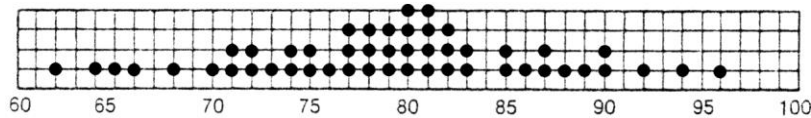


2. LA FORMA DE LA DISTRIBUCIÓN

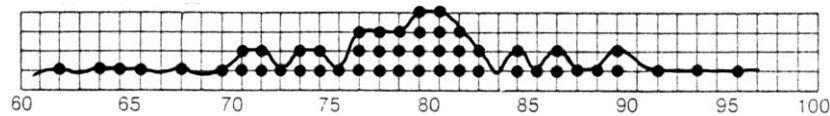
Una **DISTRIBUCIÓN** es la organización de los valores observados en una variable, del menor al mayor; por ejemplo, la frecuencia cardíaca de 50 estudiantes:

Pulsaciones por minuto de 50 estudiantes: 62, 64, 65, 66, 68, 70, 71, 71, 72, 72, 73, 74, 74, 75, 75, 76, 77, 77, 77, 78, 78, 78, 79, 79, 79, 80, 80, 80, 80, 81, 81, 81, 81, 82, 82, 82, 83, 83, 85, 85, 86, 87, 87, 88, 89, 90, 90, 92, 94, 96

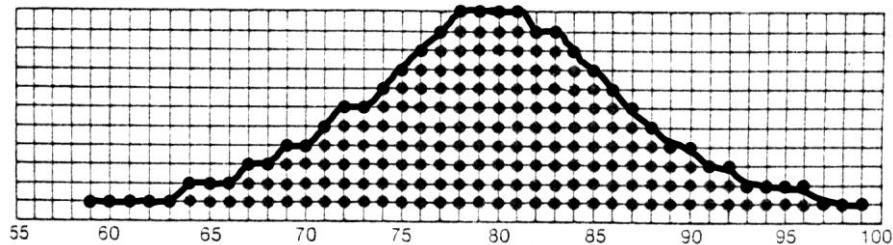
Una **DISTRIBUCIÓN DE FRECUENCIAS** es una gráfica que muestra la frecuencia con que fue observado cada valor, como la siguiente:



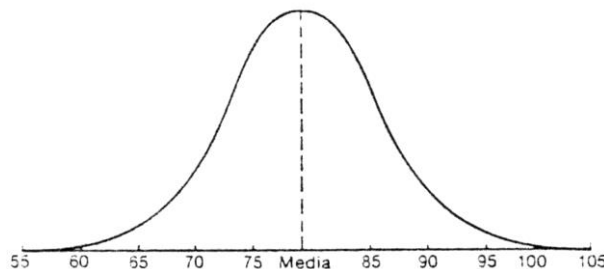
Si se unen los puntos superiores de la gráfica de puntos de una determinada variables, se obtiene una **CURVA DE DISTRIBUCIÓN DE FRECUENCIAS**:



Si se aumenta el tamaño de la muestra, la curva se va haciendo más suave (sin pequeñas elevaciones y depresiones), por ejemplo a 200 estudiantes:

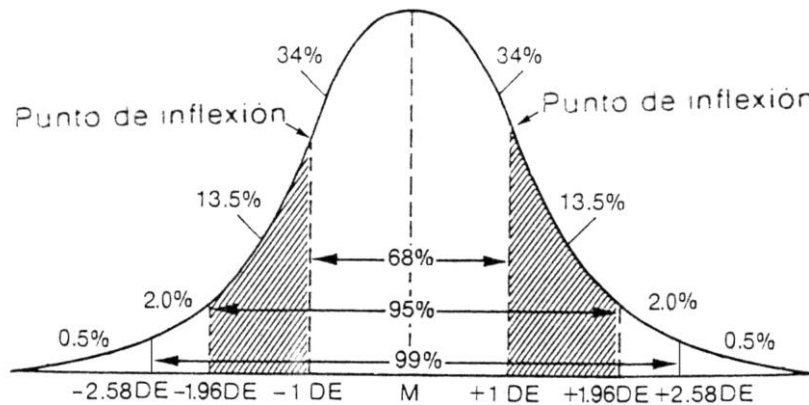


Si se miden miles de estudiantes, se tendrá una curva como la siguiente:

