1 Estadística descriptiva

1. Los resultados obtenidos al lanzar un dado 200 veces vienen reflejados en la tabla

Número	1	2	3	4	5	6
Repeticiones	?	32	35	33	?	35

   (a) Determina las frecuencias que faltan sabiendo que la puntuación media es 3.6.
   (b) Calcula la mediana, la moda y la desviación típica.

2. Dada la siguiente distribución de frecuencias absolutas acumuladas

Edad	$[0-1]$	$[2-3]$	$[4-5]$	$[6-7]$	$[8-9]$
F	4	11	24	34	40

   (a) Calcula la media aritmética y la desviación típica.
   (b) ¿Qué tanto por ciento tiene una edad inferior a 5 años?.

3. El peso medio de 5 chicas es de 52,6 Kg y el peso medio de 7 chicos es 62,8 Kg. Halla el peso medio del grupo total.

4. Una cadena hotelera tiene cinco hoteles de diferente número de plazas cada uno de ellos. Los ingresos totales y el rendimiento por habitación de cada hotel son

Hotel	1	2	3	4	5
Ingreso (euros)	1.200	2.000	2.000	1.500	1.050
Rendimiento (euros/hab.)	6	5	4	5	7

   Determina el rendimiento medio por habitación para el total de hoteles de la cadena.

5. Un médico dentista examina 150 niños que habitualmente consumieron caramelos y anota el número de piezas careadas. Los resultados fueron

Número piezas	0	1	2	3	4	5	6	7	8
Número niños	0	1	6	13	50	33	30	17	0

   Realiza una estadística descriptiva completa con estos datos.

6. Las calificaciones de 40 estudiantes fueron (sobre 100): 68 84 75 82 68 90 62 88 76 93 73 79 88 73 60 93 71 59 85 75 61 65 75 87 74 62 95 78 63 72 66 78 82 75 94 77 69 74 68 60. Agrupa los datos en intervalos, represéntalos, calcula la media aritmética y la geométrica, la varianza, la desviación típica, la cuasivarianza y la cuasidesviación típica, la moda, la mediana, los cuartiles y los percentiles.

7. Se mide repetidamente la longitud de un objeto (en milímetros) obteniéndose los siguientes resultados: 6,191; 6,199; 6,215; 6,189; 6,228; 6,203; 6,200; 6,225; 6,217; 6,205. Realiza una estadística descriptiva completa con estos datos. Estima la longitud de dicho objeto y el error de la medida.

2 Probabilidad

1. Se lanzan dos monedas y un dado. Se pide
   (a) Describir el espacio muestral.
   (b) Expresar explícitamente los siguientes sucesos:
   A = "salir dos caras y un número par"
   B = "salir un dos"
   C = "salir exactamente una cara y un número primo"
   (c) Expresar explícitamente los siguientes sucesos: A y B ocurren, sólo ocurre B, B ó C ocurren,
2. Sean A y B sucesos tales que $P(A) = 1/2$, $P(A \cup B) = 3/4$, $P(\bar{B}) = 5/8$. Calcula $P(A \cap B)$, $P(\bar{A} \cap \bar{B})$, $P(\bar{A} \cup \bar{B})$ y $P(\bar{A} \cap B)$.
3. En una ciudad se publican dos periódicos, A y B. El 40 % de la población lee A, el 30 % lee B y el 10 % lee los dos. Se pide calcular:
   (a) El porcentaje de personas que leen al menos uno de los dos periódicos.
   (b) El porcentaje que lee sólo A.
4. Supongamos que estamos en la última semana de la liga, el Barcelona juega en casa y el Madrid fuera de casa. Si la probabilidad de ganar en casa es 1/2, de ganar fuera de casa es 1/6 y de empatar es siempre 1/3, se pide:
   (a) Describir el espacio muestral de la última jornada de liga para Madrid y Barca. (Ayuda: un elemento del espacio muestral sería "Madrid gana y Barca gana")
   (b) Calcula las probabilidades de todos los elementos del espacio muestral.
   (c) Si el F.C. Barcelona llevase dos puntos de ventaja al Real Madrid ¿Qué probabilidad tendría el Madrid de ganar la liga? (si un equipo gana recibe 3 puntos, si empata 1 punto y si pierde ninguno) ¿Y si la ventaja fuera de 1 punto?
5. En una reunión hay 25 personas. Calcula la probabilidad de que celebren su cumpleaños el mismo día al menos dos personas.
6. Un examen consta de 5 temas numerados. Para elegir un tema al azar se propone lanzar un dado. Si sale de uno a cinco, el número del tema es el resultado dado; si sale un seis se vuelve a tirar hasta que salga de 1 a 5. Demostrar que la probabilidad de elección de cada tema es 1/5.
7. Una urna contiene 8 bolas blancas y 7 negras; hacemos una extracción de 2 bolas; si sabemos que una de ellas es negra, ¿cuál es la probabilidad de que la otra también lo sea?
8. ¿Cuál es la probabilidad de que al lanzar n veces una moneda obtengamos n caras? ¿Y n caras o menos?
9. Una urna contiene x bolas blancas e y negras. Una persona saca k bolas ¿Cuál es la probabilidad de que z sean blancas y k-z negras?
10. Una urna A contiene 2 bolas blancas y 3 negras; otra urna B está compuesta de 4 blancas y 3 negras. Se saca una bola al azar de la urna A y sin verla se echa en B; a continuación se saca una bola de B que resultó ser negra ¿Cuál es la probabilidad de que la bola pasada de A a B fuese blanca?
11. Hay 6 cajas que contienen 12 tornillos buenos y malos; una tiene 8 buenos y 4 defectuosos; dos cajas 6 buenos y 6 malos y tres 4 buenos y 8 malos. Se elige una caja al azar y se extraen 3 tornillos, sin reemplazamiento de dicha caja; de éstos 2 son buenos y 1 defectuoso ¿Cuál es la probabilidad de que la caja elegida contenga 6 buenos y 6 malos?
12. Una baraja española de 48 cartas se ha dividido en dos partes: pares e impares. Lanzamos un dado y extraemos una carta del grupo de las pares o de las impares, según que salga 6 ó no. Si resulta ser figura, hallar la probabilidad de que sea un 11.

