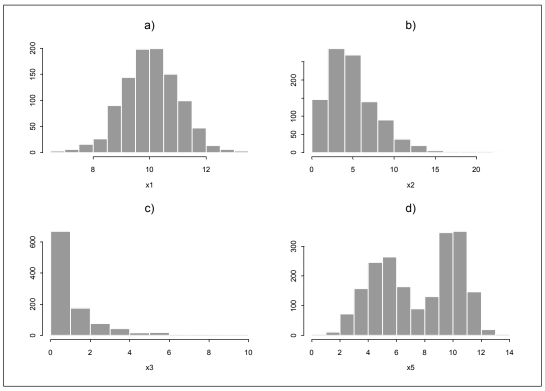
Dpto. Matemática Aplicada y Estadística

Ingeniero Técnico de Minas
Asignatura: **Estadística**
Hoja 1. Estadística Descriptiva

1. Describir las características de los cuatros histogramas siguientes, y razonar cuál es la medida de centralización y de dispersión más adecuada para la distribución correspondiente.



2. Consideramos un conjunto de datos con media 7 y varianza 3. Entre ellos hay dos datos iguales a 7. Supongamos que eliminamos del conjunto uno de los datos iguales a 7. La varianza del nuevo conjunto de datos ¿aumenta, disminuye o sigue siendo igual a 3? Justifica tu respuesta.

3. Los siguientes datos se refieren al crecimiento de una colonia de bacterias en un medio de cultivo:

x	3	6	9	12	15	18
y	115000	147000	239000	356000	579000	864000

siendo x los días desde la inoculación e y el número de bacterias.

   (a) Representar y en función de x para verificar que es razonable ajustar a una curva exponencial. b) Ajustar a una curva exponencial estos datos. c) Estimar, usando el modelo ajustado en el apartado anterior, el número de bacterias al término de 20 días.

4. El técnico responsable del funcionamiento de una empaquetadora automática la ajustó, en principio, para 450g. Media hora después del principio de la producción, se apartaron diez paquetes para verificar su peso. Los resultados son:

Peso (g)	448	450	453	451	447	449	446	451	448	447

a) ¿Cuál es el peso medio de esa muestra?. Calcular la varianza y la desviación típica, así como la mediana y los cuartiles.

b) Se considera que la empaquetadora funciona correctamente si la media de una muestra de diez paquetes se sitúa en el intervalo $448 \leq \overline{X} \leq 452$. ¿Cuál es la conclusión en el caso de la muestra anterior?. ¿Te parece correcta la elección de tal método de decisión?. ¿Alguna idea para mejorar?.

5. El responsable en control industrial de una empresa somete a un test de fiabilidad 50 dispositivos electrónicos idénticos y anota su duración en horas hasta su fallo, es decir, hasta el fin de la aptitud del dispositivo para cumplir con su utilidad. La recogida de los datos lleva a la distribución de frecuencias siguientes :

Duración (h)	Núm. dispositivos	Duración (h)	Núm. dispositivos
$0 \leq X < 200$	17	$800 \leq X < 1000$	6
$200 \leq X < 400$	9	$1000 \leq X < 1200$	2
$400 \leq X < 600$	7	$1200 \leq X < 1400$	1
$600 \leq X < 800$	7	$1400 \leq X < 1600$	1

a) Identificar la variable estadística y su unidad de medida. Representar el histograma. Establecer la tabla de las frecuencias relativas y de las frecuencias relativas acumuladas. Por la forma del histograma, escoja una medida de centralización para este conjunto de datos.

b) ¿Qué porcentaje de dispositivos tienen una duración inferior a 800h? ¿qué porcentaje de dispositivos tienen una duración superior o igual a 200h? ¿y a 600h?. ¿Qué porcentaje de dispositivos tienen una duración comprendida en el intervalo $200 \leq X < 400$?, ¿y en el intervalo $200 \leq X < 800$?

c) Calcular la duración media de la muestra, la varianza, la desviación típica y el coeficiente de variación. ¿En qué unidades se expresan esas dos últimas cantidades ?. ¿En qué clase se encuentra la mediana de los datos?

6. Se ha realizado un estudio para investigar el efecto de un determinado proceso térmico en la dureza superficial de una determinada pieza. Once piezas se seleccionaron para el estudio. Antes del tratamiento se realizaron pruebas de dureza para determinar la dureza de cada pieza. Después, las piezas fueron sometidas a un proceso térmico de templado con el fin de mejorar su dureza. Al final del proceso, se realizaron nuevamente pruebas de dureza y se obtuvo una segunda lectura. Se recogieron los siguientes datos (Kg. de presión):

Pieza	1	2	3	4	5	6	7	8	9	10	11
Dureza previa.	182	232	191	200	148	249	276	213	241	480	262
Dureza post.	198	210	194	220	138	220	219	161	210	313	226

a) Calcular la media, la desviación típica, la mediana y los cuartiles para la dureza antes y después del proceso.

b) Realizar un diagrama de caja y bigotes donde aparezcan las dos variables ¿Qué conclusiones puede sacar de estos diagramas ? ¿Se puede afirmar que el proceso de templado mejora la dureza de las piezas?

c) Se quiere ahora buscar un modelo para explicar la dureza posterior en función de la dureza previa. ¿Cuál es el método adecuado? Constuir este modelo y comentar la bondad del ajuste. ¿Cuál sería el valor de la dureza posterior para una dureza previa igual a 195?

