**EJERCICIOS – UD 5: INFERENCIA parte 1**

# DISTRIBUCIONES EN EL MUESTREO

**1)** Una muestra de tamaño N se extrae aleatoriamente de una población Normal con una desviación típica $\sigma$=3. Calcular el mínimo valor de N de modo que la probabilidad de obtener una diferencia mayor que 1 (en valor absoluto) entre la media muestral y la media poblacional sea menor del 1%.

**2)** Un parámetro de calidad sigue una distribución N(5; 2). Si se toma aleatoriamente una muestra de tamaño 4, ¿cuál es la probabilidad de obtener una media muestral mayor a 6? ¿Cuál es la probabilidad si el tamaño muestral es 25?

**3)** Se extraen dos muestras aleatorias de una población Normal con varianza 9, y con tamaños muestrales $n_1$=22 y $n_2$=30. Si se calcula la diferencia entre las dos medias muestrales, ¿cuál es la probabilidad de que ésta sea mayor que 1?

**4)** Un parámetro de calidad medido en un cierto proceso de fabricación tiene una distribución N(m; 3). Determinar el tamaño de muestra que habría que tomar para estimar la media poblacional de este parámetro de modo que la media muestral difiera de la poblacional en menos de 1 cm con una probabilidad del 95%.

**5)** Una empresa produce circuitos electrónicos. El tiempo de operación (horas) sigue una distribución N(2000; 200). Después de ciertos cambios en el proceso de fabricación, el tiempo de operación tiene una distribución N(2200; 200). Se toma una muestra de 10 circuitos antes del cambio, y otros 30 circuitos después del cambio. Si se calcula la diferencia de ambas medias muestrales, ¿cuál es la probabilidad de que esté entre 195 y 205?

**6)** Calcular *a* y *b* de modo que $P(a < s^2 < b) = 0{,}8$ siendo $s^2$ la varianza muestral de una muestra aleatoria de tamaño 16 tomada de una población con distribución N(8; 2).

**7)** Las desviaciones horizontales (X) y verticales (Y) (ambas medidas en cm) de ciertos disparos con respecto al centro de una diana fluctúan independientemente según una distribución N(0; 2). ¿Qué porcentaje de los disparos estarán situados a una distancia mayor de 4 cm con respecto al centro de la diana?

**8)** Un hospital quiere estimar el tiempo de espera medio en los servicios de urgencia mediante entrevistas a una muestra aleatoria de pacientes que son atendidos en estos servicios. En un estudio previo, la desviación típica de este tiempo se estimó en 12 minutos. Asumiendo un error máximo de ± 5 minutos y un nivel de confianza del 95%, ¿cuántos pacientes deberían entrevistarse?

**9)** Si se toman 16 individuos de una población N(0; 5), ¿cuál es la probabilidad de obtener una media muestral mayor a 3, en valor absoluto?

**10)** El tiempo de transferencia (milisegundos) de un cierto tipo de ficheros de datos a través de una red informática sigue una distribución N(m, $\sigma$=4). Si se toma una muestra aleatoria de tamaño 25, ¿cuál es la probabilidad de obtener una diferencia entre la media poblacional y la media muestral mayor a 2, en valor absoluto?

**11)** Si los tres componentes de un vector son valores distribuidos independientemente según una distribución Normal de media cero y varianza 4, ¿cuál es el valor medio del cuadrado de la longitud de este vector?

**12)** En una distribución chi-cuadrado con 20 grados de libertad, calcular $P(\chi^2 > 34{,}17)$ asumiendo que una $\chi^2_{20}$ se puede aproximar por medio de una distribución Normal.

**13)** Si una muestra de tamaño 10 se extrae de una población Normal de varianza 9, ¿cuál es la probabilidad de obtener una varianza muestral mayor de 18?

**14)** Una muestra aleatoria simple de tamaño 15 se extrae de una población $N(m=8, \sigma=4)$. Obtener un intervalo con una probabilidad del 99% de contener la varianza muestral.

**15)** En una distribución t de Student con 12 grados de libertad, ¿cuál es el intervalo que comprende el 80% de los valores?

**16)** El tiempo de respuesta en un sistema informático sigue una distribución Normal de media m=5. Se llevan a cabo 20 ensayos para medir este tiempo. Si la varianza muestral es $s^2=2{,}5$ obtener un intervalo que contenga el valor de la media muestral con una probabilidad del 95%.

