Problemas
Interpretar un caso
Vicente Manzano-Arrondo, 2012-2014

Problemas de cálculo

Ejercicio 1 resuelto

Tras preguntar a un grupo de personas cuántos días han pasado desde la última vez que tomaron un café en un bar, la tabla de frecuencias resultante es (d2 representa la distancia cuadrática a la media):

Xi	fi	Xifi	d2fi
2	4	8	16,00
3	9	27	9,00
4	1	4	0,00
6	3	18	12,00
7	2	14	18,00
9	1	9	25,00
Σ	20	80	80,00

Con estos datos, demuestra que la suma de distancias estandarizadas tiene el valor 0.

Solución

Para resolver el problema hemos de calcular primero la columna de distancias estandarizadas, para lo que es necesario obtener primero los valores de la media aritmética y la desviación tipo:

$$\bar{X} = \frac{\sum X_i f_i}{n} = \frac{80}{20} = 4 \qquad S = \sqrt{\frac{\sum \left(X_i - \bar{X}\right)^2 f_i}{n}} = \sqrt{\frac{80}{20}} = 2$$

Aplicando entonces la expresión $Z_i = \dfrac{X_i - \bar{X}}{S}$ se obtiene:

Xi	fi	Xifi	d2fi	Zi
2	4	8	16,00	-1,00
3	9	27	9,00	-0,50
4	1	4	0,00	0,00
6	3	18	12,00	1,00
7	2	14	18,00	1,50
9	1	9	25,00	2,50
Σ	20	80	80,00	

La suma de la columna Zi no resuelve el problema. Sabes que hay 20 datos, aunque solo 6 valores. Sumar los 6 valores distintos de distancia estandarizada no es sumar las 20 distancias que realmente existen. De hecho, si se te ocurre hacerlo, verás que esa columna suma 3,5. Observa que, por ejemplo, no hay una Zi=-1, sino 4 datos que se alejan de la media en 1 desviación tipo por debajo. Luego, habrá que sumar -1 cuatro veces. Esto se soluciona del mismo modo que hemos hecho con la columna Xifi que nos ha permitido sumar los 20 datos:

Xi	fi	Xifi	d2fi	Zi	Zifi
2	4	8	16,00	-1,00	-4,00
3	9	27	9,00	-0,50	-4,50
4	1	4	0,00	0,00	0,00
6	3	18	12,00	1,00	3,00
7	2	14	18,00	1,50	3,00
9	1	9	25,00	2,50	2,50
Σ	20	80	80,00		0,00

n	Media	D.t.
20	4,00	2,00

Ejercicio 2 resuelto

Hemos preguntado a un grupo de personas cuántas veces han accedido a su centro de trabajo en transporte público durante las dos últimas semanas. Los resultados son:

4, 5, 4, 4, 3, 3, 4, 9, 4, 1, 4, 2, 4, 3, 4, 10, 4, 4, 3, 2, 3

1) Construye una tabla de frecuencias con las siguientes columnas: valores, frecuencias y distancias estandarizadas. Y 2), ordena los valores según cuán difieren de la media, utilizando las distancias estandarizadas.

Solución

Primeramente, construimos la tabla con las dos columnas básicas, a la que añadimos dos más para facilitar los cálculos:

Xi	fi	Xifi	d2fi
1	1	1	9,00
2	2	4	8,00
3	5	15	5,00
4	10	40	0,00
5	1	5	1,00
9	1	9	25,00
10	1	10	36,00
Σ	21	84	84,00

Para la columna de distancias estandarizadas habrá que calcular antes la media aritmética y la desviación tipo. Para el caso de la media aritmética:

$$\bar{X}=\frac{\sum X_i f_i}{n}=\frac{1\cdot 1+2\cdot 2+3\cdot 5+4\cdot 10+\ldots}{21}=\frac{84}{21}=4$$

Para el caso de la desviación tipo:

$$S=\sqrt{\frac{\sum (X_i-\bar{X})^2 f_i}{n}}=\sqrt{\frac{(1-4)^2\cdot 1+(2-4)^2\cdot 2+(3-4)^2\cdot 5+\ldots}{21}}=2$$

Con estos resultados y sabiendo que $Z_i=\frac{X_i-\bar{X}}{S}$ entonces:

$$Z_1=\frac{1-4}{2}=-1{,}5;\quad Z_2=\frac{2-4}{2}=-1;\quad \ldots \quad ;Z_7=\frac{10-4}{2}=3$$

y la tabla queda como sigue:

Xi	fi	Zi
1	1	-1,50
2	2	-1,00
3	5	-0,50
4	10	0,00
5	1	0,50
9	1	2,50
10	1	3,00
Σ	21	