7. Por exigencias de la normativa sobre la emisión de gases contaminantes, los fabricantes de automóviles deben controlar la cantidad de distintos contaminantes emitidos por los tubos de escape de sus modelos. Dos de los principales contaminantes son el monóxido de carbono (CO)

y los óxidos de nitrógeno (NO_x). En la tabla siguiente, aparecen los valores correspondientes a las emisiones de 9 vehículos del mismo modelo. Las variables se miden en gramos por kilómetro recorrido.

CO	2.3	5.0	7.5	10.1	14.9	5.9	8.1	6.3	4.0
NO_x	1.8	1.5	1.1	0.7	0.6	1.3	1.0	1.2	1.4

a) Realice el diagrama de dispersión correspondiente. ¿Cuál es el tipo de relación entre las dos variables? ¿Se trata de una asociación positiva o negativa?.

b) Obtener la correspondiente recta de regresión utilizando la técnica de los mí-nimos cuadrados.

c) Hallar el coeficiente de determinación r^2 e interpretar su valor.

d) Un periodista especializado en coches ha escrito que " Cuando un motor está bien construido y bien ajustado, emite pocos contaminantes. Si está mal ajustado, aumenta la emisión de casi todos los contaminantes principales. Por consiguiente, puede decidirse si un motor es contaminante, midiendo uno solo de estos contaminantes. Si este valor es aceptable, todas las demás emisiones serán aceptables."

Razonar si el estudio realizado confirma o no la opinión de este periodista.

8. La hidrólisis de un cierto producto químico tiene lugar en medio ácido según un proceso cinético de primer orden. Partiendo de una concentración inicial desconocida del ester, se han medido las concentraciones del mismo a diferentes tiempos obteniéndose los resultados siguientes.

t(min)	3	5	10	15	20	30	40	50	60	75	90
c.10^3(M)	25.5	23.4	18.2	14.2	11	6.7	4.1	2.5	1.5	0.7	0.4

a) Realize una nube de puntos de las dos variables. La teoría cinética de este tipo de reacciones nos indica que la evolución de la concentración del producto químico en función del tiempo se rige por $C_t = C_0 e^{-k \cdot t}$, donde C_0 es la concentración inicial. ¿Qué transformación de los datos nos lleva a un modelo lineal? Realizar esta transformación y obtener la concentración inicial C_0 y la velocidad k de disparición del producto químico.

b) Suponemos ahora que nos comunican que la concentración inicial del producto químico era C_0 =3.10^{-2} M. ¿Cómo incorporar esta información a nuestro análisis anterior? Obtener el nuevo valor de k.

9. Las materias primas empleadas en la producción de una fibra sintética son almacenadas en un lugar en donde no se tiene control de la humedad. La siguiente tabla refleja en porcentajes la humedad relativa del almacén (X) y la humedad observada en la materias primas (Y) durante un estudio que tuvo lugar durante 9 días.

X	41	53	59	65	71	78	50	65	74
Y	1.6	13.6	19.6	25.6	31.6	33.2	14.7	21.2	28.3

a) Analice los datos construyendo un modelo para explicar la evolución de Y en función de X.

b) Contestar a la pregunta que os hace una persona sin conocimientos de estadística : ¿Qué estimación podemos dar para la humedad en las materias primas (Y), si se consigue una humedad relativa de 35?

Dpto. Matemática Aplicada y Estadística

Ingeniero Técnico de Minas
Asignatura: **Estadística**
Hoja 2. Probabilidad

1. Un dado está trucado de forma que la probabilidad de sacar 2 es doble que la de obtener 1, la de sacar 3 es triple que la de obtener 1 y así sucesivamente. ¿cuál es la probabilidad de sacar 4?.

2. Consideremos dos sucesos A y B. Sobre un conjunto de 50 elementos se ha obtenido $p(A \cap B) = 10h$; $p(A \cap \overline{B}) = 15h$; $p(\overline{A} \cap B) = 20h$; $p(\overline{A} \cap \overline{B}) = 5h$.

   Calcular el valor de la constante h. ¿Cuántos elementos de A no son de B?

3. Una urna contiene cinco bolas numeradas del 1 al 5. Probabilidad de que al sacar dos bolas sin reposición la suma de los puntos sea impar.

4. Se lanzan cinco dados sobre una mesa, ¿cuál es la probabilidad de que salgan sólo numeros pares?, ¿y de qué salga al menos un 6?

5. En cada intento un atleta tiene el 70% de posibilidades de superar una prueba. En competición dispone de tres intentos, ¿cuál es la probabilidad de superar dicha prueba en la competición?

6. Se hacen tres disparos con tres cañones independientes siendo la probabilidad de alcanzar el objetivo 0'1, 0'2, 0'3 respectivamente. Calcular la probabilidad de cada uno de los números posibles de blancos. Calcular la probabilidad de obtener al menos un blanco.

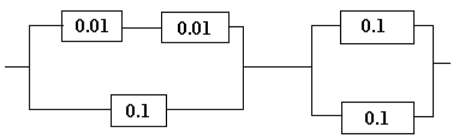
7. *[Junio 98]* Sean A y B dos sucesos cualesquiera. Se sabe que $\Pr(A) = \dfrac{1}{3}$, $\Pr(B) = \dfrac{2}{5}$ y que $\Pr(A \cup B) = \dfrac{3}{5}$. Se pide:

   a)Calcular la siguientes probabilidades : $\Pr(A \cap B); \Pr(A^c); \Pr(A^c \cap B^c)$ y $\Pr(A \mid B)$, donde A^c y B^c representan los sucesos complementarios de A y B.

   b)Indicar de forma razonada si pueden considerarse independientes. ¿Son incompatibles ambos sucesos?.

