BERNARDO MONDRAGÓN
BROZON

HISTOGRAMAS

Suponga que se observará una variable cuantitativa que tomará valores entre a y b con $a, b \in \mathbb{R}$.

Sea $a = x_0 < x_1 < \dots < x_{n-3} < x_{n-2} < x_{n-1} < x_n$ una partición del intervalo $[a,b]$ de manera que se obtienen n subintervalos o clases $C_1, C_2, \dots, C_{n-1}, C_n$ en las que la clase C_i agrupa o clasifica a todos los valores observados pertenecientes al conjunto $[x_{i-1}, x_i]$.

Entonces $\{x_i\}_{i=1,\dots,n}$ es el conjunto conformado por los límites superiores de cada subintervalo o clase.

Si la partición es regular, la longitud de cada subintervalo está dada por $h = \frac{b-a}{n}$ donde n es el número de clases, o subintervalos, o elementos de la partición y $x_i = a + hi$ para $i \in \{1, \dots, n\}$.

La frecuencia $frec_{C_i}$ de la clase C_i está dada por número de observaciones que contiene, de manera que

$$frec_{C_i} = \sum_{j=1}^{N} I_{(x_{i-1}, x_i]}(y_j) \quad y \quad frec.rel. c_i = \frac{\sum_{j=1}^{n} I_{(x_{i-1}, x_i]}(y_j)}{n}$$

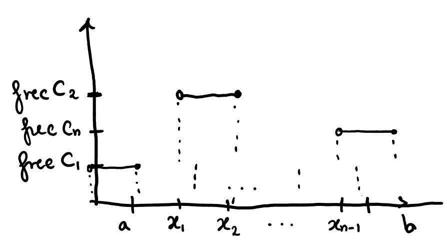
En donde N es el total de observaciones y y_j es la j-ésima observación $j = 1, \dots, N$.
Notar que

$$\sum_{i=1}^{n} frec.C_i = N \quad y \quad \sum_{i=1}^{n} frec.rel\, C_i = 1$$

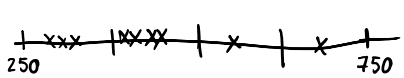
El histograma está dado por la gráfica de la siguiente función:

$$f(x) = \begin{cases} frec. C_1 & si \quad x \in (x_0, x_1) \\ frec. C_2 & si \quad x \in (x_1, x_2) \\ & \vdots \\ frec\, C_n & si \quad x \in (x_{n-1}, x_n) \end{cases}$$

Entonces el histograma tiene el siguiente aspecto:



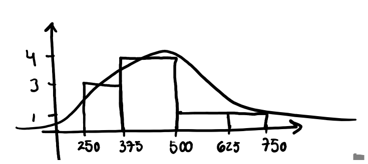
Ejemplo: Supóngase que se observará valores 250 y 750 :



y n=4

$\therefore h = \frac{750-250}{4} = \frac{500}{4} = 125$ Entonces

Clase	frec.	frec. Ac.
C_1: 250–375	3	3
C_2: 375–500	4	7
C_3: 500–625	1	8
C_4: 625–750	1	9



C_2: 375-500	4	7
C_3: 500-625	1	8
C_4: 625-750	1	9