. ¿Cuál es la probabilidad de que más del 56% de los componentes de una muestra de 150 españoles adultos tenga esa opinión?

12.- Las rentabilidades mensuales de cierto tipo de acciones son independientes unas de otras, y siguen una distribución normal con desviación típica 1,7. Se toma una muestra de 12 meses.

a) Hallar la probabilidad de que la desviación típica muestral sea menor que 2.5.

b) Hallar la probabilidad de que la desviación típica muestral sea mayor que 1.

13.- El número de horas que dedican a ver la televisión los estudiantes en la semana anterior a los exámenes finales sigue una distribución normal con una desviación típica de 4.5 horas. Se toma una muestra aleatoria de 30 estudiantes.

a) La probabilidad de que la desviación típica muestral sea mayor que 3.5 horas, ¿es mayor que 0.95?

b) La probabilidad de que la desviación típica muestral sea menor que seis horas, ¿es mayor que 0.95?

---

[9] Sol.: $n \geq 757$
[10] Sol.: a) 0.0506; b) 0.0648; c) 0.0154
[11] Sol.: 0.0708
[12] Sol.: a) 0.98; b) 0.975

14.- Se extrae una muestra aleatoria de 15 economistas y se les pregunta acerca de su predicción sobre la tasa de inflación para el próximo año. Supongamos que las predicciones para la población completa de economistas sigue una distribución normal con una desviación típica de 1.8.

a) 0.01 es la probabilidad de que la desviación típica sea mayor que ¿qué número?

b) 0.025 es la probabilidad de que la desviación típica sea menor que ¿qué número?

c) Encontrar una par de números, a y b, tales que la probabilidad de que la desviación típica muestral se encuentre entre ellos sea 0.9.

[13] Sol.: a) Sí; b) Sí

[14] Sol.: a) 2.5969; b) 1.1415; c) 1.2331; 2.341

Estimación puntual

15.- Se toma una muestra de ocho lotes de un producto químico para comprobar la concentración de impurezas. Los niveles porcentuales de impurezas encontrados en la muestra fueron

3.2 4.3 2.1 2.8 3.2 3.6 4.0 3.8

a) Hallar la media y la varianza muestrales. Hallar la proporción muestral de lotes con nivel porcentual de impurezas mayor que 3.75%.

b) ¿Para qué parámetros poblacionales se han hallado en la parte a) estimadores por procedimientos insesgados?

16.- Sea $\hat{\theta}_1$ un estimador insesgado de θ_1, y $\hat{\theta}_2$ un estimador insesgado de θ_2.

a) Probar que $(\hat{\theta}_1 + \hat{\theta}_2)$ es un estimador insesgado de $(\theta_1 + \theta_2)$.

b) Probar que $(\hat{\theta}_1 - \hat{\theta}_2)$ es un estimador insesgado de $(\theta_1 - \theta_2)$.

17.- Sea X_1 y X_2 una muestra aleatoria de dos observaciones independientes de una población con media μ y varianza σ^2. Considerar los siguientes tres estimadores puntuales de μ:

$$\overline{X} = \tfrac{1}{2}X_1 + \tfrac{1}{2}X_2$$
$$\hat{\mu}^{(1)} = \tfrac{1}{4}X_1 + \tfrac{3}{4}X_2$$
$$\hat{\mu}^{(2)} = \tfrac{1}{3}X_1 + \tfrac{2}{3}X_2$$

a) Probar que los tres estimadores son insesgados.

b) ¿Cuál de los tres estimadores es más eficiente?

c) Hallar la eficiencia relativa de $\overline{X}$ con respecto a los otros estimadores.

18.- A una clase de estadística asisten estudiantes de Informática de Gestión y de Sistemas. En una muestra de diez estudiantes de Gestión se observaron las siguientes calificaciones en el examen final

62 57 85 59 64 63 71 58 77 72

En una muestra independiente de ocho estudiantes de Sistemas se observaron las siguientes calificaciones en el mismo examen

73 79 85 73 62 51 60 57

[15]Sol.: a) $\overline{X} = 3.375$; $S_X^2 = 0.4993$; $\hat{p} = 0.3557$; b) para todos.