8. Consideremos dos sucesos A y B, con $P(A) = 0.5$ y $P(A \cup B) = 0.7$. Entonces: (a) Calcular $P(B)$ suponiendo que A y B son independientes.(b)Calcular $P(B)$ suponiendo que A y B son mutuamente excluyentes. (c) Calcular $P(B)$ sabiendo que $P(A/B) = 0.5$.

9. El siguiente circuito trabaja si, y solo si, existe un camino de dispositivos en funcionamiento de izquierda a derecha. Supongamos que los dispositivos fallan de manera independiente. En la figura se indica la probabilidad de fallo de cada uno de ellos. Calcular la probabilidad de que el circuito trabaje.



10. Se tiene una moneda trucada de forma que al lanzarla la probabilidad de obtener cara es 2/3. Se lanza la moneda al aire, y si sale cara se toma al azar un número del 1 al 9, si sale cruz se toma al azar un número del 1 al 5. Calcular la probabilidad de que el número escogido sea par.

11. En la fabricación de un cierto artículo se encuentra que se presenta un tipo de defecto con una probabilidad de 0.1 y defectos de un segundo tipo con probabilidad 0.05. Sabiendo que ambos tipos de defectos son independientes. ¿Cuál es la probabilidad de que: a) un artículo tenga ambas clases de defectos? b) un artículo sea defectuoso? c) suponiéndo que el artículo es defectuoso, sólo tenga un tipo de defecto?

12. A un almacén llega la producción de tres fábricas. La producción de la 1^a constituye el 20%, la de la 2^a el 46% y de la 3^a el 34%. El porcentaje de piezas no standard de la 1^a fábrica es del 3%, 2% en la 2^a y 1% en la 3^a. Hallar la probabilidad de que una pieza tomada al azar se haya producido en la 1^a fábrica si ha resultado no standard.

13. (*Junio 98*) Una pieza producida en una empresa puede tener dos tipos de defectos. El 8% de la producción presenta el defecto de tipo A, el 5% de la producción presenta el defecto de tipo B, y se supone que no hay piezas que tengan los dos tipos de defectos. Después de ser producida cada pieza es sometida de manera automática a un test de ruptura, con las siguientes posibilidades: Si la pieza tiene el defecto de tipo A, tiene una probabilidad de 0.9 de romperse. Si la pieza tiene el defecto de tipo B, tiene una probabilidad de 0.95 de romperse. Finalmente, si la pieza no tiene ningún tipo de defecto, tiene una probabilidad de 0.01 de romperse.

    a)Si el experimento aleatorio consiste en escoger al azar un pieza de la producción, traducir los datos del enunciado, después de haber introducido los sucesos convenientes.

    b)¿Cuál es la probabilidad de que una pieza escogida al azar en la producción se vaya a romper durante el test?

    c)Si una pieza escogida al azar se ha roto durante el test, ¿cuál es la probabilidad de que no fuese defectuosa?

14. En un sistema de alarma, la probabilidad de que se produzca un peligro es 0'1. Si este se produce, la probabilidad de que la alarma funcione es 0'95. La probabilidad de que funcione la alarma sin haber habido peligro es 0'03. Calcular:

    a)Probabilidad de que habiendo funcionado la alarma, no haya habido peligro. b)Probabilidad de que haya un peligro y la alarma no funcione. c)Probabilidad de que no habiendo funcionado la alarma, haya un peligro.

Dpto. Matemática Aplicada y Estadística

Ingeniero Técnico de Minas
Asignatura: **Estadística**
Hoja 3. Variables Aleatorias

1. Probar si las siguientes funciones pueden definir funciones puntuales de probabilidad asociadas a las variables que se indican.

$$a) f(x) = \frac{x}{15} \; para \; x = 0, 1, 2, 3, 4, 5. \qquad b) f(x) = \frac{5 - x^2}{6} \; para \; x = 0, 1, 2, 3.$$
$$c) f(x) = 0.25 \; para \; x = 3, 4, 5, 6. \qquad d) f(x) = \frac{x - 2}{2} \text{ para } x = 1, 2, 3, 4.$$

2. Suponiendo que $f(x) = k/2^x$ es una función puntual de probabilidad para una variable que puede tomar los valores x=0,1,2,3 y 4. Calcular el valor de la constante k y la función de distribución $F(x)$.

3. Calcular el valor de la constante k para que $f(x)$ sea una función puntual de probabilidad:

(a) $f(x) = \begin{cases} kx, & x = 1, 2, 3, 4, 5, 6 \\ 0, & \text{en otro caso} \end{cases}$ **(b)** $f(x) = \begin{cases} kx, & x = 1, 2, ..., n \\ 0, & \text{en otro caso} \end{cases}$

4. Consideremos el experimento aleatorio de lanzar 3 veces una moneda. Definamos la v.a. X como la diferencia entre el número de caras y el de cruces. Hallar la función puntual de probabilidad y la función de distribución de la v.a. X suponiendo que la moneda está trucada de tal forma que una cara tiene dos veces más de probabilidad de ocurrir que una cruz.

5. Consideremos el experimento aleatorio de seleccionar 4 bolas al azar sin reemplazamiento, de una urna que contiene 7 bolas rojas y 3 azules. Definamos la v.a. X como *"Número de bolas rojas extraídas".* Calcular:

   (a) La función puntual de probabilidad y la función de distribución de la v.a. X

   (b) $P(1 \leq X \leq 3)$, $P(1 < X \leq 3)$ y $P(1 \leq X < 3)$

   (c) Contestar las cuestiones anteriores cuando el experimento se realiza con devolución.

