

INTRODUCCIÓN A LA ESTADÍSTICA. (ADE). SEPTIEMBRE 2008. Examen reserva
(Código de la asignatura 202. Código de la carrera 42)

1.- Dada la distribución de frecuencias, comente la utilización y representatividad de las siguientes medidas de posición: media aritmética, mediana y moda

X	Y
-1	5
1	5

Respuesta.-

Nota previa: supondremos que la Y se refiere a la frecuencia de la variable X, ya que si consideramos que la Y es una variable, la distribución sería bidimensional y el ejercicio carece de sentido.

Con dicha hipótesis, se tiene:

$$\text{media} = \frac{(-1)5 + 1 \cdot 5}{10} = 0; \quad \text{mediana} = 0; \quad \text{moda} = \{-1, 1\} \text{ (es bimodal).}$$

Así pues, dado que la dispersión de la variable es considerable (los valores ocupan exclusivamente los extremos del recorrido), la media y la mediana no son representativas y, en este caso, la medida de posición central más adecuada sería la moda.

2.-Frecuencias marginales en una tabla de correlación.

Respuesta.-

Consideramos la tabla de correlación de una distribución bidimensional de las variables $(X, Y) = \{(x_i, y_j), i=1 \dots r, j=1 \dots s\}$, donde n_{ij} es la frecuencia del par (x_i, y_j) .

Si, para cada x_i , hallamos la suma $n_{i\bullet} = \sum_{j=1}^s n_{ij}$ obtenemos la distribución (unidimensional) de la variable X. Las frecuencias así obtenidas se denominan frecuencias marginales de la variable X.

Análogamente, si para cada y_j hallamos la suma $n_{\bullet j} = \sum_{i=1}^r n_{ij}$ obtenemos la distribución (unidimensional) de la variable Y. Las frecuencias así obtenidas se denominan frecuencias marginales de la variable Y

Y \ X	y_1	y_j	y_s	
x_1	n_{11}	n_{1j}	n_{1s}	$n_{1\bullet}$
•	•		•		•	•
•	•		•		•	•
•	•		•		•	•
•	•		•		•	•
x_i	n_{i1}	n_{ij}	n_{is}	$n_{i\bullet}$
•	•		•		•	•
•	•		•		•	•
•	•		•		•	•
•	•		•		•	•
x_r	n_{r1}	n_{rj}	n_{rs}	$n_{r\bullet}$
	$n_{\bullet 1}$	$n_{\bullet j}$	$n_{\bullet s}$	

3.-Índice de precios de Laspeyres. Ventajas e inconvenientes.

Respuesta.-

Sean p_{i0} y p_{it} , $i = 1, 2, \dots, N$, los precios de N bienes en un periodo base (periodo cero) y en otro periodo cualquiera (t). El índice de precios de Laspeyres es la media aritmética ponderada (multiplicada por 100) de los índices simples $\frac{p_{it}}{p_{i0}}$, usando como ponderadores los valores $w_i =$

$$p_{i0}q_{i0}, \text{ donde } q_{i0} \text{ es la cantidad del bien } i \text{ en el periodo } 0. \text{ Es decir: } P_L = \frac{\sum_{i=1}^N \frac{p_{it}}{p_{i0}} \cdot p_{i0}q_{i0}}{\sum_{i=1}^N p_{i0}q_{i0}} \cdot 100 = \frac{\sum_{i=1}^N p_{it}q_{i0}}{\sum_{i=1}^N p_{i0}q_{i0}} \cdot 100$$

Entre las ventajas está su cálculo relativamente fácil simple pues los ponderadores no cambian en cada periodo. Entre los inconvenientes puede citarse que su representatividad disminuye al alejarse del periodo base y también que no es un verdadero deflactor (aunque se use para ello)

4.- Axiomas de Kolmogorov para definir la probabilidad.

Respuesta.-

Consideremos un espacio muestral E y Ω una familia de sucesos de E que cumpla las propiedades:

1) si $A \in \Omega$ entonces también el suceso contrario $A' \in \Omega$; 2) dada una colección de sucesos $A_i \in \Omega$, $i = 1, 2, 3, \dots, \infty$, entonces $\bigcup_{i=1}^{\infty} A_i \in \Omega$. Decimos que Ω es una σ -álgebra.

Sea pues la σ -álgebra Ω . Una aplicación $P: \Omega \rightarrow \mathbb{R}$ es una probabilidad si cumple los siguientes axiomas (de Kolmogorov):

