e-mail: imozas@elx.uned.es
www.telefonica.net/web/imm

INTRODUCCIÓN A LA ESTADÍSTICA. FEBRERO 2010 Segunda semana. Código de carrera 43. Código de asignatura 203.

PRIMERA PARTE: PREGUNTAS TEÓRICAS

1. ¿En qué casos es preferible, por ser más representativa, utilizar la media geométrica en vez de la media aritmética?

Respuesta.-

Cuando se trata de promediar valores que no sean de naturaleza aditiva como índices, tipos de interés, porcentajes, etc.

2. Sí a una variable X_i la sometemos al mismo tiempo a un cambio de origen O_t y a un cambio de escala C ¿cuál o cuáles de las siguientes afirmaciones son falsas y correctas? Razone la respuesta:
   - los cambios de origen afectan a la media aritmética
   - los cambios de escala afectan a la media aritmética
   - la varianza y la desviación típica sólo se ven afectadas por los cambios de escala.

Respuesta.-

Las tres afirmaciones son ciertas.

Si $Y_i = \dfrac{X_i - O_t}{C}$, se cumple que $\overline{Y} = \dfrac{\overline{X} - O_t}{C}$, $\mathrm{Var}(Y) = \dfrac{\mathrm{Var}(X)}{C^2}$ y $DT(Y) = \dfrac{DT(X)}{|C|}$

3. ¿Qué momentos bidimensionales son siempre nulos?: los momentos bidimensionales respecto a la media de orden 1, los momentos bidimensionales respecto al origen de orden 1 o los momentos bidimensionales respecto al origen de orden 2. Justifique la respuesta

Respuesta.-

Los momentos bidimensionales respecto a la media de orden 1, a saber, $m_{10} = \dfrac{1}{N}\sum_{i=1}^{r}(x_i - a_{10})n_{i\cdot}$ y $m_{01} = \dfrac{1}{N}\sum_{j=1}^{s}(y_j - a_{01})n_{\cdot j}$, ya que :

$$m_{10} = \frac{1}{N}\sum_{i=1}^{r}(x_i - a_{10})n_{i\cdot} = \frac{1}{N}\sum_{i=1}^{r}x_i n_{i\cdot} - \frac{1}{N}\sum_{i=1}^{r}a_{10}n_{i\cdot} = a_{10} - a_{10}\frac{1}{N}\sum_{i=1}^{r}n_{i\cdot} = a_{10} - a_{10} = 0 .$$

Análogamente para m_{01}.

4. Explique los métodos para la determinación de variaciones estacionales en series temporales

Respuesta.-

Se construye un índice de variación estacional para cada uno de los periodos en que se divide el año (meses, trimestres...). Una vez obtenido se desestacionaliza la serie dividiendo cada valor por el índice correspondiente (y multiplicando por 100 si el índice se expresa en %).

Si la serie tiene carácter multiplicativo, para construir los índices puede emplearse el <u>método de la razón a la media móvil</u>, que consiste en dividir cada valor por la correspondiente media móvil (de periodo 1 año) centrada en él. Se obtienen a continuación las medias de cada periodo y el promedio de estas medias. Se construye el índice dividiendo cada media por este promedio.

Si la serie tiene carácter aditivo, se emplea el <u>método de la tendencia por ajuste mínimo cuadrático</u> que consiste en ajustar una recta de regresión a la distribución bidimensional formada por los años y las medias anuales correspondientes. Además se calculan las medias de cada periodo M_i que corregiremos restándole la correspondiente fracción de la pendiente de la recta de regresión b:

$$M'_i = \frac{(i-1)b}{\text{nº de estaciones}}$$

e-mail: imozas@elx.uned.es
www.telefonica.net/web/imm

Finalmente se obtiene el índice dividiendo cada M'_i por el promedio de todos ellos.

5. Defina las medidas de simetría y apuntamiento de una distribución de frecuencias.

Respuesta.-

Coeficiente de asimetría de Fisher: $g_1 = \frac{m_3}{\sigma^3} = \frac{\frac{1}{N}\sum_{i=1}^{r}(x_i - \bar{x})^3 n_i}{\sigma^3}$

Si $g_1>0$, hay asimetría positiva o la derecha; si $g_1 < 0$ hay asimetría negativa o la izquierda; si la distribución es simétrica, entonces $g_1 = 0$ (aunque no recíprocamente)

Coeficiente de curtosis de Fisher: $g_2 = \frac{m_4}{\sigma^4} - 3 = \frac{\frac{1}{N}\sum_{i=1}^{r}(x_i - \bar{x})^4 n_i}{\sigma^3} - 3$

Si $g_2>0$, la distribución se dice leptocúrtica (más apuntada de lo normal); si $g_2 = 0$, mesocúrtica (tiene un apuntamiento normal); si $g_2 < 0$ platicúrtica (apuntamiento menor que el normal).

