

# Descripción

*Carlos Javier Rincón*

*versión:2019*

## Contents

Introducción	1
VARIABLES CUANTITATIVAS	1
VARIABLES CATEGÓRICAS	3

## Introducción

En esta guía revisaremos como obtener estadísticas descriptivas de una base de datos. Seguiremos trabajando con la base de recién nacidos vista en la guía 1.

Cargamos la base de datos:

```
library(rio)
```

```
## Warning: package 'rio' was built under R version 3.6.1
```

```
RN<-import("RN.xlsx")
```

## VARIABLES CUANTITATIVAS

Para variables cuantitativas comenzamos con:

- mínimo, máximo y rango

```
min(RN$peso,na.rm=TRUE)
```

```
## [1] 0.22
```

```
max(RN$peso,na.rm=TRUE)
```

```
## [1] 14.85
```

```
range(RN$peso, na.rm = TRUE)
```

```
## [1] 0.22 14.85
```

```
max(RN$peso,na.rm=TRUE) - min(RN$peso,na.rm=TRUE)
```

```
## [1] 14.63
```

- medidas de tendencia central:

```
mean(RN$peso,na.rm = TRUE)
```

```
## [1] 6.438997
```

```
median(RN$peso,na.rm=TRUE)
```

```
## [1] 6.12
```

```
#install.packages("modeest")
```

```
#library(modeest)
```

```
#mfv(RN$peso)
```

---

Ejercicio 1: Calcular para la variable **talla** y **controles** las estadísticas: mínimo, máximo, rango, media, mediana, moda.

---

- percentiles:

```
quantile(RN$peso,0.25,na.rm = TRUE)
```

```
## 25%
```

```
## 5.48
```

```
quantile(RN$peso,0.75,na.rm = TRUE)
```

```
## 75%
```

```
## 8.35
```

```
quantile(RN$peso,c(0.25,0.75),na.rm = TRUE)
```

```
## 25% 75%
```

```
## 5.48 8.35
```

```
fivenum(RN$peso,na.rm=TRUE)
```

```
## [1] 0.22 5.48 6.12 8.35 14.85
```

```
library(Hmisc)
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Warning: package 'survival' was built under R version 3.6.1
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
```

```
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## format.pval, units
```

```
describe(RN$peso)
```

```
## RN$peso
```

```
##      n missing distinct      Info      Mean      Gmd      .05      .10
```

```
##    369      0      288         1     6.439     3.028     1.320     2.418
```

```
##     .25     .50     .75     .90     .95
```

```
##    5.480     6.120     8.350     9.824    10.580
```

```
##
## lowest : 0.22 0.25 0.27 0.45 0.52, highest: 12.50 13.00 13.03 13.10 14.85
```

- medidas de dispersión

```
sd(RN$peso,na.rm=TRUE)
```

```
## [1] 2.712548
```

```
var(RN$peso,na.rm=TRUE)
```

```
## [1] 7.357917
```

```
(sd(RN$peso,na.rm=TRUE)/mean(RN$peso,na.rm = TRUE))*100
```

```
## [1] 42.12687
```

- estadísticas de forma - coeficiente de asimetría y curtosis:

```
#install.packages("moments")
```

```
library(moments)
```

```
skewness(RN$peso,na.rm=TRUE)
```

```
## [1] -0.1584826
```

```
kurtosis(RN$peso,na.rm=TRUE)
```

```
## [1] 2.969169
```

---

Ejercicio 2: Calcular para las variables **talla** y **controles** las estadísticas: Percentiles 25, 50 y 75, la desviación estándar y varianza, coeficiente de variación, coeficiente de asimetría y curtosis.

---

- estadísticas por grupos

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.1
```

```
summarise(RN,media=mean(peso,na.rm = TRUE),de=sd(peso,na.rm = TRUE),n=n())
```

```
##      media      de      n
```

```
## 1 6.438997 2.712548 369
```

## Variables categóricas

- Frecuencias absolutas y relativas

Una variable - sexo

```
table(RN$sexo)
```

```
##
```

```
##  H  M
```

```
## 156 213
```

```
prop.table(table(RN$sexo))
```

```
##
##           H           M
## 0.4227642 0.5772358
```

```
round(prop.table(table(RN$sexo))*100,2)
```

```
##
##           H           M
## 42.28 57.72
```

Dos variables - estrato y sexo

```
table(RN$estrato,RN$sexo)
```

```
##
##           H M
## E1 21 29
## E2 32 39
## E3 52 54
## E4 23 48
## E5 19 27
## E6  9 16
```

```
round(prop.table(table(RN$estrato,RN$sexo)),2)
```

```
##
##           H M
## E1 0.06 0.08
## E2 0.09 0.11
## E3 0.14 0.15
## E4 0.06 0.13
## E5 0.05 0.07
## E6 0.02 0.04
```

```
round(prop.table(table(RN$estrato,RN$sexo),1),2)
```

```
##
##           H M
## E1 0.42 0.58
## E2 0.45 0.55
## E3 0.49 0.51
## E4 0.32 0.68
## E5 0.41 0.59
## E6 0.36 0.64
```