[17]Sol.: b) $\overline{X}$; c) $\frac{Var(\overline{X})}{Var(\hat{\mu}^{(1)})} = 0.8$; $\frac{Var(\overline{X})}{Var(\hat{\mu}^{(2)})} = 0.9$

a) Utilizar un método de estimación insesgado para obtener una estimación puntual de la diferencia de las calificaciones medias entre los estudiantes de Gestión y los de Sistemas. (Indicación: Utilizar el problema 151)

b) Utilizar un método de estimación insesgado para obtener una estimación puntual de la diferencia entre la proporción poblacional de estudiantes que obtuvieron una calificación mayor que 70 en el grupo de estudiantes de Gestión y el grupo de Sistemas. (Indicación: Utilizar el problema 151)

19.- Se toma una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una población con media μ y varianza σ^2. Se considera el siguiente estimador de μ :

$$\hat{\mu} = \frac{2}{n(n+1)}(X_1 + 2X_2 + 3X_3 + \ldots + nX_n)$$

a) Probar que $\hat{\mu}$ es un estimador insesgado de μ.

b) Hallar la eficiencia relativa de $\hat{\mu}$ respecto a $\overline{X}$, la media muestral.

20.-

a) (Examen junio 2003) Calcular el estimador máximo verosímil (MLE[1]) para el parámetro λ de una población que sigue una ley $Exp(\lambda)$ para una muestra aleatoria simple de tamaño n.

b) (Examen septiembre 2004) Calcular el MLE para el parámetro λ de una población que sigue una ley $Po(\lambda)$ para una muestra aleatoria simple de tamaño n.

c) Calcular el MLE para el parámetro μ de una población que sigue una ley $N(\mu, \sigma^2)$ para una muestra aleatoria simple de tamaño n.

d) Calcular el MLE para el parámetro σ^2 de una población que sigue una ley $N(\mu, \sigma^2)$ para una muestra aleatoria simple de tamaño n.

e) Estudiar si los estimadores MLE de los apartados anteriores son insesgados.

Estimación por intervalos

21.- De una población de barras de hierro se extrae una muestra de 64 barras y se calcula la resistencia a la rotura por tracción se obtiene que $\overline{X} = 1012\ Kg/cm^2$. Se sabe por experiencia que en este tipo de barras $\sigma = 25$. Calcular un intervalo de confianza para μ al nivel 0.95.

[18]Sol.: a) 0.2444; b)$-\frac{1}{10}$

[19]Sol.: b) $Var(\hat{\mu}) = \left(\frac{2}{n(n+1)}\right)^2 \frac{n(n+1)(2n+1)}{6}\sigma^2$; $Eff.rel = \frac{Var(\hat{\mu})}{Var(\overline{X})} = \frac{2\,(1+2\,n)}{3\,(1+n)}$.

[1]Del inglés "Maximun Likelihood Estimator"

[21]Sol.: $(1005.88, 1018.13)$

22.- Para investigar el C.I. medio de una cierta población de estudiantes, se realiza un test a 400 estudiantes. La media y la desviación típica muestrales obtenidas son $\overline{x} = 86$ y $\tilde{s}_X = 10.2$. Calcular un intervalo para μ con un nivel de significación del 98%.

23.- Para investigar un nuevo tipo de combustible para cohetes espaciales, se disparan cuatro unidades y se miden las velocidades iniciales. Los resultados obtenidos, expresados en Km/h, son :19600, 20300, 20500, 19800. Calcular un intervalo para la velocidad media μ con un nivel de confianza del 95%, suponiendo que las velocidades son normales. 20718.3

24.- Un fabricante de cronómetros quiere calcular un intervalo de estimación de la desviación típica del tiempo marcado en 100 horas por todos los cronómetros de un cierto modelo. Para ello pone en marcha 10 cronómetros del modelo durante 100 horas y encuentra que $\tilde{s}_X = 50$ segundos. Encontrar un intervalo para el parámetro σ^2 con $\alpha = 0.01$, suponiendo que la población del tiempo marcado por los cronómetros es normal.