Para responder a la segunda pregunta, hay que atender a la cuantía de la distancia estandarizada, no a su signo. El signo indica dónde se orienta la distancia: hacia arriba o hacia abajo, pero no su importancia. Así que, la tabla ordenada quedaría así (el orden de los valores 3 y 5 es arbitrario, puesto que ambos se encuentran a una misma distancia de la media, que en términos estandarizados tiene el valor absoluto 0,5)

Xi	fi	Zi
10	1	3,00
9	1	2,50
1	1	-1,50
2	2	-1,00
5	1	0,50
3	5	-0,50
4	10	0,00
Σ	21	

Ejercicio 3 resuelto

Con los mismos datos del problema 1, obten la columna de percentiles o porcentajes acumulados.

Solución

Para pasar de valores a percentiles, calculamos la columna de porcentajes acumulados y corregimos el último valor para adaptarlo a P_{99}:

Xi	fi	%i	%ai	Pi
2	4	20	20	20
3	9	45	65	65
4	1	5	70	70
6	3	15	85	85
7	2	10	95	95
9	1	5	100	99
Σ	20	100		

Para pasar de percentiles a valores, partimos de la tabla donde los identificadores de los percentiles se reparten en filas (dígito de las decenas) y columnas (dígito de las unidades). Cada casilla se corresponde con el valor de cada percentil, leyendo la tabla. Así por ejemplo, el valor 4 ocupa las posiciones de porcentaje acumulado que van desde la 66 hasta la 70, ambas inclusive. Luego:

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0		2	2	2	2	2	2	2	2	2
10	2	2	2	2	2	2	2	2	2	2
20	2	3	3	3	3	3	3	3	3	3
30	3	3	3	3	3	3	3	3	3	3
40	3	3	3	3	3	3	3	3	3	3
50	3	3	3	3	3	3	3	3	3	3
60	3	3	3	3	3	3	4	4	4	4
70	4	6	6	6	6	6	6	6	6	6
80	6	6	6	6	6	6	7	7	7	7
90	7	7	7	7	7	7	9	9	9	9

Ejercicios propuestos

1. Resuelve la misma situación que el ejercicio resuelto 2 con los siguientes datos:

1, 3, 3, 1, 5, 1, 3, 5, 3, 3, 7, 1, 3, 1, 7, 1, 5, 1, 3, 1, 3, 7, 3, 1, 1, 5

2. Calcula las distancias estandarizadas, los tres cuartiles y los percentiles 15 y 87 del siguiente conjunto de datos:

Xi	fi
1	5
2	6
4	9
6	11
9	9
Σ	40

3. ¿Cuál es el valor de la mediana de un conjunto de datos donde sabemos que $Q_1=8$, $D_6=17$, $P_{38}=12$, $Q_3=20$ y $P_{46}=17$?

Ejercicios de exámenes

7. Necesitamos seleccionar a un experto en cocina marítima. Se presentan cinco candidatos, con sus respectivos informes. Cada uno realizó un examen de aptitud en su país de origen. La puntuación que obtuvo cada participante, directa (Xi) y estandarizada (Zi) y el coeficiente de variación de Pearson del grupo de procedencia (CV) figuran en la tabla siguiente. Con estos datos, responde a las siguientes tres preguntas (pon en cada casilla el número de participante que proceda) *{6 puntos}*

	Participante				
	1	**2**	**3**	**4**	**5**
Xi	8	6	6	7	9
Zi	1,3	-0,4	0,2	1,1	1,0
CV	10	34	15	8	20

A) ¿Qué participante ha obtenido la puntuación más sobresaliente o excepcional?
B) ¿Qué grupo ha contado con puntuaciones más variadas?
C) ¿Qué participante ha obtenido la puntuación más común en su grupo?

6. En una competición de sumo participan 150 deportistas, que provienen de varios países. Hemos pesado a cada participante (**X**), obteniendo el percentil (**P**) y la puntuación tipo (**Z1**) que le corresponde en la competición. Leyendo en su ficha, hemos obtenido la puntuación tipo en su país (**Z2**). El resultado, para cinco de ellos, es:

	Ernestiño	Yakamoto	Ahmayed	McNamara	Manolo
X	110	120	128	136	152
P	40	46	50	65	87
Z1	-0,5	0,0	0,4	0,8	1,6
Z2	0,2	-0,7	1,2	0,4	0,9

Responde a las siguientes preguntas (8 puntos):

Pregunta	Respuesta
A) Mediana en la competición	
B) Media en la competición	
C) Competidor más sobresaliente de los cinco	
D) De los cinco, competidor más raro en su país	