6. Un juego consiste en lanzar un dado, ganando el triple de lo apostado si sale un 6, ganando el doble si sale un 5, perdiéndose lo apostado en otro caso. ¿Conviene jugar?

7. Consideremos el juego de la ruleta : pueden salir de manera equiprobable cualquier número del 0 al 36, hay que saber además que la mitad de los números del 1 al 36 son rojos y la otra mitad negros. El cero tiene otro color. Se puede apostar por un número particular, en este caso, si sale el número por el se ha apostado, se recupera 36 veces su apuesta, si no sale, se pierde la apuesta. Por otra parte, se puede también apostar por un color, en este caso, si sale el color elegido, se recupera dos veces la apuesta mientras que si no sale, se pierde la apuesta. Sea X la variable aleatoria "beneficio obtenido al apostar p pesetas a un número cualquiera" y sea Y la variable "beneficio obtenido al apostar p pesetas al rojo". Calcular la media y la varianza de ambas variables. Comentar los resultados obtenidos.

8. Calcular la esperanza así como la desviación típica de los beneficios para una compañía de seguros al hacer un seguro cuya prima anual es R, la probabilidad de siniestro es p y la cantidad asegurada es M. Comentar los resultados obtenidos para los valores concretos $p = 0.001$, $M = 100.000pts$ y $R = 1000pts$.

9. (Diciembre 99) En un proceso de fabricación de copas de cristal, las bases se sellan calentándolas con una llama. La temperatura X de ésta varía de manera aleatoria, siguiendo la distribución de probabilidad siguiente:

Temperatura	540^oC	545^oC	550^oC	555^oC	560^oC
Probabilidad	0.1	0.25	0.3	0.25	?

   (a) Encontrar el valor que falta en la tabla
   (b) Calcular la media y la desviación típica de X.
   (c) La temperatura óptima es de 550^oC. ¿Cuál es la media y la desviación típica de la diferencia entre X y la temperatura óptima?
   (d) Para un informe, imagina que debes pasar los resultados de grados Centígrados a grados Farenheit, siendo la relación
   $$Y = \frac{9}{5}X + 32$$
   ¿Cuál es la media y la desviación típica de la temperatura expresada en grados Farenheit?

10. Estudiar si las siguientes funciones pueden ser funciones de densidad y calcular las funciones de distribución asociadas.

   (a) $f(x) = x^2$ si $x \in (0, 2)$ (cero en el resto).
   (b) $f(x) = k(3x + 1)$ si $x \in (1, 3)$ (cero en el resto).
   (c) $f(x) = k(x^2 + 2x + 0.5)$ si $x \in (0, 2)$ (cero en el resto).
   (d) $f(x) = kx^{-3}$ si $x > 1$ (cero en el resto).
   (e) $f(x) = \begin{cases} x & \text{si } 0 < x < 1 \\ 2 - x & \text{si } 1 < x < 2 \end{cases}$ (cero en el resto).

11. Si la función de distribución de una variable aleatoria está dada por:
$$F(x) = \begin{cases} 1 - \dfrac{4}{x^2} & \text{si } x > 2 \\ \\ 0 & \text{en otro caso} \end{cases}$$
calcular las probabilidades de que esta variable aleatoria asuma un valor: a)menor que 3; b) entre 4 y 5. Calcular la función de densidad asociada a dicha variable.

12. Dada la v.a. X con función de distribución:
$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x^n & \text{si } 0 < x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$
Calcular la función de densidad y $P\left(X \geq \dfrac{3}{4}\right)$.

13. Sea

$$f(x) = \begin{cases} k(1-x^3) & \text{si } 0 < x < 1 \\ 0 & \text{en otro caso} \end{cases}$$

Calcular:

(a) el valor de k para que $f(x)$ sea una función de densidad.

(b) su función de distribución.

(c) $\Pr\left(X \leq \dfrac{1}{2}\right)$, $\Pr\left(\dfrac{1}{2} \leq X \leq \dfrac{3}{4}\right)$ y $\Pr\left(X > \dfrac{3}{4}\right)$.

Dpto. Matemática Aplicada y Estadística

Ingeniero Técnico de Minas
Asignatura: **Estadística**
Hoja 4. Modelos de Distribuciones Univariantes

1. El kilometraje (en miles de kilómetros) que los automovilistas logran de cierto tipo de neumáticos, es una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} \dfrac{1}{20}e^{-x/20} & \text{si } x > 0 \\ & \\ 0 & \text{en otro caso} \end{cases}$$

   Calcular las probabilidades de que un neumático dure: a) a lo sumo 10.000 km; b) entre 16.000 y 24.000 km; c) al menos 30.000 km. d) la duración media esperada.

2. (**Sept. 98**) Sabiendo que el número de gotas de agua que caen en un grifo mal cerrado cada 15 minutos es una distribución de Poisson de media 4. Calcular:

   (a) Probabilidad de que caigan más de 5 gotas en un cuarto de hora.

   (b) Probabilidad de que en tres cuartos de hora caigan entre 13 y 18 gotas.

   (c) Probabilidad de que en un cuarto de hora caigan 3 gotas y en el siguiente cuarto de hora otras tres.

3. Una empaquetadora automática se programa para producir paquetes de 500g. Un estudio concluye que el peso en gramos de un paquete de la producción es una variable X normal de media 498g y de varianza 16. Sabemos que producir un gramo de producto cuesta a la empresa 0.5pts mientras que lo vende 0.9pts. Llamemos B la variable "beneficio de la empresa por paquete vendido".

   (a) Expresar la relación que existe entre la variable B y la variable X. ¿Cuál es el beneficio promedio realizado por la empresa por paquete?

