

Ejercicio resuelto 3.1

La demanda de cierto artículo en 48 días fue 1, 4, 1, 0, 2, 1, 1, 3, 2, 1, 1, 0, 3, 2, 4, 3, 4, 1, 2, 1, 1, 2, 2, 2, 1, 3, 3, 3, 1, 4, 4, 0, 2, 1, 4, 0, 3, 1, 3, 3, 4, 2, 2, 1, 0, 1, 2, 4, mientras que en otros 48 días hubo una demanda de 1 artículo en 13 de ellos, de 2 artículos en 12, de 3 en 10, de 4 en 9 y el resto de los días no hubo demanda. ¿Podrías decir en qué conjunto de días hubo mayor demanda diaria?

Solución:

El **objetivo** es comparar la demanda diaria de ambas muestras. Gran parte de este problema está planteado en el **Ejercicio resuelto 1.2**. Sólo cambia que ahora se tienen 2 muestras. Se denotará por x_i los datos de la primera muestra y por y_j los de la segunda. Las frecuencias de ambas muestras se recogen en la Tabla 3.1.

x_i	n_i	y_j	n_j
0	5	0	4
1	15	1	13
2	11	2	12
3	9	3	10
4	8	4	9
Total	48	Total	48

Tabla 3.1: Tablas de frecuencias.

Método y justificación: en principio, parece que no tiene sentido hacer esa comparación, porque la demanda diaria es variable y habrá días en que sea mayor en una muestra y días en que sea mayor en la otra. Además, a simple vista no es sencillo comparar tantos valores a la vez. Sin embargo, si se logra resumir mediante un único valor cómo es “aproximadamente” la demanda diaria en cada conjunto de días, se podrían comparar esos dos valores.

Cálculos: aunque se pueden hacer las operaciones sustituyendo directamente en la fórmula, lo habitual para ilustrar todos los cálculos es que se añada una nueva columna en la tabla de frecuencias donde se vayan calculando los sumandos $x_i n_i$, que se completaría con la suma final x_T (ver Tabla 3.2).

x_i	n_i	$x_i n_i$	y_j	n_j	$y_j n_j$
0	5	$0 \times 5 = 0$	0	4	$0 \times 4 = 0$
1	15	$1 \times 15 = 15$	1	13	$1 \times 13 = 13$
2	11	$2 \times 11 = 22$	2	12	$2 \times 12 = 24$
3	9	$3 \times 9 = 27$	3	10	$3 \times 10 = 30$
4	8	$4 \times 8 = 32$	4	9	$4 \times 9 = 36$
Total	$N=48$	$x_T=96$	Total	$N=48$	$y_T=103$

Tabla 3.2: Tablas para las medias.

De esta forma se tendría que

$$\bar{x} = \frac{\text{suma de todos los artículos demandados}}{\text{número de días}} = \frac{96}{48} = 2 \text{ artículos,}$$

$$\bar{y} = \frac{\text{suma de todos los artículos demandados}}{\text{número de días}} = \frac{103}{48} = 2,1458 \text{ artículos.}$$

Para comparar esas dos cantidades se podría utilizar la diferencia y decir que la demanda media en el segundo conjunto de días fue 0,1458 artículos superior a la media del primer conjunto de días. Sin embargo, la demanda diaria está medida en escala de razón, y en estos casos resulta mucho más informativo utilizar el cociente, porque los resultados no dependen de las magnitudes. El cociente sería $2,1458/2 = 1,0729$.

Conclusión: en el primer conjunto de datos la demanda diaria varió alrededor de 2 artículos, mientras que en el segundo conjunto de datos dicha demanda diaria variaba alrededor de 2,1458 artículos. Aunque no se puede comparar la demanda diaria en las dos muestras porque varía de día en día, sí se puede decir que en media la demanda diaria fue un 7,29% superior en el segundo conjunto de días.

Ejercicio resuelto 3.2

De las devoluciones mensuales que se realizan a cierto almacén el 10 % tienen un importe de 360€ a lo sumo. El 40 % son de un importe inferior o igual a 660€. En la mitad de ellas no se superan las 1385 mientras que el 30 % están entre 1385 y 2410€.

- a) Si se sabe que en dicho almacén el importe máximo de las devoluciones es de 3000€, calcula e interpreta el importe medio de las devoluciones.
- b) ¿Podrías calcular el importe medio si no se supiera cuál es el importe máximo de las devoluciones?

