

Datos ordinales

Santiago Pérez Moncada

17/6/2020

Los **Datos Ordinales** son parecidos a los datos cualitativos, en el sentido de que son cualidades de los individuos u objetos.

La diferencia existe entre los datos cualitativos y los ordinales reside en las características que expresan. En el caso de los ordinales, éstas tienen un orden natural que permite “acumular” observaciones.

Frecuencias para datos ordinales.

Frecuencia acumuladas

Al trabajar con datos ordinales, el orden de los niveles de los datos nos permite calcular no solo frecuencias absolutas y relativas, sino también **frecuencias acumuladas**.

Es decir, podemos contar cuantas veces hemos observado un dato menor o igual a este.

Ejemplo 1

Suponer que tenemos una muestra de 15 estudiantes de los cuales sabemos su nota en el examen de Estadística. Clasificamos todos estos resultados en Suspenso(S), Aprobado(A), Notable(N) y Excelente(E_x) y consideramos su orden natural $S < A < N < E_x$.

Las notas obtenidas han sido las siguientes.

$S, A, N, E_x, S, S, E_x, E_x, N, A, A, A, A, N, S$

Como recordaras, para saber cuantas hay de cada nivel(frecuencia absoluta), utilizamos la función `table()`. Para crear un factor ordenado debemos usar la función `ordered()`.

Frecuencias Absolutas f_i

```
notas = ordered(c("S", "A", "N", "Ex", "S", "S", "Ex", "Ex", "N", "A", "A", "A", "A", "N", "S", "N", "S"), levels = c("S", "A", "N", "Ex"))

table(notas)
```

```
## notas
##  S  A  N Ex
##  4  5  3  3
```

Frecuencias absolutas acumuladas.

$$N_i = n_1 + n_2 + \dots + n_i$$

En lo referente a **Frecuencias absolutas acumuladas**, hay

- 4 estudiantes con S o menos. Ello implica que la frecuencia acumulada de S es 4.
- 9 estudiantes que han obtenido A o menos. Entonces, la frecuencia acumulada de A es 9.
- 12 estudiantes de los cuales han obtenido N o menos. Así la frecuencia acumulada de N es 12.
- 15 estudiantes (todos) que han obtenido E_x o menos. De este modo, la frecuencia acumulada de E_x es 15, o sea, el total.

Sea.

$$n_1 = f_i(S), \quad n_2 = f_i(A), \quad n_3 = f_i(N), \quad n_4 = f_i(E_x)$$

Nivel	f_i	N_i
S	4	4
A	5	9
N	3	12
E_x	3	15

Frecuencia relativa acumulada F_i

$$F_i = \frac{N_i}{N}$$

Es dividir la frecuencia Absoluta acumulada de cada nivel entre el total de la población. Así, las frecuencias relativas acumuladas respectivas son-

Nivel	f_i	N_i	F_i
S	4	4	$\frac{4}{15}$
A	5	9	$\frac{9}{15}$
N	3	12	$\frac{12}{15}$
E_x	3	15	$\frac{15}{15}$

En general, supongamos que realizamos n observaciones

$$x_1, \dots, x_n$$

de un cierto tipo de datos ordinales, cuyos posibles niveles ordenados son

$$l_1 < l_2 < \dots < l_k$$

por tanto, cada una de las observaciones x_j es igual a algún l_i . Diremos que todas estas observaciones forman una **variable ordinal**. En nuestro ejemplo anterior, los 4 niveles eran.

$$S < A < N < E_x$$

Además, nuestro $n = 15$ y nuestros x_1, \dots, x_{15} son las calificaciones obtenidas por los alumnos.

De este modo, con notaciones

- Las definiciones de frecuencias absolutas n_j y las relativas f_j , para cada nivel l_j son las mismas que en una variable cualitativa.
- La frecuencia absoluta acumulada del nivel l_j en esta variable ordinal es el número N_j de observaciones x_i tales que $x_i \leq l_j$. Es decir,

$$N_j = \sum_{i=1}^j n_i$$

- La frecuencia relativa acumulada del nivel l_j en esta variable ordinal es la fracción en tanto por 1 F_j de observaciones $x_i \leq l_j$. Es decir.

$$F_j = \frac{N_j}{n} = \sum_{i=1}^j \frac{f_i}{n}$$

Ejemplo 2

En un estudio, a un grupo de clientes de un restaurante se les hizo la siguiente pregunta: “¿Estás contento con el trato ofrecido por los trabajadores del establecimiento?” Las posibles respuestas forman una escala ordinal con $1 < 2 < 3 < 4 < 5$.