25.- Un auditor informático quiere investigar la proporción de rutinas de un programa que presentan una determinada irregularidad. Para ello observa 120 rutinas, resultando que 30 de ellas presentan alguna irregularidad. Con estos datos buscar unos límites de confianza para la proporción p de rutinas de la población que presentan esa irregularidad con probabilidad del 95%.

26.- (Examen septiembre 2003) Una infección por un virus puede haber perjudicado a muchos ordenadores con *Windwos*. Desde el Centro de Alerta Temprana (CAT) se quiere calcular la proporción de ordenadores infectados. El jefe del centro os pide que calculéis el tamaño de una muestra para que el intervalo de confianza de la proporción muestral de ordenadores infectados tenga amplitud de a lo sumo 0.01 con una probabilidad del 90%.

27.- (Examen junio 2003) Se han medido los siguientes valores (en miles de personas) para la audiencia de un programa de televisión en distintos días (supuestos igualmente distribuidos e independientes):

$$521, 742, 593, 635, 788, 717, 606, 639, 666, 624.$$

Construir un intervalo de confianza del 90%, para la audiencia poblacional media y otro para la varianza poblacional, bajo la hipótesis de que la población de audiencias sigue una ley normal.

Nota Suma de las audiencias=6531, Suma de los cuadrados de las audiencias=4320401.

28.- Supongamos que la empresa para la que trabajamos está en un proyecto de investigación, financiado con fondos de la Comunidad Europea, que pretende extender una nueva aplicación de las TIC. Una de las tareas del proyecto es realizar una encuesta de opinión sobre el grado de aceptación que tendría esta nueva tecnología en el mercado europeo. De

[22] Sol.: $(84.8117, 87.1883)$
[23] Sol.: $(19381.7, 20718.3)$
[24] Sol.: $(953.834, 12968.3)$
[25] Sol.: $(0.1725, 0.3275)$

entre todas las universidades y empresas participantes en el proyecto, es a tu empresa a la que le toca hacer el protocolo de la encuesta, llevarla a cabo y redactar esta parte del informe final. Como eres el último que llegó a la empresa y el resto de miembros del equipo no se acuerda de la estadística que vio en la carrera, te toca a ti cargar con la responsabilidad. Claro que el coste de la encuesta depende del número n de entrevistas que se realicen y el error de las proporciones de las contestaciones disminuye cuando n aumenta. Como no sabes cuánto dinero está dispuesto a gastar tu jefe, tabula los tamaños muestrales para los errores $\pm 5\%, \pm 3\%, \pm 2\%, \pm 1\%$, y para niveles de confianza 0.95 y 0.99, suponiendo el peor caso. Añade un comentario para que el equipo de dirección del proyecto, en el que hay componentes ignorantes en materia de encuestas, vea como quedarían redactado los datos técnicos de la encuesta, y pueda decidir el tamaño muestral leyendo tu informe.

29.- El número de reservas semanales de billetes de cierto vuelo de una compañía aérea sigue una distribución aproximadamente normal. Se toma un muestra aleatoria de 81 observaciones de números de reservas de este vuelo: el número medio de reservas muestral resulta ser 112, mientras que la desviación típica muestral es 36. Además de estos 81 vuelos, 30 llegaron a su destino con un retraso de más de 15 minutos.

a) Calcular un intervalo de confianza del 95% para el número medio poblacional de reservas en este vuelo.

b) Calcular un intervalo de confianza de 95% para la varianza poblacional de las reservas.

c) Calcular un intervalo de confianza del 95% para la proporción poblacional de vuelos que llegan con un retraso de más de 15 minutos.

d) Calcular el tamaño muestral que asegura un intervalo de confianza de amplitud 0.1 para la proporción de vuelos que llegan con un retraso de más de 15 minutos al nivel de confianza 95%.

