

UNIDAD 2: Práctica 09-Análisis de una variable bidimensional categórica. Usando la interfaz gráfica (R-Commander)

1 CUALITATIVA VR CUALITATIVA

Para ilustrar como realizar un análisis estadístico bivariado usando la interfaz gráfica de R, se utilizará la información contenida en el archivo “demo.sav”; el cual contiene información de variables cualitativas y cuantitativas. Se ilustrará en este documento como realizar un análisis estadístico bivariado cuando las dos variables son cualitativas.

1) Activa tu directorio de trabajo.

```
getwd()

## [1] "C:/Users/user/OneDrive/Paquete R/PRACTICAS-S4"

setwd("C:/Users/user/OneDrive/Paquete R/PRACTICAS-S4")
```

2) Lectura del conjunto de datos. El procedimiento para cargar el conjunto de datos es el que se ha venido mencionando. Lo primero que debemos hacer es elegir la opción Importar datos del Menú Datos. y dentro de éste elegir la opción desde datos SPSS... tal y como se muestra en la figura. Debemos simplemente elegir el archivo demo.sav.

```
library(Rcmdr)

## Warning: package 'Rcmdr' was built under R version 4.1.3
## Loading required package: splines
## Loading required package: RcmdrMisc
## Warning: package 'RcmdrMisc' was built under R version 4.1.3
## Loading required package: car
## Warning: package 'car' was built under R version 4.1.3
## Loading required package: carData
## Warning: package 'carData' was built under R version 4.1.3
## Loading required package: sandwich
## Warning: package 'sandwich' was built under R version 4.1.3
## Loading required package: effects
## Warning: package 'effects' was built under R version 4.1.3
## lattice theme set by effectsTheme()
## See ?effectsTheme for details.
## La interfaz R-Commander sólo funciona en sesiones interactivas
##
## Attaching package: 'Rcmdr'
## The following object is masked from 'package:base':
##
##      errorCondition

library(car)
library(RcmdrMisc)

demo <- readSPSS("C:/Users/user/OneDrive/Paquete R/PRACTICAS-S4/demo.sav", rownames=FALSE, stringsAsFactors=TRUE)
```

Se trabajará con la variable “marital”, que representa la situación marital de las personas (solamente se distinguen entre Casadas y no Casadas); y con la variable “inccat”, la cual representa la categoría del ingreso en miles de dólares

3) Crea una tabla de contingencia o de doble entrada.

El procedimiento para realizar una tabla de contingencia en la interfaz gráfica es el siguiente: en Menú Estadísticos se elige la opción Tablas de contingencia, y dentro de este se selecciona Tabla de doble entrada.

```
# Tablas de contingencia de dos vías marital, inccat
local({
  .Table <- xtabs(~marital+inccat, data=demo)
  cat("\nFrequency table:\n")
  print(.Table)
  .Test <- chisq.test(.Table, correct=FALSE)
  print(.Test)
})

##
## Frequency table:
##           inccat
## marital  $25 - $49 $50 - $74 $75+ Under $25
##   Married      1160      568   852      596
##   Unmarried     1228      552   866      578
##
## Pearson's Chi-squared test
##
## data:  .Table
## X-squared = 2.1951, df = 3, p-value = 0.5329
```

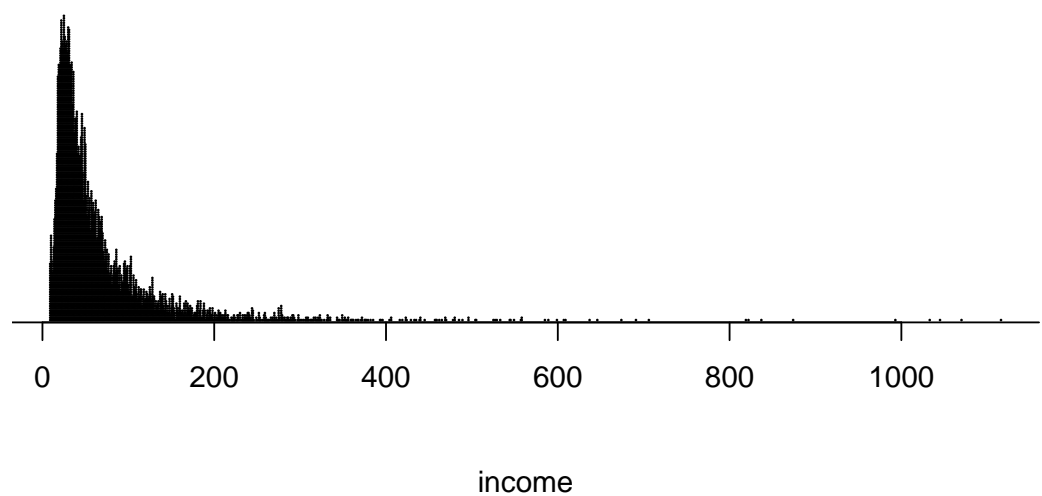
2 CUALITATIVA VR CUANTITATIVA.