   (b) Cuál es la proporción de paquetes entre la producción para los cuales la empresa realiza un beneficio mayor de 200pts?

4. En un proceso de fabricación, la probabilidad de que una pieza sea defectuosa es de 0.01. Si la producción diaria es de 10000 piezas y se empaquetan en lotes de 100 unidades:

   (a) Calcular la probabilidad de que en un lote haya por lo menos dos piezas defectuosas.

   (b) Si un lote es rechazado cuando contiene más de 5 piezas defectuosas, ¿Cuantos lotes son rechazados diariamente por término medio?

5. Un cierto sistema electrónico contiene 10 componentes. Supongamos que la probabilidad de que un componente falle es 0.3 y que los componentes fallan de manera independiente unos de otros. Dado que al menos uno de los componentes ha fallado, ¿cuál es la probabilidad de que fallen al menos dos de los componentes?

6. Una partida de bujías con alta proporción de inservibles (20%) sale al mercado en paquetes de 4 unidades y en cajas de 10 paquetes. Calcular la probabilidad de que:
   (a) Elegido un paquete al azar contenga 2 o más bujías inservibles.
   (b) Elegida una caja al azar, contenga más de 10 bujías inservibles.
   (c) Elegida una caja al azar contenga tres paquetes sin bujías inservibles.

7. Supóngase que una máquina de fabricación de tornillos produce por término medio 2 tornillos defectuosos por cada 85 y que los tornillos se empaquetan en cajas de 170 unidades. Calcular cuál es la probabilidad de que tomadas al azar 7 cajas, en solo 2 de ellas no haya ningún tornillo defectuoso.

8. Si X sigue una distribución *normal* de media 10 y desviación típica 2, se pide:

   (a) $P(X \leq 12)$; **(b)** $P(|X| \leq 3)$; **(c)** $P(|X - 10| \leq 1)$;
   (d) El valor de a para que $P(|X - 10| \leq a) = 0.95$.

9. El peso (en gramos) de las galletas que fabrica una máquina sigue una distribución *normal* de media 2 y de desviación típica $\left(\frac{5}{89}\right)$ gramos. Las galletas se venden en paquetes de 15. Consideremos una galleta correcta cuando su peso no difiere en más de 0.1 gramos del peso medio.

   (a) ¿Cuál es la probabilidad que un paquete contenga por lo menos 2 galletas defectuosas?
   (b) Si queremos que una galleta sea defectuosa con probabilidad de 0.006, calcular a qué desviación típica tendremos que ajustar la máquina.
   (c) Con la desviación típica obtenida en (b), calcular la probabilidad de que en 100 paquetes por lo menos haya 20 galletas defectuosas.

10. (Junio 99) En una fábrica que envasa agua mineral, se ha establecido que el volumen envasado por la máquina automática sigue una distribución normal de media 150cl y de desviación típica 2 cl.

    (a) Los criterios de calidad de la empresa implican que no se venda una botella que contenga menos de 147cl. ¿Cuál es la proporción de botellas en la producción que no se puede vender?
    (b) Las botellas se empaquetan por 6 unidades, ¿cuál es la probabilidad de que un paquete contenga al menos una botella con menos de 147cl?
    (c) En un día se producen 10000 botellas ¿cuál es la probabilidad de que haya en un día más de 600 botellas invendibles?
    (d) Utilizando el apartado anterior, ¿cuál es, en un mes de 30 días, el número medio de días en los que se producen más de 600 botellas invendibles?

11. Después de un estudio médico-laboral de varios meses, se estima en un 60% el porcentaje de los trabajadores que realizan una cierta tarea en una central siderúrgica y que presentan un proceso irritativo en los ojos después de una jornada laboral. Si observamos a 1100 trabajadores de este tipo,

(a) Calcular la probabilidad de que, entre los 1100 trabajadores, hayan más de 680 que padezcan irritabilidad ocular.

(b) Calcular la probabilidad de que el número de trabajadores con irritabilidad en los ojos sea mayor que 670 y menor que 675.

(c) ¿Cuántos podemos esperar que padezcan ese síntoma?

12. Eres el encargado de un proceso de fabricación, dispones de un informe de control de calidad del proceso que estima en un 2% el porcentaje de artículos defectuosos.

    (a) Te mandan producir en un día 1000 artículos, calcula la probabilidad de que, ese día, se produzcan entre 20 y 30 artículos defectuosos, y ¿que se produzcan exactamente 20 artículos?

    (b) Te informan que la semana que viene tendrás que producir 8000 artículos, ¿cuántos artículos defectuosos esperas que sean producidos en esa semana?

13. Consideremos un experimento consistente en lanzamientos independientes de una moneda trucada de manera que la probabilidad de obtener cara es de $1/3$. Calcular la probabilidad de necesitar exactamente 5 lanzamientos hasta obtener la primera cara. ¿Cuál es el número esperado de lanzamientos que se necesitan hasta obtener la primera cara?

14. Suponiendo que la probabilidad de que un cliente sea atendido en un servicio telefónico es de 0.75:

    (a) Calcular la probabilidad de que el cliente necesite tres llamadas para ser atendido.

    (b) ¿Cuál será el número esperado de llamadas para ser atendido?

15. El tiempo de curación de una determinada enfermedad sigue una distribución exponencial de media 4.5 días.

    (a) Calcular la probabilidad de curarse en los 3 primeros días.

    (b) Si llevamos dos días enfermos, ¿cuál será la probabilidad de curarnos en los próximos tres días?

