

JUNIO 2008. SEGUNDA SEMANA

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. Defina el concepto de número índice y cite sus principales propiedades.

Respuesta.-

El cociente entre el valor x_{t_1} que toma una determinada variable económica (precios, salarios, cantidades producidas, etc) en un periodo t_1 , y el valor x_{t_0} que toma en otro periodo t_0 , se denomina índice simple de esa variable en el periodo t_1 , con base en el periodo t_0 , pudiendo expresarse, o no, como porcentaje.

Si se dispone de un conjunto de variables, entonces el índice complejo se obtiene dividiendo entre sí las medias, u otra agregación, de los valores de las variables en los distintos periodos. Según la agregación que se utilice se obtiene un modelo u otro de índice.

Las propiedades de los índices son:

- a) Existencia: Todo número índice ha de tener un valor finito distinto de cero.
- b) Identidad: Si se hacen coincidir el período base y el período actual el valor del índice tiene que ser igual a la unidad (o a 100 si se elabora en porcentajes)
- c) Inversión: el índice del año 0 calculado con la base del año t, ha de ser igual al inverso del índice del año t calculado en base del año 0.
- d) Proporcionalidad: Si en el período actual todas las magnitudes experimentan una variación proporcional, el número índice tiene que experimentar también dicha variación.
- e) Homogeneidad: Un número índice no puede estar afectado por los cambios que se realicen en las unidades de medida.

2. Medidas de Asimetría de una distribución de frecuencias unidimensional. Tipos de Asimetría. ¿Qué afirmación podemos hacer si $m_3 = 0$?

Respuesta.-

Entre otras medidas de asimetría es destacable el coeficiente de asimetría de Fisher:

$$g_1 = \frac{m_3}{\sigma^3} = \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^3 n_i}{\left(\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 n_i \right)^{\frac{3}{2}}}$$

Si el coeficiente es positivo decimos que hay asimetría a la derecha, si es negativo, hay asimetría a la izquierda.

Si la distribución es simétrica, entonces $g_1 = 0$ aunque el recíproco no es cierto.

Por tanto, del hecho que $m_3 = 0$ se deduce que $g_1 = 0$, pero no significa necesariamente que la distribución sea simétrica.

3. Defina el concepto de probabilidad condicionada. Ponga un ejemplo.

Respuesta.-

El espacio muestral E de un experimento aleatorio está formado por todos los sucesos elementales asociados a dicho experimento. Un suceso cualquiera $A \subset E$ puede también considerarse como espacio muestral, formado por todos los sucesos elementales de E que pertenezcan a A. Si $B \subset E$ es otro suceso cualquiera, entonces el suceso $B \cap A$ es tanto un suceso del espacio muestral E como del espacio muestral

A. En el primer caso escribimos simplemente $B \cap A$ y en el segundo caso escribimos B/A y le llamamos suceso B condicionado por A.

Sea P la función de probabilidad definida en E y supongamos que $p(A) > 0$. Se define la probabilidad condicionada:

$$P(B/A) = \frac{P(B \cap A)}{P(A)}$$

Un ejemplo: en una bolsa hay 15 bolas del mismo tamaño numeradas del 1 al 15. Extraemos una bola al azar y consideramos los sucesos $A = \text{"extraer bola par"}$ y $B = \text{"extraer bola múltiplo de 3"}$.

Entonces el suceso $B \cap A = \{6, 12\}$, luego $P(B \cap A) = \frac{2}{15}$, mientras que $A = \{2, 4, 6, 8, 10, 12, 14\}$,

luego $P(A) = \frac{7}{15}$. Por tanto $P(B/A) = \frac{2/15}{7/15} = \frac{2}{7}$.

Puesto que $B = \{3, 6, 9, 12, 15\}$ se tiene que $P(B) = \frac{5}{15} = \frac{1}{3} \neq P(B/A)$. Decimos entonces que A y B son dependientes.

4. Defina los conceptos de Índice de Gini y curva de Lorenz. Explique el significado de los distintos valores que puede tomar el Índice de Gini.

Respuesta.-

Consideramos la variable estadística $\{(x_i, n_i), i=1, 2, \dots, r\}$ donde cada valor x_i es la renta de los n_i individuos, siendo $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_r$. Sea $u_j = \sum_{i=1}^j x_i n_i$ es decir la renta total de los individuos con renta $\leq x_j$ y sea q_j el porcentaje que dicho total representa respecto de la renta total, a saber $q_j = \frac{u_j}{u_r} \cdot 100$.

Por otra parte, sea p_j el porcentaje de individuos con renta $\leq x_j$, es decir $p_j = \frac{\sum_{i=1}^j n_i}{\sum_{i=1}^r n_i} \cdot 100$. Se define el

índice de concentración de Gini:

$$I_G = \frac{\sum_{i=1}^{r-1} (p_i - q_i)}{\sum_{i=1}^{r-1} p_i}$$

Puesto que $\frac{u_j}{\sum_{i=1}^j n_i} \leq \frac{u_r}{\sum_{i=1}^r n_i} \Leftrightarrow \frac{u_j}{u_r} \leq \frac{\sum_{i=1}^j n_i}{\sum_{i=1}^r n_i} \Leftrightarrow q_j \leq p_j$, luego $\sum_{i=1}^{r-1} (p_i - q_i) \leq \sum_{i=1}^{r-1} p_i$ de donde se

deduce que $0 \leq I_G \leq 1$.