13. Una urna se ha llenado tirando una moneda al aire dos veces y poniendo una bola blanca por cada cara y una bola negra por cada cruz. Se extrae una bola que es blanca. Halla la probabilidad de que la otra bola también lo sea.

14. Tenemos dos barajas españolas completas. A una de ellas le falta una carta y no sabemos cuál es. Elegimos una baraja al azar y sacamos una carta. Halla la probabilidad de que sea de oros.

15. Se sabe que 5 de cada 100 hombres y 25 de cada 10.000 mujeres son daltónicos. Si se elige un daltónico al azar, ¿cuál es la probabilidad de que sea varón?

16. El 70 % de los diputados son del partido A y el resto del B. El 60 % de los primeros y el 10 % de los segundos tienen más de 50 años. Si se elige un diputado al azar y tiene más de 50 años, ¿cuál es la probabilidad de que pertenezca al partido B?

17. En el interior de un círculo se selecciona un punto al azar. ¿Cuál es la probabilidad de que dicho punto esté más cerca del centro que de la circunferencia? Si se inscribe un círculo en un cuadrado ¿cuál es la probabilidad de que el punto esté fuera del círculo?

18. Cuando la concentración en sangre de una determinada sustancia en un individuo sobrepasa cierto valor se dice que el individuo pertenece a cierto grupo de riesgo. Supongamos que el 5 % de los habitantes de una población pertenece al grupo de riesgo. Determinar la probabilidad de que de tres individuos elegidos al azar, uno esté en el grupo de riesgo.

19. En una universidad el 4 % de los chicos y el 1 % de las chicas tienen una altura superior a 180 cm. El 60 % de los alumnos son chicas. Se toma un alumno al azar y se comprueba que mide más de 180 cm. Hallar la probabilidad de que tal alumno sea chico.

20. El volumen de ventas en un concesionario de coches es de 500 unidades (al año) para el deportivo, 1.000 para el familiar y 2.000 para el utilitario. Se sabe que el porcentaje de coches defectuosos es de un 2 % para el primero, un 1 % para el segundo y un 3 % para el tercero. Se pide calcular la probabilidad de los siguientes sucesos:

    (a) Elegido uno al azar, que éste no sea defectuoso.

    (b) Habiendo elegido uno defectuoso, que este automóvil pertenezca al grupo de los familiares.

21. Se dispone de dos urnas U_1 y U_2. La urna U_1 contiene el 70 % de las bolas blancas y el 30 % de negras, y en la urna U_2 hay un 30 % de blancas y un 70 % de negras. Se selecciona una urna al azar (se supone que ambas tienen la misma probabilidad de ser elegidas) y se toman 10 bolas una tras otra con reemplazamiento. El resultado fue el suceso S=bnbbbbnbbb, siendo b bola blanca y n bola negra. ¿Cuál es la probabilidad de que el suceso S proceda de la urna U_1?

22. Se tienen 10 urnas. Nueve de ellas contienen 2 bolas blancas y dos negras y una contiene cinco blancas y una negra. Se elige una urna de forma aleatoria y se extrae una bola resultando ser blanca. ¿Cuál es la probabilidad de que la urna elegida sea la que posee 4 blancas y una negra?

23. La probabilidad de que tres jugadores de dardos hagan diana es 4/5, 3/4 y 2/3. Supongamos que tiran los tres y se producen dos aciertos sobre la diana. ¿Cuál es la probabilidad de que haya fallado el tercero?

24. Dada una prueba diagnóstica para una enfermedad con las propiedades:

    (a) La probabilidad de que el test dé positivo teniendo la enfermedad es del 95 %.

    (b) La probabilidad de que el test dé negativo cuando no se tiene la enfermedad es del 95 %.

    (c) La probabilidad de que una persona padezca la enfermedad es del 0,5 %.

    i) Calcula la probabilidad de que el test dé positivo. ii) Calcula la probabilidad de que se padezca la enfermedad cuando el test ha dado positivo.

3 Variables aleatorias

1. En el experimento que consiste en lanzar tres monedas y anotar el número de caras obtenidas, calcula:

   (a) La función de probabilidad y su representación.

   (b) La función de distribución y su gráfica.

   (c) La media y la varianza de la distribución.

   (d) Si X es la variable que expresa el número de caras obtenidas, halla $P(1 < X < 3)$.

2. Sea X una variable aleatoria cuya función de probabilidad viene dada por $P(X = r) = 1/8$ ($r = 2, 3, ..., 9$). Se pide hallar:

   (a) La función de distribución y su gráfica

   (b) $P(X < -3)$ y $P(4 < X < 7)$.

3. Una compañía de bebidas anuncia premios en los tapones asegurando que en cada 1000 tapones hay 500 con "inténtelo otra vez", 300 con premio de 5 euros, 150 con premio de 10 euros, 40 con premio de 50 euros y 10 con premio de 100 euros. Un individuo al que no le gusta esa bebida decide comprar una botella cuyo coste es 10 euros. Caracterizar su ganancia mediante una variable aleatoria y calcular su esperanza. Calcula su probabilidad de perder dinero.