16. Se sabe que el tiempo de vida de un componente electrónico es una variable aleatoria de media 350h. y desviación típica 35h.

    (a) Obtener un intervalo centrado en la media que contenga al menos el 80% de los tiempos de vida de los componentes producidos.

    (b) ¿Es muy probable observar un componente con vida mayor que 455h. o menor que 245h.?

17. Una determinada máquina produce piezas de longitud media 22 mm y desviación típica 0.3 mm. Si el comprador acepta como buenas aquellas piezas cuya longitud esté comprendida entre 20 y 24 mm, obtener una cota superior para el porcentaje de piezas defectuosas.

1. Una fábrica de gaseosa utiliza una envasadora automática para rellenar botellines de plástico. Cada botella debe contener 300ml pero en realidad los contenidos varían según una distribución normal con media $\mu = 298ml$ y desviación típica $\sigma = 3ml$.

   (a) ¿Cuál es la probabilidad de que una botellín individual contenga menos de 295ml?

   (b) ¿Cuál es la probabilidad de que el contenido promedio de las botellas en un paquete de 6 contenga menos de 295ml?

2. Se realiza una medición de peso en un laboratorio, sabiendo que la desviación típica de las medidas es $\sigma = 10mg$. El operador repite la medición 3 veces y proporciona como valor del peso la media $\bar{x}$ de sus tres mediciones.

   (a) ¿Cuál es la desviación típica del resultado proporcionado?

   (b) ¿Cuantas veces habría que repetir la medición para reducir la desviación típica del resultado proporcionado hasta 5?

3. El resultado de una encuesta de opiniones fue que el 59% de la población española piensa que la situación económica es buena o muy buena. Supongamos, extrapolando los resultados del sondea a la población entera, que la proporción de todos los españoles con esta opinión es efectivamente 0.59.

   (a) Muchos de los sondeos tienen un "margen de error" del orden de ±3 puntos, ¿Cuál es la probabilidad de que una muestra aleatoria de 300 españoles presente una proporción muestral que no se aleje en más de 0.03 de la proporción poblacional auténtica $p = 0.59$?

   (b) Conteste a la pregunta anterior para una muestra de 600 individuos y otra de 1200 individuos. ¿Cuál es el efecto de aumentar el tamaño muestral?

4. Un aparato de medición es exacto (el valor proporcionado medio es el valor auténtico de la señal) y la desviación típica del valor medido es 0.1 unidades. La distribución del valor medido es aproximadamente normal. ¿Cuál es la probabilidad de que el valor de una medición se aleje de la señal auténtica en más de 0.1 unidades? Y si se repite la medición 5 veces y se toma la media de los 5 valores obtenidos?

5. En condiciones normales, una máquina produce piezas con una tasa de defectuosas del 1%. Para controlar que la máquina sigue bien ajustada, se escogen al azar cada día 100 piezas en la producción y se le somete a un test. ¿Cuál es la probabilidad de que, si la máquina está bien ajustada, haya, en una de esas muestras, más del 2% de piezas defectuosas? Si un día, 3 piezas resultan defectuosas, ¿qué conclusiones sacaría sobre el funcionamiento de la máquina?

6. Un ascensor limita el peso de sus cuatro ocupantes a 300Kg. Si el peso de un individuo sigue una distribución N(71,7), calcular la probabilidad de que el peso de 4 indviduos supere los 300Kg.

Dpto. Matemática Aplicada y Estadística

Ingeniero Técnico de Minas
Asignatura: **Estadística**
Hoja 6. Estimación Paramétrica

1. De una población normal con $\sigma = 2.5$, se extrae una muestra aleatoria simple de tamaño 9 cuyos valores se presentan en la siguiente tabla: 165 162 166 164 165 170 169 165 168

   Obtener el intervalo de confianza para la media poblacional con un nivel de confianza del 98%.

2. Un estudiante lee un artículo donde se afirma que un intervalo de confianza al 95% para la altura media de los españoles mayores de 18 años es $[1.66, 1.72]$. Se le pregunta por el significado de esta afirmación y contesta que quiere decir que el 95% de los españoles mayores de 18 años tiene una altura comprendida entre estos dos números. ¿Crees que su respuesta es correcta? Razónalo.

3. Para calibrar un instrumento de medición, se mide repetidas veces un patrón cuyo peso sabemos igual a 10 gramos. Los valores medidos siguen una distribución normal de media desconocida. En cambio, por estudios anteriores sabemos que podemos considerar la desviación típica igual a 0.0002 gramos.

   a)¿Qué representa la media poblacional de los valores medidos? ¿Qué valdría ésta si el instrumento de medición fuera perfecto?

   b)Se mide el patrón cinco veces. La media de las cinco mediciones es igual a 10.0023 gramos. Construir de manera detallada un intervalo de confianza para la media poblacional al 98% de confianza.

   c)¿Cuántas mediciones debemos realizar para conseguir con la media muestral un margen de error de ± 0.0001 con 98% de confianza?

4. Investigadores plantan 15 parcelas con una nueva variedad de maiz. Las cantidades cosechadas aparecen en la siguiente tabla:

138.0	139.1	113.0	132.5	140.7	109.7	118.9	134.8
109.6	127.3	115.6	130.4	130.2	111.7	105.5	

   a)Suponiendo que $\sigma = 10$, encontrar, después de haber descrito la variable y la población de interés, el intervalo de confianza al 90% para la cosecha promedio por parcela.

   b)Conteste a la pregunta anterior para los niveles de confianza del 95% y del 99%. ¿Cuál es el efecto del incremento del nivel de confianza sobre el intervalo de confianza?

   c)Al desconocer la distribución de la variable en este problema, los intervalos anteriores son intervalos aproximados. ¿En qué sentido? ¿Cuál es el resultado matemático que nos asegura que, aun siendo la distribución de la variable no normal, podemos utilizar el mismo intervalo que para el caso normal?

   d)¿Cuál es el número de parcelas que se deberían plantar para conseguir estimar la cosecha promedia por parcela con 90% de confianza, y un margen de error menor de 4?

