



# Medidas de Dispersión



Dr. Servio Interiano  
Dra. Claudeth Recinos  
2010



# Descripción de un grupo



- Las medidas de tendencia central nos permiten describir de forma general la localización de un grupo de datos en la escala de medición.

# Descripción de un grupo



- Sin embargo para tener una descripción más adecuada del grupo, es necesario complementarlas con las medidas de dispersión.

# Descripción de un grupo



- Sin embargo para tener una descripción más adecuada del grupo, es necesario complementarlas con las medidas de dispersión.
- “Qué tan compacto es el grupo”

- 
- 
- 

# Dispersión

- Si todos nuestros datos tuvieran el mismo valor, producirían una gráfica muy delgada.

- 
- 
- 

# Dispersión

- Si todos nuestros datos tuvieran el mismo valor, producirían una gráfica muy delgada.
- Por ejemplo, si todos pesaran 150 lbs.



- 
- 
- 

- Sin embargo esto casi nunca lo encontramos en la vida real.
- Si tomamos los pesos de un grupo de personas, estos variarán dependiendo del individuo.

- 
- 
- 

- Sin embargo esto casi nunca lo encontramos en la vida real.

- Si tomamos los pesos de un grupo de personas, estos variarán dependiendo del individuo.



- 
- 
- 



Dándonos como resultado un gráfico  
mucho más disperso.



- 
- 
- 

## Medidas de dispersión

- En estadística, las medidas que nos describen lo compacto o disperso que se encuentra un grupo de datos, son:

“las medidas de dispersión”.

- 
- 
- 

## Medidas de dispersión

- En estadística, las medidas que nos describen lo compacto o disperso que se encuentra un grupo de datos, son:

“las medidas de dispersión”.

- Rango
- Rango intercuartil
- Desviación cuartil
- Desviación media
- Varianza
- Desviación estándar

- 
- 
- 

## Medidas de dispersión

- Dispersión en base a la mediana
- Rango
- Rango intercuartil
- Desviación cuartil

- 
- 
- 

# Medidas de dispersión

- Desviación media
- Varianza
- Desviación estándar
- Dispersión en base a la media aritmética

- 
- 
- 

## Medidas de dispersión

- Dispersión en base a la mediana
  - Rango
  - Rango intercuartil
  - Desviación cuartil
- Dispersión en base a la media aritmética
  - Desviación media
  - Varianza
  - Desviación estándar

- 
- 
- 

## Rango

- En toda distribución, los límites inferior y superior determinan el recorrido de los datos.
- Por esto el rango (diferencia entre el  $L_s$  y  $L_i$ ) es una medida de dispersión de los datos.

- 
- 
- 

## Rango

- Es la medida de dispersión más fácil de obtener, sin embargo es poco utilizada debido a que es muy influenciado por valores extremos de poca frecuencia.

X	A	B
40	1	2
41	2	9
42	3	7
43	18	6
44	10	5
45	4	4
46	3	3
47	0	2
48	0	1
49	0	1
50	1	2
	42	42

- Estas series de datos tienen la misma media aritmética,

$$\bar{x} = \sum fx/n = 1830/42$$

$$\bar{x} = 43.57$$

X	A	B
40	1	2
41	2	9
42	3	7
43	18	6
44	10	5
45	4	4
46	3	3
47	0	2
48	0	1
49	0	1
50	1	2
	42	42

- Estas series de datos tienen la misma media aritmética,

$$\bar{x} = \sum fx/n = 1830/42$$

$$\bar{x} = 43.57$$

- Tienen también el mismo rango,

$$R = 50 - 40 = 10$$

X	A	B
40	1	2
41	2	9
42	3	7
43	18	6
44	10	5
45	4	4
46	3	3
47	0	2
48	0	1
49	0	1
50	1	2
	42	42

- $\bar{x} = 43.57$
- $R = 50 - 40 = 10$
- Sin embargo si eliminamos el valor extremo en el grupo A, veremos que realmente son diferentes.

X	A	B
40	1	2
41	2	9
42	3	7
43	18	6
44	10	5
45	4	4
46	3	3
47	0	2
48	0	1
49	0	1
50	<del>1</del>	2
	42	42

- $\bar{x} = 43.57$
- $R = 50 - 40 = 10$
- Sin embargo si eliminamos el valor extremo en el grupo A, veremos que realmente son diferentes.

- 
- 
- 

# Cuartiles y percentiles

- En una serie de datos, la mediana ocupa el punto central de la serie, dividiéndola en dos partes iguales



- 
- 
- 

## Cuartiles y percentiles

- Si dividimos cada mitad de la serie en dos partes iguales, tendremos nuestra serie dividida en cuatro segmentos iguales, separados por los llamados “cuartiles”.



