

1. Un índice cuantifica la variación conjunta de un período t respecto de otro período base 0 de ciertas magnitudes (habitualmente de precios, cantidades o valores) medidas en una serie de bienes. Al calcular el valor de un índice conjunto realizamos una suma en la que cada sumando se corresponde con un bien concreto. Por ejemplo, si hacemos un índice tipo media ponderada de 3 bienes de un período t_1 con base 0, tendríamos que:

$$I_0^{t_1} = \frac{I_0^{t_1}(1)w_1}{w_1 + w_2 + w_3} + \frac{I_0^{t_1}(2)w_2}{w_1 + w_2 + w_3} + \frac{I_0^{t_1}(3)w_3}{w_1 + w_2 + w_3} = \text{aportación del bien 1} + \text{aportación del bien 2} + \text{aportación del bien 3}$$

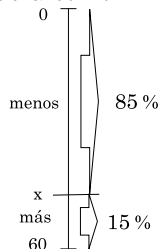
Vemos que cada sumando es la aportación del bien al índice. De la misma forma, la variación absoluta al pasar de un período t_1 a otro t_2 base 0 se puede descomponer también en las partes que se corresponden de cada bien. Ya que calculamos la variación absoluta restando el valor de los índices, para saber qué parte se debe a cada bien debemos restar las aportaciones de cada bien, es decir:

$$\text{aportación del bien } i \text{ en } I_0^{t_2} - \text{aportación del bien } i \text{ en } I_0^{t_1} = \frac{I_0^{t_2}(i)w_i}{\sum_{i=1}^N w_i} - \frac{I_0^{t_1}(i)w_i}{\sum_{i=1}^N w_i} = \frac{(I_0^{t_2}(i) - I_0^{t_1}(i))w_i}{\sum_{i=1}^N w_i}$$

Esta fórmula se conoce como **repercusión del bien i en la variación del índice base 0 al pasar de t_1 a t_2** y se denota por $R_0^{t_1 \rightarrow t_2}(i)$. La repercusión del bien i nos indica cuántos **puntos** aumentó o disminuyó el índice base 0 al pasar de t_1 a t_2 a causa de dicho bien i . Si sumamos las repercusiones de todos los bienes, obtenemos la variación absoluta del índice.

2. **Planteamiento:** el experimento consiste en observar el tiempo que permanecen en la sala de espera (variable 1) y el tiempo de consulta (variable 2) de cada cliente (individuo) de cierto bufete de abogados. La población la constituirán entonces todos los clientes y tenemos una muestra de 46.

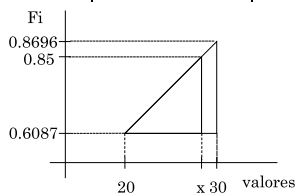
El **Apartado a)** se refiere a la variable 1 (tiempo de espera). Nos preguntan cuánto tiempo permanece en la sala de espera como mínimo un cliente que está entre el 15 % de los que más esperan, lo que gráficamente se representaría así:



Lo mínimo que espera una persona que está en el 15 % superior es justo el valor x . Nosotros preferimos realizar los cálculos acumulando de menor a mayor para poder utilizar la columna F_i . En el gráfico vemos que el valor x es el percentil 85, ya que el valor que deja por encima el 15 % es el mismo que deja por debajo el 85 %. Para hallar x necesitamos la tabla de frecuencias del tiempo de espera: las frecuencias absolutas marginales n_i se obtienen sumando todos los clientes que hay en cada clase y para obtener las F_i debemos ir acumulando y dividiendo por $N = 46$:

clase	n_i	F_i
[0,10]	4+8+2=14	$\frac{14}{46} = 0,3043$
(10,20]	5+3+6=14	$\frac{14+14}{46} = 0,6087$
(20,30]	8+3+1=12	$\frac{14+14+12}{46} = 0,8696$
(30,60]	1+2+3=6	1
Total	46	

La primera vez que F_i supera 0.85 es en el intervalo (20,30]. Vemos que el 60.87 % de los clientes esperan a lo sumo 20 minutos, mientras que el 86.96 % esperan como mucho 30 minutos, luego el valor x que buscamos tiene que estar entre 20 y 30 minutos. Si suponemos que los clientes están más o menos uniformemente repartidos en ese intervalos, para afinar más podemos utilizar triángulos semejantes.



