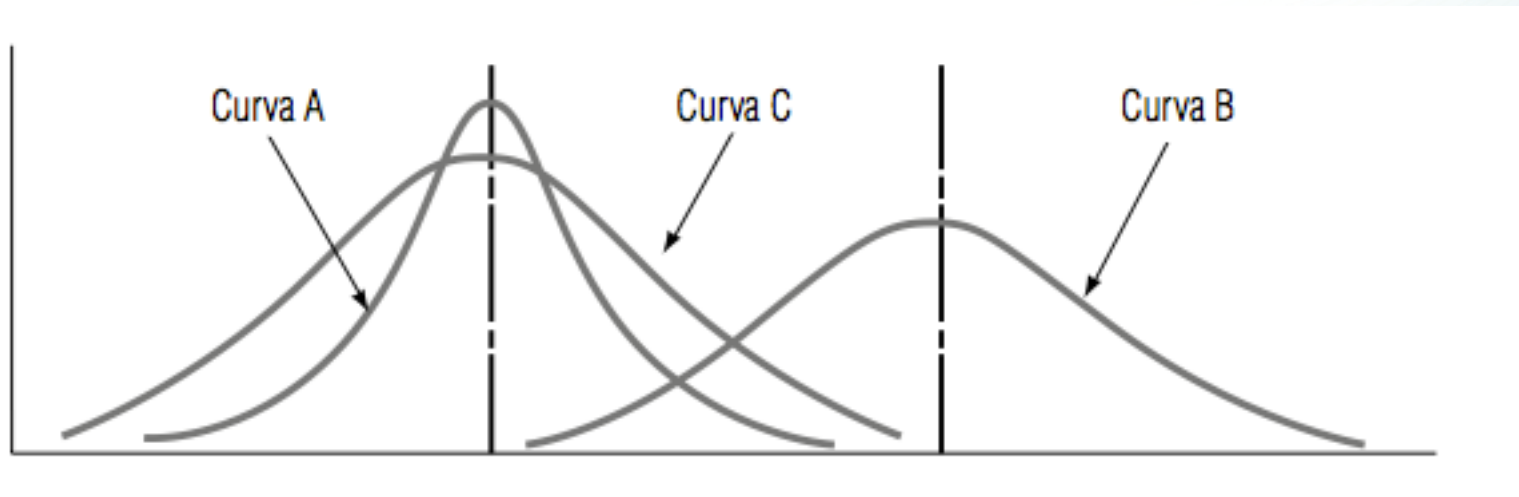


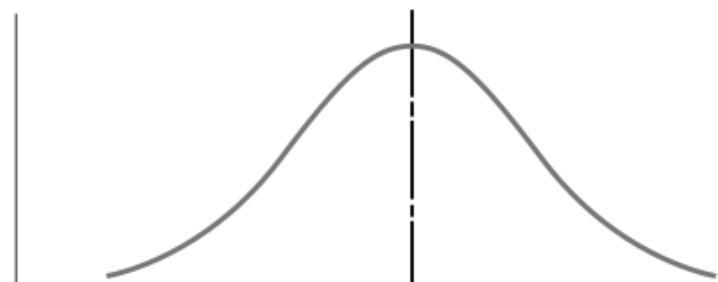
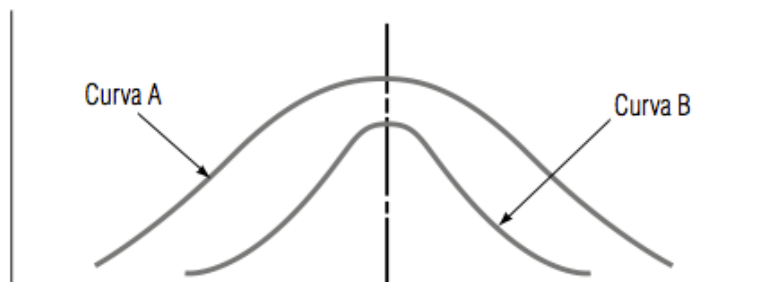
# MEDIDAS NUMÉRICAS

- Tendencia central** La tendencia central se refiere al punto medio de una distribución. Las medidas de tendencia central se conocen también como *medidas de posición*.



