PROBLEMAS DE MUESTREO

Extracción de muestras

1.- El departamento de control de calidad tiene como empleados a 5 técnicos en el turno matutino. A continuación aparece el número de veces que cada técnico indicó al supervisor de producción que interrumpiera el proceso durante la última semana

Técnico	Interrupciones
A	4
B	3
C	5
D	3
E	2

a) Enumerar todas las muestras de dos observaciones que se pueden tomar sin reemplazamiento

b) Calcular la distribución muestral del número medio de interrupciones en la muestra

c) Calcular la media y la desviación típica de dicha distribución y compararlos con los valores poblacionales

2.- Una empresa cuenta con 6 representantes de ventas en una sucursal. A continuación aparece el número de refrigeradores vendidos por cada representante el último mes

Representante	Refrigeradores vendidos
A	54
B	50
C	52
D	48
E	50
F	52

a) Listar todas las muestras posibles de tamaño 2 tomadas sin reemplazamiento

b) Calcular la distribución muestral del número medio de refrigeradores vendidos

c) Calcular la media y la desviación típica de dicha distribución y compararlos con los valores poblacionales

Método exacto

3.- El puente aéreo Madrid-Barcelona sólo lo realizan dos compañías: Iberia y Aviaco. El director de Iberia afirma que el 80% de los usuarios viajan en su compañía, mientras que el director de Aviaco dice que sólo lo hacen el 75%. Para confirmar su teoría, el C.I.S. realiza una encuesta a 1000 usuarios. Si al menos 775 de los entrevistados vuelan con Iberia le dará la razón a su director y en caso contrario al director de Aviaco.

Calcula la probabilidad de que:

a) Teniendo razón el director de Aviaco, se la dé al de Iberia
b) Teniendo razón el director de Iberia, se la dé al de Aviaco

4.- En un proceso de fabricación se sabe que el número de unidades defectuosas producidas diariamente viene dado por una v.a. cuya distribución es Poisson de media 10.

Determinar la probabilidad de que en 150 días el nº de unidades defectuosas producidas supere las 1480, suponiendo que las producciones diarias son independientes.

5.- Un alumno en la clase de Estadística que no tiene nada mejor que hacer, se dedica a contar el número de palabras en las 100 primeras páginas de su libro de texto. Concluye afirmando que el número de palabras por página es una variable aleatoria con distribución normal de media 236 y desviación típica 55. Supongamos que selecciona 10 páginas al azar del resto del libro. ¿Cuál es la probabilidad de que el número medio de palabras en esas 10 páginas sea al menos 250?

6.- La longitud de unas determinadas piezas, que produce la Morton Metalworks, está distribuida según una normal de media 30 cm. con desviación estándar 0.1. Un inspector acaba de tomar una muestra aleatoria simple de 4 piezas y ha obtenido una media muestral de 29.875 cm. ¿Cuál es la probabilidad de obtener una media muestral igual o menor que dicho valor? ¿Qué conclusiones se pueden sacar?

7.- Una empresa petrolera informa que el precio medio por galón de gasolina normal es 3,26€ por litro con una desviación típica de 0,18€. Seleccionamos al azar una muestra de 40 estaciones de gasolina y calculamos el coste medio de gasolina normal. Si la distribución de dicho coste es normal, calcular

a) Probabilidad de que la media de la muestra oscile entre 3,24€ y 3,28€
b) Probabilidad de que la diferencia entre la media de la muestra y la media poblacional sea inferior a 0,01€
c) Probabilidad de que la media de la muestra sea superior a 3,34€

8.- Una empresa de transportes afirma que el peso medio de sus camiones cuando se encuentran completamente cargados es 6000 libras y la desviación típica de 150 libras. Se seleccionan al azar 40 camiones y se pesan. Suponiendo normalidad en la distribución del peso de un camión ¿entre qué limites se presentará el 95% de las medias de la muestra? (elegir el intervalo más pequeño)

9.- El tiempo que tiene que pasar un cliente en la cola de un supermercado se distribuye exponencialmente con una media de 5 minutos. Calcular la probabilidad de que

a) Se tarde más de 55 minutos en atender a 10 clientes

b) Se tarde entre 20 y 30 minutos en atender a 5 clientes

Muestreo artificial

10.- En una ciudad hay 25 moteles. El número de habitaciones de cada motel es el siguiente:

