

Estadística Descriptiva con Datos Cualitativos

Ramon Ceballos

25/1/2021

ESTADÍSTICA MULTIDIMENSIONAL. TABLAS DE FRECUENCIAS MULTIDIMENSIONALES

1. Datos multidimensionales

En general, se pueden calcular tablas de frecuencia de cualquier n° de variables. El manejo de estas tablas multidimensionales es similar al manejo de tablas bidimensionales. Hemos de tener en cuenta que cuantas más variables haya en una tabla de frecuencias, mayor será el n° de frecuencias marginales.

2. Análisis estadístico de tres variables. Tres dimensiones

Ejemplo con tres dimensiones*

Se definen tres variables cualitativas distintas en las que se recogen las respuestas, el género de las personas que dan las respuestas y los lugares en los que se dió dicha respuesta.

```
#PRIMERA VARIABLE
#Respuestas con valores "sí" y "no"
ans = sample(c("Si", "No"), size = 100, replace = TRUE)

#SEGUNDA VARIABLE
#Género de las personas que contestan; "H" y "M" (hombre y mujer)
sex = sample(c("H", "M"), size = 100, replace = TRUE)

#TERCERA VARIABLE
#Lugar donde se realiza la respuesta
place = sample(c("San Francisco", "Barcelona",
                 "Valencia", "Cobiya", "Asturias"), size = 100, replace = TRUE)
```

Frecuencias absolutas

Se realiza con la función **table()** una tabla de **frecuencias absolutas** para el ordenamiento siguiente: 1-género; 2-respuestas; 3-lugar.

Se obtienen cinco tablas de frecuencias bidimensionales, una para cada lugar. Por tanto, se obtiene una lista de tablas bidimensionales, separando la población según los niveles de la tercera variable (en este caso, lugar).

#FRECUENCIAS ABSOLUTAS

```
table(sex, ans, place)
```

```
## , , place = Asturias
##
##   ans
## sex No Si
##   H  7  9
##   M  7  5
##
## , , place = Barcelona
##
##   ans
## sex No Si
##   H  6  6
##   M  6  4
##
## , , place = Cobija
##
##   ans
## sex No Si
##   H  5  4
##   M  2  4
##
## , , place = San Francisco
##
##   ans
## sex No Si
##   H  6  7
##   M  3  1
##
## , , place = Valencia
##
##   ans
## sex No Si
##   H  7  3
##   M  4  4
```

Se puede obtener esta misma tabla de frecuencias en formato plano con la función **ftable()** que viene de **flat** table. Este formato junta la población sin separarlo en tablas bidimensionales, siendo más difícil de leer.

En esta función **ftable()** se puede especificar que variable queremos que aparezca como fila y que variable como columna. Para ello aparezcan en las columnas las variables género y respuesta asignamos `col.vars = c("sex", "ans")`.

```
ftable(sex, ans, place)
```

```
##           place Asturias Barcelona Cobija San Francisco Valencia
## sex ans
## H   No           7           6           5           6           7
##     Si           9           6           4           7           3
## M   No           7           6           2           3           4
##     Si           5           4           4           1           4
```

```
fable(sex, ans, place, col.vars = c("sex", "ans"))
```

```
##           sex  H      M
##           ans No Si No Si
## place
## Asturias      7  9  7  5
## Barcelona      6  6  6  4
## Cobiya         5  4  2  4
## San Francisco  6  7  3  1
## Valencia      7  3  4  4
```

Filtrar las variables de una tabla de frecuencias tridimensional

Se deben poner entre corchetes una condición para cada una de las variables.

```
#para mujeres que responden si en san francisco
table(sex, ans, place)["M", "Si", "San Francisco"]
```

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

```
#Para todos los géneros que responden si en valencia
table(sex, ans, place)[ , "Si", "Valencia"]
```

```
## H M
## 3 4
```

```
#para toda la gente que dice que no en cualquier lugar
table(sex, ans, place)[ , "No", ]
```

```
##      place
## sex Asturias Barcelona Cobiya San Francisco Valencia
##  H          7          6          5          6          7
##  M          7          6          2          3          4
```

```
#para las mujeres que respondieron en cobija con independencia de la respuesta
table(sex, ans, place)["M", , "Cobiya"]
```

```
## No Si
##  2  4
```

Frecuencias relativas

Se podría hacer la frecuencia relativa marginal de las respuestas de la combinación de género y lugar.