# MEDIDAS NUMÉRICAS

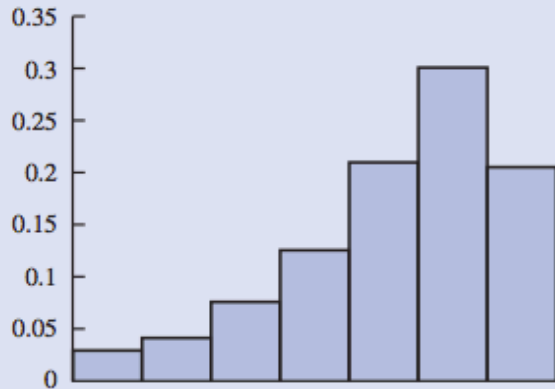
- **Dispersión** La *dispersión* se refiere a la separación de los datos en una distribución, es decir, al grado en que las observaciones se separan.



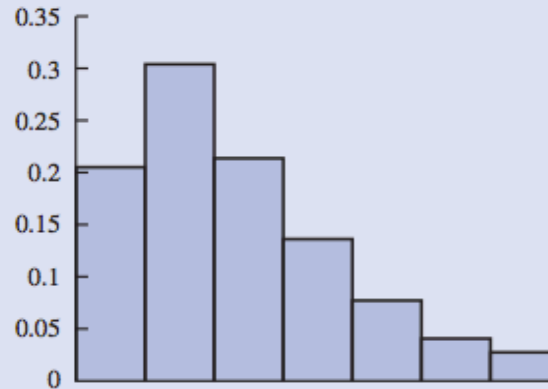
- **Sesgo** Las curvas que representan los datos puntuales de un conjunto de datos pueden ser simétricas o sesgadas.

# MEDIDAS NUMÉRICAS

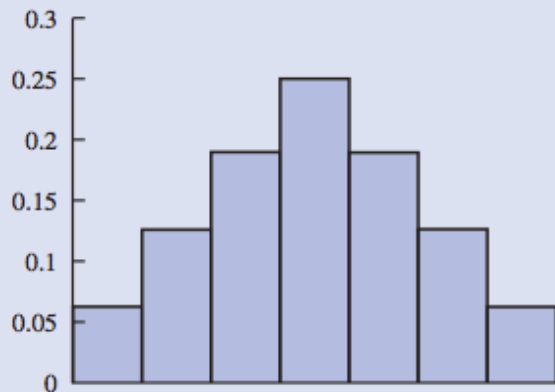
**Histograma A: Moderadamente sesgado a la izquierda**



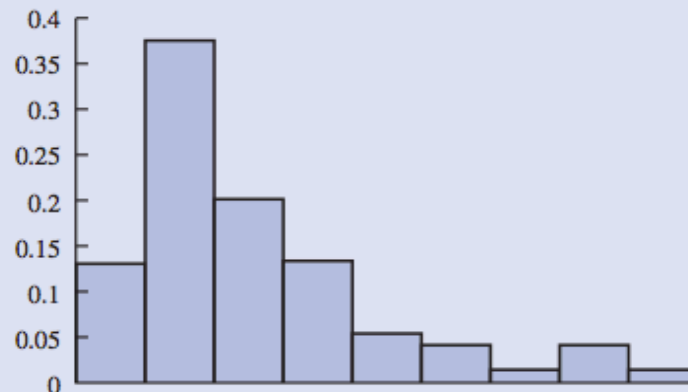
**Histograma B: Moderadamente sesgado a la derecha**



**Histograma C: Simétrico**

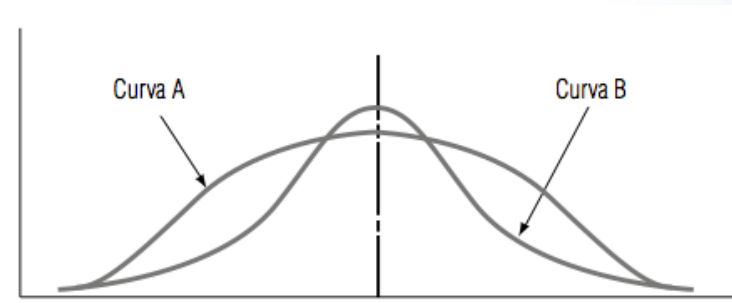
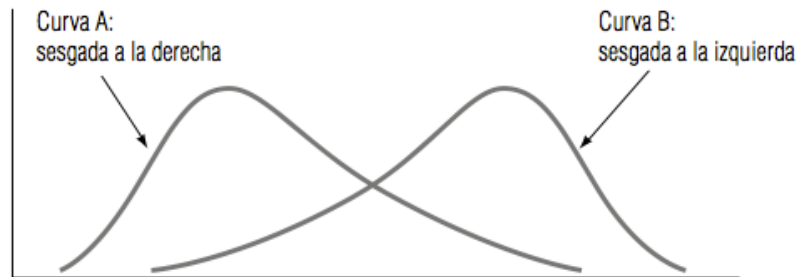


**Histograma D: Muy sesgado a la derecha**



# MEDIDAS NUMÉRICAS

- **Curtosis** Cuando medimos la *curtosis* de una distribución, estamos midiendo qué tan puntiaguda es.



