

Tablas de contingencia

Diego Delgado Palomares

17/9/2020

Tablas de Contingencia

```
datos <- factor(sample(c("H","M"), size = 8, replace = T))
table(datos)
```

```
## datos
## H M
## 4 4
```

```
table(datos)[ "M" ]
```

```
## M
## 4
```

```
sum(table(datos))
```

```
## [1] 8
```

Tablas de frecuencias unidimensionales

```
x <- sample(1:5, size = 12, replace = T)
respuestas <- sample(c("sí","no"), size = 15, replace = T)
prop.table(x)
```

```
## [1] 0.04347826 0.10869565 0.08695652 0.08695652 0.10869565 0.10869565
## [7] 0.10869565 0.04347826 0.02173913 0.10869565 0.06521739 0.10869565
```

```
prop.table(table(x))
```

```
## x
##      1      2      3      4      5
## 0.08333333 0.16666667 0.08333333 0.16666667 0.50000000
```

```
prop.table(table(respuestas))
```

```
## respuestas
##      no      sí
## 0.6666667 0.3333333
```

```
names(which(table(x)==max(table(x))))
```

```
## [1] "5"
```

```
datos <- c(rep("H",6),rep("M",14))
table(datos)
```

```
## datos
## H M
## 6 14
```

```
prop.table(table(datos))
```

```
## datos  
##   H   M  
## 0.3 0.7
```

```
names(which(table(datos)==max(table(datos))))
```

```
## [1] "M"
```

Frecuencias Relativas

$$f_i = \frac{n_i}{n}$$

```
prop.table(table(datos))
```

```
## datos  
##   H   M  
## 0.3 0.7
```

```
table(datos)/length(datos)
```

```
## datos  
##   H   M  
## 0.3 0.7
```

```
moda= function(d){  
  c(names(which(table(d)==max(table(d))))," y aparece ",sum(max(table(d))),"veces")  
}  
m_d <- moda(datos)
```

la moda del data frame es M, y aparece , 14, veces

Tablas de frecuencias Bidimensionales

```
sexo <- sample(c("H","M"), size = length(respuestas), replace = T)  
sexo
```

```
## [1] "H" "H" "M" "M" "M" "H" "M" "M" "M" "M" "M" "M" "H" "M" "M"
```

```
table(respuestas, sexo)
```

```
##           sexo  
## respuestas H M  
##           no 3 7  
##           sí 1 4
```

Tabla de frecuencia relativa global

```
prop.table(table(respuestas,sexo))
```

```
##           sexo  
## respuestas      H      M  
##           no 0.2000000 0.4666667  
##           sí 0.0666667 0.2666667
```

Tabla de frecuencias marginales

Con `margin = 1` obtenemos la tabla de frecuencias marginales por fila

```
prop.table(table(respuestas,sexo), margin = 1)
```

```
##           sexo
## respuestas  H    M
##           no 0.3 0.7
##           sí 0.2 0.8
```