30.- Una empresa cervecera sabe que las cantidades de cerveza que contienen sus latas sigue una distribución normal con desviación típica poblacional 0.03 litros.

a) Se extrae una muestra aleatoria de 25 latas y, a partir de la misma, un experto en estadística construye un intervalo de confianza para la media poblacional del contenido en litros de las latas que discurre entre 0.32 y 0.34 ¿Cuál es el nivel de confianza de este intervalo?

b) Un gerente de esta empresa exige un intervalo de confianza del 99% que tenga una amplitud máxima de 0.03 litros a cada lado de la media muestral ¿Cuántas observaciones son necesarias, como mínimo, para alcanzar este objetivo?

[29]Sol.: a) $(104.16, 119.84))$; b) $(972.343, 1814.08))$; c) $(0.265, 0.475))$; d) $n = 385$

[30]Sol.: a) 90.3%; b) $n = 7$

Contraste de hipótesis

31.- Siendo $\overline{x} = 63.5$ la media de una muestra aleatoria simple de tamaño 36 extraída de una población normal con $\sigma^2 = 144$, poner a prueba, con un nivel de significación $\alpha = 0.05$, la hipótesis nula $\mu = 60$ y decir si se rechaza en favor de la alternativa $\mu < 60$. Calcular el p-valor.

32.- Siendo $\overline{x} = 72.5$ la media de una muestra aleatoria simple de tamaño 100 extraída de una población normal con $\sigma^2 = 900$, poner a prueba, con un nivel de significación $\alpha = 0.10$, la hipótesis nula $\mu = 77$ y decir si se rechaza en favor de las hipótesis alternativas $\mu \neq 70$, $\mu > 70$, $\mu < 70$. Calcular el p-valor en cada caso.

33.- En un contraste bilateral, con $\alpha = 0.01$, ¿para qué valores de $\overline{X}$ rechazaríamos la hipótesis nula $H_0 : \mu = 70$, a partir de una muestra aleatoria simple de tamaño 64 extraída de una población normal con $\sigma^2 = 256$?

34.- El salario anual medio de 1600 personas, elegidas aleatoria e independientemente de cierta población de economistas con $\sigma = 20000$ euros, ha valido 45000 euros ¿Es compatible con este resultado la hipótesis nula, $H_0 : \mu = 43500$, suponiendo $\alpha = .01$? ¿Cuál es el intervalo de confianza para μ? Calcular el p-valor.

35.- Con los datos del ejercicio anterior , ¿son compatibles con el resultado obtenido los siguientes contrastes?:

a) $\begin{cases} H_0 : \mu = 44000 \\ H_1 : \mu > 44000 \end{cases}$

b) $\begin{cases} H_0 : \mu = 46250 \\ H_1 : \mu > 46250 \end{cases}$

36.- El peso medio de los paquetes de café puestos a la venta por la casa comercial CAFEINASA es supuestamente de 1 Kg. Para comprobar esta suposición, elegimos una muestra aleatoria simple de 100 paquetes y encontramos que su peso medio es de 0.978 Kg. y su desviación típica $s = 0.10$ kg. Siendo $\alpha = 0.05$ ¿es compatible este resultado con la hipótesis nula $H_0 : \mu = 1$ frente a $H_1 : \mu \neq 1$? ¿Lo es frente a $H_1 : \mu > 1$? Calcular el p-valor.

37.- El fabricante de la marca de tornillos FDE afirma que el diámetro medio de sus tornillos vale 20 mm. Para comprobar dicha afirmación, extraemos aleatoria e independientemente 16 tornillos , y vemos que la media de sus diámetros es 22 mm. y la desviación típica 4 mm. ¿Podemos aceptar la pretensión del fabricante, suponiendo $\alpha = 0.05$ y siendo el contraste bilateral? Calcular el p-valor.