Para ilustrar como realizar un análisis estadístico bidimensional entre una variable cualitativa y una cuantitativa se trabajará con la variable “marital”, que representa la situación marital de las personas (solamente se distinguen entre Casadas y no Casadas); y con la variable “income”, la cual representa el ingreso económico.

1) Dibuja un gráfico horizontal de puntos para los tres procesos.

Podemos realizar un gráfico de puntos, en el cual podemos observar gráficamente si la variable income se comporta de manera diferente en cada uno de los niveles de la variable marital. El procedimiento para realizar el gráfico es el siguiente.

```
# Gráfico de puntos income
with(demo, Dotplot(income, bin=FALSE))
```



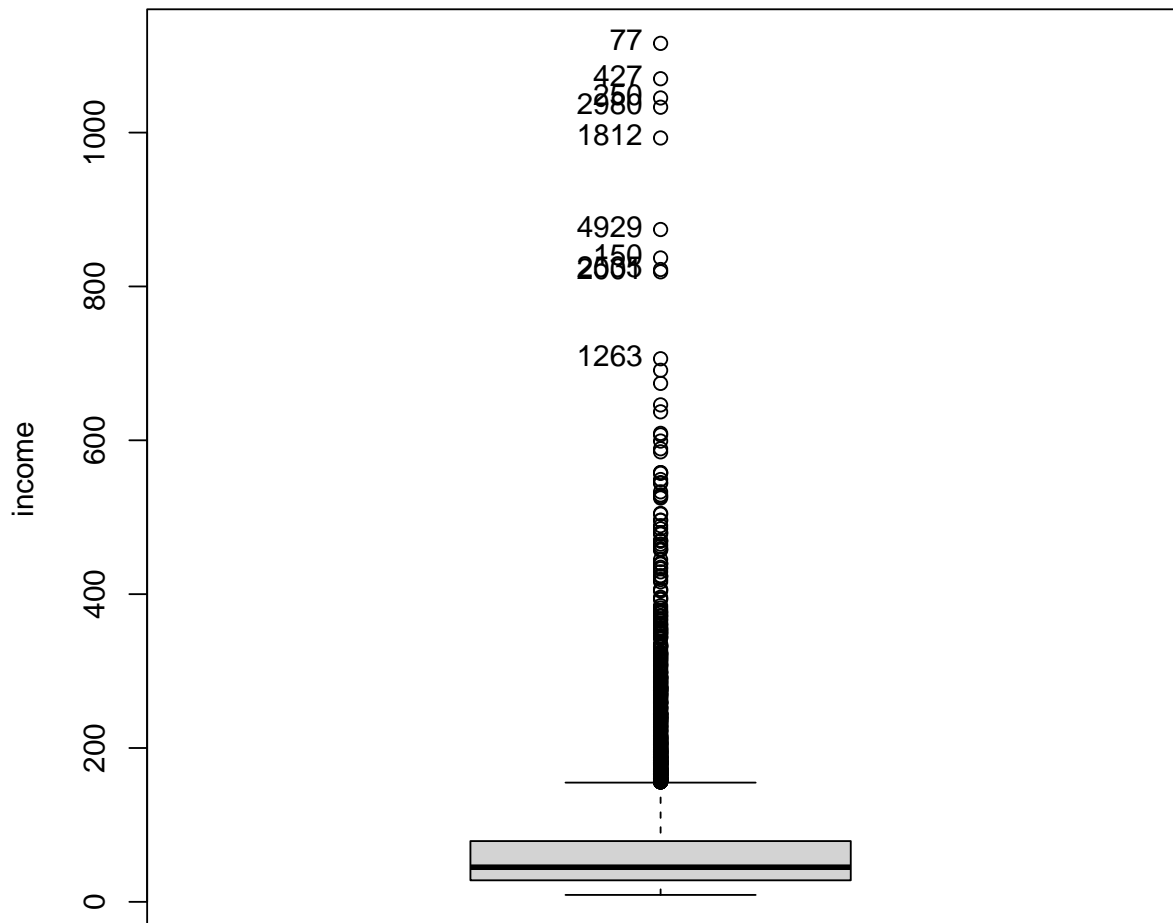
2) Muestra un resumen estadístico para los estados maritales. El procedimiento es como sigue: en el Menú Estadísticos seleccionar la opción Resúmenes, y dentro del sub Menú que aparecerá seleccionar la opción Resúmenes numéricos

```
# Resúmenes numéricos: demo
numSummary(demo[, "age", drop=FALSE], statistics=c("mean", "sd", "IQR", "quantiles"), quantiles=c(0,.25,.5,.75,1))
```

