

## DISTRIBUCIÓN DE FRECUENCIAS

(Tomado de: <http://mx.geocities.com/fracosta11/dfrec.html>)

Una distribución de frecuencias es una herramienta estadística muy útil para organizar un grupo de observaciones.

**Distribución de frecuencias:** Es una serie de datos agrupados en categorías, en las cuales se muestra el número de observaciones que contiene cada categoría.

Los pasos para la construcción de una distribución de frecuencias son mejor explicados con un ejemplo.

Ejemplo:

Los siguientes datos son el número de meses de duración de una muestra de 40 baterías para coche.

22	41	35	45	32	37	30	26
34	16	31	33	38	31	47	37
25	43	34	36	29	33	39	31
33	31	37	44	32	41	19	34
47	38	32	26	39	30	42	35

### El rango

$$\text{Rango} = \text{Dato mayor} - \text{Dato menor} = 47 - 16 = 31$$

### Número tentativo de los intervalos de clase

El número de intervalos (nic) puede ser como mínimo 5 y como máximo 15 de acuerdo a la fórmula  $2^{\text{nic}} \leq n$

Número de intervalos (nic)	Número máximo de datos
5	32
6	64
7	128
8	256
9	512
10	1024

### Tamaño de los Intervalos de Clase (tic)

$$\text{TIC} = \frac{\text{Rango}}{\text{NIC}} = \frac{31}{6} = 5,16 \text{ se redondea a } 5$$

Para facilitar la clasificación de los datos, el Tic se redondea a una cifra mas o menos cerrada.

### **Primer límite inferior**

Usualmente, el límite inferior del primer intervalo de clase es un múltiplo del tamaño del intervalo (tic), en este problema el tic es de 5, entonces el primer límite inferior será **el mayor múltiplo de 5 pero inferior o igual al dato menor, el 15.**

### **Límite inferior y superior de cada intervalo**

Se comienza en el primer límite inferior al que se le va sumando el tic, hasta llegar a un valor máximo antes de superar el valor del dato mayor. Los límites superiores se calculan con la siguiente fórmula:

$$\text{LS} = \text{LI} + \text{TIC} - 1$$

LI	LS
15	19
20	24
25	29
30	34
35	39
40	44
45	49

### **Límites Reales**

Los límites anteriores son llamados límites “oficiales” pero no son los reales. Los límites reales son el punto medio entre el límite superior y el límite inferior del siguiente intervalo.

LI	LS	LSR
15	19	19.5
20	24	24.5
25	29	29.5
30	34	34.5
35	39	39.5
40	44	44.5
45	49	49.5

$$\text{LSR} = \frac{\text{LS} + \text{Lsig}}{2}$$

### **Marca de clase (x)**

La marca de clase, también llamada punto medio del intervalo es la mitad de la distancia entre los límites inferior y superior de cada intervalo. La marca de clase es el valor más representativo de los valores del intervalo.

LI	LS	LSR	X
15	19	19.5	17
20	24	24.5	22
25	29	29.5	27
30	34	34.5	32
35	39	39.5	37
40	44	44.5	42
45	49	49.5	47

$$X = \frac{LI + LS}{2}$$

### Clasificación de los datos y conteo de frecuencias

Clasificar las observaciones en los intervalos. La práctica usual es marcar con una línea (/) que representa una observación. En el ejemplo la observación 22 se clasifica en el intervalo 20 – 24 porque se encuentra entre el 20 y el 24 inclusive. Una vez clasificados todos los datos se cuentan las líneas de cada intervalo y el resultado es la frecuencia de cada intervalo de clase.

LI	LS	cuenta	F
15	19	//	2
20	24	/	1
25	29	////	4
30	34	//// //	15
35	39	//// //	10
40	44	////	5
45	49	///	3

### Distribución de frecuencia relativa

Se pueden convertir las frecuencias de clase en **frecuencias relativas de clase** para mostrar los porcentajes de observaciones en cada intervalo de clase. Para convertir una distribución de frecuencia en una distribución de frecuencia relativa cada una de las frecuencias de clase se dividen entre el número total de observaciones.