38.- Para evitar basarse en su intuición los jefes de admisión de personal de las grandes empresas discriminan mediante un test diseñado por un gabinete de psicólogos, supuestamente especializado en selección de personal, a los aspirantes a trabajar en la empresa. La varianza del test de selección solía venir siendo 100. Aplicando un nuevo test a una muestra aleatoria simple de tamaño $n = 31$, se obtiene que $S = 129$. Suponiendo que la población se distribuye normalmente, ¿es compatible la hipótesis nula $H_0 : \sigma^2 = 100$, frente a la alternativa

$H_1 : \sigma^2 > 100$, con $\alpha = 0.01$? Calcular el p-valor.

39.- Una máquina produce cierto tipo de piezas mecánicas. El tiempo en producirlas se distribuye normalmente con varianza desconocida σ^2. Elegida una muestra aleatoria simple de 21 de dichas piezas $(x_1, \ldots, x_{21})$, se obtiene que $\overline{x} = 30$ y $\sum_{i=1}^{21} x_i^2 = 19100$. Comprobar si es compatible la hipótesis nula $H_0 : \sigma^2 = 22$ frente $H_1 : \sigma^2 \neq 22$, con $\alpha = 0.1$, y construir un intervalo de confianza del $(1-\alpha)100\%$ para el verdadero valor de σ^2. Calcular el p-valor.

40.- A partir de las puntuaciones 15, 22, 20, 21, 19 ,23, construir el intervalo de confianza de σ^2 y decir si es compatible con estos resultado la hipótesis $H_0 : \sigma = 2$, siendo $\alpha = 0.01$ contra una H_1 bilateral. Decir si se utiliza alguna hipótesis adicional. Calcular el p-valor.

41.- Sabiendo que con $\hat{p} = 0.52$ ha sido rechazada $H_0 : p = 0.50$, al nivel de significación $\alpha = 0.05$, ¿cuál ha tenido que ser el tamaño mínimo de la muestra mediante la cual fue rechazada H_0

a) frente a $H_1 : p \neq 0.5$?

b) frente a $H_1 : p > 0.5$?

42.- Lanzamos una moneda al aire 10 veces consecutivas . ¿Con qué número de caras rechazaremos la hipótesis nula de que la moneda está bien balanceada, siendo $\alpha = 0.05$?

43.- Un fabricante de productos farmacéuticos tiene que mantener un estándar de impurezas en el proceso de producción de sus píldoras. Hasta ahora el número medio poblacional de impurezas es correcto pero está preocupado porque las impurezas en algunas de las partidas se salen del rango admitido de forma que provocan devoluciones y posibles reclamaciones por daños a la salud. El gabinete de control de calidad afirma que si la distribución de las impurezas es normal y que si el proceso de producción mantiene una varianza inferior a 1 no tendría que existir ningún problema pues las píldoras tendrían una concentración aceptable. Preocupado por esta tema la dirección encarga una prueba externa en la que se toma una muestra aleatoria de 100 de las partidas obteniéndose $S^2 = 1.1$. ¿Puede aceptar el director de la prueba externa que el proceso de producción cumple la recomendación del gabinete de control?

44.- Un IAP está preocupado por su estándar de calidad y quiere compararlo con el medio europeo. El estándar medio europeo dice que una empresa de este sector tiene una calidad aceptable si tiene un número de quejas que no excede del 3%.

Se sabe que la varianza de las quejas es 0.16. Examinando 64 clientes escogidos al azar se encuentra con que el porcentaje de quejas es del 3.07%. Calcular el p-valor.

a) Contrastar al nivel de significación del 5%, la hipótesis nula de que la media poblacional del porcentaje de quejas es del 3% frente a la alternativa de que es superior al 3%.

b) Hallar el p-valor del contraste.

c) Supongamos que la hipótesis alternativa fuese bilateral en lugar de unilateral(con hipótesis nula $H_0 : \mu = 3$). Deducir, sin hacer ningún cálculo, si el p-valor del contraste sería mayor, menor o igual que el del apartado anterior. Construir un gráfico para ilustrar el razonamiento.

d) En el contexto de este problema, explicar por qué una hipótesis alternativa unilateral es más apropiada que una bilateral.