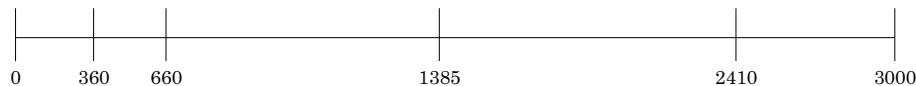
Solución:

El **objetivo del Apartado a)** es calcular el importe medio de las devoluciones.

Planteamiento: el experimento consiste en seleccionar devoluciones (individuos) y observar su importe (variable), luego la población son todas las devoluciones (no dan ningún tamaño muestral, parece que la información se refiere a toda la población). Los datos son numéricos, el 0 significa que no hay devolución y, en principio, cualquier importe (385, 385,6, etc.) es posible, luego la variable es cardinal, de razón y continua. No se tienen datos aislados, si no rangos (datos agrupados) y frecuencias.

Método y justificación: se calculará la media del importe, porque ese es el objetivo. Más adelante se justificará por qué esta medida es la mejor para determinar el centro de una variable de este tipo.

Cálculos: se realizarán a partir de la tabla de frecuencias. Como se conocen sólo los rangos, lo primero es localizar las clases de datos agrupados. Para ello es útil ir marcando en una recta los valores que determinan los rangos, así quedarían ya ordenadas dichas clases:



Sólo se tienen frecuencias relativas, no absolutas. En la Tabla 3.3 se han incluido los datos que proporciona directamente el enunciado y se han utilizado letras para representar los datos que no conocemos directamente.

Para calcular la media se necesita la columna de f_i completa, por lo que se debe completar la tabla. F_i representa las frecuencias acumuladas, por

lo que $F_2 = 0,4$ tiene que ser F_1 más lo que corresponda a f_2 . Matemáticamente sería $0,4 = 0,1 + a$, por lo que $a = 0,3$. De igual forma $b = f_3$ tiene que ser lo que falta para pasar de $F_2 = 0,4$ a $F_3 = 0,5$, es decir $b = 0,5 - 0,4 = 0,1$.

clases	f_i	F_i	x_i
$[0, 360]$	0,1	0,1	180
$(360, 660]$	a	0,4	510
$(660, 1385]$	b	0,5	1022,5
$(1385, 2410]$	0,3	c	1897,5
$(2410, 3000]$	d	e	2705
Total	f	—	—

Tabla 3.3: Tabla incompleta de frecuencias.

Además $c = F_4$ tiene que ser todo lo acumulado hasta ese momento. Matemáticamente sería $c = F_3 + 0,3 = 0,5 + 0,3 = 0,8$ y el último F_i siempre es 1, porque al final se tiene ya todo acumulado, por lo que $e = F_5 = 1$. Por el mismo motivo, $f = 1$. Por último, como para que la columna de f_i sume 1 falta 0,2, se tiene que $d = 0,2$. Las frecuencias completas y la columna necesaria para calcular la media se representan en la Tabla 3.4.

clases	f_i	F_i	x_i	$x_i f_i$
$[0, 360]$	0,1	0,1	180	$0,1 \times 180 = 18$
$(360, 660]$	0,3	0,4	510	$0,3 \times 510 = 153$
$(660, 1385]$	0,1	0,5	1022,5	$0,1 \times 1022,5 = 102,25$
$(1385, 2410]$	0,3	0,8	1897,5	$0,3 \times 1897,5 = 569,25$
$(2410, 3000]$	0,2	1	2705	$0,2 \times 2705 = 541$
Total	1	—	—	1383,5

Tabla 3.4: Tabla de frecuencias.

De este modo, se tiene que $\bar{x} = \sum_{i=1}^k x_i f_i = 1383,5\text{€}$.

Conclusión: el importe medio de las devoluciones mensuales es de 1383,5€, lo que significa que el importe de las devoluciones en ese almacén oscila alrededor de 1383,5€.

El **objetivo del Apartado b)** es calcular la media si no se sabe el importe máximo de las devoluciones. El **planteamiento** es el mismo que el del Apartado a).

Método, justificación, cálculos y conclusión: no se puede decir hasta dónde llega el último intervalo y eso en matemáticas se indica utilizando el símbolo ∞ (infinito). Es decir, nuestro último intervalo sería $(2410, \infty)$ y el resto de la tabla sería igual. Como no se puede calcular el punto medio de este último intervalo, no se puede calcular la media aritmética.