Supongamos que recogieron las siguientes respuestas de 50 técnicos:

```
set.seed(2018)
clientes = sample(1:5, 50, replace = TRUE)
clientes

## [1] 3 4 5 2 5 1 3 4 2 4 3 3 1 1 5 3 1 3 3 5 1 4 2 5 3 4 5 1 2 2 1 5 5 2 1 2 5 5
## [39] 2 1 2 1 3 2 1 2 3 3 1 2
```

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

En este caso tenemos 5 niveles ($k = 5$) y 50 observaciones ($n = 50$) que forman una variable ordinal a la que hemos de llamar clientes.

```
clientes = ordered(clientes)
Absoluta = table(clientes)
Relativa = prop.table(table(clientes))
Acumulada.Rel = cumsum(Relativa)
Acumulada.Abs = cumsum(Absoluta)
```

```
Absoluta
```

```
## clientes
## 1 2 3 4 5
## 12 12 11 5 10
```

```
Relativa
```

```
## clientes
##      1      2      3      4      5
## 0.24 0.24 0.22 0.10 0.20
```

```
Acumulada.Rel
```

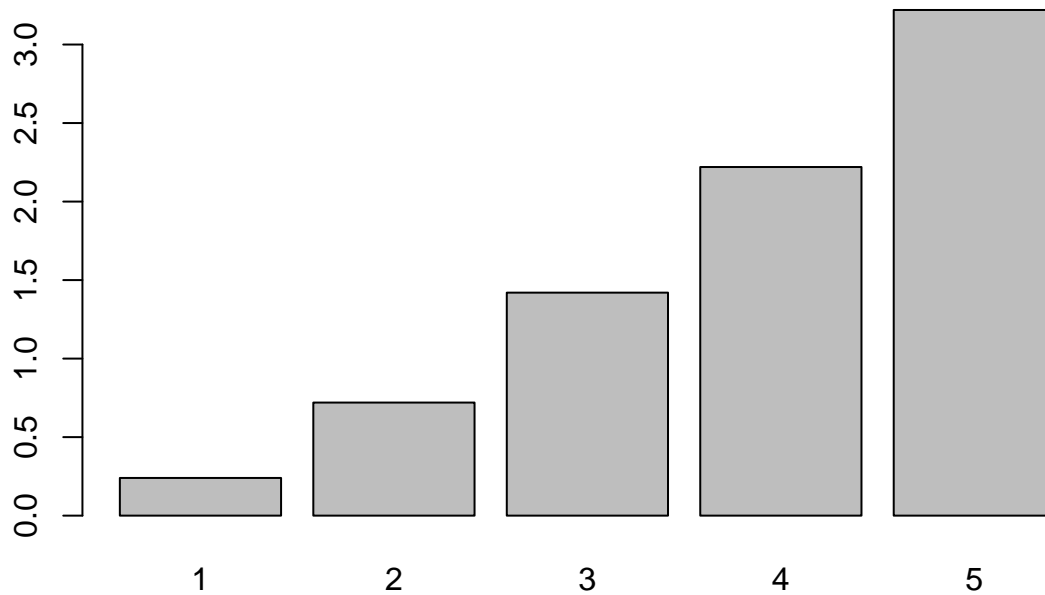
```
##      1      2      3      4      5
## 0.24 0.48 0.70 0.80 1.00
```

```
Acumulada.Abs
```

```
##  1  2  3  4  5
## 12 24 35 40 50
```

```
barplot(cumsum(Acumulada.Rel), main = "Diagrama de barras frecuencias relativas acumuladas")
```