4. En una calle hay un semáforo que está verde para los coches durante un minuto y rojo durante 15 segundos. Suponiendo un automovilista llega al semáforo con igual probabilidad en cualquier instante, calcula el tiempo medio de espera.

5. En un puesto de feria se ofrece la posibilidad de lanzar un dardo a unos globos. Si se consigue reventar un globo, se recibe un premio igual a una cantidad oculta tras el globo. Supongamos que la probabilidad de acertar con algún globo es 1/3. Los premios se distribuyen de la siguiente manera:

   - 40 % de premios de 0.50 euros
   - 30 % de premios de 1 euro
   - 20 % de premios de 2 euros
   - 10 % de premios de 6 euros

   Si cada lanzamiento cuesta 1 euro, ¿cuál es la ganancia esperada por el dueño del puesto en cada lanzamiento?

6. La función de densidad de una variable aleatoria continua viene definida por

$$f(x) = \begin{cases} 2x, & \text{si } 0 \leq x \leq 1, \\ 0, & \text{en el resto.} \end{cases}$$

   Halla su función de distribución, la media y la desviación típica. Si la variable aleatoria X representa el nivel en sangre de una sustancia G en una población. Calcula la probabilidad de que tomando un individuo al azar su nivel de sustancia G sea mayor de 5 u (unidades).

7. Dada la función de distribución de una variable alatoria continua X, definida de la forma:

$$F(x) = \begin{cases} 0, & \text{si } 0 \leq x, \\ \text{sen} x, & \text{si } 0 < x \leq \frac{\pi}{2}, \\ 1 & \text{si } x > \frac{\pi}{2}. \end{cases}$$

   Calcula la función de densidad, su desviación típica y la probabilidad de que X esté comprendido entre $-\pi$ y π. Si X representa el contenido de aluminio en una muestra de un determinado producto industrial, halla la probabilidad de que tomando una unidad de producto su contenido en aluminio sea mayor que 0.4 u. Determina el contenido de aluminio más probable.

8. La variable ξ tiene como función de densidad:

$$f(x) = \begin{cases} x, & \text{si } x \in [a, b], \\ 0, & \text{si } x \notin [a, b]. \end{cases}$$

Determina el valor de a y b si sabemos que $P(a < \xi < 2a) = 0.375$. Si ξ representa el nivel de calcio en los dientes en una población, ¿qué nivel de calcio es el más esperado? halla la desviación típica.

9. Una variable aleatoria X tiene como función de densidad:

$$f(x) = \begin{cases} \frac{k}{\sqrt{x}}, & \text{si } 0 < x < 500, \\ 0, & \text{en el resto.} \end{cases}$$

Determina: a) el valor de k; b) la función de distribución; c) $P(0.5 < X < 1.5)$; d) la desviación típica; e) la esperanza matemática; f) si representa el número de personas que asiste diariamente a una sala de reuniones, ¿cuántas sillas han de estar disponibles en la sala para poder atender a estas personas con una probabilidad no menor de 0.90?

10. El tiempo de vida (en minutos) de un determinado virus es una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} \frac{1}{1000} e^{-x/1000}, & \text{si } x > 0, \\ 0, & \text{en el resto.} \end{cases}$$

a) Halla la probabilidad de que el tiempo de vida sea superior a 100 minutos e inferior a 1000 minutos. b) Observamos el virus a los 500 minutos y comprobamos que ha muerto. ¿Cuál es la probabilidad de que estuviese vivo a los 100 minutos?

11. Un autobús pasa por una parada cada 20 minutos. Para un viajero que llega a dicha parada de improviso, la variable aleatoria "tiempo de espera" sigue una función de densidad de la forma

$$f(x) = \begin{cases} \dfrac{k}{20} & \text{si } 0 \leq x \leq 20 \\ 0 & \text{en el resto.} \end{cases}$$

Calcula: a) la constante k y la función de distribución de dicha variable aleatoria. b) La probabilidad de que espere al autobús menos de 7 minutos. c) La esperanza y la varianza de la variable aleatoria. d) Si la persona sabe que en los 5 minutos anteriores a su llegada no pasó y decide esperar como mucho 10 minutos, ¿cuál es la probabilidad de que coja el autobús?

12. Sea la variable aleatoria X cuya función de densidad es

$$f(x) = \begin{cases} x & \text{si } x \in [a, b], \\ 0 & \text{en el resto,} \end{cases}$$

donde a y b son números reales tales que $0 < a < b$.

   (a) Calcula el valor de a y el de b sabiendo que $P(a \leq X \leq 2a) = 3/8$.
   (b) Determina la función de distribución de X.
   (c) Si llamamos A al suceso $A = \{\frac{a}{2} \leq X \leq \frac{3a}{2}\}$ y B al suceso $B = \{a \leq X \leq 2a\}$, calcula $P(A/B)$.

13. El periodo de hospitalización en días, para pacientes que siguen un tratamiento para un cierto tipo de desorden renal es una variable aleatoria X que tiene como función de densidad

$$f(x) = \begin{cases} \dfrac{k}{(x+4)^3} & \text{si } x > 0, \\ 0 & \text{en el resto.} \end{cases}$$

a) Halla k y la función de distribución de la variable X y la probabilidad de que el periodo de hospitalizacin para un paciente que sigue este tratamiento supere los 8 días.