Ejercicio resuelto 3.3

Un estudiante ha realizado 1 examen que constaba de 3 partes: una teórica, otra de problemas y otra de prácticas de informática. El profesor le da el doble de importancia a los problemas que a la teoría y el triple a las prácticas. Si ha obtenido una calificación de 5,8 sobre 10 en teoría, 6,4 sobre 10 en problemas y 7,9 sobre 10 en prácticas, ¿cuál crees que será su calificación final en el examen?

Solución:

El **objetivo** es calcular la calificación final del examen.

Planteamiento: el experimento consiste en seleccionar cada parte de examen (individuo) y observar la nota del alumno en esa parte (variable). Luego la población serían las 3 partes del examen y la muestra sería igual a la población. La variable es cardinal (porque las notas son números), de intervalo (porque la escala es subjetiva, el 0 no significa ausencia de nota, ni un 10 representa el doble de conocimientos que un 5) y continua (en principio, se puede obtener notas de 6,7 y 6,789, etc.).

Método y justificación: la nota final debería ser la nota media de todas las partes teniendo en cuenta la importancia que tiene cada una de esas partes.

Cálculos: la media ponderada se calculará también a partir de la tabla de frecuencias. Aunque sólo haya hecho 1 examen de cada tipo, se le da el doble de importancia a un examen que a otro, lo que implica que a la hora de calcular la media, ese examen debería tener el doble de peso, es decir, quedaría multiplicado por 2 (así, las ponderaciones jugarían el papel de las frecuencias en la fórmula matemática). La tabla de frecuencias sería entonces:

x_i	n_i	w_i	$w_i x_i$
5,8	1	1	$1 \times 5,8 = 5,8$
6,4	1	2	$2 \times 6,4 = 12,8$
7,9	1	3	$3 \times 7,9 = 23,7$
Total	$N=3$	6	42,3

Por lo tanto, $\bar{x}_w = \frac{42,3}{6} = 7,05$ puntos.

Conclusión: la nota final del examen sería de 7,05 puntos, que representa la media de todas las partes teniendo en cuenta su importancia.

Ejercicio resuelto 3.4

En un supermercado incrementaron el precio de uno de sus productos un 10 % en el mes de julio y en septiembre lo han vuelto a incrementar en un 30 %, ¿cuál ha sido el incremento total?, ¿y el incremento medio?

Solución:

El **objetivo** es calcular el incremento total de precios y el incremento medio.

Planteamiento: el experimento consiste en observar el incremento de precios (variable) en cada subida (individuos), luego la población la constituirían todas las subidas de precios y se tiene una muestra de 2 subidas. La variable es cardinal, de razón y continua (porque los posibles valores son números, un incremento de 0 significa que no hay subida y, en principio, se puede fijar cualquier subida: 25 % o 25,36 %, etc.).

Método y justificación: como los incrementos de precios son *variaciones acumulativas*, no se puede sumar, porque, si un producto costase 1€, después de la primera subida costaría 1€ + 10 % de 1 = 1 + 0,1 = 1,1€ y después de la segunda subida costaría 1,1€ + 30 % de 1,1 = 1,1 + 0,3 × 1,1 = 1,43€. Es decir, la subida total sería del 43 %, y no del 40 % que saldría sumando. Esta diferencia se debe a que la segunda subida viene ya afectada por la primera (se sube sobre el precio incrementado ya la primera vez).

Los **cálculos** en el ejemplo serían:

Tasas de variación	n_i	$t_i^{n_i}$
1,1	1	1,1
1,3	1	1,3
Total	2	1,43

Tasa de variación total=1,43

$\bar{x}_G$ = media geométrica= tasa media de variación= $\sqrt{1,43} = 1,1958$.

Conclusión: se produjo una subida total del 43 % y una subida media por incremento del 19,58 % (esto es, las subidas variaron alrededor del 19,58 %).

Ejercicio resuelto 3.5

Se ha hecho un estudio de los precios de venta del agua y de la leche en cierto hipermercado. Respecto al agua, se verificó que 3 marcas cuestan 0,27€, 2 cuestan 0,21€, 1 cuesta 0,24€, y otra más cuesta 0,3€. En cuanto a la leche, se constató que 4 de las marcas costaban 0,79€, 3 costaban 0,61€, 2 costaban 0,69€, otras 2 costaban 0,82€ y una costaba 0,73€. Calcula e interpreta el valor de la mediana del precio del agua y del precio de la leche.

Solución:

El **objetivo** es calcular el precio mediano del agua y de la leche. Se comenzará con los precios del agua.