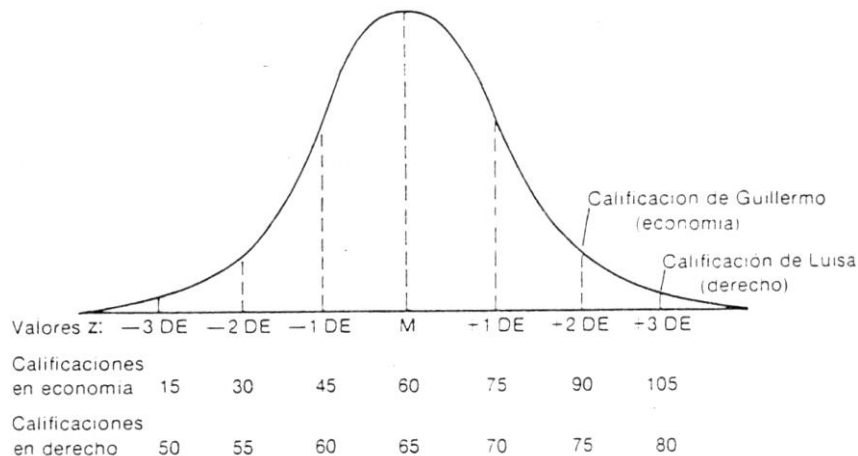


Esta curva de distribución tiene la forma de la **CURVA NORMAL**, que se caracteriza porque: (1) tiene forma de campana, (2) es perfectamente simétrica, (3) la moda, mediana y media están en el centro. (4) La curva normal no es la más común, sino una **curva ideal**, definida por una ecuación matemática, con la cual comparar las que se obtienen de la realidad.

(5) La característica más útil de la curva normal es que puede establecerse con exactitud el porcentaje de población comprendido entre dos valores de la variable; para esto sólo se requiere conocer la media ($M = \bar{x}$ o " μ ") y desviación estándar ($DE = "s"$ o " σ ", de un muestra o de la población respectivamente) de la variable dada. Entre el punto ($M-1DE$) y el punto ($M+1DE$), está el 68% de los datos; estos puntos corresponden al "punto de inflexión", donde la convexidad pasa a concavidad. Los dos puntos más utilizados en el análisis estadístico de son: ($M \pm 1.96DE \rightarrow 95\%$ de los datos) y ($M \pm 2.58DE \rightarrow 99\%$ de los datos).



El número de desviaciones estándar, llamado "unidades estándar" o "unidades Z", se utiliza para comparar valores de distribuciones diferentes. Por ejemplo, un estudiante de economía que obtuvo como calificación: 90 y un estudiante de derecho que obtuvo 80; las medias fueron 60 y 65, y las desviaciones estándar 15 y 5, respectivamente. En la gráfica se señala la media, las desviaciones estándar y el dato correspondiente a cada grupo, así:



Cualquier valor de una distribución puede convertirse en unidades estándar o Z, sustrayéndole la media y dividiendo esta diferencia entre la desviación estándar, así:

$$Z = \frac{x - M}{DE}$$

$$\text{Economía: } Z = \frac{90 - 60}{15} = 2$$

$$\text{Derecho: } Z = \frac{80 - 65}{5} = 3$$

Es evidente que 3 DE a la derecha de la media es mucho más extremo (está por encima del 99.7% de los demás datos) que 2 DE (que está por encima del 95% de los demás); o sea, que la nota de derecho es relativamente mejor que la de economía.