4 Distribuciones

1. Suponiendo que la probabilidad de que un niño que nace sea varón es 0,51, halla la probabilidad de que una familia de 6 hijos tenga:
   (a) por lo menos una niña,
   (b) por lo menos un niño,
   (c) por lo menos dos niños y una niña.
2. Una compañía de seguros con 10.000 asegurados halla que el 0,005 % de la población fallece cada año de un cierto tipo de accidente.
   (a) Halla la probabilidad de que la compañía tenga que pagar a más de tres asegurados por dicho accidente en un año determinado.
   (b) ¿Cuál es el número medio de accidentes por año?
3. La probabilidad de que un individuo tenga una reacción alérgica al inyectarle un suero es 0,001. Halla la probabilidad de que, entre 2.000 individuos, tengan reacción alérgica a) exactamente tres, b) más de dos.
4. Una máquina produce varillas metálicas. Las longitudes siguen una normal con $\mu = 19,8$ cm y $\sigma = 5$ mm. La normativa exige que la longitud de las varillas se sitúe entre 19,5 y 20,5 cm. ¿Qué porcentaje de las varillas satisface la normativa?
5. La probabilidad de error en la transmisión de un bit por un canal de comunicación es $p = 10^{-4}$. ¿Cuál es la probabilidad de que se produzcan más de dos errores al transmitir un bloque de 1.000 bits?
6. El coeficiente de inteligencia es una variable aleatoria que se distribuye según una $N(100; 16)$. Calcula:
   (a) La probabilidad de que un individuo elegido al azar tenga un coeficiente superior a 120.
   (b) Suponiendo que un individuo con carrera universitaria debe tener un coeficiente superior a 110, halla la probabilidad de que un licenciado tenga un coeficiente superior a 120.
7. Consideramos un curso con 10 alumnos, la variable $X =$"número de alumnos que han nacido en primavera" es una variable binomial $X \sim B(10; p)$, donde p es la probabilidad de que una persona cualquiera nazca en primavera. Si $p = 0.25$, halla a) el número esperado de alumnos nacidos en primavera, b) la varianza de X, c) la probabilidad de que exactamente 3 alumnos hayan nacido en primavera, d) la probabilidad de que hayan nacido a lo sumo 2.
8. Sea una clase con 50 alumnos. Llamamos X a la variable aleatoria "número de alumnos nacidos en enero". Suponiendo que la probabilidad de nacer en cada mes del año es la misma, $p = 1/12$. Halla a) el número esperado de alumnos nacidos en primavera, b) la varianza de X, c) la probabilidad de que exactamente 3 alumnos hayan nacido en primavera, d) la probabilidad de que hayan nacido a lo sumo 2. Realiza el problema suponiendo que X sigue una distribución $B(50; p)$ y lo mismo suponiendo que sigue una distribución de Poisson.
9. Si la probabilidad de acertar en un blanco es $1/5$ y se hacen 10 disparos en forma independiente, ¿cuál es la probabilidad de acertar por lo menos dos veces? Describe el problema con una variable aleatoria. ¿Qué tipo de distribución sigue? Calcula el número esperado de aciertos.
10. Si hay, en promedio, un 1 por ciento de zurdos, ¿cuál es la probabilidad de tener por lo menos 4 zurdos entre 200 personas? Describe el problema con una variable aleatoria. ¿Qué tipo de distribución sigue? Calcula el número esperado de zurdos.
11. En una población el 1 por ciento padece daltonismo, ¿cuál es la probabilidad de que en una muestra de 100 personas, a) ninguna padezca daltonismo, b) dos o más lo padezcan? ¿cuál es el número esperado de daltónicos? ¿qu tamaño debería tener la muestra para asegurar que hay un daltónico con probabilidad mayor o igual a 0.95?

12. Una compañía aérea sabe por experiencia que el 12 % de las reservas telefónicas de plazas no se llevan a efecto, de modo que reserva más plazas de las que dispone. Si en un vuelo hay 150 plazas, ¿cuántas reservas puede hacer para que la probabilidad de cubrir al menos 145 plazas sea del 99 %?

13. Una universidad pública oferta 120 plazas para el primer curso de los estudios de Química, recibiendo 800 solicitudes y siendo el único criterio de admisión la nota del examen de selectividad. Sabiendo que las notas de selectividad siguen una dsitribución normal de media $7,3$ y desviación típica $0,7$, determina la nota mínima necesaria para conseguir una de las 120 plazas ofertadas.

14. La presión de la sangre arterial en reposo en escolares de edades comprendidas entre 10 y 13 años, es una variable normal de media 120 mm de Hg y desviación típica 15 mm. Determina el porcentaje de escolares

    (a) Con presión inferior a 104 mm.

    (b) Con presión superior a 110 mm.

    (c) Con presión entre 100 mm y 120 mm.

    Determina la presión por debajo de la cual se encuentra el 80 % de la población.

15. En una ciudad hay epidemia de gripe de modo que la probabilidad de que un individuo aislado (sin considerar posibles situaciones de contagio directo) contraiga la enfermedad es 0,2. Considerados 8 individuos, determina

    (a) Probabilidad de que contraiga la enfermedad uno sólo.

    (b) Probabilidad de que al menos dos caigan enfermos.

    (c) Probabilidad de que ninguno enferme.

16. Repetir el ejercicio anterior con un colectivo de 100 individuos vacunados, de manera que la probabilidad de enfermar ha descendido a 0,02.

17. Un botánico ha observado que la anchura, X, de las hojas del álamo sigue una distribución normal con $\mu = 6$ cm., y que el 90 % de las hojas tiene una anchura inferior a 7,5 cm. a) Halla σ. b) Halla la probabilidad de que una hoja mida más de 8 cm.

18. La duración, en minutos, de un proceso textil sigue una distribución $N(\mu, \sigma)$. El 60 % de las veces dura más de 40 minutos. El 55 % de ellas dura menos de 50 minutos. Halla μ y σ.

