

JUNIO 2008. PRIMERA SEMANA

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. Explique las principales ventajas e inconvenientes de la mediana.

Respuesta.-

Ventajas:

Es la medida más representativa en el caso de variables que sólo admiten la escala ordinal.

En la mediana sólo influyen los valores centrales de la distribución y es insensible a los valores extremos. Puede calcularse en distribuciones en las que los valores extremos son desconocidos siempre y cuando tengamos información sobre sus frecuencias (casos de intervalos iniciales y finales de naturaleza abierta).

Inconvenientes:

En su determinación no intervienen todos los valores de la variable (Aunque puede ser una ventaja cuando son desconocidos los valores extremos o existe una enorme dispersión entre los mismos que invalidan las medias como medidas de posición central al no ser representativas del conjunto de la distribución por la enorme influencia que ejercen los mencionados valores extremos en su cálculo)

2. Concepto y componentes de una serie temporal.

Respuesta.

Una serie temporal es un conjunto de datos de una determinada magnitud económica, ordenados a lo largo del tiempo.

Cada uno de los valores de una serie temporal puede considerarse el resultado de componer (mediante suma o multiplicación) una serie de valores entre los cuales los más importantes son:

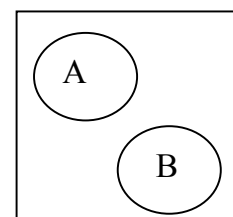
- componente secular o tendencia que refleja la evolución de la serie a largo plazo;
- componente estacional, recoge las oscilaciones periódicas de periodo igual o inferior a un año;
- componente cíclica, recoge las oscilaciones periódicas no regulares de la serie, de periodo superior al año;

-componente accidental, recoge las oscilaciones ocasionales que se producen por causa de fenómenos imprevisibles.

3. ¿Qué significa que 2 sucesos son disjuntos? Ponga un ejemplo.

Respuesta.-

Aquellos que no se pueden verificar simultáneamente. Por ejemplo si lanzamos tres dados entonces los sucesos $A = \text{''obtener una puntuación menor que 7''}$ y $B = \text{''obtener una puntuación mayor que 10''}$ son disjuntos (o incompatibles).



4. Si la Covarianza entre dos variables x e y es 0, ¿Podemos afirmar que son independientes? Razone su respuesta.

No. Por ejemplo, dada la distribución:

$x_i \backslash y_j$	1	2	3	n_i	$x_i \cdot n_i$
1	0	1	0	1	1
2	1	1	1	3	6
3	0	1	0	1	3
n_j	1	3	1	5	10
$y_j \cdot n_j$	1	6	3	10	

podemos comprobar que las variables x e y no son independientes pues, por ejemplo,

$$\frac{n_{.1} \cdot n_{1.}}{N \cdot N} = \frac{1 \cdot 1}{5 \cdot 5} = \frac{1}{25} \neq \frac{n_{11}}{5} = 0. \text{ Sin embargo la covarianza es cero:}$$

$$\left. \begin{aligned} a_{10} &= \frac{10}{5} = 2 \\ a_{01} &= 2 \\ a_{11} &= \frac{2+2+4+6+6}{5} = \frac{20}{5} = 4 \end{aligned} \right\} \rightarrow m_{11} = 4 - 4 = 0$$

PROBLEMAS

1.- Dada la siguiente distribución de sueldos de una empresa:

Sueldos	[500-1500)	[1500-2000)	[2000-3000)	[3000-5000]
Nº Empleados	35	50	22	3

- Calcular el Índice de Gini.
- Dibujar la curva de Lorenz.

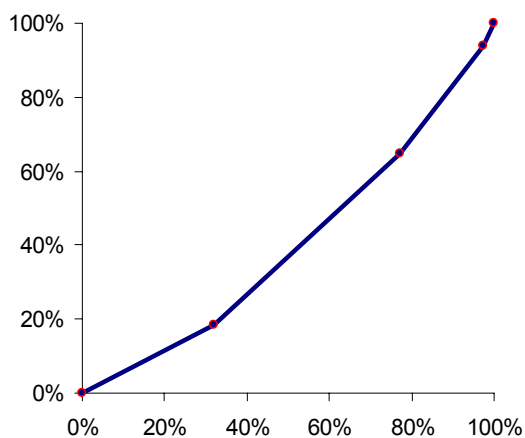
Solución.-

Representando cada clase de sueldos por su marca de clase, construimos la tabla:

x_i (marca de clase)	n_i	$x_i n_i$	u_i	q_i	N_i	p_i	$p_i - q_i$
1000	35	35000	35000	18,47%	35	31,82%	13,35%
1750	50	87500	122500	64,64%	85	77,27%	12,63%
2500	22	55000	177500	93,67%	107	97,27%	3,61%
4000	3	12000	189500		110		
	110	189500		18,47%	35	206,36%	29,58%

De donde el índice de Gini:
$$I_G = \frac{\sum_{i=1}^3 p_i - q_i}{\sum_{i=1}^3 p_i} = \frac{29,58}{206,36} = 0,14$$

La curva de Lorenz:



2.- Dada la siguiente distribución bidimensional:

	2	3	5
1	2	3	1
4	3	5	5
8	0	2	1
12	1	0	3

Calcular:

- Las distribuciones marginales de frecuencias
- La covarianza entre x e y
- El coeficiente de correlación lineal simple

Solución.-

a) Las distribuciones marginales de frecuencias, obtenidas sumando las frecuencias, por filas o por columnas, respectivamente, serían:

<u>x_i</u>	<u>n_i</u>		<u>y_j</u>	<u>n_j</u>
1	6		2	6
4	13	y	3	10
8	3		5	10
12	4			

b) En la tabla de correlación efectuamos los cálculos correspondientes:

<u>x_i</u> \ <u>y_j</u>	2	3	5	<u>n_i</u>	<u>x_i·n_i</u>	<u>x_i²·n_i</u>	Calculamos también los productos x _i y _j n _{ij} :	<u>x_i</u> \ <u>y_j</u>	2	3	5	
1	2	3	1	6	6	6		1	4	9	5	
4	3	5	5	13	52	208		4	24	60	100	
8	0	2	1	3	24	192		8	0	48	40	
12	1	0	3	4	48	576		12	24	0	180	
n_j	6	10	10	26	130	982						494
y_j·n_j	12	30	50	92								
y_j²·n_j	24	90	250	364								

De donde obtenemos los momentos:

$$a_{11} = 19$$

$$a_{10} = 5$$

$$a_{01} = 3,54$$

Así pues, la covarianza: $m_{11} = 19 - 5 \cdot 3,54 = 1,31$.

c) Calculamos las varianzas:

$$m_{20} = 12,77$$

$$m_{02} = 1,48$$

Luego el coeficiente de correlación lineal simple $r = \frac{1,31}{\sqrt{12,77 \cdot 1,48}} \cong 0,301$