Diagrama de barras frecuencias relativas acumuladas



Ejemplo 3

Se ha evaluado el tamaño de los cuellos de 100 jirafas. Los niveles que se han utilizado se los considera ordenados de la siguiente manera:

Muy corto < Corto < Normal < Largo < Muy largo

Los valores obtenidos en dicho estudio han sido los siguientes:

```
longitud = ordered(sample(c("Muy corto", "Corto", "Normal", "Largo", "Muy largo"),
                           size = 100, replace = TRUE))
```

```
Fr.Abs = table(longitud)
Fr.Abs
```

```
## longitud
##      Corto      Largo Muy corto Muy largo      Normal
##         17         26         13         19         25
```

```
Fr.Rel = prop.table(Fr.Abs)
Fr.Rel
```

```
## longitud
##      Corto      Largo Muy corto Muy largo      Normal
##      0.17      0.26      0.13      0.19      0.25
```

```
Fr.Acum= cumsum(Fr.Abs)
Fr.Acum
```

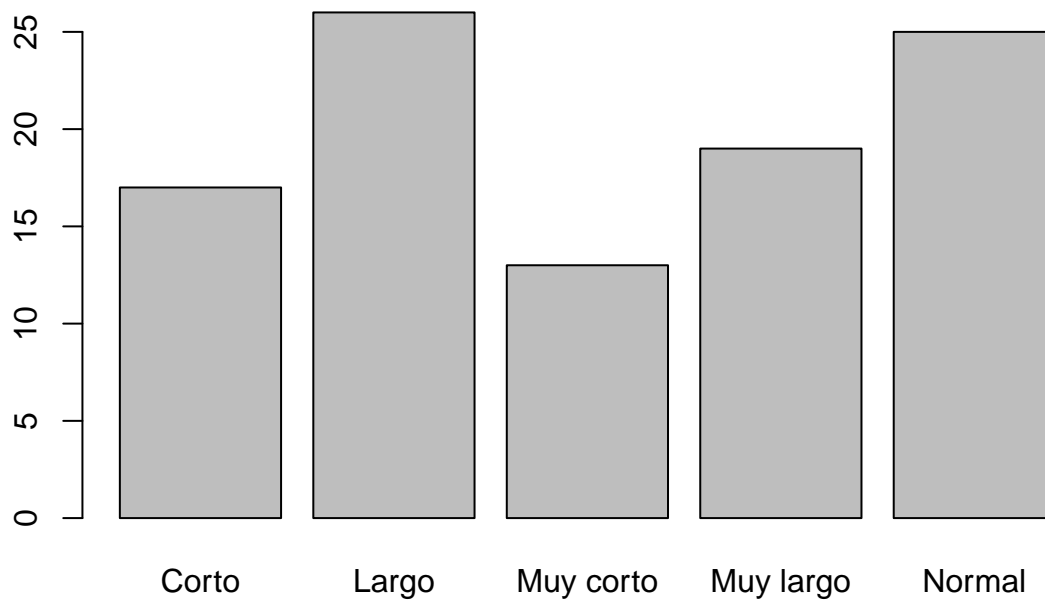
```
##      Corto      Largo Muy corto Muy largo      Normal
##         17         43         56         75         100
```

```
Fr.RAcum = cumsum(Fr.Rel)
Fr.RAcum
```

```
##      Corto      Largo Muy corto Muy largo      Normal
##      0.17      0.43      0.56      0.75      1.00
```

