

Tablas de frecuencia

Jocelyn Sánchez Campos

2023-10-17

Existen dos tipos de tablas de frecuencia, **agrupadas** y **no agrupadas** que se determinan según la marca de clase. **Agrupadas:** Empieza a construir los intervalos de clase. **No agrupadas:** Identifica las marcas de clase. Ambas nos ayudan a calcular: -Frecuencia absoluta -Frecuencia absoluta acumulada -Frecuencia relativa -Frecuencia relativa acumulada ## Tablas de distribución de frecuencias## **Datos no agrupados** Permiten identificar la tendencia y variabilidad de los datos **Frecuencia absoluta:** Indica el número de veces que se repite cada una de las observaciones. **Frecuencia absoluta acumulada** Es la suma de las frecuencias absolutas $N_j = n_1 + n_2 + \dots + n_k$ **Frecuencia relativa** Es la fracción de elementos que pertenecen a una marca de clase (X_j) **Frecuencia relativa acumulada** Es la frecuencia absoluta acumulada dividida entre el número de datos totales (n) ## Datos agrupados## **Intervalos de clase** Son el conjunto de datos que integran una clase. Se calcula el rango (R), número de intervalos (K) y la amplitud (A) ## Practica## Se trabajara con la matriz de datos (iris) En donde se encuentran 3 especies de flores del iris (Versicolor, Setosa y Virginica) 1. Exportamos la matriz de datos "Iris"

```
data(iris)
```

2. Exploramos la matriz

```
dim(iris)
```

```
## [1] 150 5
```

3. Nombramos las columnas

```
colnames(iris)
```

```
## [1] "Sepal.Length" "Sepal.Width" "Petal.Length" "Petal.Width" "Species"
```

4. Vemos los tipos de variables

```
str(iris)
```

```
## 'data.frame': 150 obs. of 5 variables:
## $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
## $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
## $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
## $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
## $ Species : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
```

5. Buscamos datos perdidos

```
anyNA(iris)
```

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

Generar tablas## 1.- Convertimos la matriz de datos a un data frame, se agrupan los valores para la variable Petal.Length y se calcula la frecuencia absoluta.

```
tabla_PL <- as.data.frame(table(PL = iris$Petal.Length))
```

2.- Frecuencia absoluta de la variable Petal.Length (PL) Se muestra la tabla de contingencia para la variable PL su respectiva frecuencia absoluta.

tabla_PL

```
##      PL Freq
## 1      1      1
## 2     1.1      1
## 3     1.2      2
## 4     1.3      7
## 5     1.4     13
## 6     1.5     13
## 7     1.6      7
## 8     1.7      4
## 9     1.9      2
## 10      3      1
## 11     3.3      2
## 12     3.5      2
## 13     3.6      1
## 14     3.7      1
## 15     3.8      1
## 16     3.9      3
## 17      4      5
## 18     4.1      3
## 19     4.2      4
## 20     4.3      2
## 21     4.4      4
## 22     4.5      8
## 23     4.6      3
## 24     4.7      5
## 25     4.8      4
## 26     4.9      5
## 27      5      4
## 28     5.1      8
## 29     5.2      2
## 30     5.3      2
## 31     5.4      2
## 32     5.5      3
## 33     5.6      6
## 34     5.7      3
## 35     5.8      3
## 36     5.9      2
## 37      6      2
## 38     6.1      3
## 39     6.3      1
## 40     6.4      1
## 41     6.6      1
## 42     6.7      2
## 43     6.9      1
```

3.- Se contruye la tabla de frecuencias completas redondeando las frecuencias absolutas a 3 decimales

```
transform(tabla_PL,
          freqAc=cumsum(Freq),
          Rel= round(prop.table(Freq),3),
          RelAc=round(cumsum(prop.table(Freq)),3))
```