19. En un examen se plantean 10 cuestiones a las que debe responderse verdadero o falso. Un alumnos aprobará el examen si, al menos, 7 respuestas son acertadas. ¿Qué probabilidad de aprobar tiene un estudiante que responde todo al azar? ¿Y uno que sabe el 30 % de la asignatura?

20. De una variable aleatoria que sigue una distribución normal se sabe que $P(X > 3) = 0,8413$ y $P(X > 9) = 0,0228$. Calcula la media y la desviación típica de esta distribución normal y $P(X < 6)$.

21. La demanda mensual de coches utilitarios en Ciudad Real sigue una distribución $N(\mu, \sigma)$. El 80 % de las veces se demandan más de 200 coches. El 60 % de ellas se demandan menos de 300. Halla μ y σ. Calcula la probabilidad de que se soliciten más de 500.

5 Inferencia. Estimación por intervalos de confianza

1. En una población se desea conocer la probabilidad de que un individuo sea alérgico al polen de los olivos. En 100 individuos tomados al azar se observaron 10 alérgicos. Halla el intervalo de confianza al 95 % para la probabilidad pedida. ¿Cuántos individuos se deberían observar para que con probabilidad 0.95, el error máximo en la estimación de la proporción de alérgicos sea de 0.01?

2. Se supone que el número de erratas por página en un libro sigue una distribución de Poisson. Elegidas al azar 95 páginas, se obtuvieron los siguientes resultados:

Número de erratas	0	1	2	3	4	5
Número de páginas	40	30	15	7	2	1

   Halla el intervalo de confianza al 90 % para el número medio de erratas por página en todo el libro.

3. Se mide el tiempo de duración (en segundos) de un proceso químico realizado 20 veces en condiciones similares, obteniéndose los siguientes resultados: 93; 90; 97; 90; 93; 91; 96; 94; 91; 91; 88; 93; 95; 91; 89; 92; 87; 88; 90; 86. Suponiendo que la duración sigue una distribución Normal, halla los intervalos de confianza al 90 % para ambos parámetros.

4. Las tensiones de rotura (en Kp.) de 5 cables de un determinado metal fueron 660, 460, 540, 580, 550. Suponiendo normalidad para las tensiones: a) Estima la tensión media de rotura mediante un intervalo de confianza al nivel 0.95. b) Estima σ^2 mediante un intervalo de confianza al nivel 0.90.

5. En una población, la altura de los individuos varones sigue una $N(\mu; \sigma = 7.5)$. Halla el tamaño de la muestra para estimar μ con un error inferior a ± 2 cm con un nivel de confianza 0.90.

6. La vida activa (en días) de cierto fármaco sigue una distribución $N(1200; 40)$. Se desea enviar un lote de medicamentos de modo que la vida media del lote no sea inferior a 1180 días, con probabilidad 0.95. Halla el tamaño del lote.

7. Se intenta estudiar la influencia de la hipertensión en los padres sobre la presión sanguínea de los hijos. Para ello se seleccionan dos grupos de niños, unos con padres de presión sanguínea normal (grupo 1) y otros con uno de sus padres hipertenso (grupo 2), obteniéndose las siguientes presiones sistólicas:

Grupo 1:	104	88	100	98	102	92	96	100	96	96
Grupo 2:	100	102	96	106	110	110	120	112	112	90

   Halla el intervalo de confianza para la diferencia de medias, suponiendo que las varianzas en las dos poblaciones de niños son iguales.

8. Una noticia del periódico dice que, de 1000 persona encuestadas sobre una cuestión, 556 se muestran a favor y 444 en contra, y concluye afirmando que el 55.6 % de la población se muestra a favor, con un margen de error de $\pm 3\%$. ¿Cuál es el nivel de confianza de esta afirmación?

9. Se quiere estudiar la proporción p de declaraciones de la renta que presenta algún defecto. En una muestra preliminar pequeña (muestra piloto) de tamaño 50 se han observado 22 declaraciones defectuosas. ¿Cuál es el tamaño muestral necesario para estimar p cometiendo un error máximo de 0.01 con una probabilidad de 0.99?

10. Se quiere comparar dos métodos, A y B, para determinar el calor latente de fusión del hielo. La siguiente tabla da los resultados obtenidos (en calorías por gramo de masa para pasar de -0.72^o C a 0^o) usando reiteradamente ambos métodos):

Método A:	79.98	80.04	80.02	80.04	80.03	80.03	80.04	79.97	80.05	80.03
Método B:	80.02	79.94	79.98	79.97	79.97	80.03	79.95	79.97		

    Se pide obtener un intervalo de confianza 0.95 para la diferencia de las mediciones medias obtenidas por ambos métodos.

11. En una gran zona ganadera se desea estimar la proporción de ovejas que sufren cierta enfermedad degenerativa. Calcula el tamaño muestral necesario para estimar esta proporción con un error menor que 0.03 a un nivel de confianza del 0.95 sabiendo que en una pequeña muestra preliminar de 30 ovejas dos padecían la enfermedad.

12. Calcula el mínimo tamaño muestral necesario para cometer un error menor que 0.05, con una probabilidad de 0.99, en la estimación de la proporción de personas que tienen sensibilidad para la feniltiocarbamina sabiendo que en una muestra de 60 personas se han encontrado 14 que son sensibles a este producto.

6 Inferencia. Contraste de hipótesis

1. Los valores observados para la tensión de ruptura de una muestra de 12 elementos de una determinada fibra sintética es: 9.9; 6; 5.2; 7.3; 11.8; 10.3; 8.2; 7.5; 6.6; 12.6; 16.8; 12.3

   (a) Determina el intervalo de confianza para la tensión media de ruptura.

   (b) Calcula el intervalo de confianza para la varianza poblacional de la variable tensión.