**17)** El voltaje de cierto circuito electrónico tiene una distribución Normal de media m=120. Se extrae una muestra de 10 circuitos, resultando una desviación típica muestral s=1,5. Calcular la probabilidad de obtener una media muestral mayor de 121.

**18)** En un grupo de 1500 personas, ¿cuántos individuos tendrán una puntuación entre 0,28 y 8,1 si dicha puntuación es una variable aleatoria que sigue una distribución $F_{8,6}$?

**19)** La temperatura de un circuito electrónico después de una hora de funcionamiento sigue una distribución Normal. Si se extraen dos muestras de la población de circuitos, de tamaños $N_1=10$ y $N_2=18$, ¿cuál es la probabilidad de obtener una varianza muestral de la primera muestra más de 2,5 veces mayor que la segunda varianza muestral?

**20)** El tiempo de la transferencia de datos (ms) a través de una red informática tiene una distribución N(m; 3). Si se toma aleatoriamente una muestra de tamaño 16, ¿cuál s la probabilidad de obtener una diferencia superior a 2, en valor absoluto, entre la media poblacional y la media muestral?

**21)** ¿Cuál es la probabilidad de obtener una varianza muestral superior a 10 si se toma una muestra de tamaño 20 de una población Normal de varianza igual a 5?

**22)** Una muestra de tamaño 9 se extrae aleatoriamente de una población con distribución N(30; 5). ¿Cuál es la probabilidad de obtener una media muestral superior a 25? ¿Cuál es la probabilidad de obtener una desviación típica muestral inferior a 3?

**23)** Para estimar la media de una población Normal de varianza 4, se extrae una muestra de tamaño N de dicha población y se calcula la media muestral. ¿Cuál es el mínimo tamaño de muestra requerido para tener un error en la estimación superior a 0,5 unidades con una probabilidad inferior al 1%?

**24)** Una variable aleatoria X está definida en una población distribuida Normalmente con desviación típica 10. ¿Cuál es el mínimo tamaño muestral requerido para garantizar con una probabilidad del 95% que la media muestral diferirá menos de una unidad de la media poblacional?

**EJERCICIOS – UD 5: INFERENCIA parte 2:**

## INFERENCIA SOBRE UNA POBLACIÓN NORMAL

**25)** En un proceso industrial de pintado de vehículos, el espesor de la capa de pintura sigue una distribución Normal de media 105 μm y desviación típica 5 μm. El control de calidad de este proceso se lleva a cabo tomando la media de 4 medidas procedentes de 4 vehículos seleccionados aleatoriamente. El proceso se considera correcto si la media obtenida es > 100 μm. ¿Cuál es la probabilidad de que el proceso no funcione correctamente?

**26)** El número de defectos de un cierto tipo de piezas sigue una distribución de Poisson, Ps($\lambda$). A partir de una muestra de tamaño 10, se quiere estudiar la hipótesis nula de que el parámetro $\lambda$ de la distribución de Poisson vale 2 frente a la hipótesis alternativa de que es superior a 2. Se aceptará la $H_0$ si la media muestral es < 2,5 y se rechazará $H_0$ en caso contrario.

a) ¿Cuál es el riesgo de tipo I de este test?
b) ¿Cuál es el riesgo de tipo II si $\lambda$ vale realmente 3?
c) ¿Cuál es el riesgo de tipo II si $\lambda$ vale realmente 4?

**27)** A partir de una muestra aleatoria de tamaño 21 tomada de una población Normal, se plantea el siguiente contraste de hipótesis: $H_0 : \sigma^2=4$ frente a $H_1 : \sigma^2=9$. La hipótesis nula se acepta si la varianza muestral $s^2$ es inferior a 6, y se rechaza si $s^2$ resulta ser superior a 7. Si $s^2$ está comprendida entre 6 y 7, la $H_0$ se acepta con una probabilidad de 0,6 y se rechaza con una probabilidad de 0,4. Obtener la probabilidad de cometer un error de tipo I ($\alpha$) y un error de tipo II ($\beta$).

**28)** Indicar cuál de las siguientes afirmaciones es verdadera.

a) Un intervalo de confianza de un cierto parámetro obtenido con un nivel de confianza del 90% contiene con una probabilidad del 90% el verdadero valor estimado del parámetro poblacional.
b) Un intervalo de confianza de un cierto parámetro obtenido con un nivel del confianza del 90% es la región de aceptación de un test de hipótesis considerando $\alpha=0,1$.
c) Ninguna de las dos anteriores.