y con `margin = 2`, por columnas

```
prop.table(table(respuestas,sexo), margin = 2)
```

```
##           sexo
## respuestas      H      M
##           no 0.7500000 0.6363636
##           sí 0.2500000 0.3636364
```

```
library(gmodels)
sexo = factor(c(sample(c("H","M"),size = 12,replace = T)))
ans <- factor(c(sample(c(0,1),size = length(sexo),replace = T)))
ans<- factor(ans, levels = c(0,1), labels = c("si","no"))
ans
```

```
## [1] no si si no si si si no si si no no
## Levels: si no
```

```
CrossTable(sexo, ans, prop.chisq = F)
```

```
##
##
##   Cell Contents
## |-----|
## |                      N |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table:  12
##
##
##           | ans
##           | si | no | Row Total |
## -----|-----|-----|-----|
##           H |    4 |    3 |          7 |
##           | 0.571 | 0.429 |    0.583 |
##           | 0.571 | 0.600 |          |
##           | 0.333 | 0.250 |          |
## -----|-----|-----|-----|
##           M |    3 |    2 |          5 |
##           | 0.600 | 0.400 |    0.417 |
##           | 0.429 | 0.400 |          |
##           | 0.250 | 0.167 |          |
## -----|-----|-----|-----|
```

```
## Column Total |          7 |          5 |          12 |
##              |        0.583 |        0.417 |              |
## -----|-----|-----|-----|
##
##
```

```
tt <- table(sexo, ans)
pt <- prop.table(table(sexo,ans))
tt
```

```
##      ans
## sexo si no
##   H  4  3
##   M  3  2
```

```
pt
```

```
##      ans
## sexo      si      no
##   H 0.3333333 0.2500000
##   M 0.2500000 0.1666667
```

```
rowSums(tt)
```

```
## H M
## 7 5
```

```
colSums(tt)
```

```
## si no
## 7 5
```

```
rowSums(pt)
```

```
##      H      M
## 0.5833333 0.4166667
```

```
colSums(pt)
```

```
##      si      no
## 0.5833333 0.4166667
```

```
prop.table(tt, margin = 1) # frecuencia relativa por sexo
```

```
##      ans
## sexo      si      no
##   H 0.5714286 0.4285714
##   M 0.6000000 0.4000000
```

```
prop.table(tt, margin = 2) # frecuencia relativa por respuesta
```

```
##      ans
## sexo      si      no
##   H 0.5714286 0.6000000
##   M 0.4285714 0.4000000
```

```
apply(tt, FUN = sum, MARGIN = 1)
```

```
## H M
## 7 5
```

```
apply(tt, FUN = sum, MARGIN = 2)
```

```
## si no  
## 7 5
```

Multivariante

Datos multidimensionales

```
ans1 <- sample(c(0,1), size = 100, replace = T)  
ans <- factor(ans1, levels = c(0,1), labels = c("si","no"))  
sex <- sample(c("hombre", "mujer"), size = length(ans), replace = T)  
places <- sample(c("CDMX","ESTADO","TLAXCALA","PUEBLA","VERACRUZ"), size = length(ans),replace = T)  
table(sex,ans,places)
```

```
## , , places = CDMX  
##  
##      ans  
## sex    si no  
## hombre 6  1  
## mujer  8  9  
##  
## , , places = ESTADO  
##  
##      ans  
## sex    si no  
## hombre 5  9  
## mujer  4  7  
##  
## , , places = PUEBLA  
##  
##      ans  
## sex    si no  
## hombre 6  1  
## mujer  1  2  
##  
## , , places = TLAXCALA  
##  
##      ans  
## sex    si no  
## hombre 4  6  
## mujer  2  4  
##  
## , , places = VERACRUZ  
##  
##      ans  
## sex    si no  
## hombre 8  7  
## mujer  6  4
```

```
table(ans,sex,places)
```

```
## , , places = CDMX  
##
```

```
##      sex
## ans  hombre mujer
##   si      6      8
##   no      1      9
##
## , , places = ESTADO
##
##      sex
## ans  hombre mujer
##   si      5      4
##   no      9      7
##
## , , places = PUEBLA
##
##      sex
## ans  hombre mujer
##   si      6      1
##   no      1      2
##
## , , places = TLAXCALA
##
##      sex
## ans  hombre mujer
##   si      4      2
##   no      6      4
##
## , , places = VERACRUZ
##
##      sex
## ans  hombre mujer
##   si      8      6
##   no      7      4
```

```
ftable(sex,ans,places)
```

```
##           places CDMX ESTADO PUEBLA TLAXCALA VERACRUZ
## sex      ans
## hombre si           6       5       6       4       8
##         no           1       9       1       6       7
## mujer  si           8       4       1       2       6
##         no           9       7       2       4       4
```

```
cumsum(prop.table(ftable(sex,ans,places)))
```

```
## [1] 0.06 0.07 0.15 0.24 0.29 0.38 0.42 0.49 0.55 0.56 0.57 0.59 0.63 0.69 0.71
## [16] 0.75 0.83 0.90 0.96 1.00
```