	mean	sd	IQR	0%	25%	50%	75%	100%	n
##	42.05859	12.28952	18	18	33	41	51	77	6400

3) Dibuja un gráfico de cajas (box-plot) para los estado maritales.
Para realizar un diagrama de caja de una variable cuantitativa en los diferentes niveles de una segunda variable la cual es cualitativa, el procedimiento es como sigue. En el Menú Gráficas seleccionamos la opción Diagrama de cajas

```
# Diagrama de caja: ~ income
Boxplot( ~ income, data=demo, id=list(method="y"))
```

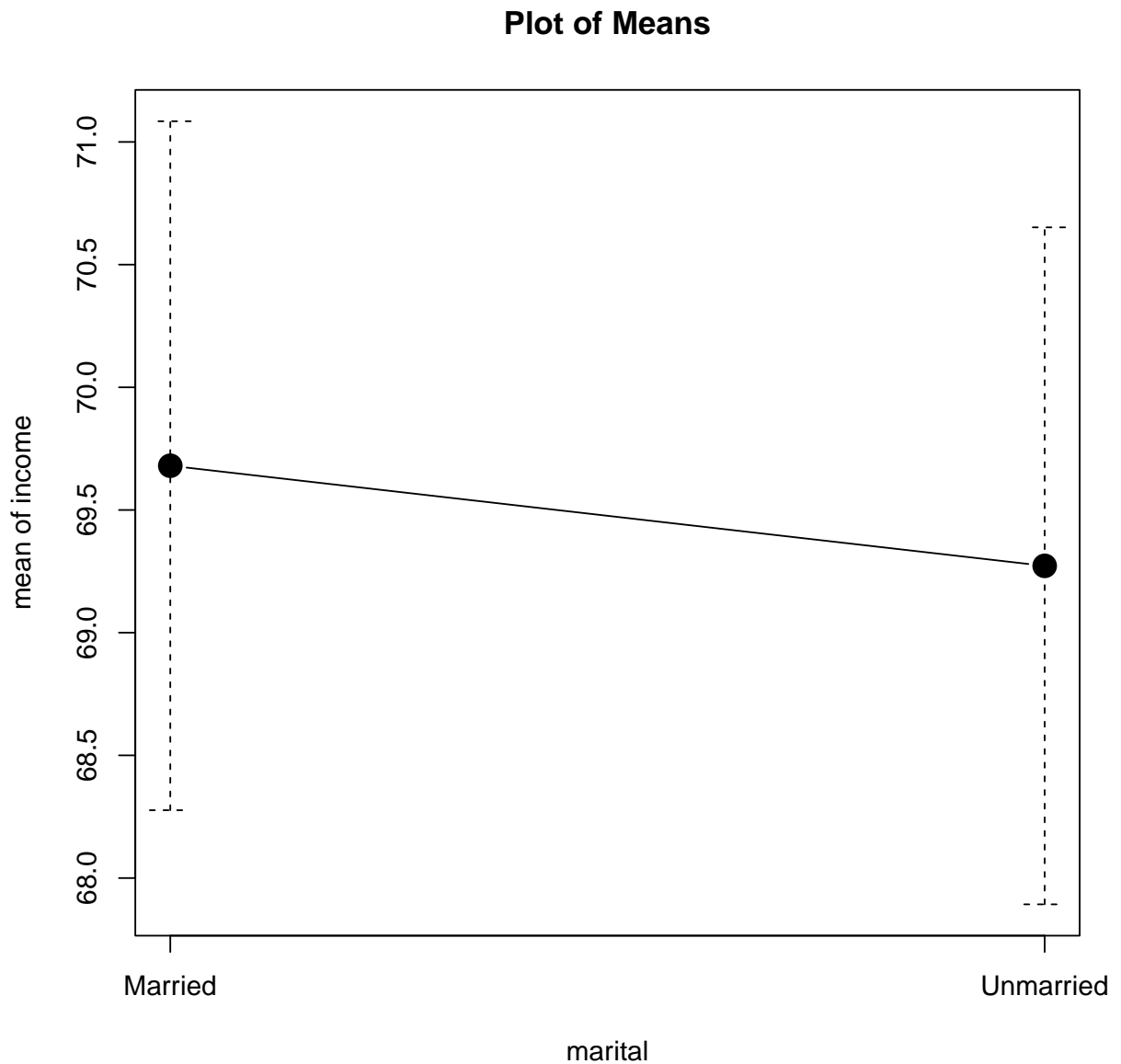


```
## [1] "77" "427" "250" "2980" "1812" "4929" "150" "2535" "2001" "1263"
```

