



Introducción al manejo del software estadístico R

Términos principales

Objeto: matriz de datos, lista, gráfica o cualquier objeto específico producido como una salida de R.

Función: procedimiento o acción representada por un nombre y una lista de argumentos entre un paréntesis. `sapply(datos1$salary, "class")`

Procedimiento 1

Paso 1: copiar al portapapeles la matriz de Universidad (seleccionar y dar Ctrl+c; ver camino de hormigas).

Paso 2: importar datos. Escribir en R

```
read.table("clipboard") -> datos1
```

Paso 3: examinar el tipo de variables (por su contenido)

```
sapply(datos1, "class")
```

Paso 4: descripción de la variable de respuesta, un histograma

```
summary(datos1$salary)  
sd(datos1$salary)
```

Paso 5: exploración univariada de la variable de respuesta, un histograma

```
hist(datos1$salary, col="yellow")
```

Paso 6: exploración univariada de la variable de respuesta, una gráfica de caja

```
boxplot(datos1$salary, col="green")
```

Paso 7: probar la normalidad de la variable de respuesta (salary)

```
shapiro.test(datos1$salary)
```

Paso 8: ¿Existe diferencia entre el salario de hombres y mujeres? (respuesta gráfica)

```
boxplot(salary~sex, data=datos1, col="orange")
```

Paso 9: ¿Existe diferencia entre el salario de hombres y mujeres? (prueba de *t* a dos colas)

```
t.test(salary~sex, data=datos1)
```

Paso 10: ¿Es menor el salario de mujeres respecto al de hombres? (prueba de *t* a una cola)

```
t.test(salary~sex, data=datos1, alternative="less")
```

Paso 11: ¿Se relaciona el salario con los años de servicio? (respuesta gráfica)

```
plot(salary~yrs2, data=datos1, pch=21, bg="red")
```

Paso 12: ¿Se relaciona el salario con los años de servicio? (prueba de correlación)

```
cor.test(datos1$salary, datos1$yrs2, method="pearson")
```

Paso 13: ¿Cuáles son las variables más importantes para predecir el salario? (regresión lineal simple)

```
summary(lm(salary~., data=datos1))
```

Paso 14: ¿Cuál podría ser el mejor modelo? (regresión lineal simple)

```
summary(lm(salary~rank, data=datos1))
```

Paso 15: ¿Cómo es la relación entre el salario y el cargo? (gráfica de caja)

```
boxplot(salary~rank, data=datos1, col="orange")
```

Paso 16: ¿Cómo es la relación entre el salario y el cargo? (análisis de varianza)

```
anova(lm(salary~rank, data=datos1))
```

```
pairwise.t.test(datos1$salary, datos1$rank, p.adj="bonf")
```

Procedimiento 2

Paso 1: copiar al portapapeles la matriz de EmpresaX (seleccionar y dar Ctrl+c; ver camino de hormigas).

Paso 2: importar datos. Escribir en R

```
read.table("clipboard") -> datos2
```

Paso 3: examinar el tipo de variables (por su contenido)

```
sapply(datos2, "class")
```

Paso 4: descripción de la variable de respuesta, un histograma

```
summary(datos2$salario)  
sd(datos2$salario)
```

Paso 5: exploración univariada de la variable de respuesta, un histograma

```
hist(datos2$salario, col="yellow")
```

Paso 6: exploración univariada de la variable de respuesta, una gráfica de caja

```
boxplot(datos2$salario, col="green")
```

Paso 7: probar la normalidad de la variable de respuesta (salary)

```
shapiro.test(datos2$salario)
```

Paso 8: ¿Existe diferencia entre el salario de hombres y mujeres? (respuesta gráfica)

```
boxplot(salario~sexo, data=datos2, col="orange")
```

Paso 9: ¿Existe diferencia entre el salario de hombres y mujeres? (prueba de *t* a dos colas)

```
t.test(salario~sexo, data=datos2)
```

Paso 10: ¿Es menor el salario de mujeres respecto al de hombres? (prueba de *t* a una cola)

```
t.test(salario~sexo, data=datos2, alternative="less")
```

Paso 11: ¿Se relaciona el salario con los años de educación? (respuesta gráfica)

```
plot(salario~educa, data=datos2, pch=21, bg="red")
```

Paso 12: ¿Se relaciona el salario con los años de educación? (prueba de correlación)

```
cor.test(datos2$salario, datos2$educa, method="kendall")
```

Paso 13: ¿Cuáles son las variables más importantes para predecir el salario? (regresión lineal simple)

```
summary(lm(salario~., data=datos2))
```

Paso 14: ¿Cuál podría ser el mejor modelo? (regresión lineal simple)

```
summary(lm(salario~sexo+educa, data=datos2))
```

Apéndice 1: Reglas para importar datos

1. Celda de la esquina superior izquierda en blanco
 2. Sin espacios
 3. Nombres de cabeceras cortos
 4. Sin celdas en blanco, si faltan datos, utilizar “NA”
- Poner atención a las mayúsculas
 - Trabajar ordenado

Apéndice 2: Sintaxis general

<u>Símbolo</u>	<u>Significado</u>
>	Podemos escribir, todo está bien.
+	Nos hizo falta cerrar una pareja de paréntesis, corchetes o comillas ((), “”)

<u>Símbolo</u>	<u>Atajo</u>	<u>Unicode</u>	<u>Uso</u>
<	Alt+60	Ctrl+Shift+u 003c Entrar	Menor que, cabeza de flecha a la izq.
>	Alt+62	Ctrl+Shift+u 003e Entrar	Mayor que, cabeza de flecha a la der.
->			Lo que está a la izq. se guardará en la der.
<-			Lo que está a la der. se guardará en la izq.
~	Alt+126		Sustituto de “=” en un argumento <u>fórmula</u>
#	Shift+3		Todo lo que está a la derecha es solo texto
{[]}	Alt gr + 7, 8, 9, 0		Uso particular a cada función
^	Alt+92	Ctrl+Shift+u 005e Entrar	Elevar a potencias

Apéndice 3: Funciones básicas adicionales

<u>Función</u>	<u>Uso</u>
ls()	Lista los objetos en el espacio de trabajo
help(<i>función</i>)	Abre la ayuda específica de una función (<i>e.g.</i> rda)
help(package= “ <i>paquete</i> ”)	Abre la ayuda general de un paquete (<i>e.g.</i> vegan)
read.table(“clipboard”)	Lee el portapapeles, el cual puede ser guardado en un objeto
read.table(pipe(“pbpaste”), sep=“\t”, header=T)	que funciona en Mac OS X
getwd()	Muestra el directorio dónde está el espacio de trabajo
setwd(“ <i>dirección</i> ”)	Establece en la dirección el directorio de trabajo (“/home/jitomate/Do16”)
install.packages(“ <i>paquete</i> ”)	Instala un paquete. Solo se necesita usar una vez. (<i>e.g.</i> nortest)
library(<i>paquete</i>)	Carga un paquete ya instalado (<i>e.g.</i> gam)
q()	Salida de R. Pregunta para guardar cambios en el espacio de trabajo.