I) $P(A) \geq 0, \forall A \in \Omega$

II) $P(E) = 1$

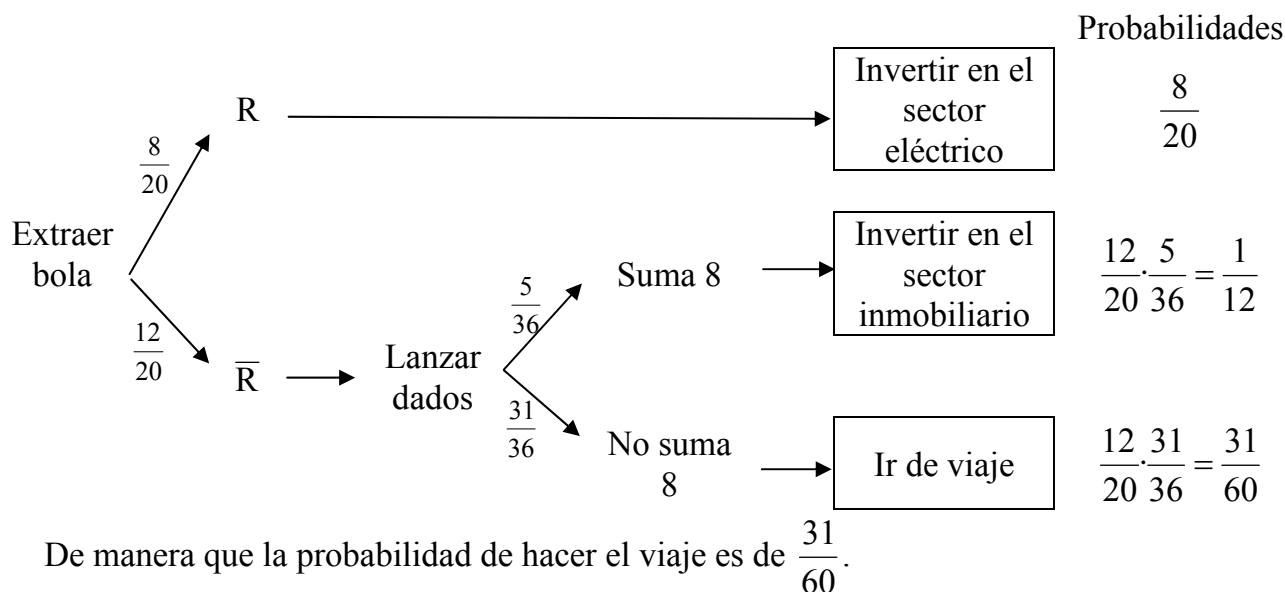
III) Si $A_i \in \Omega$, $i = 1, 2, 3, \dots, \infty$, es una sucesión de sucesos mutuamente incompatibles, entonces $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Ejercicios prácticos

1.- El Sr. Fernández duda entre dedicar sus ahorros a un viaje o invertir en renta variable. Le ofrecen dos alternativas atractivas para invertir pero el Sr Fernández confía al azar su decisión. Invertirá en el sector eléctrico si saca una bola roja de una urna que contiene 20 bolas, de las cuales 8 son rojas, 3 verdes y 9 negras. Si la bola no es roja lanzará dos dados y si obtiene una suma de 8 entre ambos invertirá en el sector inmobiliario; en caso contrario se decidirá por el viaje. ¿Cuál es la probabilidad de que finalmente disfrute del viaje?

Solución.-

La probabilidad de obtener bola roja es $\frac{8}{20}$. La probabilidad de obtener 8 al lanzar dos dados es $\frac{\text{casos favorables}}{\text{casos posibles}} = \frac{5}{36}$. Podemos resumir el problema mediante el siguiente diagrama:



2.-A partir de una muestra de 100 familias se hace un estudio de la relación entre los ingresos (X) y el ahorro (Y). Los datos obtenidos expresados en miles de Unidades monetarias por año son:

X \ Y	40 – 80	80 – 150	160 - 250
0 - 10	30	20	0
0 - 50	3	14	18
50 - 80	0	0	15

a) compruebe si el ahorro puede relacionarse con los ingresos según un modelo lineal. b) Ajuste dicha función y comente los resultados. c) ¿qué ocurriría con el ahorro si todas las familias aumentaran sus ingresos en 12 U.M.

Solución.-

Observaciones previas: 1) el segundo intervalo de la variable Y contiene al primero. Seguramente se tratará de una errata y deba ser 10 – 50. 2) Cuando se agrupan datos en intervalos, estos debe cubrir todo el recorrido de la variable de forma que el extremo superior de un intervalo coincida con el extremo inferior del siguiente (no pueden quedar espacios vacíos). Corregiremos el último intervalo de la X, que pasará a ser 150 – 250.

Efectuando estas correcciones, sustituyamos cada intervalo por su marca de clase:

X \ Y	60	115	200
5	30	20	0
30	3	14	18
65	0	0	15

a) Calcularemos el coeficiente de determinación para establecer la bondad de un ajuste lineal, efectuando los cálculos marginales adecuados:

$\begin{matrix} X \\ Y \end{matrix}$	60	115	200	n_i	$y_i \cdot n_i$	$y_i^2 \cdot n_i$
5	30	20	0	50	250	1250
30	3	14	18	35	1050	31500
65	0	0	15	15	975	63375
n_j	33	34	33	100	2275	96125
$x_j n_j$	1980	3910	6600	12490		
$x_j^2 n_j$	118800	449650	1320000	1888450		

Construimos también la tabla de los productos $x_j y_i n_{ij}$ para poder hallar a_{11} :

$y_i x_j n_{ij}$	60	115	200
5	9000	11500	0
30	5400	48300	108000
65	0	0	195000
	14400	59800	303000
			377200

Así pues tenemos:

$$a_{10} = 124,9$$

$$a_{01} = 22,75$$

$$a_{11} = 3772$$

$$a_{20} = 18884,5$$

$$a_{02} = 961,25$$

de donde se
obtienen:

$$m_{20} = 3284,49$$

$$m_{02} = 443,6875$$

$$m_{11} = 930,525$$

Luego el coeficiente de determinación $R^2 = \frac{m_{11}^2}{m_{20} m_{02}} = 0,5942$, que resulta ser demasiado

pequeño para que la distribución dada se ajuste adecuadamente a un modelo lineal.

b) La recta de regresión de Y/X: $y - 22,75 = \frac{930,525}{3284,49}(x - 124,9) \Leftrightarrow y = 0,2833x - 12,635$

c) El coeficiente de regresión (0,2833) representa el aumento del ahorro por unidad de aumento de los ingresos. Luego si los ingresos aumentan 12 UM, el ahorro aumentaría $12 \cdot 0,2833 \approx 3,4$ UM