



UNIVERSIDAD
CATÓLICA DE
TEMUCO

DEPARTAMENTO DE
INGENIERÍA INFORMÁTICA
FACULTAD DE INGENIERÍA

Taller Formativo 9

Mario Castillo Sanhueza

Docente: Dr. Julio Rojas Mora

Departamento de Ingeniería Informática
Facultad de Ingeniería
Universidad Católica de Temuco

Agosto 18, 2022

1. Problemática.

- Compruebe que el problema propuesto en la lámina 106 y resuelto con el código de la lámina 108 cumple con las propiedades de una distribución normal (lámina 105). Aumente el número de datos a 10000, construya nuevamente la tabla de frecuencias y analice si el ajuste mejora.
- ¿El ajuste es bueno? ¿Qué tal si aumenta el número de datos a 10000?.
- ¿Qué pasaría si cambia la distribución de los datos?.

2. Desarrollo

- Podemos entender de la muestra de 100 datos que no es un ajuste bueno dado que no cumple con las propiedades descritas tales como si al ubicarnos en la media $\pm 3s.d$ tendremos 99.7% de los datos, pero una vez la muestra aumenta a 10000 datos esta propiedad se va ajustando a lo esperado, entonces podríamos inferir que a medida que la muestra de datos crezca nuestros análisis serán más fidedignos y, por lo tanto, tendremos una muestra representativa.
- Al cambiar el tipo de distribución de los datos ya no podríamos utilizar la propiedad anteriormente descrita para $\pm 1s.d.$, $\pm 2s.d.$, $\pm 3s.d.$ para comprobar la agrupación de mis datos.

3. Código

```
library(Hmisc)
library(TTR)
#We generate the seed
set.seed(1121)
x<-rnorm(n=100, mean=10,sd=2)
y<- rnorm(n=10000, mean=10,sd=2)

#n=100
data<-data.frame(x=x)
```

```

n<-nrow( data )
#Sturges
m<-1+ceiling (3.322*log10(n))
C<-2
cuts<-c(-Inf,4,6,8,10,12,14,16,Inf)
data$xFact <-cut2( data$x, cuts , digits=5)
frcTable<-data.frame(x=levels( data$xFact))
frcTable$f <- table( data$xFact)
frcTable$fr <- frcTable$f/n
frcTable$F <- cumsum( frcTable$f)
frcTable$Fr <- cumsum( frcTable$fr)

#n=10000
datay<-data.frame(y=y)
n<-nrow( datay )
#Sturges
m<-1+ceiling (3.322*log10(n))
C<-2
cuts2<-c(-Inf,4,6,8,10,12,14,16,Inf)
datay$yFact <-cut2( datay$y, cuts2 , digits=5)
frcTable2<-data.frame(y=levels( datay$yFact))
frcTable2$f <- table( datay$yFact)
frcTable2$fr <- frcTable2$f/n
frcTable2$F <- cumsum( frcTable2$f)
frcTable2$Fr <- cumsum( frcTable2$fr)

```

```

frcTable
frcTable2

```

```

> frcTable

```

	x	f	fr	F	Fr
1	[-Inf , 4)	0	0.00	0	0.00
2	[4, 6)	2	0.02	2	0.02
3	[6, 8)	12	0.12	14	0.14
4	[8, 10)	32	0.32	46	0.46
5	[10, 12)	42	0.42	88	0.88

```

6 [ 12, 14) 12 0.12 100 1.00
7          14 0 0.00 100 1.00
8 [ 16, Inf] 0 0.00 100 1.00
> frcTable2
      y      f      fr      F      Fr
1 [-Inf, 4)   11 0.0011    11 0.0011
2 [  4, 6)  189 0.0189   200 0.0200
3 [  6, 8) 1411 0.1411  1611 0.1611
4 [  8, 10) 3455 0.3455  5066 0.5066
5 [ 10, 12) 3404 0.3404  8470 0.8470
6 [ 12, 14) 1324 0.1324  9794 0.9794
7 [ 14, 16)  193 0.0193  9987 0.9987
8 [ 16, Inf]   13 0.0013 10000 1.0000
>

```