45.- A partir de una muestra aleatoria se contrasta:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

y se acepta la hipótesis nula al nivel de significación del 5%.

a) ¿Implica esto necesariamente que μ_0 está contenido en el intervalo de confianza del 95% para μ?

b) Si la media muestral observada es mayor que μ_0, ¿implica necesariamente que μ_0 está contenido en el intervalo de confianza del 90% para μ?

46.- Una compañía que se dedica a la venta de franquicias afirma que, por término medio, los delegados obtienen un redimiendo del 10% en sus inversiones iniciales. Una muestra aleatoria de diez de estas franquicias presentaron los siguientes rendimientos el primer año de operación:

6.1, 9.2, 11.5, 8.6, 12.1, 3.9 , 8.4, 10.1, 9.4, 8.9

Asumiendo que los rendimientos poblacionales tienen distribución normal, contrastar la afirmación de la compañía.

47.- Una distribuidora de bebidas refrescantes afirma que una buena fotografía de tamaño real de un conocido actor, incrementará las ventas de un producto en los supermercados en una media de 50 cajas semanales. Para una muestra de 20 supermercados, el incremento medio fue de 41.3 cajas con una desviación típica de 12.2 cajas. Contrastar al nivel de significación $\alpha = 0.05$, la hipótesis nula de que la media poblacional del incremento en las ventas es al menos 50 cajas, indicando cualquier supuesto que se haga. Calcular el p-valor del contraste e interpretarlo.

Contrastes de dos parámetros.

Comparación de medias.

Los siguientes problemas tratan de contrastes de parámetros entre dos muestras. Para cada uno de los enunciados contrastar contra las hipótesis unilaterales y bilaterales. Calcular también el intervalo de confianza para la diferencia o el cociente de los parámetros. Tomar finalmente la decisión más correcta. Calcular todos los test e intervalos de confianza para $\alpha = 0.05$. Calcular el p-valor en cada caso.

48.- Para comparar la producción media de dos procedimientos de fabricación de cierto elemento se toman dos muestras, una con los elementos fabricados durante 25 días con el primer método y otra con los producidos durante 16 días con el segundo método. Por experiencia se sabe que la varianza del primer procedimiento es $\sigma_1^2 = 12$ y al del segundo $\sigma_2^2 = 10$. De las muestras obtenemos que $\overline{X}_1 = 136$ para el primer procedimiento y $\overline{X}_2 = 128$ para el segundo.

49.- Estamos interesados en comparar la vida media, expresada en horas de dos tipos de componentes electrónicos. Para ello se toma una muestra de cada tipo y se obtiene:

Tipo	tamaño	$\overline{X}$	S
1	50	1260	20
2	100	1240	18

Suponer si es necesario las poblaciones aproximadamente normales.

50.- Para reducir la concentración de ácido úrico en la sangre se prueban dos drogas. La primera se aplica a un grupo de 8 pacientes y la segunda a un grupo de 10. Las disminuciones observadas en las concentraciones de ácido úrico de los distintos pacientes expresadas en tantos por cien de concentración después de aplicado el tratamiento son:

droga 1	20	12	16	18	13	22	15	20		
droga 2	17	14	12	10	15	13	9	19	20	11

Suponer que las reducciones de ácido úrico siguen una distribución normal son independientes y de igual varianza. Ídem pero suponiendo que las varianza son distintas.

51.- Para comparar la dureza media de dos tipos de aleaciones (tipo 1 y tipo 2) se hacen 5 pruebas de dureza con la de tipo 1 y 7 con la de tipo 2. Obteniéndose los resultados siguientes:

$$\overline{X}_1 = 18.2, \quad S_1 = 0.2 \text{ y}$$

$$\overline{X}_2 = 17.8; \quad S_2 = 0.5$$

Suponer que la población de las durezas es normal y que las desviaciones típicas no son iguales. Hacer lo mismo si las varianzas son distintas.