LI	LS	LSR	X	F	FR
15	19	19.5	17	2	2 / 40 = .05
20	24	24.5	22	1	1 / 40 = .025
25	29	29.5	27	4	4 / 40 = .1
30	34	34.5	32	15	15 / 40 = .375
35	39	39.5	37	10	10 / 40 = .25
40	44	44.5	42	5	5 / 40 = .125
45	49	49.5	47	3	3 / 40 = .075
N =				40	

### Distribuciones de frecuencia acumulada

Si queremos responder preguntas como ¿cuántas observaciones (baterías) están por debajo de 40?, o ¿qué porcentaje de observaciones (baterías) caen por debajo de 30?. Las respuestas a estas preguntas pueden ser contestadas con una **distribución de frecuencia acumulada** o por una **distribución de frecuencia relativa acumulada**.

Las distribuciones de **frecuencia acumulada** se usan cuando queremos determinar cuantas observaciones, o que porcentaje de observaciones están debajo de cierto valor.

La distribución de frecuencia acumulada de cierto intervalo se calcula sumando las frecuencias de clase desde el primer intervalo hasta la frecuencia de clase del intervalo de interés. Si queremos la frecuencia acumulada del intervalo 25 – 29, sumamos las frecuencias de clase  $2 + 1 + 4 = 7$ .

La distribución de frecuencia relativa acumulada de cierto intervalo se calcula dividiendo la frecuencia acumulada entre el número total de observaciones.

LI	LS	LSR	X	F	FR	FA	FRA
15	19	19.5	17	2	.05	2	.05
20	24	24.5	22	1	.025	3	.075
25	29	29.5	27	4	.1	7	.175
30	34	34.5	32	15	.375	22	.55
35	39	39.5	37	10	.25	32	.8
40	44	44.5	42	5	.125	37	.925
45	49	49.5	47	3	.075	40	1

---

N =

40

## **Presentación Gráfica**

Tres gráficas que nos ayudarán a representar gráficamente una distribución de frecuencias son el histograma, el polígono de frecuencia y la ojiva de frecuencia acumulada.

### **El histograma**

El histograma es una de las gráficas mas ampliamente utilizadas y una de las mas fáciles de entender. Un histograma describe una distribución de frecuencia utilizando una serie de rectángulos adyacentes donde la altura de cada rectángulo es proporcional a la frecuencia de clase que representa.

### **Polígono de frecuencia**

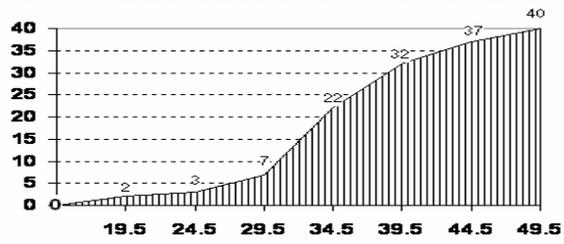
El polígono de frecuencia consiste de segmentos de línea conectando los puntos formados por la intersección de las marcas de clase y las frecuencias de clase.

El polígono de frecuencia relativa es similar al anterior solo que en este se muestran porcentajes, es decir las frecuencias relativas de cada clase.

### **Ojivas de Frecuencia Acumulada**

Las distribuciones de frecuencia acumulada y frecuencia relativa acumulada se presentan gráficamente con las ojivas de frecuencia acumulada y frecuencia relativa acumulada, que es una gráfica de segmentos de línea que une los puntos donde se cruzan los límites reales con las frecuencias acumulada y relativa acumulada de cada intervalo de clase.

### OJIVA DE FRECUENCIA ACUMULADA



### OJIVA DE FRECUENCIA RELATIVA ACUMULADA

