

TUTORIAL R STUDIO

COMANDOS

Cargar paquetes:

```
install.packages(c("summarytools","modeest","ggplot2","BioestadisticaR2","car"))
library(summarytools)
library(modeest)
library(ggplot2)
library(BioestadisticaR2)
library(car)
```

Operaciones Aritméticas: Suma: + Resta: - Multiplicación: * División: / Exponente: ^ Raíz: `sqrt(x)`

Cargar datos:

```
x=c(12,0,14,10,8,9,18,20,15,0)
y=c(7,9,6,4,3,6,7,7,9,8)
z=c(1,1,1,1,1,2,2,2,2,2)

zcuali=factor(z,labels=c("A","B"))

base=data.frame(x,y,z,zcuali)

attach(base)
```

Recodificar variables:

```
x.sino=vector()
x.sino[x=="0"]=0
x.sino[x!="0"]=1
x.sino=factor(x.sino,labels=c("No","Sí"))
base$x.sino=x.sino

x.sinoR=vector()
x.sinoR[x.sino=="No"]=2
x.sinoR[x.sino!="No"]=1
x.sinoR=factor(x.sinoR,labels=c("Si","No"))
levels(x.sinoR)
base$x.sinoR=x.sinoR
levels(x.sino)
levels(x.sinoR)

ycuali=vector()
ycuali[y<5]=1
ycuali[y>=5 & y<7]=2
ycuali[y>=7 & y<9]=3
ycuali[y>=9 & y<=10]=4
ycuali=factor(ycuali,labels=c("Suspendido","Aprobado","Notable","Excelente"))
base$ycuali=ycuali
```

Tabla:

```
frec=table(ycuali)
frec.acum=cumsum(table(ycuali))
frec.acum=cumsum(frec)
prop=prop.table(table(ycuali))
prop=prop.table(frec)
prop.acum=cumsum(prop.table(table(ycuali)))
prop.acum=cumsum(prop)
tabla.frecuencias=cbind(frec,frec.acum,prop,prop.acum)
tabla.frecuencias
```

```
frec(ycuali)
```

Estadísticos:

```
mfv(x)
mean(x)
median(x)
quantile(x)
var(x)
sd(x)
max(x)
min(x)
summary(x)

descr(x,stats=c("mean","sd","q1","med","q3","min","max","cv"))
```

Gráficos:

```
ggplot(base,aes(x=ycuali))+geom_bar(width=0.5,colour="blue",fill="blue")
ggplot(base,aes(x=factor(1),fill=ycuali))+geom_bar()+coord_polar("y")+theme_void()
ggplot(base,aes(x=y))+geom_histogram(binwidth=1)
ggplot(base,aes(x=0,y=y))+geom_boxplot()

ggplot(base,aes(x=zcuali,fill=ycuali))+geom_bar(width=0.5,position="dodge")
ggplot(base,aes(x=zcuali,y=y,fill=zcuali))+stat_boxplot(geom="errorbar",width=0.20)+geom_boxplot()
```

Distribuciones:

```
dbinom(x,n,p)  pbinom(q,n,p)  qbinom(p,n,p)
dpois (x,λ)  ppois (q,λ)  qpois (q,λ)
pnorm(q,μ,σ)  pnorm(q,μ,sd,lower.tail=FALSE)  qnorm(p,μ,σ)
pt(t,gl)  pt(t,gl,lower.tail=FALSE)  qt(p,gl)
pchisq(x,gl)  pchisq(x,gl,lower.tail=FALSE)  qchisq(p,gl)
pf(f,gl1,gl2)  pf(f,gl1,gl2,lower.tail=FALSE)  qf(p,gl1,gl2)
```

Intervalos de confianza de medias

```
icm(n=100,m=6,s=2,conf=.95)
icm(y,conf=.95)
```

Intervalos de confianza de proporciones

```
icp(x=40,n=100,conf=.95)
icp(zcuali,level="A")
```

Contraste de diferencia de medias independientes:

```
G1=c(7,9,8,5)
G2=c(5,5,4,1)

tapply(x,zcuali,shapiro.test)
var.test(x~zcuali)

testt(m1=G1,m2=G2)
testt(m=x,grupos=as.factor(zcuali))
testwx(m=x,grupos=zcuali)
```

