ESTADÍSTICA E INTRODUCCIÓN A LA ECONOMETRÍA
Curso 2022/23
Problemas del Tema 1

OBSERVACIÓN: En los problemas 2 a 11, supondremos que la muestra de que se dispone es una muestra aleatoria simple.

1.- Un investigador dispone de un listado con los nombres de 450 directivos de una gran empresa de la Comunidad Valenciana; el listado incluye también la provincia de residencia de cada directivo (200 viven en Valencia, 150 en Alicante y 100 en Castellón), y su sexo (250 son hombres y 200 son mujeres). El investigador quiere analizar el número de hijos que tienen los directivos de la empresa, y para ello ha decidido seleccionar una muestra de 12 directivos, y preguntar a cada uno de ellos cuántos hijos tiene. En cada uno de los casos siguientes, indica si el procedimiento seguido para obtener la muestra proporcionará una muestra aleatoria simple o no (justifica tu respuesta).

a) Selecciona a los 4 primeros del listado con provincia de residencia Valencia, a los 4 primeros del listado con provincia de residencia Alicante, y a los 4 primeros del listado con provincia de residencia Castellón.

b) Selecciona a los 6 primeros del listado que son hombres y a los 6 primeros del listado que son mujeres.

c) Al primer directivo del listado le asigna el número 1, al segundo le asigna el número 2, y así sucesivamente hasta llegar al último, al que asigna el número 450; a continuación, selecciona aleatoriamente un número natural entre 1 y 450 utilizando un programa informático que genera números aleatorios, y escoge como primer directivo de su muestra al que tiene el número seleccionado. Después, repite exactamente este mismo procedimiento 11 veces más.

2.- En una gran población, se selecciona al azar una persona adulta y se considera la variable aleatoria (v.a.) discreta X que indica el número de teléfonos móviles que tiene esta persona. Supondremos que la función de probabilidad de X es:

x	1	2
$f_X(x)$	$\frac{3}{4}$	$\frac{1}{4}$

a) Calcula $E(X)$ y $\text{Var}(X)$.

b) Una investigadora desconoce la función de probabilidad de X , pero dispone de una muestra de tamaño 3 de X para estimar características de esta variable.

b1) Determina todos los posibles valores de la muestra, y calcula la probabilidad de cada uno de ellos.

b2) Determina la función de probabilidad de $\bar{X}$.

b3) A partir de la función de probabilidad obtenida en el subapartado anterior, calcula $E(\bar{X})$, e indica si este valor coincide con $E(X)$.

b4) A partir de la función de probabilidad obtenida en el subapartado **b2**, calcula $\text{Var}(\bar{X})$, e indica si este valor coincide con $\text{Var}(X)$.

b5) Calcula la probabilidad de que la media muestral sea inferior a la media poblacional.

3.- La duración (en años) de un componente electrónico fabricado por una empresa es una v.a. continua X con función de densidad:

$$f(x) = \begin{cases} \frac{3}{8}x^2 & \text{si } 0 \leq x \leq 2 \\ 0 & \text{en caso contrario} \end{cases}$$

a) Calcula de la media de X .

b) Una investigadora no conoce la función de densidad de X , y desea estimar su media; para ello, dispone de una muestra con la duración de 5 componentes fabricados por la empresa, y calcula la media muestral $\bar{X}$.

b1) ¿Cuál es la esperanza de $\bar{X}$?; ¿cuál es su varianza?

b2) Explica cómo cambiarían los resultados del subapartado anterior si la investigadora dispusiera de una muestra con la duración de 100 componentes.

b3) Las duraciones de la muestra de la investigadora son: 1.55, 1.30, 1.80, 0.95, 1.70. Calcula la media muestral de esta muestra, y explica qué relación hay entre este valor y el calculado en el apartado **a**.

4.- El peso (en gramos) de una galleta producida por nuestra empresa es una v.a. con distribución normal de media 60 y desviación típica 5.

a) Seleccionamos una muestra de 16 galletas producidas por nuestra empresa. Calcula:

a1) La probabilidad de que el peso medio de la muestra sea inferior a 58.5 gramos.

a2) La probabilidad de que el peso medio de la muestra esté entre 59 y 61 gramos.

b) ¿Qué probabilidades se habrían obtenido en el apartado **a** si la muestra hubiera estado formada por 100 galletas? ¿Qué ocurre con estas dos probabilidades conforme aumenta el tamaño de la muestra? Explica por qué.

c) Considerando ahora una muestra formada por 64 galletas, determina el valor c para que la siguiente condición sea cierta: la probabilidad de que el peso medio de la muestra sea inferior a c es 0.02.

5.- Utilizando el Teorema Central del Límite, calcula aproximadamente las probabilidades que se indican a continuación:

a) La probabilidad de que la media muestral sea menor que 0.1, si se dispone de una muestra de tamaño 350 de una v.a. discreta X con la siguiente función de probabilidad:

x	0	1	2
$f_X(x)$	0.90	0.08	0.02

b) La probabilidad de que la media muestral esté entre 0.49 y 0.51, si se dispone de una muestra de tamaño 525 de una v.a. continua X , y la distribución de X es uniforme entre 0 y 1.