   (c) De acuerdo con los datos, ¿es aceptable la afirmación según la cual dicha fibra soporta por término medio una tensión de ruptura igual a 12?

2. Una marca de cigarrillos de tabaco negro declara en cada paquete un contenido medio de nicotina de 0.8 mg por cigarrillo. La Asociación de Fumadores Empedernidos de Tabaco Alto en Nicotina (AFETAN) mide el contenido de nicotina de 30 cigarrillos con la finalidad de comprobar que no les dab menos nicotina de la prometida, obteniendo una media muestral de 0.75 mg con una desviación típica de 0.08 mg. ¿Estaría justificada una reclamación de la AFETAN en el sentido de que el contenido real de nicotina es menor que el que se asegura?

3. El tiempo de cicatrización de una herida con un tipo de pomada en 20 pacientes es:

   10; 15; 15; 16; 17; 17; 17; 18; 20; 21; 21; 21; 21; 21; 22; 22; 25; 28;29; 30

   (a) Determina el intervalo de confianza para el tiempo medio de curación.

   (b) Calcula el intervalo de confianza para la varianza poblacional de la variable tiempo de curación.

   (c) De acuerdo con los datos, ¿es aceptable la afirmación según la cual el tiempo medio de curación de una herida con dicha pomada es igual a 20 días?

4. Se considera buena la edición de un libro si el número medio de erratas por página no supera el 0,1 (H_0). Dadas las pruebas de imprenta, se eligen 10 páginas al azar, y se rechazan las pruebas si se observan 2 ó más erratas. Se supone que el número de erratas por página sigue una Poisson. ¿Qué nivel de significación tiene el contraste? ¿Con qué probabilidad aceptaremos un libro si realmente tiene una media de 0,2 erratas por página?

5. En una piscifactoría se desea contrastar la hipótesis (H_0) de que el porcentaje de peces adultos que miden menos de 20 cm. es, como máximo, del 10 %. Para ello, se va tomar una muestra de 6 peces y rechazaremos H_0 si encontramos más de un pez con longitud inferior a 20 cm. ¿Cuál es el nivel de significación del contraste?

6. Se recibe un envío de latas de conserva de las que se afirma que le peso medio es de 150 gr. Examinada una muestra de 6 latas se obtiene un peso medio de 145 gr. con una cuasivarianza de $s^2 = 2,3$. Al nivel de confianza del 95 % ¿se puede aceptar que el peso medio es 150 gr.?

7. Un dentista afirma que el 30 % de los niños de 9 años presenta caries dental. Tomada una muestra de 100 niños, se observó que 24 presentaban indicios de caries, ¿se puede aceptar al nivel de confianza del 90 % la hipótesis del dentista?

8. La proporción de gente que votó a un partido en unas elecciones es el 30 %. Se toma hoy una muestra de 500 electores y se obtiene el 25 % de votantes de dicho partido. ¿Hay evidencia de un cambio en el número de votos?

9. Se espera que la resistencia en Kg/cm^2 de cierto material suministrado por un proveedor se distribuya normalmente. Se toma una muestra de 9 elementos, obteniendo: 203, 229, 215, 220, 233, 208, 228, 209. Contrastar la hipótesis de que esta muestra proviene de una población con media 220. Contrastar la hipótesis de que esta muestra proviene de una población con $\sigma = 7,75$.

10. La proporción de defectos en un lote de 100 carpetas de un proveedor es 0,04, mientras que dicha proporción en 150 carpetas de otro proveedor es 0,07. ¿Hay evidencia suficiente de diferencias entre los proveedores?

11. Queremos comparar las duración de las cintas de video de dos marcas. Para esto obtenemos dos muestras aleatorias con los siguientes resultados:

    Marca A: 230 235 238 242 242 246

    Marca B: 232 234 239 245 248 253

    Aceptando normalidad e igualdad de varianzas, ¿se puede considerar estadísticamente probado (al nivel 0,10) que la duración media de las cintas de la marca B es superior a las de la marca A?

12. En un estudio sobre la influencia del sexo en la aptitud matemática se tomaron dos muestras independientes de tamaño 10 de los alumnos de cierta escuela y se obtuvieron las siguientes puntuaciones en un examen:

    chicos: 92 84 93 91 93 90 86 89 91 88

    chicas: 88 85 82 90 81 93 87 92 86 85

    ¿proporcionan estos datos suficiente evidencia estadística, al nivel 0,05, para afirmar que las aptitudes matemáticas son distintas en ambos sexos?

13. Se supone que las calificaciones de un examen de Estadística se distribuyen normalmente con un valor medio desconocido y varianza igual 0.5. Se selecciona al azar una muestra de cinco estudiantes cuyas calificaciones han sido 5,5; 6,2; 8,5; 4,0; 3,5. Verifica la hipótesis nula $H_0 : \mu = 5$ con niveles de significación de 0.05 y 0.01 y calcula el intervalo de confianza al 95 % y 99 %.

14. Realiza el ejercicio anterior suprimiendo la suposición de que la varianza es 0.5. Calcula el intervalo de confianza al 95 % para la varianza.

15. En otro grupo del mismo curso se eligió una muestra de seis estudiantes cuyas calificaciones fueron 5,0; 4,5; 6,0; 7,5; 9,0; 8,0. ¿Se puede decir a un nivel del 0.01 que es mejor el segundo grupo?

7 Contrastes χ^2

1. Después de lanzar un dado 300 veces se han obtenido las siguientes frecuencias:

Cara:	1	2	3	4	5	6
Frecuencia:	43	49	56	45	66	41

    Al nivel de significación 0,05, ¿se puede afirmar que el dado es regular?

