

# Aplicaciones del análisis multivariante con R

## Estadística Multivariante - Universidad de Granada

Laura Gómez Garrido  
Miguel Lentisco Ballesteros  
Antonio Martín Ruiz  
Daniel Pozo Escalona  
Francisco Javier Sáez Maldonado

22 de enero de 2020

- 1 Introducción: R
- 2 R en el análisis multivariante
  - Distribución Normal Multivariante. Ejemplos
  - Scatterplots
- 3 Para ampliar

## 1 Introducción: R

## 2 R en el análisis multivariante

- Distribución Normal Multivariante. Ejemplos
- Scatterplots

## 3 Para ampliar

# ¿Qué es R?

R es un entorno y lenguaje de programación enfocados a la computación estadística y de gráficos. Surge como una reimplementación libre del lenguaje y entorno S. Proporciona una amplia variedad de funcionalidades estadísticas y gráficas y es altamente extensible.

R está disponible como software libre bajo los términos de la GNU General Public License de la Free Software Foundation en forma de código fuente. Puede ser compilado y ejecutado en una gran cantidad de plataformas UNIX, Windows y MacOS.

R forma parte de un proyecto colaborativo y abierto donde sus propios usuarios pueden publicar paquetes.

Utiliza la forja de desarrollo R-Forge.

En Diciembre de 2019, el repositorio oficial tenía disponibles 15.315 paquetes organizados en vistas o temas.





- Bayesian
- ChemPhys
- ClinicalTrials
- Cluster
- Databases
- DifferentialEquations
- Distribution
- Genetics
- MachineLearning
- Multivariate



## 1 Introducción: R

## 2 R en el análisis multivariante

- Distribución Normal Multivariante. Ejemplos
- Scatterplots

## 3 Para ampliar



# Nuestro dataset

Característica del data set	Multivariante	Nº de Instancias	178
Características de los atributos	Enteros, Reales	Nº de Atributos	13
Área	Física	Donado	01/07/1991

Fuente: Machine Learning Repository

Propietarios Originales:

*Forina, M. et al, PARVUS -  
An Extendible Package for Data Exploration, Classification  
and Correlation.  
Institute of Pharmaceutical and Food Analysis and Technologies,  
Via Brigata Salerno, 16147 Genoa, Italy.*



```
wine <- read.table("http://archive.ics.uci.edu/ml
/machine-learning-databases/wine/wine.data",
sep=",")
```

In [4]: wine

A data.frame: 178 x 14

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<int>
1	14.23	1.71	2.43	15.6	127	2.80	3.06	0.28	2.29	5.64	1.04	3.92	1065	
1	13.20	1.78	2.14	11.2	100	2.65	2.76	0.26	1.28	4.38	1.05	3.40	1050	
1	13.16	2.36	2.67	18.6	101	2.80	3.24	0.30	2.81	5.68	1.03	3.17	1185	
1	14.37	1.95	2.50	16.8	113	3.85	3.49	0.24	2.18	7.80	0.86	3.45	1480	

- *read.csv* o *read.csv2*
- *read.delim*

```
sapply(wine[2:14], mean)
```

```
V2 13.0006179775281
```

```
V3 2.33634831460674
```

```
...
```

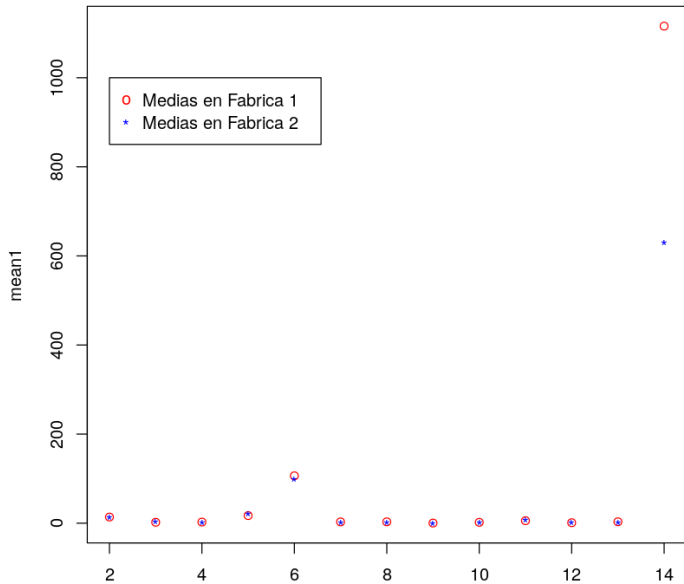
```
V13 2.61168539325843
```

```
V14 746.893258426966
```

Contents split  
into two lines

## Ejemplo de selección de datos

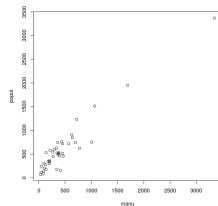
```
selection1 <- wine[wine$V1 == "1",]  
selection3 <- wine[wine$V1 == "3",]  
mean1 <- sapply(selection1[2:14], mean)  
mean2 <- sapply(selection3[2:14], mean)  
chemical <- c(2,3,4,5,6,7,8,9,10,11,12,13,14)  
plot(chemical, mean1, col = "red")  
points(chemical, mean2, col="blue", pch="*")  
legend(2,1000, legend=c("Medias en Fabrica 1",  
                        "Medias en Fabrica 2"), col=c("red", "blue"),  
      pch=c("o", "*")  
      )
```



# Scatterplots

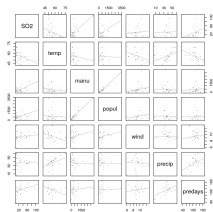
Podemos realizar scatterplots de la siguiente manera

```
plot(popul  
     ~ manu, data = USairpollution)
```



O, dibujar todas las parejas posibles a la vez, y añadir las rectas de regresión

```
pairs(USairpollution,
      panel = function(x, y, ...) {
        points(x, y, ...)
        abline(lm(y ~ x), col = "grey")
      }, pch = ".", cex = 1.5)
```





# Implementación Teórica de una DNM

```
DNM <- setRefClass("DNM",  
  fields = list(p = "numeric",  
               media = "matrix",  
               ,  
               cov = "matrix"))
```

## 1 Introducción: R

## 2 R en el análisis multivariante

- Distribución Normal Multivariante. Ejemplos
- Scatterplots

## 3 Para ampliar

*Computing Machinery and Intelligence* Alan Turing (1950)

*Artificial Intelligence: A Modern Approach* Stuart J. Russell y Peter Norvig

*Concrete Problems in AI Safety* Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, Dan Mané

*The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* Miles Brundage, Shahar Avin et al.