Se utiliza **prop.table()** respecto a **table()** que da la tabla de frecuencias absolutas, y se iguala el **margin** a la dimensión o combinación de dimensiones respecto a la que se quiere calcular la **frecuencia relativa marginal**. Si no se especifica el **margin** se obtendrán las **frecuencias relativas globales**.

```
prop.table(table(sex, ans, place))#Frec. Rel. Globales
```

```
## , , place = Asturias
##
##      ans
## sex    No    Si
##  H 0.07 0.09
##  M 0.07 0.05
##
## , , place = Barcelona
##
##      ans
## sex    No    Si
##  H 0.06 0.06
##  M 0.06 0.04
##
## , , place = Cobija
##
##      ans
## sex    No    Si
##  H 0.05 0.04
##  M 0.02 0.04
##
## , , place = San Francisco
##
##      ans
## sex    No    Si
##  H 0.06 0.07
##  M 0.03 0.01
##
## , , place = Valencia
##
##      ans
## sex    No    Si
##  H 0.07 0.03
##  M 0.04 0.04
```

#Da la frecuencia relativa de cada lugar

```
prop.table(table(sex, ans, place), margin = 3) # Frec. Rel. Marginal por Lugar/ciudad
```

```
## , , place = Asturias
##
##      ans
## sex          No          Si
##  H 0.25000000 0.32142857
##  M 0.25000000 0.17857143
##
## , , place = Barcelona
##
##      ans
## sex          No          Si
##  H 0.27272727 0.27272727
```

```
## M 0.27272727 0.18181818
##
## , , place = Cobija
##
## ans
## sex      No      Si
## H 0.33333333 0.26666667
## M 0.13333333 0.26666667
##
## , , place = San Francisco
##
## ans
## sex      No      Si
## H 0.35294118 0.41176471
## M 0.17647059 0.05882353
##
## , , place = Valencia
##
## ans
## sex      No      Si
## H 0.38888889 0.16666667
## M 0.22222222 0.22222222
```

```
#Da el % de mujeres dentro de mujeres y por otro lado el % hombres dentro
#de hombres que respondieron si o no en el lugar concreto
prop.table(table(sex, ans, place), margin = c(1, 3)) # Frec. Rel. Marg. por Sexo/genero y País
```

```
## , , place = Asturias
##
## ans
## sex      No      Si
## H 0.4375000 0.5625000
## M 0.5833333 0.4166667
##
## , , place = Barcelona
##
## ans
## sex      No      Si
## H 0.5000000 0.5000000
## M 0.6000000 0.4000000
##
## , , place = Cobija
##
## ans
## sex      No      Si
## H 0.5555556 0.4444444
## M 0.3333333 0.6666667
##
## , , place = San Francisco
##
## ans
## sex      No      Si
## H 0.4615385 0.5384615
## M 0.7500000 0.2500000
```

```
##
## , , place = Valencia
##
##   ans
## sex      No      Si
##  H 0.7000000 0.3000000
##  M 0.5000000 0.5000000
```

Se puede aplicar la función **ftable()** para una tabla de frecuencias relativas. Queda toda la información resumida dentro de una tabla única.

```
fable(prop.table(table(sex, ans, place)))
```

```
##           place Asturias Barcelona Cobija San Francisco Valencia
## sex ans
## H  No           0.07          0.06  0.05              0.06       0.07
##   Si           0.09          0.06  0.04              0.07       0.03
## M  No           0.07          0.06  0.02              0.03       0.04
##   Si           0.05          0.04  0.04              0.01       0.04
```

A medida que la dimensión crece, es más difícil de describir los datos de manera coherente.