

Tablas de frecuencia

Kathya Valdez Gómez

2022-05-16

a) *Exploración de la matriz Iris*

1.- Importación de la matriz

```
data(iris)
```

2.- Exploración de la matriz

```
dim(iris)
```

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

Dimensión de la matriz: Iris tiene 150 individuos (observaciones) y 5 variables

3.- Nombre de las columnas (variables)

```
colnames(iris)
```

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

4.- Tipo 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.- Visualización de una variable específica

```
iris$Petal.Length
```

```
## [1] 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 1.5 1.6 1.4 1.1 1.2 1.5 1.3 1.4
## [19] 1.7 1.5 1.7 1.5 1.0 1.7 1.9 1.6 1.6 1.5 1.4 1.6 1.6 1.5 1.5 1.4 1.5 1.2
## [37] 1.3 1.4 1.3 1.5 1.3 1.3 1.3 1.6 1.9 1.4 1.6 1.4 1.5 1.4 4.7 4.5 4.9 4.0
## [55] 4.6 4.5 4.7 3.3 4.6 3.9 3.5 4.2 4.0 4.7 3.6 4.4 4.5 4.1 4.5 3.9 4.8 4.0
```

```
## [73] 4.9 4.7 4.3 4.4 4.8 5.0 4.5 3.5 3.8 3.7 3.9 5.1 4.5 4.5 4.7 4.4 4.1 4.0
## [91] 4.4 4.6 4.0 3.3 4.2 4.2 4.2 4.3 3.0 4.1 6.0 5.1 5.9 5.6 5.8 6.6 4.5 6.3
## [109] 5.8 6.1 5.1 5.3 5.5 5.0 5.1 5.3 5.5 6.7 6.9 5.0 5.7 4.9 6.7 4.9 5.7 6.0
## [127] 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1 5.6 6.1 5.6 5.5 4.8 5.4 5.6 5.1 5.1 5.9
## [145] 5.7 5.2 5.0 5.2 5.4 5.1
```

6.- En busca de datos perdidos (NA)

```
anyNA(iris)
```

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

b) *Generación de tablas*

1.- Convertimos la matriz de datos en un data frame.

Se agrupan los valores para la variable **Petal.Length**. Indico que el nombre me lo abrevie en **PL**. Lo que resulte quiero que me lo ponga en formato tabla. Lo que resulte quiero que adquiera formato data.frame. A partir de eso, se genera una nueva variable (objeto) llamada **tabla_PL**

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

2.- Frecuencia absoluta de la variable Petal.Length (PL)

Se visualiza la tabla de contingencia para la variable PL con su respectiva frecuencia absoluta (no. de observaciones de la variable PL).

```
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

```
tabla_PL1<-transform(tabla_PL,
                      freqAc=cumsum(Freq),
                      Rel= round(prop.table(Freq),3),
                      RelAc=round(cumsum(prop.table(Freq)),3))
```

3.1- Visualizar la variable tabla_PL1

```
tabla_PL1
```

```
##      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
```

3.2- Formato tabla

Abrir librería knitr

```
library(knitr)
```

Formato tabla

```
kable(tabla_PL1)
```

PL	Freq	freqAc	Rel	RelAc
1	1	1	0.007	0.007
1.1	1	2	0.007	0.013
1.2	2	4	0.013	0.027
1.3	7	11	0.047	0.073
1.4	13	24	0.087	0.160
1.5	13	37	0.087	0.247
1.6	7	44	0.047	0.293
1.7	4	48	0.027	0.320
1.9	2	50	0.013	0.333
3	1	51	0.007	0.340
3.3	2	53	0.013	0.353
3.5	2	55	0.013	0.367
3.6	1	56	0.007	0.373
3.7	1	57	0.007	0.380
3.8	1	58	0.007	0.387
3.9	3	61	0.020	0.407

PL	Freq	freqAc	Rel	RelAc
4	5	66	0.033	0.440
4.1	3	69	0.020	0.460
4.2	4	73	0.027	0.487
4.3	2	75	0.013	0.500
4.4	4	79	0.027	0.527
4.5	8	87	0.053	0.580
4.6	3	90	0.020	0.600
4.7	5	95	0.033	0.633
4.8	4	99	0.027	0.660
4.9	5	104	0.033	0.693
5	4	108	0.027	0.720
5.1	8	116	0.053	0.773
5.2	2	118	0.013	0.787
5.3	2	120	0.013	0.800
5.4	2	122	0.013	0.813
5.5	3	125	0.020	0.833
5.6	6	131	0.040	0.873
5.7	3	134	0.020	0.893
5.8	3	137	0.020	0.913
5.9	2	139	0.013	0.927
6	2	141	0.013	0.940
6.1	3	144	0.020	0.960
6.3	1	145	0.007	0.967
6.4	1	146	0.007	0.973
6.6	1	147	0.007	0.980
6.7	2	149	0.013	0.993
6.9	1	150	0.007	1.000

c) *Generación de tablas agrupadas*

1.- Agrupación de la variable PL.

Agruparemos las variables en 8 clases y se calculará la frecuencia absoluta

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

2.- Visualización de 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
```

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

```
tabla_PL2<-transform(tabla_clases,  
                      freqAc=cumsum(Freq),  
                      Rel= round(prop.table(Freq),3),  
                      RelAc=round(cumsum(prop.table(Freq)),3))
```

4.-Visualización de la tabla (Organización visual de la tabla (variable Petal.length))

```
tabla_PL2
```

```
##   Petal.length Freq freqAc   Rel RelAc  
## 1 (0.994,1.74]   48    48 0.320 0.320  
## 2 (1.74,2.48]    2    50 0.013 0.333  
## 3 (2.48,3.21]    1    51 0.007 0.340  
## 4 (3.21,3.95]   10    61 0.067 0.407  
## 5 (3.95,4.69]   29    90 0.193 0.600  
## 6 (4.69,5.43]   32   122 0.213 0.813  
## 7 (5.43,6.16]   22   144 0.147 0.960  
## 8 (6.16,6.91]    6   150 0.040 1.000
```

4.1.-Formato tabla

```
kable(tabla_PL2)
```

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