**29)** Cierta característica X sigue una distribución de Poisson Ps($\lambda$). Se toma una observación (es decir, tamaño muestral=1) para contrastar la hipótesis nula $\lambda=1$ frente a la alternativa $\lambda=2$, y la $H_0$ se rechaza si $x>3$. Obtener la probabilidad del error de tipo I ($\alpha$) y del error de tipo II ($\beta$).

**30)** Una fábrica de circuitos electrónicos compra cada mes una gran cantidad de cierto tipo de pieza. Para contrastar la calidad de dichas piezas antes de la compra, se toma una muestra de 9 piezas y se acepta la partida si todas las piezas en la muestra son correctas. Obtener la probabilidad del error de tipo II en caso de una partida con un 10% de piezas defectuosas.

**31)** Una empresa fabrica bombillas de 40 W. El filamento de esta bombillas tiene una resistencia eléctrica que sigue una distribución N(605; 1,18) ohms. La bombilla se considera correcta cuando la resistencia está comprendida entre 607,86 y 600,96 ohms. Para contrastar si la media poblacional es 605, que es el valor deseado, se toman periódicamente muestras de 7 bombillas. El proceso se considera fuera de control si la media muestral es superior a 606,338 o inferior a 603,662.

a) Obtener la probabilidad de cometer un error de tipo I
b) Obtener la probabilidad de error de tipo II si la media poblacional es 606,5.

**32)** El número de defectos por metro cuadrado en cierto tipo de mármol sigue una distribución de Poisson de parámetro $\lambda$=0,3 defectos/m$^2$. La calidad del mármol se considera adecuada (es decir, $\lambda$=0,3) si se encuentran menos de 3 defectos en una superficie de 4 m$^2$ elegida al azar. De lo contrario, el número de defectos se considera inapropiado.

a) Obtener el valor $\alpha$ de este test.
b) Si $\lambda$=0,4 defectos/m$^2$, ¿cuál será el valor $\beta$?
c) ¿Cuántos metros cuadrados deberán examinarse para obtener $\alpha$=0,05?

**33)** El director del servicio de metro en París ha determinado que el valor medio del retraso con respecto al tiempo esperado de llegada del metro es de 15 segundos, con una desviación típica de 10 segundos. En verano, el director piensa que el tiempo de retraso podría ser diferente. Para estudiar esta cuestión, se mide el tiempo de retraso en 20 casos, obteniéndose los siguientes valores (segundos): { 10, 0, 3, -2, -4, 14, 20, 4, 30, 9, 3, 3, 6, 23, -10, 21, 2, 5, 23, -10 }. Asumiendo que los datos se distribuyen normalmente,

a) Obtener un nivel de confianza considerando $\alpha$=0,1 para el tiempo medio de retraso a nivel poblacional, basándose exclusivamente en la información de la muestra.
b) Considerando un nivel de significación del 1%, ¿hay suficiente evidencia para afirmar que el tiempo medio de retraso en verano no es de 15 segundos?
c) Calcular la probabilidad del error de tipo II de este test si la media poblacional del tiempo de retraso en verano fuese de 10 segundos.

**34)** Para estudiar el espesor de una capa de pintura en una cadena de fabricación de automóviles, se toman 10 muestras de distintos vehículos, resultando los siguientes valores medidos en micras: { 45, 40, 44, 43, 45, 43, 46, 45, 44, 45 }. Asumiendo que los datos se distribuyen Normalmente,

a) Obtener un intervalo de confianza del espesor medio considerando $\alpha$=0,05.
b) Obtener un intervalo de confianza para la varianza del espesor considerando $\alpha$=0,05.

**35)** En la fabricación de ciertos componentes electrónicos, el tiempo de funcionamiento hasta el fallo es el parámetro de calidad más importante. Una muestra de 16 componentes de este tipo ha presentado un tiempo medio de 734 horas. Se asume que los datos se distribuyen Normalmente.

a) ¿Puede aceptarse que la media poblacional es de 740 horas con un nivel de significación de 0,05 considerando $\sigma$ = 12 horas?
b) ¿Puede aceptarse que la media poblacional es de 740 horas con un nivel de significación de 0,05 considerando s = 12 horas?

