

Medidas de Posición

Santiago Pérez Moncada

23/6/2020

Las **medidas de posición** estiman qué valores dividen las observaciones en una determinadas proporciones.

Los valores que determinan estas posiciones son conocidos como los **cuantiles**.

Pensándolo de este modo, la mediana puede interpretarse como una medida posición, debido a que divide la variable cuantitativa en dos mitades.

Dada una proporción $p \in (0, 1)$, el **cuantil de orden p** de una variable cuantitativa, Q_p , es el valor más pequeño x que su frecuencia relativa acumulada es mayor o igual a p .

Dicho de otro modo, si tenemos un conjunto de observaciones x_1, \dots, x_n y los ordenamos de menor a mayor, entonces Q_p será el número más pequeño que deja a su izquierda (incluyéndose a sí mismo) como mínimo a la fracción p de datos. Es decir, $p \cdot n$.

Así, ahora es más claro ver que la mediana vendría a ser $Q_{0.5}$, el cuantil de orden 0.5

Ejemplo

Consideremos un experimento en el que lanzamos 50 veces un dado de 4 caras y obtenemos los siguientes resultados.

```
set.seed(260798)
dado = sample(1:4, 50, replace = TRUE)
set.seed(NULL)
length(dado)
```

```
## [1] 50
```

```
dado = sort(dado) ## Los ordenamos de menor a mayor
dado
```

```
## [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 4 4
## [39] 4 4 4 4 4 4 4 4 4 4 4 4 4 4
```

```
df.dado = data.frame(Puntuacion = 1:4,
                      Fr.abs = as.vector(table(dado)),
                      Fr.rel = as.vector(prop.table(table(dado))),
                      Fr.acum = as.vector(cumsum(table(dado))),
                      Fr.Rel.Acum = as.vector(round(cumsum(prop.table(table(dado))), 2)))
df.dado
```

##	Puntuacion	Fr.abs	Fr.rel	Fr.acum	Fr.Rel.Acum
## 1	1	16	0.32	16	0.32
## 2	2	15	0.30	31	0.62
## 3	3	5	0.10	36	0.72
## 4	4	14	0.28	50	1.00

Si nos piden calcular $Q_{0.3}$, sabemos que este es el primer elemento de la lista cuya frecuencia relativa acumulada es mayor o igual a 0.3. Este corresponde a la puntuación 1.

También podríamos hallarlo de otro modo: Fijándonos en la lista ordenada de puntuaciones, el cuantil $Q_{0.3}$ sería el primer elemento de dicha lista tal que fuera mayor o igual que, como mínimo, el 30% de los datos. Si calculamos el 30% de 50, obtenemos que es 15. Esto lo que nos dice es que el cuantil que buscamos es el número que se encuentra en la quinceava posición de la lista ordenada.

```
dato[15]
```

```
## [1] 1
```

Cuantiles

Algunos cuantiles tienen nombre propio:

- Los **Cuartiles** son los cuantiles $Q_{0.25}, Q_{0.5}, Q_{0.75}$. Respectivamente, son los llamados primer, segundo y tercer cuartil. El primer cuartil, $Q_{0.25}$ será el menor valor que es mayor o igual a una cuarta parte de las observaciones y $Q_{0.75}$ el menor valor que es mayor o igual a tres cuartas partes de los datos observados.
- El **cuantil** $Q_{0.5}$ es la mediana.
- Los **Deciles** son los cuantiles Q_p con p un múltiplo de 0.1.
- Los **Percentiles** son los cuantiles Q_p con p un múltiplo de 0.01

La definición de cuantil anteriormente dada es orientativa. La realidad es que exceptuando el caso de la mediana, no hay consenso sobre cómo deben calcularse los cuantiles. En verdad, existen diferentes métodos que pueden dar lugar a soluciones distintas.

Al fin y al cabo, nuestro objetivo no es el de encontrar el primer valor de una muestra cuya frecuencia relativa acumulada en la variable sea mayor o igual a p , sino estimar de esta cantidad para el total de la población.

Para calcular los cuantiles de orden p de una variable cualitativa x con R, se utiliza la instrucción `quantile(x,p)`, la cual dispone de 9 métodos diferentes que se especifican con el parámetro `type`. El valor por defecto es `type=7` y no hace falta especificarlo, como veremos en el siguiente ejemplo.

```
set.seed(0)
```

```
dato2 = sample(1:6,15, replace = TRUE)
dato2
```

```
## [1] 6 1 4 1 2 5 3 6 2 3 3 1 5 5 2
```

```
set.seed(NULL)
```

```
quantile(dato2,0.25) ##primer cuartil
```

```
## 25%  
## 2
```

```
quantile(dado2,0.8)
```

```
## 80%  
## 5
```