- **Percentil, cuartiles y deciles** aporta información acerca de la dispersión de los datos en el intervalo que va del menor al mayor valor de los datos, acumulados en porcentajes de 1%, 25% y 10%, respectivamente.



# MEDIDAS NUMÉRICAS

- **Media o valor Promedio:** Es la medida de localización más importante de una variable. La media proporciona una medida de localización central de los datos. Su ecuación está dada por:

- Media Poblacional:  $\mu = \frac{1}{N} \sum_{i=1}^N x_i$  , donde

- **N:** Número de elementos de la población
- **X<sub>i</sub>:** valores de todas las observaciones

- Media Muestral:  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  , donde

- **n:** Número de elementos de la Muestra
- **X<sub>i</sub>:** valores de todas las observaciones

# MEDIDAS NUMÉRICAS

- **Ejemplo:** Sueldos mensuales iniciales en una muestra de 12 recién egresados de la carrera de Ingeniería.

Egresado	Sueldo mensual inicial (\$)	Egresado	Sueldo mensual inicial (\$)
1	3450	7	3490
2	3550	8	3730
3	3650	9	3540
4	3480	10	3925
5	3355	11	3520
6	3310	12	3480

# MEDIDAS NUMÉRICAS

- **Media aritmética de una muestra con datos agrupados**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n f \cdot x_i \quad , \quad \text{donde:}$$

- $\bar{x}$  : media de la muestra
- $f$  : frecuencia (número de observaciones) de cada clase  $x$
- $x_i$  : punto medio de cada clase en la muestra ( marca de clase)
- $n$  : número de observaciones en la muestra

# MEDIDAS NUMÉRICAS

- **Ejemplo:** La siguiente distribución de frecuencias representa el tiempo en segundos que los cajeros de BullsEye Discount Store necesitaron para servir a una muestra de clientes en diciembre de 2011.

Tiempo (en segundos)	Frecuencia
20- 29	6
30- 39	16
40- 49	21
50- 59	29
60- 69	25
70- 79	22
80- 89	11
90- 99	7
100-109	4
110-119	0
120-129	2



# MEDIDAS NUMÉRICAS

- La **media ponderada** nos permite calcular un promedio que toma en cuenta la *importancia* de cada valor con respecto al total.

$$\bar{x}_w = \frac{\sum(w \times x)}{\sum w}$$

- $x_w$  = símbolo para la media ponderada\*
- $w$  = peso o valor asignado a cada observación
- $x$  = *observaciones*

**Ejemplo: ...**

# MEDIDAS NUMÉRICAS

- **Mediana** es el valor de enmedio en los datos ordenados de menor a mayor (en forma ascendente).
  - Ordenar los datos de menor a mayor (en forma ascendente).
    1. Si el número de observaciones es impar, la mediana es el valor de enmedio.
    2. Si el número de observaciones es par, la mediana es el promedio de las dos observaciones de enmedio.
- **Ejemplo:** Suponga que también desea calcular la mediana del salario inicial de los 12 recién egresados de la carrera de ingeniería.
  - Primero ordena los datos de menor a mayor:

3310 3355 3450 3480 3480 3490 3520 3540 3550 3650 3730 3925

# MEDIDAS NUMÉRICAS

- **Cálculo de la mediana a partir de datos agrupados:**

$$m = \left( \frac{\left( \frac{n+1}{2} \right) - (F+1)}{f_m} \right) w + L_m \quad \text{donde,}$$

- $m$  = mediana de la muestra
- $n$  = número total de elementos de la distribución
- $F$  = suma de todas las frecuencias de clase *hasta*, pero *sin incluir*, la clase de la mediana
- $f_m$  = frecuencia de la clase de la mediana
- $w$  = ancho de intervalo de clase
- $L_m$  = límite inferior del intervalo de clase de la mediana

# MEDIDAS NUMÉRICAS

- **Ejemplo:** para la siguiente distribución de frecuencias determine la mediana e interprétela.