90	72	75	60	75	72	84	72	88	74	105	115	68	74	80	64	104	82	48	58	60	80	48	58	100

a) Seleccionar una muestra aleatoria de 5 moteles de esta población sin reemplazamiento

b) Obtener una muestra sistemática seleccionando un punto de partida aleatorio entre los primeros 5 moteles

c) Los últimos 10 moteles son de tarifa rebajada. Utilizando el muestreo estratificado elegir una muestra con afijación proporcional de tamaño 5

Método asintótico y TCL

11.- La venta diaria de un grupo de concesionarios de automóviles se encuentra entre 20 y 40 coches, y se puede suponer que la distribución de las ventas es una variable aleatoria uniforme discreta. Después de 282 días de ventas ¿cuál es la probabilidad que tiene dicho grupo de concesionarios de haber vendido más de 8670 automóviles? (Suponer las ventas diarias independientes y que se puede aplicar el TCL)

12.- Se diseña un ascensor cuyo límite de carga es de 3150 kilogramos. Se indica que su capacidad máxima es de 40 personas. Si los pesos de todas las personas que utilizan el ascensor se suponen que están distribuidos alrededor de 75 kilogramos con una desviación típica de 22 kilogramos. ¿Cuál es la probabilidad de que un grupo de 38 personas exceda el límite de carga del ascensor? (Suponer que se puede aplicar el TCL)

13.- La cantidad de dinero que las personas de cierta ciudad llevan en sus bolsillos tiene una distribución con media 9 euros y una desviación típica de 2.5 euros. Suponiendo que el dinero que lleva cada persona es independiente del que lleven los demás, ¿cuál es la probabilidad de que un grupo de 225 individuos lleve una cantidad superior a 2100 euros? (Suponer que se puede aplicar el TCL)

Tamaño muestral

14.- Suponga que se desea precisar el tamaño que ha de tener una muestra tomada sin reemplazamiento para determinar el porcentaje de piezas defectuosas de una población (en total 10.000 unidades), supuesta esta proporción no superior al 5% y operando con una confianza del 95%. Calcular el tamaño muestral necesario si queremos que la proporción no difiera en más del 2% del porcentaje real.

15.- Las compañías de auditoria generalmente seleccionan una muestra aleatoria de los clientes de un banco y verifican los balances contables reportados por el banco. Si una compañía de este tipo se encuentra interesada en estimar la proporción de cuentas para las cuales existe una discrepancia entre el cliente y el banco, ¿cuántas cuentas deberán seleccionarse de manera tal que con una confianza del 99% la proporción muestral se encuentre a no más de un 2% de la proporción real?

16.- Unos grandes almacenes tienen 1000 empleados en uno de sus centros. Calcular el tamaño muestral necesario para estimar su salario anual medio con un error máximo de 100 euros para un nivel de confianza del 95%. (Por estudios anteriores se sabe que el salario anual sigue una distribución normal con desviación típica de 1000 euros).

17.- Una tienda se interesa en estimar su volumen de ventas diarias. Supóngase que el valor de la desviación típica es de 50 euros. Si el volumen de ventas se puede modelizar por una distribución normal, ¿cuál debe ser el tamaño de la muestra para que con una confianza del 95% la media muestral se encuentre a no más de 20 euros del verdadero volumen medio de ventas?

18.- La *International Rectifier Corporation* está estudiando el proceso de producción de microprocesadores. ¿Cuántos microprocesadores tendrán que considerar para asegurar con un 95% de confianza que la proporción muestral de defectuosos no difiere de la proporción poblacional en más del 4%? (Se espera obtener una proporción poblacional del 20%).