##	PL	Freq	freqAc	Rel	RelAc
## 1	1	1	1	0.007	0.007
## 2	1.1	1	2	0.007	0.013
## 3	1.2	2	4	0.013	0.027
## 4	1.3	7	11	0.047	0.073
## 5	1.4	13	24	0.087	0.160
## 6	1.5	13	37	0.087	0.247
## 7	1.6	7	44	0.047	0.293
## 8	1.7	4	48	0.027	0.320
## 9	1.9	2	50	0.013	0.333
## 10	3	1	51	0.007	0.340
## 11	3.3	2	53	0.013	0.353
## 12	3.5	2	55	0.013	0.367
## 13	3.6	1	56	0.007	0.373
## 14	3.7	1	57	0.007	0.380
## 15	3.8	1	58	0.007	0.387
## 16	3.9	3	61	0.020	0.407
## 17	4	5	66	0.033	0.440
## 18	4.1	3	69	0.020	0.460
## 19	4.2	4	73	0.027	0.487
## 20	4.3	2	75	0.013	0.500
## 21	4.4	4	79	0.027	0.527
## 22	4.5	8	87	0.053	0.580
## 23	4.6	3	90	0.020	0.600
## 24	4.7	5	95	0.033	0.633
## 25	4.8	4	99	0.027	0.660
## 26	4.9	5	104	0.033	0.693
## 27	5	4	108	0.027	0.720
## 28	5.1	8	116	0.053	0.773
## 29	5.2	2	118	0.013	0.787
## 30	5.3	2	120	0.013	0.800
## 31	5.4	2	122	0.013	0.813
## 32	5.5	3	125	0.020	0.833
## 33	5.6	6	131	0.040	0.873
## 34	5.7	3	134	0.020	0.893
## 35	5.8	3	137	0.020	0.913
## 36	5.9	2	139	0.013	0.927
## 37	6	2	141	0.013	0.940
## 38	6.1	3	144	0.020	0.960
## 39	6.3	1	145	0.007	0.967
## 40	6.4	1	146	0.007	0.973
## 41	6.6	1	147	0.007	0.980
## 42	6.7	2	149	0.013	0.993
## 43	6.9	1	150	0.007	1.000

4.- Agruparemos las variables en 8 clases y se calcula la frecuencia absoluta

```
tabla_clases<-as.data.frame(table (Petal.length = factor
                                   (cut(iris$Petal.Length,
                                         breaks = 8))))
```

5.- Visualizamos la tabla de clases

```
tabla_clases
```

```
##   Petal.length Freq
## 1 (0.994,1.74]   48
## 2 (1.74,2.48]    2
## 3 (2.48,3.21]    1
## 4 (3.21,3.95]   10
## 5 (3.95,4.69]   29
## 6 (4.69,5.43]   32
## 7 (5.43,6.16]   22
## 8 (6.16,6.91]    6
```

6.- Contrucción de la tabla de frecuencias completa redondeando las frecuencias relativas a 3 decimales

```
tabla<-transform(tabla_clases,
                  freqAc=cumsum(Freq),
                  Rel= round(prop.table(Freq),3),
                  RelAc=round(cumsum(prop.table(Freq)),3))
```

Para la tabla 7.- Organización visual de la tabla (variable Petal.length) 7.1.- Instalamos la librería knitr

```
install.packages("knitr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

7.2.- Se abre la librería

```
library(knitr)
```

7.3.- Se visualiza la tabla

```
kable(tabla)
```

Petal.length	Freq	freqAc	Rel	RelAc
(0.994,1.74]	48	48	0.320	0.320
(1.74,2.48]	2	50	0.013	0.333
(2.48,3.21]	1	51	0.007	0.340
(3.21,3.95]	10	61	0.067	0.407
(3.95,4.69]	29	90	0.193	0.600
(4.69,5.43]	32	122	0.213	0.813
(5.43,6.16]	22	144	0.147	0.960
(6.16,6.91]	6	150	0.040	1.000