Clase	Frecuencia
100-149.5	12
150-199.5	14
200-249.5	27
250-299.5	58

Clase	Frecuencia
300-349.5	72
350-399.5	63
400-449.5	36
450-499.5	18



# MEDIDAS NUMÉRICAS

- La **moda** es una medida de tendencia central diferente de la media, pero un tanto parecida a la mediana.
- La **moda** es *el valor que más se repite en el conjunto de datos*.
- *Para daddos agrupados se calcula con la siguiente fiormula:*

$$M_o = L_{M_o} + \left( \frac{d_1}{d_1 + d_2} \right) w$$

- $L_{MO}$  = límite inferior de la clase modal
- $d_1$  = frecuencia de la clase modal menos la frecuencia de la clase que se encuentra *inmediatamente menor que ella*.
- $d_2$  = frecuencia de la clase modal menos la frecuencia de la clase *inmediatamente mayor que ella*
- $w$  = ancho del intervalo de la clase modal

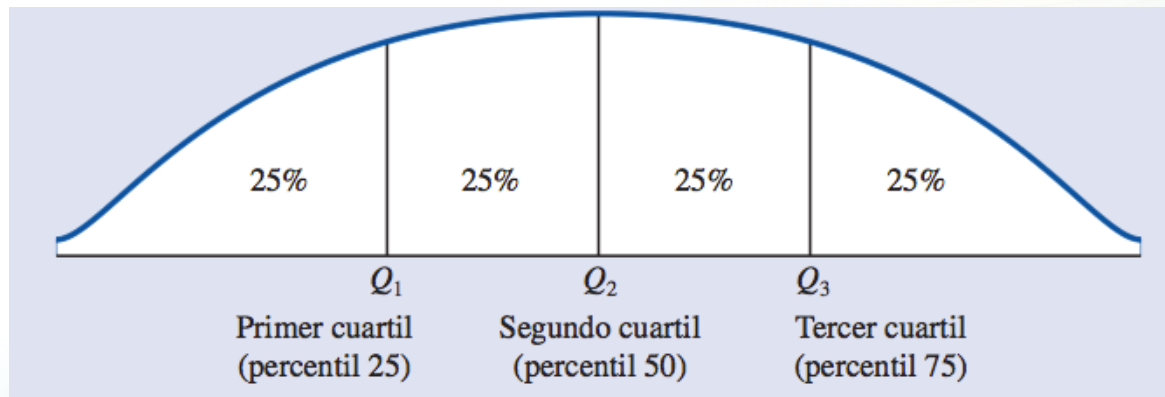
# MEDIDAS NUMÉRICAS

- **Ejemplo:** calcular la moda de la siguiente distribución de datos.

Clase	Frecuencia
100-149.5	12
150-199.5	14
200-249.5	27
250-299.5	58
300-349.5	72
350-399.5	63
400-449.5	36
450-499.5	18

# MEDIDAS NUMÉRICAS

- Unas de las medidas mas utilizadas son los Percentiles, Cuartiles y deciles, estos muestran la posición de la distribución de de los datos.
- **Percentil  $p$**  es un valor tal que por lo menos  $p$  por ciento de las observaciones son menores o iguales que este valor y por lo menos  $(100 - p)$  por ciento de las observaciones son mayores o iguales que este valor.
- Los **Cuartiles  $Q$**  sólo son percentiles determinados; así que los pasos para calcular los percentiles también se emplean para calcular los cuartiles.



# MEDIDAS NUMÉRICAS

## CÁLCULO DEL PERCENTIL $p$

– **Paso 1.** Ordenar los datos de menor a mayor (colocar los datos en orden ascendente).

– **Paso 2.** Calcular el índice  $i$  
$$i = \left( \frac{p}{100} \right) n$$

donde  $p$  es el percentil deseado y  $n$  es el número de observaciones.

– **Paso 3.** (a) Si  $i$  no es un número entero, debe redondearlo. El primer entero mayor que  $i$  denota la posición del percentil  $p$ . (b) Si  $i$  es un número entero, el percentil  $p$  es el promedio de los valores en las posiciones  $i$  e  $i + 1$ .



# MEDIDAS NUMÉRICAS

*Ejemplo:*

**Paso 1.** Ordenar los datos de menor a mayor

3310 3355 3450 3480 3480 3490 3520 3540 3550 3650 3730 3925

**Paso 2.**

$$i = \left( \frac{p}{100} \right) n = \left( \frac{85}{100} \right) \cdot 12 = 10.2$$

**Paso 3.** Como  $i$  no es un número entero, se debe *redondear*. La posición del percentil 85 es el primer entero mayor que 10.2, es la posición 11.