19.- Se quiere llevar a cabo un sondeo para determinar el ingreso medio familiar en un área rural. En una muestra piloto de 10 familias la desviación típica de la muestra fue de 500 €.

a) ¿A cuántas familias se debe incluir en la muestra si, con un nivel de confianza del 95% el margen de error máximo permitido es de 100 €? (Suponer que se puede aplicar el TCL)

b) Repetir a) si se sabe, además, que en dicho área viven 1000 familias

c) En las hipótesis de a) y si al final se ha decidido entrevistar a 150 familias ¿cuál sería el margen de error de la estimación con un nivel de confianza del 99%?

20.- Se quiere estimar la proporción de contadores públicos que cambiaron de empresa en los últimos 3 años con un margen de error del 3% y un nivel de confianza del 95%. Un estudio realizado hace varios años reveló que el porcentaje de contadores públicos que cambiaron de compañía fue del 21%

a) Para actualizar el estudio, ¿cuál es el número mínimo de expedientes de contadores públicos que se deben estudiar?

b) ¿Con cuántos contadores públicos es necesario ponerse en contacto si no se considera fiable el estimador de la población poblacional obtenido anteriormente?

c) Resolver b) si, además, se sabe que el número total de contadores públicos es 5000

Problemas varios

21.- Se ha estimado que el número medio de personas que acuden a una determinada sala de cine en cada sesión es de 100, con una desviación típica de 50. El propietario del cine quiere hacer unas reformas cuyos gastos ascienden a 18000 euros y le obligarán a tener el cine cerrado durante un mes. El precio de las entradas es de 4 euros y sólo hay una sesión diaria, pero el propietario quiere que en la reapertura del cine se cubran los gastos ocasionados por la reforma, en un plazo máximo de dos meses (60 días).

a) ¿Puede mantener el mismo precio de las entradas para lograr su propósito con unas garantías del 95%, si los gastos diarios son de 200 euros? Si no es así, ¿a qué precio deberá vender las entradas?

b) Si mantiene el precio de 4 euros ¿cuántos días necesitará para cubrir los gastos con las mismas garantías anteriores?

22.- Una compañía de Alimentación envasa un producto que es inspeccionado periódicamente por el Ministerio de Sanidad, el cual rige que dicho producto no puede contener más de 2 gr. de cierto tipo de conservante.

Según un estudio anterior la compañía afirma que sus envases tienen una cantidad aleatoria de dicho conservante, cuya distribución es normal de media 1.25 y desviación estándar 0.5.

a) ¿Cuál es la probabilidad de que un envase supere el límite impuesto?
b) ¿Y de que un envase contenga entre 1’75 y 2 gr. de dicho conservante?
c) Un equipo de inspectores piensa que la empresa les engaña y deciden tomar una muestra de 25 envases ya comercializados y si la cantidad media de dicho conservante es superior a 1’8 gr., ellos cerrarán la empresa y realizarán una inspección a fondo. ¿Cuál es la probabilidad de que suceda esto? ¿Te parece adecuada la forma de actuar de los inspectores?

23.- Una empresa produce pilas alcalinas cuyo tiempo de vida es aleatorio y se distribuye según una exponencial de media 100 horas. Una linterna funciona con una sola pila y cuando se agota su carga ponemos otra en funcionamiento.

¿Cuántas pilas debemos tener disponibles para asegurar con una confianza del 95% que la linterna funciona más de 10000 horas? (Suponer que se puede aplicar el TCL)

24.- El tiempo de duración de un anuncio se puede suponer una v. a. exponencial de media 1/4 de minuto. Suponiendo independencia en la duración de los anuncios, calcular:

a) La probabilidad de que 50 anuncios duren menos de 12 minutos.
b) El número de anuncios emitidos diariamente por la empresa JAS en la cadena Alfa se distribuye como una normal con desviación típica igual a 8. Si se quiere calcular el número medio de anuncios emitidos en la cadena A con un error máximo de 6 anuncios y un nivel de confianza del 99%, ¿cuántos días se tendrá que muestrear el nº de anuncios emitidos? (Suponer independencia entre los anuncios emitidos cada día).