Filtrar tablas

```
ttt<-(table(sex,ans,places))
```

Frecuencias relativas

```
prop.table(ttt) # frecuencia relativa global
```

```
## , , places = CDMX
##
##      ans
## sex      si    no
##  hombre 0.06 0.01
##  mujer  0.08 0.09
##
## , , places = ESTADO
##
##      ans
## sex      si    no
##  hombre 0.05 0.09
##  mujer  0.04 0.07
##
## , , places = PUEBLA
##
##      ans
## sex      si    no
##  hombre 0.06 0.01
##  mujer  0.01 0.02
##
## , , places = TLAXCALA
##
##      ans
## sex      si    no
##  hombre 0.04 0.06
##  mujer  0.02 0.04
##
## , , places = VERACRUZ
##
##      ans
## sex      si    no
##  hombre 0.08 0.07
##  mujer  0.06 0.04
```

```
prop.table(ttt, margin = 1) #frecuencia relativa marginal respectp a sexo
```

```
## , , places = CDMX
##
##      ans
## sex      si      no
##  hombre 0.11320755 0.01886792
##  mujer  0.17021277 0.19148936
##
## , , places = ESTADO
##
##      ans
## sex      si      no
##  hombre 0.09433962 0.16981132
##  mujer  0.08510638 0.14893617
##
## , , places = PUEBLA
##
##      ans
## sex      si      no
```

```
## hombre 0.11320755 0.01886792
## mujer 0.02127660 0.04255319
##
## , , places = TLAXCALA
##
##      ans
## sex      si      no
## hombre 0.07547170 0.11320755
## mujer 0.04255319 0.08510638
##
## , , places = VERACRUZ
##
##      ans
## sex      si      no
## hombre 0.15094340 0.13207547
## mujer 0.12765957 0.08510638
```

```
prop.table(ttt, margin = 2) #frecuencia relativa marginal respecto a respuesta
```

```
## , , places = CDMX
##
##      ans
## sex      si      no
## hombre 0.12 0.02
## mujer 0.16 0.18
##
## , , places = ESTADO
##
##      ans
## sex      si      no
## hombre 0.10 0.18
## mujer 0.08 0.14
##
## , , places = PUEBLA
##
##      ans
## sex      si      no
## hombre 0.12 0.02
## mujer 0.02 0.04
##
## , , places = TLAXCALA
##
##      ans
## sex      si      no
## hombre 0.08 0.12
## mujer 0.04 0.08
##
## , , places = VERACRUZ
##
##      ans
## sex      si      no
## hombre 0.16 0.14
## mujer 0.12 0.08
```



```
prop.table(ttt, margin = 3) #frecuencia relativa marginal respecto a estado
```

```
## , , places = CDMX
##
##      ans
## sex      si      no
## hombre 0.25000000 0.04166667
## mujer  0.33333333 0.37500000
##
## , , places = ESTADO
##
##      ans
## sex      si      no
## hombre 0.20000000 0.36000000
## mujer  0.16000000 0.28000000
##
## , , places = PUEBLA
##
##      ans
## sex      si      no
## hombre 0.60000000 0.10000000
## mujer  0.10000000 0.20000000
##
## , , places = TLAXCALA
##
##      ans
## sex      si      no
## hombre 0.25000000 0.37500000
## mujer  0.12500000 0.25000000
##
## , , places = VERACRUZ
##
##      ans
## sex      si      no
## hombre 0.32000000 0.28000000
## mujer  0.24000000 0.16000000
```

```
prop.table(ttt, margin = c(1,2)) # frecuencia relativa marginal por sexo y respuestaa
```

```
## , , places = CDMX
##
##      ans
## sex      si      no
## hombre 0.20689655 0.04166667
## mujer  0.38095238 0.34615385
##
## , , places = ESTADO
##
##      ans
## sex      si      no
## hombre 0.17241379 0.37500000
## mujer  0.19047619 0.26923077
##
## , , places = PUEBLA
```

```

##
##      ans
## sex      si      no
##  hombre 0.20689655 0.04166667
##  mujer  0.04761905 0.07692308
##
## , , places = TLAXCALA
##
##      ans
## sex      si      no
##  hombre 0.13793103 0.25000000
##  mujer  0.09523810 0.15384615
##
## , , places = VERACRUZ
##
##      ans
## sex      si      no
##  hombre 0.27586207 0.29166667
##  mujer  0.28571429 0.15384615

```