# MEDIDAS NUMÉRICAS

## CÁLCULO DEL CUARTIL Q

– **Paso 1.** Ordenar los datos de menor a mayor (colocar los datos en orden ascendente).

– **Paso 2.** Calcular el índice  $i$  
$$i = \left( \frac{Q}{4} \right) n$$

donde  $Q$  es el Cuartil deseado y  $n$  es el número de observaciones.

– **Paso 3.** (a) Si  $i$  no es un número entero, debe redondearlo. El primer entero mayor que  $i$  denota la posición del cuartil  $Q$ . (b) Si  $i$  es un número entero, el cuartil  $Q$  es el promedio de los valores en las posiciones  $i$  e  $i + 1$ .

# MEDIDAS NUMÉRICAS

- Los **deciles** son diez valores que distribuyen la serie de datos, ordenando en forma ascendente o descendente, en diez tramos iguales, en los que cada uno de ellos concentra el 10% de los resultados.

## CÁLCULO DEL DECIL $d$

- **Paso 1.** Ordenar los datos de menor a mayor (colocar los datos en orden ascendente).

- **Paso 2.** Calcular el índice  $i$  
$$i = \left( \frac{d}{10} \right) n$$

Donde  $d$  es el decil deseado y  $n$  es el número de observaciones.

- **Paso 3.** (a) Si  $i$  no es un número entero, debe redondearlo. El primer entero mayor que  $i$  denota la posición del percentil  $d$ . (b) Si  $i$  es un número entero, el percentil  $d$  es el promedio de los valores en las posiciones  $i$  e  $i + 1$ .

# MEDIDAS NUMÉRICAS

## CÁLCULO DE PERCENTILES, CUARTILES Y DECILES PARA DATOS AGRUPADOS.

• **Percentiles:**

$$P_i = \left( \frac{\frac{i(n+1)}{100} - (F+1)}{f_m} \right) w + L_m$$

- $P_i$  = percentil deseado.
- $i = 1, 2, 3, \dots, 100$
- $n$  = número total de elementos de la distribución
- $F$  = suma de todas las frecuencias de clase *hasta*, pero *sin incluir*, la clase donde se encuentra el percentil.
- $f_m$  = frecuencia de la clase del percentil.  
 $w$  = ancho de intervalo de clase
- $L_m$  = límite inferior del intervalo de clase donde se encuentra el percentil.



## CÁLCULO DE PERCENTILES, CUARTILES Y DECILES PARA DATOS AGRUPADOS.

- **Cuartiles :**

$$Q_i = \left( \frac{\frac{i(n+1)}{4} - (F+1)}{f_m} \right) w + L_m$$

- $Q_i$  = Cuartil deseado.
- $i = 1, 2, 3, \dots, 4$
- $n$  = número total de elementos de la distribución
- $F$  = suma de todas las frecuencias de clase *hasta*, pero *sin incluir*, la clase donde se encuentra el cuartil
- $f_m$  = frecuencia de la clase del cuartil
- $w$  = ancho de intervalo de clase
- $L_m$  = límite inferior del intervalo de clase donde se encuentra el cuartil

# MEDIDAS NUMÉRICAS

## CÁLCULO DE PERCENTILES, CUARTILES Y DECILES PARA DATOS AGRUPADOS.

- **Deciles :**

$$d_i = \left( \frac{\frac{i(n+1)}{10} - (F+1)}{f_m} \right) w + L_m$$

- $d_i$  = Decil deseado.
- $i = 1, 2, 3, \dots, 10$
- $n$  = número total de elementos de la distribución
- $F$  = suma de todas las frecuencias de clase *hasta*, pero *sin incluir*, la clase donde se encuentra el decil
- $f_m$  = frecuencia de la clase del decil
- $w$  = ancho de intervalo de clase
- $L_m$  = límite inferior del intervalo de clase donde se encuentra el decil

# MEDIDAS NUMÉRICAS

- **Ejemplo:** La distribución de frecuencias siguiente muestra los precios de las 30 acciones del Promedio Industrial Dow Jones (The *Wall Street Journal*, 16 de enero de 2006).

Precio por acción	Frecuencia
\$20–29	7
\$30–39	6
\$40–49	6
\$50–59	3
\$60–69	4
\$70–79	3
\$80–89	1

- Calcular:  $P_{17}$ ,  $Q_3$ ,  $P_{28}$ ,  $d_{50} = Q_2 = P_{50}$ , e interpretar cada uno.