2. En el trascurso de dos horas, el número de llamadas por minuto solicitadas a centralita telefónica fue:

Número de llamadas/minuto:	0	1	2	3	4	5	6
Frecuencia:	6	18	32	35	17	10	2

    ¿Se puede aceptar que el número de llamadas por minuto sigue una distribución de Poisson?

3. Con objeto de controlar la producción de una máquina que produce láminas de madera, se inspeccionan 100 láminas al azar, con un espesor medio de 9,7 mm y una cuasivarianza de la muestra de 1,05. Los resúmenes de los resultados muestrales se indican a continuación:

    20 láminas con espesor inferior a 9 mm.

    38 láminas con espesor entre 9 y 10 mm.

    25 láminas con espesor entre 10 y 11 mm.

    17 láminas con espesor superior a 11mm.

¿Se puede afirmar que el espesor de los datos se ajusta a una distribución normal con un nivel de confianza del 95 %?

4. Se desea estudiar el número de accidentes por día que se producen en cierto regimiento. Para ello se toman al azar los partes de 200 días dentro de los últimos 5 años, encontrando los siguientes resultados:

Número de accidentes/día	0	1	2	3	4	5	6
Número de días	58	75	44	18	3	1	1

¿Se puede aceptar, con nivel de confianza del 90 %, que el número de accidentes por día sigue una distribución de Poisson?

5. La demanda diaria, en unidades de un producto, durante 20 días fue: 65 25 71 34 62 35 38 42 30 58 60 61 66 28 50 67 72 60 70 36. Prueba si es cierta la hipótesis nula (al nivel 0,05) de que estos datos se encuentran normalmente distribuidos con media 50 y desviación típica 10.

6. Los siguientes datos indican el tiempo de espera (en minutos) de 15 clientes en una sucursal bancaria: 2 5 7 1 3.5 15 8 2 1.5 4.8 3 3 4.6 2.1 2.9. Prueba si es cierta la hipótesis nula (al nivel 0,05) de que estos datos se encuentran siguen una distribución $N(4,2)$.

7. En un proceso de producción se toma una muestra aleatoria diaria de 100 artículos y se inspecciona para encontrar artículos defectuosos, se observó el siguiente número de unidades defectuosas: el lunes 12, el martes 7, el miércoles 6, el jueves 5 y el viernes 10. Se puede decir (al nivel de significación 0,05) que la poporción de artículos defectuosos es del 8 %?

8 Regresión y correlación

1. Ajustar una recta a los datos siguientes:

X:	21	21	33	21	27	35	25	37	25	18	21	37	45	27	18	35
Y:	5	6	6	8	6	7	5	8	7	4	7	9	10	5	5	8

   (a) Halla el coeficiente de correlación lineal e interpreta el resultado.
   (b) Estima el valor de Y para $X = 30$.
   (c) Halla la tabla ANOVA correspondiente y deduce si hay una dependencia lineal entre las variables X e Y (con un nivel de significación 0,05).

2. La concentración de dos sustancias en la sangre parece estar relacionada. Para estudiar esta posible relación se miden estas cantidades en 30 personas, obteniéndose los siguientes resultados:
$\sum x_i = 41,2$, $\sum y_1 = 63,8$, $\sum x_i y_i = 118,7$, $\sum x_i^2 = 188,2$, $\sum y_i^2 = 296,4$.

   (a) Halla la recta de regresión de Y sobre X.
   (b) Halla el coeficiente de correlación lineal e interpreta el resultado.
   (c) Estima el valor de Y para $X = 30$.
   (d) Halla la tabla ANOVA correspondiente y deduce si hay una dependencia lineal entre las variables X e Y (con un nivel de significación 0,05).

3. En un estudio sobre la resistencia a bajas temperaturas del bacilo de la fiebre tifoidea, se expusieron cultivos del bacilo durante diferentes periodos de tiempo a -5^o C. Los siguientes datos representan X= tiempo de exposición (en semanas), Y= porcentaje de bacilos supervivientes.

   X: 0 0,5 1 2 3 5 9 15

   Y: 100 42 14 7,5 0,4 0,11 0,05 0,002

   $\sum x_i = 35,5$, $\sum y_i = 164,062$, $\sum \log y_i = 0,664$, $\sum x_i^2 = 345,25$, $\sum y_i^2 = 12.016,42$, $\sum(\log y_i)^2 = 99,52$, $\sum x_i y_i = 52,23$, $\sum x_i \log y_i = -125,394$.

   (a) Ajusta a una recta y una exponencial los datos.

(b) Calcula los distintos coeficientes de correlación e interpreta los resultados.

(c) Estima el porcentaje de bacilos supervivientes al cabo de 7 semanas con los dos modelos.

(d) Realiza el análisis de la varianza para ambos ajustes (al nivel 0,01) e interpreta los resultados.

4. En un estudio de laboratorios se han medido para una especie canina las variables peso (X) y concentración en sangre de cierta sustancia (Y). Los datos resumidos son los siguientes:

   n=7, $\sum x_i = 13,5$, $\sum y_i = 11,7$, $\sum \frac{1}{x_i} = 3,7281$, $\sum x_i^2 = 26,75$, $\sum y_i^2 = 19,83$, $\sum \frac{1}{x_i^2} = 2,0374$, $\sum x_i y_i = 22,23$, $\sum \frac{x_i}{y_i} = 6,3206$.

   (a) Calcula el coeficiente de correlación entre X e Y.

   (b) Ajusta los datos a una curva de la forma $Y = a + b\frac{1}{x}$.

   (c) Estima el contenido de dicha sustancia en sangre para un perro de 2 Kg. de peso.