Como base pequeña es a altura pequeña lo mismo que base grande es a altura grande, tenemos:

$$30-20=10 \quad \text{—} \quad 0.8696-0.6087=0.2609$$

$$x-20 \quad \text{—} \quad 0.85-0.6087=0.2413$$

por lo que $x-20 = 0,2413 \times 10 / 0,2609$, es decir $x = 29,2488$ minutos. En consecuencia, un cliente que está entre el 15 % de los que más esperan, permanecerá como mínimo 29.2488 minutos en la sala de espera.

En el **Apartado b)** nos preguntan si podemos predecir de manera fiable el tiempo de consulta a partir de un valor del tiempo de espera. Lo que tenemos que predecir es lo que llamamos habitualmente Y (en este caso, tiempo de consulta) y lo que conocemos para hacer la predicción es X (que será entonces, el tiempo de espera). Las predicciones se pueden realizar mediante regresión, aunque únicamente si el valor que conocemos está entre el mínimo y el máximo de los que tenemos en la muestra (para no extrapolar). En este caso como 12 minutos está entre 0 y 60, sí se cumple. Sin ayuda del SPSS sólo podemos utilizar la regresión lineal, y para saber si la predicción es fiable o no, lo primero que debemos hacer es calcular el coeficiente de determinación $R^2 = \frac{S_{xy}^2}{S_x^2 S_y^2}$ a partir de la tabla bidimensional:

$X \setminus Y$	5	15	25	
5	4	8	2	14
15	5	3	6	14
25	8	3	1	12
45	1	2	3	6
	18	16	12	$N = 46$

En esta tabla hemos sustituido las clases por su marca para realizar las operaciones y hemos calculado las distribuciones marginales sumando respectivamente las frecuencias de las filas y de las columnas. Operando tenemos que:

$$\bar{x} = \frac{5 \times 14 + \dots + 45 \times 6}{46} = 18,4783 \quad \overline{x^2} = \frac{5^2 \times 14 + \dots + 45^2 \times 6}{46} = 503,2609 \quad S_x^2 = \overline{x^2} - \bar{x}^2 = 503,2609 - 18,4783^2 = 161,8133$$

$$\bar{y} = \frac{5 \times 18 + 15 \times 16 + 25 \times 12}{46} = 13,6957 \quad \overline{y^2} = \frac{5^2 \times 18 + 15^2 \times 16 + 25^2 \times 12}{46} = 251,0870 \quad S_y^2 = \overline{y^2} - \bar{y}^2 = 63,5009$$

$$\overline{xy} = \frac{5 \times 5 \times 4 + 5 \times 15 \times 8 + \dots + 45 \times 25 \times 3}{46} = 259,7826 \quad S_{xy} = \overline{xy} - \bar{x}\bar{y} = 259,7826 - 18,4783 \times 13,6957 = 6,7093$$

Luego $R^2 = \frac{6,7093^2}{161,8133 \times 63,5009} = 0,0044$, lo que significa que únicamente el 0,44 % de la variabilidad del tiempo de consulta se explica por la regresión lineal con el tiempo de espera, así que la predicción no es en absoluto fiable, por lo que no merece la pena calcularla.