## MEDIDAS DE DISPERSIÓN

- La **varianza** es una medida de variabilidad que utiliza todos los datos. La varianza está basada en la diferencia entre el valor de cada observación ( $x_i$ ) y la media

– **Varianza de población**

$$\sigma^2 = \frac{\sum (x - \mu)^2}{N} = \frac{\sum x^2}{N} - \mu^2$$

$\sigma^2$  = Varianza de la Población

$x$  = elemento u observación

$\mu$  = media de la Población

$N$  = numero total de datos.



# MEDIDAS NUMÉRICAS

## – Varianza de una muestra:

$$s^2 = \frac{\sum (x - \bar{x})^2}{n-1} = \frac{\sum x^2}{n-1} - \frac{n\bar{x}^2}{n-1}$$

$s^2$  = Varianza de la muestra

$x$  = elemento u observación

$\bar{x}$  = media de la muestra

$n$  = numero total de datos.

# MEDIDAS NUMÉRICAS

- **Varianza de datos agrupados**

- **Varianza de población**

$$\sigma^2 = \frac{\sum f \cdot (x - \mu)^2}{N} = \frac{\sum fx^2}{N} - \mu^2$$

$\sigma^2$  = Varianza de la Población

$x$  = elemento u observación

$\mu$  = media de la Población

$f$  = frecuencia de cada una de las clases

$N$  = numero total de datos.

## Varianza de datos agrupados

### – Varianza de una muestra:

$$s^2 = \frac{\sum f \cdot (x - \bar{x})^2}{n - 1} = \frac{\sum f \cdot x^2}{n - 1} - \frac{n \bar{x}^2}{n - 1}$$

$s^2$  = Varianza de la muestra

$x$  = elemento u observación

$\bar{x}$  = media de la muestra

$f$  = frecuencia de cada una de las clases

$n$  = numero total de datos.

# MEDIDAS NUMÉRICAS

- La **desviación estándar** se define como la raíz cuadrada positiva de la varianza.

– **Desviación Estandar Poblacional** :  $\sigma = \sqrt{\sigma^2}$

– **Deviación Estandar Muestral**:  $s = \sqrt{s^2}$

– **Coeficiente de Variabilidad** :  $cv = \left(\frac{\sigma}{\mu}\right) * 100$  o  $cv = \left(\frac{s}{x}\right) * 100$



# MEDIDAS NUMÉRICAS

- **Ejemplo 1:** Los siguientes datos son una muestra de la tasa de producción diaria de botes de fibra de vidrio de la Hy-droport, Ltd., un fabricante de Miami: 17 21 18 27 17 21 20 22 18 23

El gerente de producción de la compañía siente que una desviación estándar de más de tres botes por día indica variaciones de tasas de producción inaceptables. ¿Deberá preocuparse por las tasas de producción de la planta?

- **Ejemplo 2:** Basart Electronics piensa emplear uno de dos programas de capacitación. Se capacitó a dos grupos para la misma tarea. El grupo 1 recibió el programa A; el grupo 2, el B. Para el primer grupo, los tiempos requeridos para capacitar a los empleados tuvieron un promedio de 32.11 horas y una varianza de 68.09. En el segundo grupo, el promedio fue 19.75 horas y la varianza fue 71.14. ¿Qué programa de capacitación tiene menos variabilidad relativa en su desempeño?

# MEDIDAS NUMÉRICAS

- **Ejemplo 3:** La edad de los estudiantes regulares que acuden a un curso en los turnos matutino y vespertino del nivel licenciatura de la Universidad Central se describe en las siguientes dos muestras:
  - **Turno matutino** 23 29 27 22 24 21 25 26 27 24
  - **Turno vespertino** 27 34 30 29 28 30 34 35 28 29

Si la homogeneidad de la clase es un factor positivo en el aprendizaje, utilice una medida de variabilidad relativa para sugerir en cuál de los dos grupos será más fácil enseñar.

# MEDIDAS NUMÉRICAS

- **Ejemplo 4:** El número de cheques cobrados diariamente en las cinco sucursales del Bank of Orange County durante el mes anterior tuvo la siguiente distribución de frecuencias:

Clase	Frecuencia
0-199	10
200-399	13
400-599	17
600-799	42
800-999	18

- Hank Spivey, director de operaciones del banco, sabe que una desviación estándar en el cobro de cheques mayor que 200 cheques diarios ocasiona problemas de personal y de organización en las sucursales, debido a la carga de trabajo dispareja. ¿Deberá preocuparse por la cantidad de empleados que van a utilizar el mes siguiente?