25.- Una empresa se dedica a la alimentación infantil. Fabrica leche en polvo que envasa en botes. La cantidad envasada en cada bote tiene una media de 900 gramos y una desviación típica de 30 gramos. Por otra parte el bote vacío tiene un peso de 100 gramos y una desviación típica de 10 gramos. Además, ambos pesos se pueden suponer independientes.

a) Si se embalan en bloques de 40 botes, ¿cuál es la probabilidad de que un bloque pese más de 40’5 kilogramos?

b) El precio de venta depende del tipo de cliente, pero se puede suponer que sigue una distribución normal y que su desviación típica no sobrepasa los dos euros. Un cliente quiere conocer el precio medio de venta para saber si a él se lo venden muy caro o no; para ello va preguntando el precio de venta a diferentes clientes que él conoce. ¿A cuántos clientes debe preguntar para estar seguro en un 95 % de que la diferencia entre lo que él calcule y el verdadero valor no supere el euro de diferencia?

FUNCIONES DE EXCEL

- **ALEATORIO():** Devuelve un número aleatorio entre 0 y 1 generado a partir de una Uniforme en (0,1)

- **ALEATORIO.ENTRE(inferior;superior)**: Devuelve un número aleatorio entre el valor inferior y superior generado a partir de una distribución uniforme discreta entre inferior y superior

- **DISTR.EXP.N(x,lambda,acum.).** Calcula la función de densidad (si acumulado = 0) o de distribución (si acumulado = 1) en x para una distribución Exp(λ).

- DISTR.GAMMA(x;p;1/a;acum)) calcula el valor de la función de densidad (si acum = 0) o de distribución (si acum =1) en el punto x para una distribución gamma con parámetros p y a

- DISTR.NORM.N(x;media;dev_estándar;acum) que calcula el valor de la función de cuantía (si acum = 0) o de distribución (si acum =1) en el punto x para una distribución normal con parámetros μ= media y σ=dev_estándar.

- **INV.GAMMA(probabilidad,alfa,beta).** Devuelve el valor F^{-1}(probabilidad) donde F es la función de distribución de una Gamma(alfa,1/beta)

- **INV.NORM(probabilidad,media,desv):** Devuelve el valor F^{-1}(probabilidad) donde F es la función de distribución de una N(media,desv)

- INV.NORM.ESTAND(alfa). Calcula $z_{1\text{-alfa}}$.

SOLUCIONES A LOS PROBLEMAS DE MUESTREO

1.- a)

X_1	X_2	$f(x_1,x_2 \mid \mu)$	$\bar{x}_2$
A	B	0.1	3.5
A	C	0.1	4.5
A	D	0.1	3.5
A	E	0.1	3
B	C	0.1	4
B	D	0.1	3
B	E	0.1	2.5
C	D	0.1	4
C	E	0.1	3.5
D	E	0.1	2.5
	Total	1	

b)

$\bar{x}_2$	$g(\bar{x}_2 \mid \mu)$
2.5	0.2
3	0.2
3.5	0.3
4	0.2
4.5	0.1
Total	1

c)

μ	3.4
$E(\bar{x}_2 \mid \mu)$	3.4
σ	1.02
$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$	0.62
$D(\bar{x}_2 \mid \mu)$	0.62

2.- a)

X_1	X_2	$f(x_1,x_2\|\mu)$	$\bar{x}_2$
A	B	0.066666667	52
A	C	0.066666667	53
A	D	0.066666667	51
A	E	0.066666667	52
A	F	0.066666667	53
B	C	0.066666667	51
B	D	0.066666667	49
B	E	0.066666667	50
B	F	0.066666667	51
C	D	0.066666667	50
C	E	0.066666667	51
C	F	0.066666667	52
D	E	0.066666667	49
D	F	0.066666667	50
E	F	0.066666667	51
	Total	1	

b)

$\bar{x}_2$	$g(\bar{x}_2 \mid \mu)$
49	0.13
50	0.20
51	0.33
52	0.20
53	0.13
Total	1.00

c)

μ	51
σ	1.91
$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$	1.21
$E(\bar{x}_2 \mid \mu)$	51
$D(\bar{x}_2 \mid \mu)$	1.21