3. Planteamiento: el experimento consiste en observar el salario mensual (variable) de los empleados de una multinacional (individuos). La población la constituirán entonces todos los empleados y tenemos una muestra de $N = 398$ (lo vemos en el gráfico de cajas). En el **Apartado a)** nos preguntan alrededor de qué valor se distribuyen los sueldos aproximadamente, así que debemos dar una medida de tendencia central. Sabemos que la mejor media de tendencia central si no hay datos extremos es la media, aunque en los gráficos vemos que sí hay casos extremos, por lo que será mejor la mediana. Además, como sólo tenemos esos gráficos, debemos contestar la pregunta con la información visual que proporcionan. Sabemos que la raya negra del gráfico de cajas es justo la mediana. Vemos que está aproximadamente alrededor de los 2700 dólares, por lo que podemos decir que los sueldos se distribuyen aproximadamente alrededor de los 2700 dólares. En el **Apartado b)** nos preguntan por la simetría. Sabemos que una variable es simétrica si los valores se distribuyen de la misma forma por encima y por debajo del centro, lo que gráficamente lo podemos observar de forma aproximada en el diagrama de tallos y hojas. Vemos que hay una ligera asimetría a la derecha: la mayor parte se concentra en el centro (filas largas) y hay algunos valores más altos que bajos, lo que viene a significar que la mayor parte de los empleados tienen sueldos moderados, aunque hay algunos más con sueldos altos que con sueldos bajos. En el **Apartado c)** nos preguntan por los casos extremos. En el gráfico de tallos y hojas leemos que hay 11 extremos por arriba, lo que significa que hay 11 empleados que cobran mucho en comparación con el resto. En el gráfico de cajas, vemos que en concreto hay 2 con sueldos extremadamente altos (asteriscos) y que los otros 9 (círculos) están bastante cerca del máximo de lo normal, que se sitúa entorno a los 4500 dólares (segmento de la pata superior). Los empleados que cobran mucho más que el resto tienen unos sueldos respectivos de aproximadamente 7000 y 8000 dólares respectivamente.

4. Planteamiento: tenemos 1 bien, que son las hipotecas y 2 magnitudes, que son las cantidades y los valores o gastos (precios por cantidades). El período base es el año 2001 y el actual es 2002. Los datos son:

	Cantidades	Valores
Año 2001	15442	1105
Año 2002	18101	1401

En el **Apartado a)** nos piden comparar el valor real de las hipotecas, por lo que debemos deflactar para eliminar la influencia de la subida generalizada de precios medida mediante el IPC. Nos dicen que el IPC aumentó un 4 % durante el año 2002, luego $IPC_{01}^{02} = 1,04$.

El Valor constante (en unidades del año 2001) se puede calcular como $\frac{\text{valor corriente año 2002}}{IPC_{01}^{02}} = \frac{1401}{1,04} = 1347,1154$ millones de euros. Eso significa que los 1401 millones invertidos en hipotecas durante en año 2002 equivalen a 1347.1154 millones de euros “del año 2001”. Como durante dicho año se habían invertido un total de 1105 millones de euros, se puede evaluar el incremento mediante el índice: $\frac{1347,1154}{1105} = 1,2191$, lo que implica que hubo un crecimiento del 21.91 % en términos reales.

En el **Apartado b)** nos dicen que el índice cuántico de las hipotecas de 2003 respecto 2002 fue $Q_{02}^{03} = 1,0315$ y nos preguntan cuantas viviendas se habrán hipotecado ese año. Sabemos que en 2002 se hipotecaron 18101 viviendas y sabemos que el valor del índice que nos dan significa que el número de hipotecas en 2003 fue el 103.15 % el número de hipotecas de 2002, luego sólo tenemos que calcular ese porcentaje: 103.15 % de 18101 = $1,0315 \times 18101 = 18671.1815$ (no nos da un número entero por cuestiones de redondeo). En conclusión, en el año 2003 se habrán hipotecado unas 18671 viviendas.

En el **Apartado c)** nos preguntan la variación del importe medio de las hipotecas entre 2001 y 2002, para lo que necesitamos calcular los importes medios en ambos años. En 2001 se hipotecaron 15442 viviendas por un total de 1105 millones, luego el importe medio de la hipoteca de cada vivienda fue de $1105/15442 = 0,0716$ millones de euros (gasto total/número de individuos), es decir, el importe medio es de 71600 euros. Análogamente, en 2002 tenemos $1401/18101 = 0,0774$ millones de euros, por lo que este año el importe medio fue de 77400 euros. La variación de 2002 respecto a 2001 se obtiene calculando el índice o precio relativo correspondiente: $P_{01}^{02} = 77400/71600 = 1,0810$, lo que significa que el importe medio de las hipotecas en 2002 aumentó un 8.1 % respecto a 2001.