5. Disponemos de los siguientes datos referentes a cinco pares de observaciones de dos variables X e Y:

   $\sum x_i = 20$, $\sum y_i = 15,70$, $\sum \log x_i = 6,5792$, $\sum x_i^2 = 90$, $\sum y_i^2 = 55,89$, $\sum(\log x_i)^2 = 9,4099$, $\sum x_i \log y_i = 29,025$, $\sum y_i \log x_i = 22,6926$.

   a) Expresa Y en función de X mediante un modelo de la forma $Y = a + b \log X$. b) Utilizando el modelo hallado, da una estimación del valor de Y cuando $X = 3,5$.

6. Los siguientes datos corresponden a la evolución del peso celular (en mgr./ml.) y la cantidad de nitrato en un cultivo de algas durante 3 días (mediciones cada 24 horas).

Tiempo (T)	Peso (X)	Nitrato (Y)
Inicio	0,07	12,5
1 día	0,19	10,4
2 días	0,52	7,8
3 días	1,07	4,5

   (a) Ajusta una recta y una exponencial a los datos peso (X) y cantidad de nitrato (Y).

   (b) Ajusta una exponencial a la evolución temporal del peso.

   (c) Mediante lo obtenido en los apartados anteriores, estima la cantidad de nitrato que había en el cultivo al cabo de 36 horas.

   (d) Halla la tabla ANOVA correspondiente para las variables X e Y y deduce si hay una dependencia lineal entre dichas variables (con un nivel de significación 0,05).

   (e) Halla la tabla ANOVA correspondiente para las variables X e $\log Y$ y deduce si hay una dependencia lineal entre dichas variables (con un nivel de significación 0,05).

7. Se desea estudiar la relación entre la intensidad de regadío X y la productividad (Y) de un cierto cultivo. Se han obtenido los siguientes resultados:

   x_i: 9 10 13 15 18 13
   y_i: 36 44 48 63 70 45

   (a) Ajusta un modelo de regresión lineal simple.

   (b) ¿Hay suficiente evidencia estadística (al nivel de significación 0,05) para afirmar que la productividad tiende a aumentar con la intensidad de regadío?

8. Uno de los índices más útiles para analizar la contaminación en las aguas es la cantidad de oxígeno que contienen en disolución (a mayor contaminación menor cantidad de oxígeno disuelto). Se ha hecho un estudio de la contaminación en un río tomando muestras en cuatro lugares diferentes del río, y midiendo las respectivas cantidades de oxígeno en disolución. Los resultados obtenidos fueron:

Lugar 1:	5,9	6,1	6,3	6,1	6
Lugar 2:	6,3	6,6	6,4	6,4	6,5
Lugar 3:	4,8	5,2	5	4,7	5,1
Lugar 4:	6	6,2	6,1	5,8	

¿Proporcionan estos datos suficiente evidencia (al nivel de significación 0,01)para indicar diferencias significativas en la cantidad media de oxígeno en disolución en los cuatro lugares?

9. Ajusta a una parábola los siguientes datos

x	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y	1.1	1.3	1.6	2.3	2.7	3.4	4.1

Calcula la varianza residual, ¿es adecuado el tipo de ajuste? Compáralo con un ajuste lineal.

10. Dados los datos de la siguiente tabla, se pide obtener un ajuste de estos datos a la expresión $y = \alpha e^{\beta x + \gamma x^2}$

x	1	2	3	4	5	6
y	4.3	8.2	9.5	10.35	12.1	13.1

11. Usando los datos de la tabla siguiente un químico necesita estimar α y β del modelo $\eta = \alpha e^{\beta \cos t}$, donde η es el rendimiento del producto obtenido y t es el tiempo de reacción. Calcula α y β por un ajuste de mínimos cuadrados

Tiempo de reacción (min)	10	20	30	40	50	60
Concentración de producto (gr/l)	0.20	0.52	0.69	0.64	0.57	0.48

12. Se tienen los siguientes datos de temperaturas máximas y mínimas del día 9 de mayo de 2005 en distintos puntos de Castilla-La Mancha de los que consideramos su latitud.

Ciudad	Latitud	Máxima	Mínima
Albacete	38,9	29,0	13,3
Ciudad Real	39,0	28,0	14,3
Tomelloso	39,2	28,7	14,7
Madridejos	39,5	28,1	12,4
Toledo	39,9	30,3	14,0
Talavera	40,0	28,0	13,1
Cuenca	40,1	25,7	11,3
Guadalajara	45,7	27,5	12,5

(a) ¿Existe una dependencia lineal entre la latitud y la temperatura máxima, la latitud y la temperatura mínima, la temperatura máxima y la mínima? ¿Existe algún otro tipo de dependencia?

(b) En caso de existir dependencia lineal estima la correspondiente recta de regresión y calcula los intervalos de confianza de la pendiente y la ordenada en el origen. En caso de no ajustarse bien a una recta busca otro tipo de ajuste que sea mejor.

(c) ¿Qué temperaturas máxima y mínima serían previsibles para Cañizares que se encuentra a 40,5 grados de latitud?

(d) ¿Hay suficiente evidencia estadística (al nivel de significación 0,05) para afirmar que la temperatura máxima disminuye cuando la latitud aumenta?

(e) Estudia la ANOVA para la dependencia entre las temperaturas máxima y mínima.

13. Los siguientes datos corresponden a la densidad de población de un organismo en un recipiente de cultivo a lo largo de 10 días:

Día	1	2	3	4	5	6	7	8	9	10
N/L	1,6	16,7	65,2	23,6	345,3	1341,6	2042,9	7427,0	15571,8	33128,5

Si llamamos X a los días e Y al número de organismos por litro, responde a las mismas preguntas que en el ejercicio 2.