52.- Se encuestó a dos muestras independientes de internautas, una en Menorca y otra en Mallorca, sobre si utilizaban telefonía por intente. La encuesta de Menorca tuvo un tamaño $n_1 = 500$ y 100 usuarios mientras que en Mallorca se encuestaron a $n_2 = 750$ y se obtuvo un resultado de 138 usuarios.

53.- Se pregunta a un grupo de 100 personas elegido al azar asiste a una conferencia sobre tecnologías de la comunicación. Antes de la conferencia se les pregunta si consideran a internet peligrosa, después de la conferencia se les vuelve a preguntar cual es su opinión. Los resultados fueron los siguientes:

		Después	
		Sí es peligrosa	No es peligrosa
Antes	Sí es peligrosa	50	30
	No es peligrosa	5	15

54.- Tenemos 10 ordenadores, deseamos optimizar su funcionamiento. Con este fin se piensa en ampliar su memoria. Se les pasa una prueba de rendimiento antes y después de ampliar la memoria. Los resultados fueron:

	Ordenador									
Muestra/Tiempo	1	2	3	4	5	6	7	8	9	10
Antes ampliación	98.70	100.48	103.75	114.41	97.82	91.13	85.42	96.8	107.76	112.94
Después ampliación	99.51	114.44	108.74	97.92	103.54	104.75	109.69	90.8	110.04	110.09

55.- Las siguientes muestras provienen de dos poblaciones independientes y supuestamente normales. Se desea comparar la igualdad de sus medias, pero antes debemos contrastar si podemos o no aceptar que sus varianza son iguales o distintas. Se pide hacer el contraste de las medias en el caso en que se se decida aceptar varianzas iguales o distintas al nivel de significación $\alpha = 0.05$.

Contrastar también la hipótesis de igualdad de medias en el otro caso (es decir si se decide varianzas distintas contrastar para iguales y viceversa).

Problemas de bondad de ajuste

56.- Una compañía de gas afirma, basándose en experiencias anteriores, que normalmente, al final del invierno, el 80% de las facturas han sido ya cobradas, un 10% se cobrará con pago aplazado a un mes, un 6% se cobrará a 2 meses y un 4% se cobrará a más de dos meses. Al final del invierno actual, la compañía selecciona una muestra aleatoria de 400 facturas, resultando que 287 de estas facturas cobradas, 49 a cobrar en un mes, 30 a cobrar en dos meses y 34 a cobrar en un periodo superior a dos meses. ¿Podemos concluir, a raíz de los resultados, que la experiencia de años anteriores se ha vuelto a repetir este invierno?

57.- El Rector de una Universidad opina que el 60% de los estudiantes consideran los cursos que realizan como muy útiles, el 20% como algo útiles y el 20% como nada útiles. Se toma una muestra aleatoria de 100 estudiantes, y se les pregunta sobre la utilidad de los cursos. Resultando que 68 estudiantes consideran que los cursos son muy útiles, 18 consideran que son poco útiles y 14 consideran que no son nada útiles. Contrastar la hipótesis nula de que los resultados obtenidos se corresponden con la opinión personal del Rector.

58.- Considérense los fondos de inversión ordenados en función de su rendimiento en el periodo 1995-99. Se realizó un seguimiento del rendimiento en los cinco años posteriores de una muestra aleatoria de 65 fondos entre el 25% más rentable del periodo 1995-99. En este segundo periodo se observó que 11 de los fondos de la muestra se hallan entre el 25% más rentable en este segundo periodo, 17 en el segundo 25%, 18 en el tercer 25% y 19 en el 25% menos rentable. Contrastar la hipótesis de que un fondo de inversión escogido azar del 25% más rentable en 1995-99 tenga la misma probabilidad de hallarse en cualquiera de las cuatro categorías de rendimiento en el periodo 2000-2004.