3.- X = nº personas que vuelan con Iberia ➔ $X \sim B(1000,p)$

a) p=0,75 ➔ $P(X \geq 775)=0{,}0357$

b) p=0,80 ➔ $P\{X \leq 774\}=0{,}023$

4.- X_i = nº unidades defectuosas en el día i-ésimo $X_i \sim P(10)$; i=1,…, 150

S_{150}=nº defectuosas en 150 días=ΣX_i ➔ $P\{S_{150}>1480\}=0{,}6916$

5.- X = nº palabras por página $\overline{X}_{10}$ = media de 10 páginas ➔ P{$\overline{X}_{10} \geq 250$}=0,2119

6.- X=longitud de una pieza X~N(30; 0,1) $\overline{X}_4$=longitud media de 4 piezas

P($\overline{X}_4 \leq 29{,}875$)=0,00621. Ha disminuido significativamente la longitud media de las piezas fabricadas

7.- X = precio medio de un galón de gasolina normal en una estación (en €) ~N(3,26;0,18)

a) P(3,24≤$\overline{X}_{40}$≤3,28) = 0,5178

b) P(|$\overline{X}_{10}$-3.26|≤0,01) = 0,2747

c) P($\overline{X}_{10}$>3,34) = 0,0025

8.- X = peso de un camión completamente cargado (en libras) ~N(6000,150)

Intervalo: (5953,51;6046,49)

9.- X = tiempo que tiene que pasar en la cola un cliente (en minutos) ~ Exp(0,2)

S_n: tiempo que tarda en atenderse a n clientes ~ Gamma(n,0.2)

a) P(S_{10}>55) =0,3405

b) P(20<S_5<30) = 0,3438

10.-

a) Utilizar el procedimiento descrito en la práctica de informática

b) k = 25/5 = 5. Se coge un número al azar entre 1 y 5, x y luego se cogen los que ocupan los lugares x, x+5, x+10,x+15, x+20

c) n = 5, N_1=15, N_2= 10. Tamaños muestrales en cada estrato n_1 = 3, n_2 = 2

11.- X_i = coches vendidos en el día i-ésimo X_i~U_D(20,40); i=1,…,n

S_{282}=venta después de 282 días ➔ P(S_{282}>8670)=0,01923

12.- X_i = peso de la persona i-ésima; i=1,…, 38

S_{38}=peso de las 38 personas ➔ P(S_{38}>3150)=0,01355

13.- X_i=dinero de la persona i-ésima; i=1,…, 225

S_{225}=dinero de un grupo de 225 personas ➔ P{S_{225}>2100}=0,02275

14.- n = 457. Corrección población finita n = 438

15.- n = 4146

16.- n = 385. Corrección pob. finita n = 278

17.- n=25

18.- n=385

19.- σ = 500

a) n = 97 b) Corrección población finita: n = 89 c) e = 105,33 €

20.- a) $n = 709$ b) $n = 1068$ c) Corrección población finita: $n = 881$

21.- X=nº personas que acuden a una sesión de cine

a) Y_i=beneficio del día i=pX_i-200 ➔ S_{60}=beneficio en 60 días

$P\{S_{60}>18000\}=0{,}95$ ➔ p=5,59 €

b) Y_i=beneficio del día i=$4X_i-200$ ➔ S_n=beneficio en n días

$P\{S_n>18000\}=0{,}95$ ➔ n=108 días

22.- X = cantidad de conservante (en gramos) $X\sim N(1{,}25;\ 0{,}5)$

a) $P(X>2) = 0{,}0668$

b) $P(1{,}75<X<2) = 0{,}0919$

c) $\overline{X}_{25}$ = media de 25 envases ➔ $P(\overline{X}_{25}>1{,}8)=0$. Si dado que, si tiene razón la empresa es imposible que se dé un peso medio superior a 1,8 gramos

23.- S_n = tiempo de funcionamiento con n pilas ≈ Normal

$P(S_n>10000)=0{,}95$ → n = 118 pilas

24.- X = tiempo de un anuncio (en minutos) ~Exp(4)

a) S_{50} = tiempo de 50 anuncios~ Gamma(50,4) $P(S_{50}<12) = 0{,}4054$

b) $n = 12$

25.- X = peso de un bote (en gramos) ~ N(1000;31,62)

a) $P\{S_{40} > 40500\} = 0{,}00621$

b) $n = 16$