6.- En un país, se sabe que el 70% de todos los votantes apoya la reforma constitucional que se ha propuesto. Se ha extraído una muestra de esta población de 150 votantes y se les ha preguntado su opinión sobre esta reforma.

a) Calcula de modo exacto la probabilidad de que el porcentaje de votantes de la muestra que apoyan la reforma sea inferior al 67%.

b) Utilizando el Teorema Central del Límite, calcula de modo aproximado la probabilidad obtenida en el apartado anterior, y explica por qué es razonable utilizar el Teorema Central del Límite para este cálculo aproximado.

c) Supongamos que en la muestra extraída 102 votantes apoyan la reforma, y los demás no la apoyan. En este caso, la proporción poblacional de votantes que apoyan la reforma ¿es menor, es mayor o es igual a la proporción muestral de votantes que apoyan la reforma?

7.- Sea p la proporción de artículos producidos por una determinada máquina que resultan defectuosos. El fabricante de la máquina afirma que $p = 0.02$. Para estudiar la veracidad de esta afirmación, hemos revisado 20 artículos producidos por esta máquina, de modo independiente, y hemos comprobado si son defectuosos o no.

a) Si la afirmación del fabricante de la máquina es cierta ¿cuál es la probabilidad de que el porcentaje de artículos defectuosos entre los 20 que hemos revisado sea igual o superior al 10%?

b) ¿Te parece adecuado utilizar el Teorema Central del Límite para obtener de modo aproximado la probabilidad calculada en el apartado anterior?

c) Supongamos que entre los artículos revisados hay un porcentaje igual o superior al 10% de defectuosos. En este caso, ¿te parece razonable decir que la afirmación del fabricante es falsa?; ¿te parece razonable decir que la afirmación del fabricante es poco verosímil?

8.- En una población, el 70% de hogares no tiene ningún ordenador fijo, el 20% tiene un ordenador fijo y el 10% tiene dos ordenadores fijos. Estamos interesados en la v.a. discreta X que indica el número de ordenadores fijos que hay en un hogar seleccionado aleatoriamente en esta población. Una investigadora no conoce los porcentajes anteriores, y quiere estimar la media de la v.a. X , que llamaremos μ , y su desviación típica, que llamaremos σ . Para su estudio la investigadora dispone de una muestra de tamaño 2, con la que calcula la media muestral $\bar{X}$ y la desviación típica muestral S .

a) Calcula la esperanza y la varianza de $\bar{X}$.

b) Calcula la función de probabilidad de S .

c) Calcula la esperanza de S , y analiza si esta esperanza es igual a σ .

d) Calcula la probabilidad de que la desviación típica muestral sea superior a la desviación típica poblacional.

9.- Sean X_1 , X_2 y X_3 variables aleatorias independientes, todas ellas con distribución normal estándar. En cada uno de los casos siguientes, identifica cuál es la distribución de la variable aleatoria W y, después, utilizando Excel o algunos de los resultados obtenidos con Excel que aparecen en la tabla que se da al final de esta hoja de problemas, calcula la probabilidad de que W sea menor que 1 y determina cuál es el número real c que hace que la igualdad $P(W > c) = 0.10$ sea cierta.

a) $W = X_1 - 2X_2 + 2X_3$

b) $W = X_1^2 + X_2^2 + X_3^2$

c) $W = (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + (X_3 - \bar{X})^2$, siendo $\bar{X} = (X_1 + X_2 + X_3)/3$

10.- Se sabe que el tiempo (en horas) que dedica semanalmente al estudio un alumno seleccionado al azar entre los alumnos de una Facultad de Letras es una v.a. con distribución normal con varianza 2 horas². Disponemos de una muestra de tamaño 20 de esta v.a., y con ella calculamos la media muestral, la varianza muestral y la desviación típica muestral. Calcula:

- a) La probabilidad de que la varianza muestral esté entre 1.9 y 2.1 horas².
- b) La probabilidad de que la desviación típica muestral sea inferior a 1.3 horas.
- c) El número real c que cumple la condición de que la probabilidad de que la varianza muestral sea superior a c horas² es 0.90.
- d) Si sabemos que la varianza muestral ha sido superior a 2.2 horas², ¿cuál es la probabilidad de que el valor absoluto de la diferencia entre la media muestral y la media poblacional sea inferior a 0.5 horas?

11.- Llamemos $\bar{X}$ a la media muestral y S^2 a la varianza muestral obtenidas con una muestra de tamaño n de una v.a. X con media μ y varianza σ^2 . Responde **razonadamente** a las siguientes preguntas:

- a) ¿Cuál es el valor de $E(\bar{X}^2)$? (expresa el resultado en función de μ , σ^2 y n).
- b) Si la distribución de X es binomial $\text{Bi}(1; p)$, ¿cuál es el valor de $\text{Var}(\bar{X})$? (expresa el resultado en función de p y n).
- c) Si la distribución de X es uniforme entre $-c$ y c , siendo c un número real positivo, ¿cuál es el valor de $E(S^2)$? (expresa el resultado en función de c).
- d) Si la distribución de X es chi-cuadrado con k grados de libertad, ¿cuál es el valor de $E(\frac{\bar{X}-1}{2})$? (expresa el resultado en función de k).
- e) Si la distribución de X es normal, y definimos W como

$$W = \left(\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \right)^2$$

¿cuál es la distribución de la v.a. W ?