59.- A una muestra aleatoria de 502 consumidores se les preguntó la importancia que se le daba al precio a la hora de elegir un ordenador. Se les pidió que valoraran entre: "ninguna importancia", "alguna importancia" y "principal importancia". El número respectivo de respuestas en cada tipo fueron 169, 136 y 197. Contrastar la hipótesis nula de que la probabilidad de que un consumidor elegido al azar conteste cualquiera de las tres respuestas es la misma.

60.- Durante cien semanas se ha venido observando el número de veces a la semana que se ha fuera de servicio un servidor de una pequeña empresa de informática, presentándose los resultados de la siguiente tabla:

Núm. Fuera Servicio	0	1	2	3	4	5 o más
Núm. Semanas	10	24	32	23	6	5

El número medio de veces que quedo fuera de servicio por semana durante este periodo fue de 2.1. Contrastar la hipótesis nula de que la distribución de averías es una Poisson.

61.- A lo largo de 100 minutos, llegaron a una web de un periódico 100 internautas. La siguiente tabla muestra la frecuencia de llegadas a lo largo de ese intervalo de tiempo.

Núm. llegadas/mín.	0	1	2	3	4 o más
Frec. Observada	10	26	35	24	5

Contrastar la hipótesis nula de que la distribución es Poisson.

KS test.

62.- Se quiere saber si el tiempo entre accesos, en una determinada franja horaria, a una cierta página web sigue una ley exponencial. Se dispone de la siguiente muestra de 25 intervalos entre tiempos de acceso:

```
x=c(140.7,13.7,67.6,7.8,49.3,128.5,59.6,234,171.1,205.8,99.3,199.8,
    100.8,13.5,12,33.9,44.1,12.3,56.4,9.4,112.1,8.2,110.5,79,55.4)
x
```

```
##  [1] 140.7  13.7  67.6   7.8  49.3 128.5  59.6 234.0 171.1 205.8  99.3
## [12] 199.8 100.8  13.5  12.0  33.9  44.1  12.3  56.4   9.4 112.1   8.2
## [23] 110.5  79.0  55.4
```

Resolver realizando los cálculos de forma manual y utilizando funciones de R las siguientes cuestiones.

a) Contrastar la hipótesis de que la distribución sigue una ley exponencial de parámetro

$\lambda = 100$, al nivel $\alpha = 0.05$.

b) Contrastar la hipótesis de que la distribución sigue una ley Poisson de parámetro $\lambda = 105$, al nivel $\alpha = 0.05$.

c) Contrastar la hipótesis de que la distribución sigue una ley Poisson de parámetro $\lambda = 110$, al nivel $\alpha = 0.05$.

d) Contrastar la hipótesis de que la distribución sigue una ley Poisson, estimando el parámetro parámetro partir de la muestra, al nivel $\alpha = 0.05$.

63.- Resolver las mismas cuestiones que en el problema anterior para la muestra (decir si se viola algunas de las condiciones del test KS, pero resolver igualmente el ejercicio):

$$69.9, 31.5, 130.2, 80.5, 236.1, 151.2, 74.8, 13.8, 54.5, 147.6$$

En esta ocasión realizar los cálculos manualmente.

64.- Nos hemos bajado un generador de números aleatorios normales de Internet. Queremos contrastar si funciona correctamente. Para ello generamos una muestra de 10 números aleatorios de una normal estándar:

$$-1.18, -0.77, -0.59, -0.27, -0.12, 0.27, 0.29, 0.40, 1.27, 1.60$$

a) Contrastar si provienen de una normal estándar al nivel de significación $\alpha = 0.05$ mediante el test KS. Decir si ha violado alguna de las suposiciones de este test.

b) Contrastar la hipótesis de normalidad contra una distribución normal de media y varianza la estimadas a partir de la muestra.

65.- Con la muestra:

$$0.60, -1.42, 1.05, -0.14, 0.57, 0.11, -0.59, 1.11, -1.55, -1.41$$

Contrastar con un test KS si los datos provienen de una distribución uniforme en el intervalo $(-2, 2)$ al nivel $\alpha = 0.05$