5. Para determinar la resistencia exacta de un componente eléctrico, se realizan, en las mismas condiciones, 5 mediciones y se obtiene los resultados siguientes : 495, 498, 493, 501, y 500 . Se supone que el error que se comete durante la medición sigue una distribución normal de media 0 y de desviación típica $\sigma \simeq 3$.

   a)¿Cuál es la distribución del valor medido de la resistencia? En particular, ¿cuál es su media? ¿Es el aparato de medición exacto? ¿Le parece suficientemente preciso?

   b)Construir un intervalo de confianza al nivel de 90% para la resistencia exacta. Presentar el resultado en forma de intervalo ($valor \pm error.$)

   c)Si quiero cometer como máximo un error de 2 con una confianza del 95%, ¿cuántas mediciones más debería realizar?

6. ¿Qué es un estimador puntual de un parámetro? ¿Qué propiedades tiene un buen estimador? Dar ejemplos.

7. Se quiere controlar la temperatura X (en grados centigrados) de la llama que sirve para sellar los pies de unas copas de crital. Por estudios anteriores se sabe que la desviación típica de esta temperatura es igual a 5.5^oC.

   a)Midiendo 5 veces la temperatura en momentos distintos, se encuentra una media de 555^oC. Constuir un intervalo de confianza al 95% de confianza para la temperatura promedio de la llama.

   b)Se quiere pasar los resultados anteriores a grados Farenheit. Si Y es la temperatura expresada en ^{o}F, se recuerda que $Y = \frac{9}{5}X + 32$.Utilizando los datos del apartado a), ¿cuál es un intervalo de confianza al 95% para la temperatura promedio de la llama en ^{o}F? Justifica tu respuesta.

8. Para una encuesta, se entrevistan a 1025 mujeres y 427 hombres al azar entre la población española. El 47% de las mujeres entrevistadas opinan que no tienen bastante tiempo libre para si mismas.

   a) El sondeo anuncia un margen de error de ± 3 puntos para un nivel de confianza de 95 % para las mujeres. Explique a alguien que no sabe nada de estadística porque no se puede afirmar que el 47% de las mujeres españolas opinan que no tienen bastante tiempo libre para sí. Explique a continuación qué quiere decir un "nivel de confianza del 95%".

   b) El margen de error para los hombres, con el mismo nivel de confianza, es ± 4 puntos. ¿Porque este margen de error es mayor para los hombres que para las mujeres?

9. De una población normal se extrae una muestra aleatoria simple de tamaño 9 cuyos valores se presentan en la siguiente tabla:

165	162	166	164	165	170	169	165	168

   Obtener el intervalo de confianza para la media poblacional con un nivel de confianza del 98%.

10. Los resultados de una muestra aleatoria de 20 personas en el examen de ingreso MIR tuvieron una puntuación media de 540 y una desviación típica de 50. Obtener un intervalo de confianza al 95% para la media poblacional del los resultados del MIR, suponiendo que éstos siguen una distribución aproximadamente normal.

11. Se desea estudiar la influencia del calcio en el cemento estándar y en el cemento contaminado con plomo, puesto que un bajo nivel de calcio indicará que el mecanismo de hidratación del cemento queda bloqueado y esto permite que el agua ataque varias partes de la estructura del cemento. Al tomar diez muestras de cemento estándar se encontró que el peso promedio del calcio es $\overline{x} = 90.0$ con una varianza muestral de $s^2 = 25.0$; los resultados obtenidos con 15 muestras de cemento contaminado con plomo fueron $\overline{x} = 87.0$ y $s^2 = 16.0$. Supóngase que el peso de calcio está distribuido de manera normal y que la varianza en las dos poblaciones es la misma. Obtener un intervalo de confianza al 95% para la diferencia de medias obtenidas. Comentar los resultados obtenidos.

12. En una empresa se quiere estudiar el rendimiento de dos operarios. Para ello, se apunta durante cinco semanas, el número de unidades producidas por semana por el operario A, y durante las cinco semanas siguientes, por el operario B. Los resultados son los siguientes.

    Operario A : 48,44,62,58,50.

    Operario B : 56,66,66,60,50

13. Para determinar el contenido exacto de carbonato de calcio de una caliza, se realizan, en las mismas condiciones, 5 mediciones y se obtiene los resultados siguientes :

    $$49.56\% \quad 49.82\% \quad 49.30\% \quad 50.16\% \quad 50.06\%$$

    Se supone que el error que se comete durante la medición sigue una distribución normal de media 0 y de varianza desconocida σ^2.

    (a) ¿Cuál es la distribución del valor medido del contenido de carbonato de calcio? En particular, ¿cuál es su media?
    (b) Construir un intervalo de confianza al nivel de 90% para el contenido exacto de carbonato de calcio de la caliza. Presentar el resultado en forma de intervalo ($valor \pm error.$)
    (c) Si quiero cometer como máximo un error de 0.2 con una confianza del 95%, ¿cuántas mediciones más debería realizar?