Si $I_G = 0$, significa que $p_i = q_i$, $i=1, 2, \dots, r-1$, es decir, los valores de q_i son los máximos posibles, o lo que es equivalente, el porcentaje p_i de individuos dispone de la máxima renta posible. La concentración de la renta es la mínima.

Si $I_G = 1$, significa que $q_i = 0$, $i = 1, 2, \dots, r-1$, es decir, el total de la renta se acumula exclusivamente en los n_r individuos de mayor renta. La concentración es máxima.

La curva de Lorenz es la poligonal que une los puntos:

$$(0, 0), (p_1, q_1), (p_2, q_2), \dots, (p_{r-1}, q_{r-1}), (100, 100)$$

PROBLEMAS

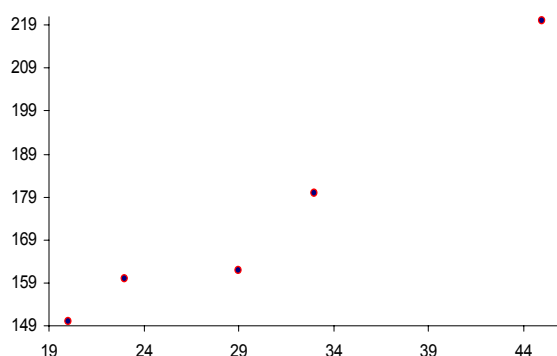
1.- Una empresa tiene, para los últimos cinco años, la siguiente distribución de empleados y volumen de ventas:

Nº Empleados	20	23	29	33	45
Ventas	150	160	162	180	220

- Establecer el modelo lineal que relacione ambas variables.
- Juzgue la bondad de dicho modelo.
- Si la empresa aumentara a 50 empleados su plantilla, ¿Qué volumen de ventas obtendría según el modelo estimado?

Solución.-

Consideramos las variables $x = \text{"nº de empleados"}$ e $y = \text{"ventas"}$. Representando la nube de puntos podemos comprobar que una recta se ajusta bien. Ajustaremos la recta de regresión de Y/X . De la tabla:



x_i	y_i	x_i^2	y_i^2	$x_i y_i$
20	150	400	22500	3000
23	160	529	25600	3680
29	162	841	26244	4698
33	180	1089	32400	5940
45	220	2025	48400	9900
150	872	4884	155144	27218

obtenemos los momentos:

$$\begin{aligned} a_{10} &= 30,00 & a_{02} &= 31028,80 & m_{11} &= 211,60 \\ a_{01} &= 174,40 & a_{20} &= 976,80 & m_{20} &= 76,80 \\ a_{11} &= 5443,60 & & & m_{02} &= 613,44 \end{aligned}$$

de donde:

a) la recta de regresión de Y/X :

$$y - 174,4 = \frac{211,6}{76,8}(x - 30) \quad \Leftrightarrow \quad y = 2,7552x + 91,744$$

b) El coeficiente de determinación: $R^2 = \frac{m_{11}^2}{m_{20} m_{02}} = \frac{211,6^2}{76,8 \cdot 613,44} \cong 0,95$. Por tanto la recta de

regresión está bastante bien ajustada.

c) Haciendo $x = 50$ en la recta de regresión, se obtiene una estimación de las ventas $y = 229,504$.

2.- Dada la siguiente distribución de frecuencias:

X	(0-4)	(5-9)	(10-15)	(15-25)
N	3	4	12	9

Calcular:

- La media aritmética.
- La mediana.
- La moda.

Solución.-

En primer lugar hay que precisar los extremos exactos de los intervalos porque no pueden quedar huecos entre dos intervalos consecutivos¹. Entre otras opciones, podemos considerar los intervalos [0, 5[, [5, 10[, [10, 15[y [15, 25]. Así pues, de la tabla:

Clases	x_i	n_i	$x_i n_i$	N_i	h_i
[0, 5[2,5	3	7,5	3	0,6
[5, 10[7,5	4	30	7	0,8
[10, 15[12,5	12	150	19	2,4
[15, 25]	20	9	180	28	0,9
		28	367,5		

- La media aritmética: $\bar{x} = \frac{367,5}{28} = 13,125$
- La mediana: $Me = 10 + \frac{14-7}{12} \cdot 5 \cong 12,917$
- La clase modal (la de mayor densidad de frecuencia) es [10, 15[, luego la moda:

$$Mo = 10 + \frac{0,9}{0,9+0,8} \cdot 5 \cong 12,647$$

¹ Consúltase el texto base, página 44, donde establece que el recorrido R debe coincidir con la suma de las amplitudes de los intervalos. De ahí que, a la hora de los cálculos, debamos establecer los **límites exactos** de las clases sin dejar huecos.