```
round(prop.table(table(RN$estrato,RN$sexo),2),2)
```

```
##
##           H M
## E1 0.13 0.14
## E2 0.21 0.18
## E3 0.33 0.25
## E4 0.15 0.23
## E5 0.12 0.13
## E6 0.06 0.08
```

Dos variables por grupos

```
by(RN,RN$lactancia, function(RN)
  round(prop.table(table(RN$estrato,RN$sexo),1)*100,1))
```

```
## RN$lactancia: No
##
##           H      M
## E1 36.4 63.6
## E2 36.7 63.3
## E3 47.1 52.9
## E4 46.4 53.6
## E5 38.1 61.9
## E6 40.0 60.0
## -----
## RN$lactancia: Si
##
##           H      M
## E1 46.4 53.6
## E2 51.2 48.8
## E3 50.9 49.1
## E4 23.3 76.7
## E5 44.0 56.0
## E6 30.0 70.0
```

Otras opción para construir tablas es utilizando el paquete `gmodels`

```
#install.packages("gmodels")
library(gmodels)
```

```
## Warning: package 'gmodels' was built under R version 3.6.1
```

```
CrossTable(RN$estrato,RN$sexo,digits = 3,expected = FALSE,prop.chisq = FALSE,
  prop.r = TRUE,prop.c = TRUE,prop.t = TRUE)
```

```
##
##
## Cell Contents
## |-----|
## |                N |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table: 369
##
##
##           | RN$sexo
## RN$estrato |      H |      M | Row Total |
## -----|-----|-----|-----|
##           E1 |      21 |      29 |      50 |
##           | 0.420 | 0.580 | 0.136 |
##           | 0.135 | 0.136 | |
##           | 0.057 | 0.079 | |
## -----|-----|-----|-----|
##           E2 |      32 |      39 |      71 |
```

```

##           |      0.451 |      0.549 |      0.192 |
##           |      0.205 |      0.183 |           |
##           |      0.087 |      0.106 |           |
## -----|-----|-----|-----|
##           |      52 |      54 |      106 |
##           |      0.491 |      0.509 |      0.287 |
##           |      0.333 |      0.254 |           |
##           |      0.141 |      0.146 |           |
## -----|-----|-----|-----|
##           |      23 |      48 |      71 |
##           |      0.324 |      0.676 |      0.192 |
##           |      0.147 |      0.225 |           |
##           |      0.062 |      0.130 |           |
## -----|-----|-----|-----|
##           |      19 |      27 |      46 |
##           |      0.413 |      0.587 |      0.125 |
##           |      0.122 |      0.127 |           |
##           |      0.051 |      0.073 |           |
## -----|-----|-----|-----|
##           |      9 |      16 |      25 |
##           |      0.360 |      0.640 |      0.068 |
##           |      0.058 |      0.075 |           |
##           |      0.024 |      0.043 |           |
## -----|-----|-----|-----|
## Column Total |      156 |      213 |      369 |
##           |      0.423 |      0.577 |           |
## -----|-----|-----|-----|
##
##

```

---

Ejercicio 3: Calcular frecuencias absolutas y relativas de las variables:  
Sexo, Lactancia, Cuidador, Sexo y cuidador, Sexo y cuidador por lactancia.

---

- “La tabla 1”

```

#install.packages("tableone")
library(tableone)

CreateTableOne(vars=c("peso", "talla", "cuidador", "controles", "sueno", "estrato"),
  data=RN)

```

```

##
##           Overall
## n           369
## peso (mean (SD)) 6.44 (2.71)
## talla (mean (SD)) 50.12 (4.99)
## cuidador (%)
##   ambos      244 (66.1)
##   madre      111 (30.1)
##   Otro        14 ( 3.8)
## controles (mean (SD)) 4.43 (2.57)
## sueno (mean (SD))    7.09 (1.76)

```

```
## estrato (%)
## E1 50 (13.6)
## E2 71 (19.2)
## E3 106 (28.7)
## E4 71 (19.2)
## E5 46 (12.5)
## E6 25 ( 6.8)
```

```
CreateTableOne(vars=c("peso", "talla", "cuidador", "controles", "sueno", "estrato"),
                strata="sexo", data=RN)
```

```
## Stratified by sexo
## H M p test
## n 156 213
## peso (mean (SD)) 6.26 (2.63) 6.57 (2.77) 0.285
## talla (mean (SD)) 50.09 (4.93) 50.15 (5.04) 0.900
## cuidador (%) 0.478
## ambos 100 (64.1) 144 (67.6)
## madre 48 (30.8) 63 (29.6)
## Otro 8 ( 5.1) 6 ( 2.8)
## controles (mean (SD)) 4.50 (2.59) 4.37 (2.56) 0.634
## sueno (mean (SD)) 6.99 (1.78) 7.17 (1.76) 0.343
## estrato (%) 0.359
## E1 21 (13.5) 29 (13.6)
## E2 32 (20.5) 39 (18.3)
## E3 52 (33.3) 54 (25.4)
## E4 23 (14.7) 48 (22.5)
## E5 19 (12.2) 27 (12.7)
## E6 9 ( 5.8) 16 ( 7.5)
```