**36)** Una resistencia eléctrica se ha medido 6 veces, resultando los siguientes valores (en ohmios): { 1,5; 1,6; 1,4; 1,5; 1,3; 1,1 }. Obtener un intervalo de confianza para el verdadero valor de la resistencia considerando $\alpha$=0,05.

**37)** Distintos estudios de investigación sugieren que el número de latidos del corazón en la población masculina entre 20 y 25 años sigue una distribución N(72, 9). Si se toma una muestra de 100 jugadores profesionales de fútbol y la media muestral resulta ser de 64 latidos por minuto, ¿podemos considerar que la diferencia observada respecto a la población es estadísticamente significativa, considerando $\alpha$=0,05?

**38)** Una compañía de vuelos de bajo coste ofrece el vuelo Madrid – Valencia. Se toma una muestra aleatoria de 30 vuelos, resultando un precio medio de 24 € y una varianza de 25 $€^2$. Esta compañía actualmente anuncia que el precio medio de este vuelo es de 20 €. Asumiendo que los precios se distribuyen normalmente y basándose en la información de la muestra, ¿podemos aceptar dicha hipótesis considerando $\alpha$=0,05?

**39)** Una empresa ha desarrollado un nuevo programa de OCR (reconocimiento óptico de caracteres). Para estudiar el funcionamiento del programa se seleccionan aleatoriamente 10 textos con 1.000 caracteres y se procesan con el programa, resultando estos porcentajes de correcto reconocimiento: 99%, 97%, 98%, 98%, 96%, 99%, 97%, 95%, 94% y 92%.

a) ¿Puede la compañía garantizar que el porcentaje medio de correcto reconocimiento es del 95%, considerando un nivel de significación del 5%?
b) Obtener un intervalo de confianza para el porcentaje medio con un nivel de confianza del 98%.

**40)** Un proceso de fabricación produce cables USB de 3 m de longitud. Para estudiar si el proceso funciona correctamente y que la longitud media poblacional es de 3 m, se toma una muestra de 20 cables, resultando una media muestral de 2,96 m y una desviación típica muestral de 0,1 m. En base a esta información, ¿debería reajustarse el proceso considerando $\alpha$=0,05?

**41)** Un programa de ordenador traduce texto del inglés al español. Para estudiar la calidad de este programa, se seleccionan al azar 14 párrafos con 100 palabras en inglés y se traducen con el programa. Después, un experto nativo cuenta el número de errores. El número de errores encontrado en la muestra de 14 párrafos es el siguiente: 10; 12; 7; 8; 6; 10; 9; 9; 10; 8; 9; 11; 10; 9.

a) Obtener un intervalo de confianza para el número medio de errores por cada 100 palabras ($\alpha$=0,05).
b) ¿Podemos afirmar con un 95% de confianza que esta media es de 8?

**42)** Un fabricante de ordenadores portátiles vende un cierto modelo de portátil afirmando que su peso medio es de 1000 gr. Se toma una muestra de 5 portátiles de dicho modelo, resultando los siguientes valores de peso: 995, 992, 1005, 998 y 1000.

a) ¿Los valores de la muestra son coherentes con la hipótesis de que m=1000, con un nivel de significación $\alpha$=10%?
b) Obtener un intervalo de confianza al 95% para la media poblacional del peso.

## SOLUCIONES

1) 60
2) 0.1587 ;  0.0062
3) 0.234
4) 36
5) 0.056
6) $a$ = 2.28, $b$ = 5.95
7) 13.53%
8) 23
9) 0.0164
10) 0.0124
11) 12
12) 0.0125
13) 0.035
14) [4.657,  35.793]
15) [-1.356,  1.356]
16) [4.26, 5.74]
17) 0.032
18) 1410
19) 0.05

25) 0.0228
26) a: 0.11  b: 0.21  c: 0.01
27) $\alpha$ = 0.042  $\beta$ = 0.23
28) a)
29) $\alpha$ = 0.019 $\beta$ = 0.8571
30) 0.3874
31) $\alpha$=0.0027, $\beta$=0.3594
32) a: 0.121  b: 0.783  c: 2.677
33) a: [3.17, 11.83]   b: yes   c: 0.633
34) a: [42.78, 45.22]   b: [1.37, 9.63]
35) a: no  b: sí
36) [1.21, 1.59]
37) sí
38) no
39) a: sí  b: [94.874, 98,126]
40) no,  $m \in$ [2.91, 3.01]
41) a: [8.24, 10.04]   b: no