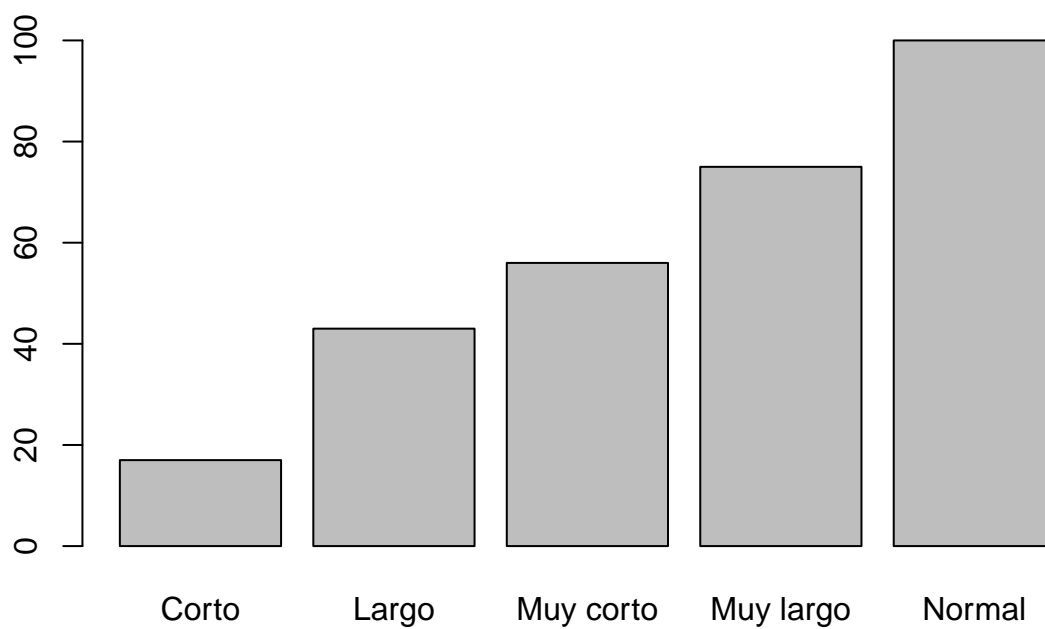
```
barplot(Fr.Abs, main = "Diagrama de Frecuencias Absolutas")
```

Diagrama de Frecuencias Absolutas



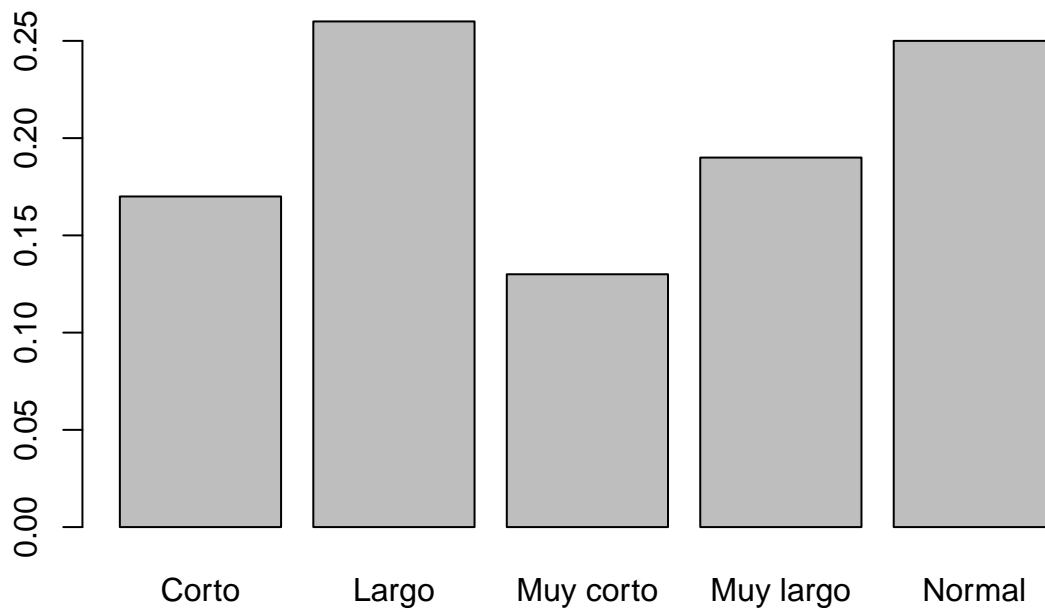
```
barplot(Fr.Acum, main = "Diagrama de Frecuencias Absolutas acumuladas")
```

Diagrama de Frecuencias Absolutas acumuladas



```
barplot(Fr.Rel, main = "Diagrama de Frecuencias Relativas")
```

Diagrama de Frecuencias Relativas



```
barplot(Fr.RAcum, main = "Diagrama de Frecuencias Relativas acumuladas")
```


Diagrama de Frecuencias Relativas acumuladas