Contraste de diferencia de medias relacionadas

```
testt(m=-1,s=2,n=5)

testt(m1=G1,m2=G2,par=TRUE)
testt(m1=x,m2=y,par=TRUE)
testwx(m1=x,m2=y,par=TRUE)
```

Contraste de diferencia de proporciones independientes:

```
testp(x1=20,n1=100,x2=40,n2=100)
testp(x=x.sino,grupos=zcuali)
```

Contraste de diferencia de proporciones relacionadas:

```
testmcnemar(pre=x.sino,post=x.sinoR)
```

Contraste chi cuadrado de independencia:

```
C1=c(3,7)
C2=c(17,13)
tabla=data.frame(C1,C2)
tablarxc(frecs=tabla,tablas=c("F","C","E"))

tablarxc(fvar=zcuali,cvar=x.sino,fcats=c("A","B"),ccats=c("No","Si"),tablas=c("F","C","E"))
tabla2x2(o=c(3,17,7,13),fcats=c("M","T"),ccats=c("S","A"),estudio=c("P"),tablas=c("F"))
```

Modelo de regresión:

```
scatterplot(y~x,smooth=FALSE,boxplots="")
cor(x,y)
cor(x,y,method="spearman")

Modelo.RLS<-lm(y~x)
summary(Modelo.RLS)
confint(Modelo.RLS, level=0.95)
fitted(Modelo.RLS)
resid(Modelo.RLS)

rls(y~x)
```

EJERCICIOS CON BASE DE DATOS

Carga la base de datos TutorialRStudioBEUGR.csv en RStudio (Environment->Import Dataset-> From text (base)) poniendo como nombre "base", y con Heading "Yes". Después, resuelve las siguientes cuestiones:

0. Describe las variables Notas_final y Sexo, con los estadísticos o gráficos que consideres adecuados.

```
attach(base)
descr(Notas_final, stats=c("mean", "sd", "q1", "med", "q3", "min", "max", "cv"))
freq(Sexo)
ggplot(base, aes(x=Sexo)) + geom_bar(width=0.5, colour="blue", fill="blue")
ggplot(base, aes(x=factor(1), fill=Sexo)) + geom_bar() + coord_polar("y") + theme_void()
ggplot(base, aes(x=Notas_final)) + geom_histogram(binwidth=1)
ggplot(base, aes(x=0, y=Notas_final)) + geom_boxplot()
```

1. Calcula la probabilidad de...

a) ...que un alumno suspenda el examen final, si sabemos que la nota en la población sigue una normal de media 6,28 y desviación típica 2,15.

```
pnorm(5, 6.28, 2.15)
```

b) ...estar por encima de una t de 1,5 con 29 grados de libertad.

```
pt(1.5, 29, lower.tail=FALSE)
```

c) ... que una chi cuadrado de 2 grados de libertad esté entre 3 y 5.

```
pchisq(3, 2) - pchisq(5, 2)
```

2. Calcula ...

a) ... el valor que en una t_{10} , deja el 99% acumulado

```
qt(0.99, 10)
```

b) ... los valores que limitan el 90% central en una normal estándar.

```
qnorm(0.05, 0, 1)
```

```
qnorm(0.95, 0, 1)
```

3. ¿Entre qué valores está la media de la nota del examen final en toda la población? ¿Y la proporción que hacen evaluación continua?

```
icm(Notas_final, conf=.95)
```

```
icp(Evaluacion, level="Ev. Continuada")
```

4. ¿Hay diferencias significativas en la nota del examen final entre los que van a academia y los que no van?

```
testt(m= Notas_final, grupos=as.factor(Academia))
```

```
testwx(m=Notas_final, grupos=Academia)
```

5. ¿Ha habido un cambio en las notas entre el examen parcial y el final?

```
testt(m1=Notas_parcial, m2=Notas_final, par=TRUE)
```

```
testwx(m1=Notas_parcial, m2=Notas_final, par=TRUE)
```

6. ¿Hay relación entre el sexo y el tipo de evaluación?

```
tablarxc(fvar=Sexo, cvar=Evaluacion, fcat=c("H", "M"), ccat=c("EC", "EU"), tablas=c("F", "E"))
```

7. Analiza la relación entre la asistencia a clase y la nota del examen final.

```
rls(Notas_final~Asistencia)
```