Planteamiento: el experimento consiste en seleccionar marcas de agua (individuos) y observar su precio (variable). La población son todas las marcas de agua que hay en ese supermercado y se dispone de una muestra (quizás sean todas las que hay) de 7 marcas. La variable es cardinal, de razón y continua (porque los precios son números, 0 significa que no cuesta nada y, en principio, cualquier precio es válido).

Método y justificación: se trata de calcular la mediana porque es lo que se pide. Es decir, se busca un valor que deje la mitad de los datos por debajo y la otra mitad por encima, una vez ordenados.

Cálculos: si se ordenan todos los datos se tiene: 0,21, 0,21, 0,24, 0,27, 0,27, 0,27 y 0,3. El primer valor igual a 0,27 deja 3 marcas por encima y otras 3 por debajo, esto es, es el valor que está en el medio, así que la mediana sería 0,27.

Calcular así la mediana cuando se tienen muchos datos resultaría demasiado tedioso, por eso se presentará una forma de hacer el cálculo *utilizando las tablas*.

En realidad, se necesita conocer el punto en el que se lleva acumulada la mitad, es decir, el 50 % de la muestra. Por eso el cálculo se basa en la columna de frecuencias acumuladas F_i .

Según la Tabla 3.5, el 42,86 % de las marcas cuestan 0,24€ o menos, mientras que el 85,72 % de las marcas cuestan 0,27€ o menos, así que se sobrepasa el 50 % justo cuando se consideran las marcas que cuestan 0,27€, luego $Me = 0,27€$. En general se debe buscar el primer valor de F_i que sobrepasa el 0,5, y el valor correspondiente x_i será la mediana.

Conclusión: el precio mediano del agua son 27 céntimos, que significa que la mitad de las marcas cuestan 0,27€ o menos y la otra mitad cuestan

x_1	n_i	f_i	F_i
0,21	2	0,2857	0,2857
0,24	1	0,1429	0,4286
0,27	3	0,1286	0,8572
0,3	1	0,1429	1
Total	7	1	–

Tabla 3.5: Tabla de frecuencias (precio del agua)

0,27€ o más (hay más formas de expresar eso, por ejemplo, se puede decir que la mitad de las marcas cuestan a lo sumo 0,27€, etc.). El precio del agua oscila alrededor de los 27 céntimos.

En relación con los precios de la leche, el **planteamiento**, el **método** y la **justificación** son similares a los del caso anterior.

Cálculos: Si se ordena la muestra de menor a mayor se tiene

$$\underbrace{0,61 \ 0,61 \ 0,61 \ 0,69 \ 0,69 \ 0,73}, \underbrace{0,79 \ 0,79 \ 0,79 \ 0,79 \ 0,82 \ 0,82}.$$

La mitad de 12 son 6, el hueco entre el 0,73 y el 0,79 separa 6 datos por debajo y 6 por encima, luego cualquier valor que esté entre esos 2 cumpliría la condición para ser mediana. Si se necesita un solo número, se puede optar por dar el número del medio: $(0,73 + 0,79)/2 = 0,76€$ (y se asegura que la mitad de las marcas cuestan 0,76 o menos y la otra mitad 0,76€ o más).

x_1	n_i	f_i	F_i
0,61	3	0,25	0,25
0,69	2	0,1667	0,4167
0,73	1	0,0833	0,5
0,79	4	0,3333	0,8333
0,82	2	0,1667	1
Total	12	1	–

Tabla 3.6: Tabla de frecuencias (precio de la leche)

Al aplicar el método de **cálculo** con tablas (ver Tabla 3.6) aparece justo el valor $F_i = 0,5$, eso significa que exactamente la mitad de las marcas cuestan 0,73€ o menos (o también 0,73€ o más), pero el mismo razonamiento es válido para todos los números entre 0,73 y 0,79. en consecuencia si el 0,5 exacto aparece en la columna de F_i , la mediana es cualquier valor

entre el x_i correspondiente y el siguiente (en particular, el punto medio podría servir de representante).

Conclusión: el precio mediano de la leche es cualquier valor entre 73 y 79 céntimos. Se puede decir que la mitad de las marcas cuestan alrededor de 76 céntimos o menos y la otra mitad cuestan alrededor de 76 céntimos o más. El precio de la leche oscila alrededor de los 76 céntimos.

Ejercicio resuelto 3.6

Un comercial dedica al 15 % de sus clientes menos de 10 minutos, al 38 % entre 10 y 30 minutos y al resto entre 30 y 60 minutos. Calcula e interpreta la mediana del tiempo de atención por cliente.

Solución:

El **objetivo** es calcular la mediana del tiempo de atención por cliente.

