

# 11-Estadísticos para datos agrupados

Adrian

19/1/2022

## Estadísticos para datos agrupados

Al tener una muestra de datos, conviene calcular los estadísticos antes de realizar los agrupamientos, puesto que de lo contrario podemos perder información.

Hay situaciones en las que las muestras ya vienen agrupadas. En estos casos, aun sigue siendo posible calcular los estadísticos y utilizarlos como aproximaciones de los estadísticos de los datos “reales”.

Los estadísticos son la media, varianza, varianza muestra, varianza típica, varianza típica muestral, moda, mediana, cuantiles...

La moda pasa a ser el intervalo modal, que es la clase con mayor frecuencia (absoluta o relativa).

La mediana pasa a ser el intervalo crítico para la mediana, y es el primer intervalo donde la frecuencia relativa acumulada sea mayor o igual que 0.5

Aproximación de los cuantiles

## Ejemplo

```
# Obtener el total de muestras
# TOT = tabla$Fr.cum.abs[10] donde 10 es el ultimo intervalo

# Obtener la media de las muestras
# anchura.media = round(sum(tabla$Fr.abs*tabla$mc)/TOT, 3)

# Obtener la varianza de las muestras
# anchura.var = round(sum(tabla$Fr.abs*tabla$mc^2)/TOT-anchura.media^2,3)

# Obtener la desviacion tipica de las muestras
# anchura.dt = round(sqrt(anchura.var), 3)

# Obtener el intervalo modal
# I.modal = tabla$intervals[which(tabla$Fr.abs == max(tabla$Fr.abs))]

# Obtener el intervalo critico para la mediana
# I.critic = tabla$intervals[which(tabla$Fr.cum.rel >= 0.5)]
# I.critic[1]

# Estimacion de la mediada de los datos "reales"
# n = TOT
# Lc = L[4]
```

```

# Lc.pos = L[5]
# Ac = L[5]-L[4]
# Nc.ant = tabla$Fr.cum.abs[3]
# nc = tabla$Fr.abs[4]
# M = Lc+Ac*((n/2)-Nc.ant)/nc
# M Es la aproximacion de la mediana

```

**Funcion que calcula los cuantiles de forma aproximada**

```

aprox.quantile.p = function(Lcrit, Acrit, n, p, Ncrit.ant, ncrit) {
  round(Lcrit+Acrit*(p*n-Ncrit.ant)/ncrit, 3)
}
# aprox.quantile.p(Lc, Ac, n, 0.25, Nc.abt, nc) primer cuartil

```