4) Gráficas de medias.

También en algunos casos es útil realizar el gráfico de las medias, el cual nos da mayor información a los diagramas de puntos. Para obtenerlo el procedimiento es: en el Menú Gráficas seleccionamos la opción Gráficas de las medias, tal y como se muestra en la figura. Al realizar este procedimiento deberá aparecernos una ventana en la cual debemos especificar la variable explicativa income (nuestra variable cuantitativa), y los factores, es decir, la variable marital (la cual es cualitativa).

```
# Gráfica de las medias: income by marital
with(demo, plotMeans(income, marital, error.bars="se", connect=TRUE))
```



5) Prueba de comparación de medias (asumiendo normalidad).

Se utiliza para contrastar las siguientes hipótesis:

$H_0 : \mu_A = \mu_B$, no existe diferencias de los ingresos para los estados maritales.

$H_1 : \mu_A \neq \mu_B$, si existe diferencia.

El procedimiento para llevar a cabo tal contraste de hipótesis es mediante la prueba t; en el Menú Estadísticos seleccionamos la opción Medias, y dentro de este seleccionamos la opción Test t para muestras independientes, tal y como se muestra en la siguiente figura. En el cuadro resultante únicamente debemos verificar si explicativa income (que es cuantitativa) y el factor marital (que es cualitativa); definimos el tipo de prueba (una o dos colas), y especificamos si las varianzas son o no iguales.

```
# Test t para muestras independientes: income~marital

t.test(income~marital, alternative='two.sided', conf.level=.95,var.equal=FALSE, data=demo)

##
```

```
## Welch Two Sample t-test
##
## data: income by marital
## t = 0.20733, df = 6394, p-value = 0.8358
## alternative hypothesis: true difference in means between group Married and group Unmarried is not
## 95 percent confidence interval:
## -3.450475 4.266641
## sample estimates:
## mean in group Married mean in group Unmarried
## 69.68042 69.27233
```

3 CUANTITATIVA VR CUALITATIVA

```
# 1) Lectura de datos
# 2) Muestra un resumen de principales estadísticos de las variables.
knitr::opts_chunk$set(comment=NA, prompt=TRUE, out.width=750, fig.height=8, fig.width=8)

estaturas<-read.table("C:/Users/user/OneDrive/Paquete R/PRACTICAS-S4/estaturas.dat",header=TRUE, str

## Error in read.table("C:/Users/user/OneDrive/Paquete R/PRACTICAS-S4/estaturas.dat", : unused
argument (n.strings = "NA")

# 3) Elabora un gráfico de dispersión para analizar alguna relación entre las variables.
scatterplot(X0.000000000e.000~X7.880000000e.001, regLine=FALSE, smooth=FALSE, boxplots=FALSE, data=es

## Error in eval(m$data, sys.frame(sys.parent())): objeto 'estaturas' no encontrado

# 4) Aplica la función lm() para encontrar el modelo lineal que se ajusta a los datos.

RegModel.2<-lm(X7.880000000e.001~X0.000000000e.000, data=estaturas)

## Error in is.data.frame(data): objeto 'estaturas' no encontrado

summary(RegModel.2)

## Error in summary(RegModel.2): objeto 'RegModel.2' no encontrado

# 5) Efectúa una análisis de variabilidad del modelo o descomposición de la varianza.
Anova(RegModel.2, type="II")

## Error in Anova(RegModel.2, type = "II"): objeto 'RegModel.2' no encontrado
```