14. En el análisis de un mineral que contiene bióxido de manganeso, se obtuvieron los siguientes datos :

    $$37.62\% \quad 38.23\% \quad 38.44\% \quad 37.62\% \quad 38.71\%$$

    Se supone que el error que se comete durante la medición sigue una distribución normal de media 0 y de varianza desconocida σ^2.

    Construir un intervalo de confianza al nivel de 90% para el contenido real en bióxido de manganeso.

15. En el departemento de control de calidad de una empresa, se quiere determinar si ha habido un descenso significativo de la calidad du su producto entre las producciones de dos semanas consecutivas a consecuencia de un incidente ocurrido durante el fin de semana. Deciden tomar una muestra de la producción de cada semana, si la calidad de cada artículo se mide en una escala de 100, obtienen los resultados siguientes :

    Semana 1 : 93 86 90 90 94 91 92 96

Semana 2 : 93 87 97 90 88 87 84 93

Suponiendo que las varianzas de la puntuación en las dos producciones son iguales, construye un intervalo de confianza para la diferencia de medias al nivel de 95%. Comentar los resultados obtenidos.

Dpto. Matemática Aplicada y Estadística

1. Se quiere estudiar la vida media μ de tubos fluorescentes producidos en una empresa. Por experimentos anteriores, se piensa que la desviación típica de la población es 120h. Al extraer una muestra de 100 tubos, se encuentra una media muestral de 1.570 h. (a) Contrastar la hipótesis de que $\mu = 1.600$ h, frente a la alternativa $\mu \neq 1.600$; (b) Idem., frente a la alternativa $\mu < 1.600$.

2. El programa Statistix dispone de un generador de números aleatorios, (Data->Transformations, función "Random") que produce numéros cuya distribución se supone es uniforme sobre el intervalo $[0, 1]$. Si el generador funciona bien, los números generados provienen de una población con media $\mu = 0.5$ y desviación típica $\sigma = 0.2887$. Generamos con Statistix 100 números aleatorios, y encontramos $\bar{x} = 0.4647$. LLevar a cabo el contraste $H_0 : \mu = 0.5$, $H_1 : \mu \neq 0.5$.

3. Una empresa estudia introducir un nuevo sistema de producción para mejorar su productividad media establecida actualmente en 42 unidades por persona y día. Se estima que el cambio no será rentable si no consigue elevar dicho número al menos a 45 u. Realizada una prueba con la nueva tecnología, aplicada a 35 personas, se obtuvo una producción media de 46.5 y no se observó ningún cambio apreciable en la dispersión que estaba establecida en σ=1.5 u. por día. Se debe efectuar el cambio tecnológico?

4. Hemos llevado a cabo distintos contrastes de hipótesis con Statistix, y el programa nos proporciona los p-valores siguientes: 1) :0.00012 2) 0.54 3) 0.028 4) 0.17

   Para cada uno de estos casos, determinar si se rechaza H_0 al 90% de confianza, al 95% y al 99% de confianza.

5. A continuación indicamos para distintas situaciones las hipótesis sobre la media poblacional de una variable normal para σ conocida que se plantearon, así con el valor del estadístico de prueba $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ que se obtuvo. Determinar los p-valores correspondientes y indicar cuál es su decisión acerca de H_0.

   a) $H_0 : \mu = 120$; $H_1 : \mu \neq 120$, estadístico de prueba: $z_0 = 2.32$

   b) $H_0 : \mu = 120$; $H_1 : \mu \neq 120$, estadístico de prueba: $z_0 = -1.88$

   c) $H_0 : \mu = 0.25$; $H_1 : \mu > 0.25$, estadístico de prueba: $z_0 = 1.48$

   d)$H_0 : \mu = 1000$; $H_1 : \mu > 1000$, estadístico de prueba: $z_0 = 1.59$

   e)$H_0 : \mu = 1000$; $H_1 : \mu < 1000$, estadístico de prueba: $z_0 = -0.23$

6. Se quiere determinar si unos detectores de radón (un gas inodoro y incoloro ligeramente radioactivo) son fiables. Para ello, se colocan 12 de estos detectores en una cámara y se exponen durante 3 días a 105 picoCuries por litro de radón. Los datos obtenidos son los siguientes:

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

Supongamos que sabemos que la desviación típica de las mediciones para este tipo de detectores es $\sigma = 9$, y que podemos utilizar una distribución normal.

a) Construir un intervalo de confianza al 95% para el valor promedio de radon proporcionado por este tipo de detectores.

b) ¿Podemos afirmar que este valor promedio difiere significativamente al 95% del valor real 105? Compare el resultado obtenido con el apartado anterior?

7. Una empresa de transportes desconfía de la afirmación de que la vida útil promedio de ciertos neumáticos es al menos de 28.000 Km. Para verificar la afirmación, se colocan 40 de estos neumáticos en sus camiones, obteniendo una vida útil promedio de 27.463 Km con una desviación estándar muestral de 1.348 Km. ¿Qué se puede concluir de ese dato si la probabilidad de cometer un error de Tipo I es a lo sumo de 0.01?

8. *(Adaptado del problema IV.2 de Diciembre 99)* Para calibrar un instrumento de medición, se mide repetidas veces un patrón cuyo peso sabemos igual a 10 gramos. Los valores medidos siguen una distribución normal de media desconocida.

   a)Se mide el patrón cinco veces. La media de las cinco mediciones es igual a 10.0023 gramos y la desviación típica muestral es igual a 0.0002 gramos. Construir un intervalo de confianza para la media poblacional al 98% de confianza.

   b) Queremos determinar si nuestro aparato de medición sobrevalora el peso real. Plantea el contraste de hipótesis correspondiente y llévalo a cabo con los datos del apartado (a).