Planteamiento: el experimento consiste en seleccionar clientes (individuos) y observar el tiempo que les dedica un comercial (variable). La población son todos sus clientes y la muestra, en este caso, coincide con la población. La variable es cardinal, de razón y continua (porque el tiempo se mide con números, 0 significa que no le dedica tiempo y podría dedicarle cualquier cantidad de tiempo).

Método y justificación: de trata de calcular la mediana porque es el objetivo. Es decir, se busca un valor que deje el 50 % de los clientes por debajo y el otro 50 % por encima.

Cálculos: en primer lugar hay que extraer la tabla de frecuencias. Para calcular la mediana se necesitan las frecuencias acumuladas.

clases	f_i	F_i
[0,10]	0,15	0,15
(10,30]	0,38	0,53
(30,60]	0,47	1

Tabla 3.7: Tabla de frecuencias.

En primer lugar se debe localizar el intervalo donde está la mediana. En la Tabla 3.7 se observa que cuando se pasa por el 10 se lleva acumulado el 15 %, en cambio al pasar por el 30 ya es el 53 %. Eso significa que el 50 %, es decir, el valor que se corresponde a la mediana, está entre 10 y 30.

La idea está en suponer que la F_i se va acumulando paulatinamente desde 0,15 hasta 0,53, es decir, siguiendo la pendiente del triángulo grande de la Figura 3.1. Así que se trata de ver cuanto se lleva acumulado hasta 0,5, es decir, lo que corresponde a la pendiente del triángulo pequeño. Siguiendo la regla de triángulos semejantes (base grande es a altura grande lo mismo que base pequeña es a altura pequeña), se tiene la siguiente regla de tres:

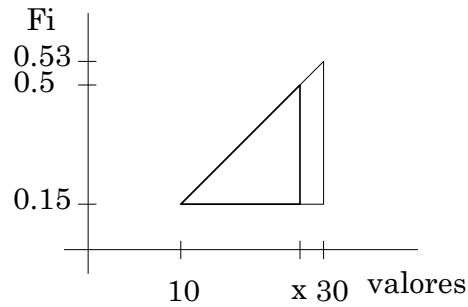


Figura 3.1: Interpolación para aproximar la mediana.

$$\begin{array}{ll} \text{base grande} = 30 - 10 & \text{altura grande} = 0,53 - 0,15 \\ \text{base pequeña} = x - 10 & \text{altura pequeña} = 0,5 - 0,15 \end{array}$$

Entonces

$$\begin{array}{ll} 20 & 0,38 \\ x - 10 & 0,35 \end{array}$$

por lo que $x - 10 = 0,35 \times 20 / 0,38 = 18,4211$, es decir $x = 18,4211 + 10 = 28,4211$.

Conclusión: $Me = 28,4211$ minutos, que significa que ese comercial le dedica a la mitad de sus clientes 28,4211 minutos o menos y a la otra mitad les dedica 28,4211 o más. Esto es una aproximación, porque se supuso que los tiempos estaban uniformemente repartidos en ese intervalo y esto no tiene porqué ser exactamente así. El tiempo que le dedica el comercial a cada cliente oscila alrededor de aproximadamente 28,4211 minutos.

Ejercicio resuelto 3.7

En una encuesta se les preguntó a 16 personas si su hogar era de alquiler o en propiedad. Las contestaciones fueron: alquiler, alquiler, propiedad, alquiler, no sabe/no contesta, alquiler, propiedad, alquiler, propiedad, propiedad, alquiler, alquiler, no sabe/no contesta, alquiler, propiedad, y alquiler. ¿Qué valor representa el centro de esa distribución?

Solución:

El **objetivo** es encontrar el “centro” de la distribución del régimen de propiedad (alquiler o propiedad).

Planteamiento: el experimento consiste en seleccionar personas y observar si su hogar es alquilado o lo tienen en propiedad. La población serán todas las personas (con hogar) y se tiene una muestra de 16 datos. La variable es nominal, porque los valores (alquiler/propiedad) son nombres que no se pueden ordenar.

Método y justificación: como la variable es nominal, no se puede utilizar ni la media ni la mediana, por lo que se utilizará la moda.

Cálculos: Según la distribución de la Tabla 3.8 el mayor n_i , que es 9, se corresponde al alquiler, luego $Mo = \text{alquiler}$.

x_i	n_i
Alquiler	9
Propiedad	5
NS/NC	2
Total	16

Tabla 3.8: Tabla de frecuencias.

Conclusión: el centro de la distribución se situaría en la modalidad de alquiler, por ser la más frecuente.