SEGUNDA PARTE: PROBLEMAS

1. Una empresa quiere realizar un estudio sobre la influencia de las campañas publicitarias en sus cifras de ventas. Para ello dispone del gasto destinado a publicidad y sus ventas (ambos en millones de euros) en los últimos 6 años:

Años	Gastos publicidad	Ventas
2001	2,2	195
2002	2,5	200
2003	2,8	221
2004	2,9	230
2005	3,1	239
2006	3,5	248

- Obtener un modelo lineal que permita predecir las ventas en función de los gastos en publicidad.
- Predecir las ventas de 2007 si se piensa invertir en publicidad 5 millones de euros.
- Valorar los errores obtenidos por la recta de regresión.

Solución.-

Construimos la tabla:

Años	Gastos publicidad x_i	Ventas y_i	x_i^2	y_i^2	$x_i y_i$
2001	2,2	195	4,84	38025	429
2002	2,5	200	6,25	40000	500
2003	2,8	221	7,84	48841	618,8
2004	2,9	230	8,41	52900	667
2005	3,1	239	9,61	57121	740,9
2006	3,5	248	12,25	61504	868
	17	**1333**	**49,2**	**298391**	**3823,7**

de donde obtenemos los momentos:

e-mail: imozas@elx.uned.es
www.telefonica.net/web/imm

a_{10}= 2,83

a_{01}= 222,17 m_{11}= 7,81

a_{11}= 637,28 m_{20}= 0,17

a_{20}= 8,20 m_{02}= 373,81

a_{02}= 49731,83

y la recta de regresión de Y/X: $y - 222{,}17 = \dfrac{7{,}81}{0{,}17}\left(x - 2{,}83\right) \quad \leftrightarrow y = 45{,}36x + 93{,}66$

Si hacemos x = 5, se obtiene $y \cong 320{,}44$.

La valoración de los errores cometidos por la recta de regresión la podemos hacer calculando la varianza residual $S_{yr}^2 = m_{02} - \dfrac{m_{11}^2}{m_{20}} \cong 19{,}53$ que. comparada con la varianza $S_y^2 = m_{02}$, resulta:

$$\frac{S_{yr}^2}{S_y^2} \cong \frac{19{,}53}{373{,}18} \cong 0{,}05$$

que es un valor muy pequeño, lo que indica que la recta se ajusta bien a la nube de puntos.

De manera equivalente, se podría haber calculado el coeficiente de determinación $R^2 = 1 - 0{,}05 = 0{,}95$ que indica un alto valor de la bondad del ajuste.

2. Un hotel cuenta con 3 habitaciones, una simple, una doble y una suite. Se muestran a continuación los precios aplicados y el número de clientes en los años t y t+1. Calcule los índices de precios de Laspeyres, Paasche y Fisher, y los precios medios de cada año.

	Año t		Año t+1	
	Precio	Nº de clientes	Precio	Nº de clientes
Habitación simple	40	65	45	70
Habitación doble	65	46	70	55
Suite	85	84	100	65

Solución.-

Índice de Laspeyres del año t+1, base t = $\dfrac{\sum p_{i,t+1}q_{it}}{\sum p_{it}q_{it}} = \dfrac{45\cdot 65 + 70\cdot 46 + 100\cdot 84}{40\cdot 65 + 65\cdot 46 + 85\cdot 84} \cong 114{,}26$

Índice de Paasche del año t+1, base t = $\dfrac{\sum p_{i,t+1}q_{i,t+1}}{\sum p_{it}q_{i,t+1}} = \dfrac{45\cdot 70 + 70\cdot 55 + 100\cdot 65}{40\cdot 70 + 65\cdot 55 + 85\cdot 65} \cong 113{,}45$

Índice de Fisher del año t+1, base t = $\sqrt{114{,}26\cdot 113{,}45} \cong 113{,}85$

Precio medio de la habitación, año t: $\dfrac{40\cdot 65 + 65\cdot 46 + 85\cdot 84}{65 + 46 + 84} \cong 65{,}28$

Precio medio de la habitación, año t+1: $\dfrac{45\cdot 70 + 70\cdot 55 + 100\cdot 65}{70 + 55 + 65} \cong 71{,}05$

A pesar de que el precio de la habitación en moneda corriente es mayor el año t+1 que el año t, si deflactamos el precio medio del año t+1 con cualquiera de los índices se obtiene un precio medio, en moneda constante del año t, aproximadamente de 62, inferior al precio medio del año t.