- 
- 
- 

## Cuartiles y percentiles

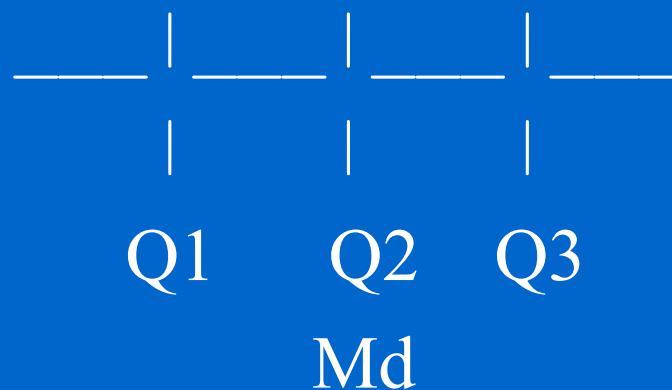
- Si dividimos cada mitad de la serie en dos partes iguales, tendremos nuestra serie dividida en cuatro segmentos iguales, separados por los llamados “cuartiles”.



- 
- 
- 

# Cuartiles y percentiles

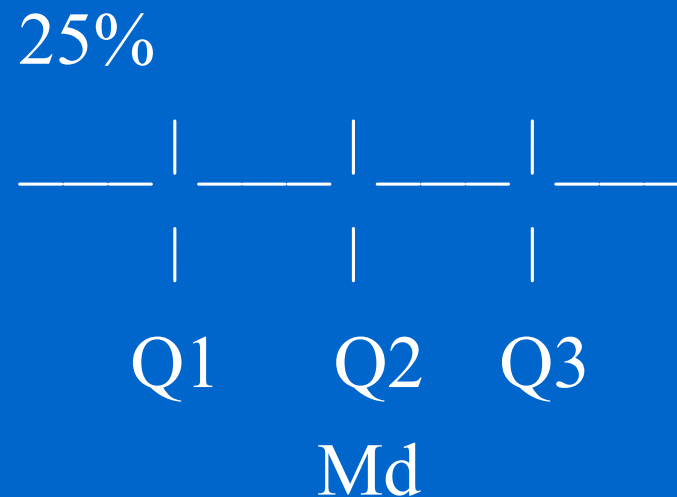
- Si dividimos cada mitad de la serie en dos partes iguales, tendremos nuestra serie dividida en cuatro segmentos iguales, separados por los llamados “cuartiles”.



- 
- 
- 

# Cuartiles y percentiles

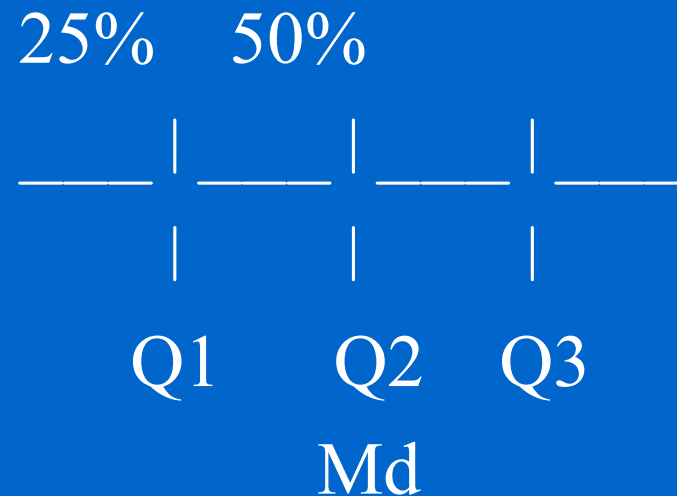
- Como podemos observar, “por debajo” del primer cuartil se encuentra el 25 % de las observaciones.



- 
- 
- 

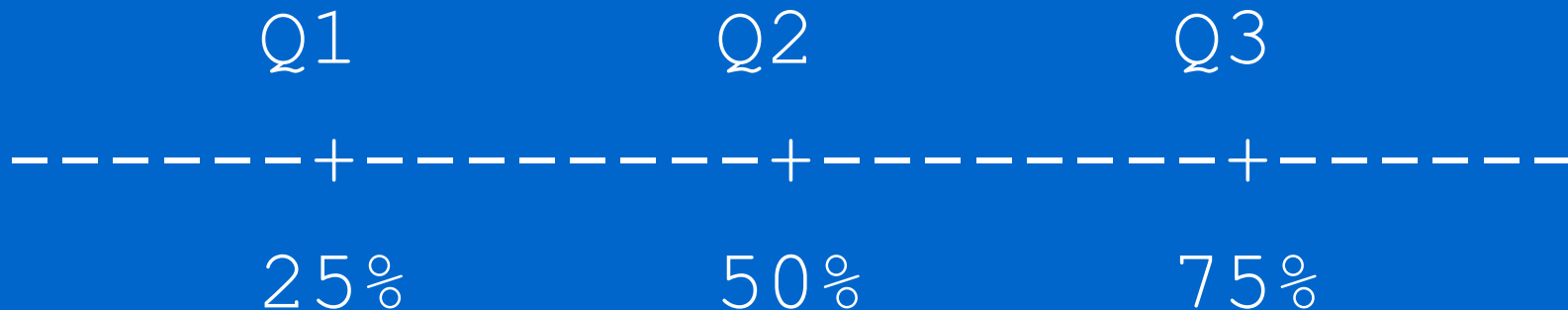
# Cuartiles y percentiles

- Como podemos observar, “por debajo” del primer cuartil se encuentra el 25 % de las observaciones.
- Debajo del Q2, están el 50% de las observaciones.



•  
•  
•

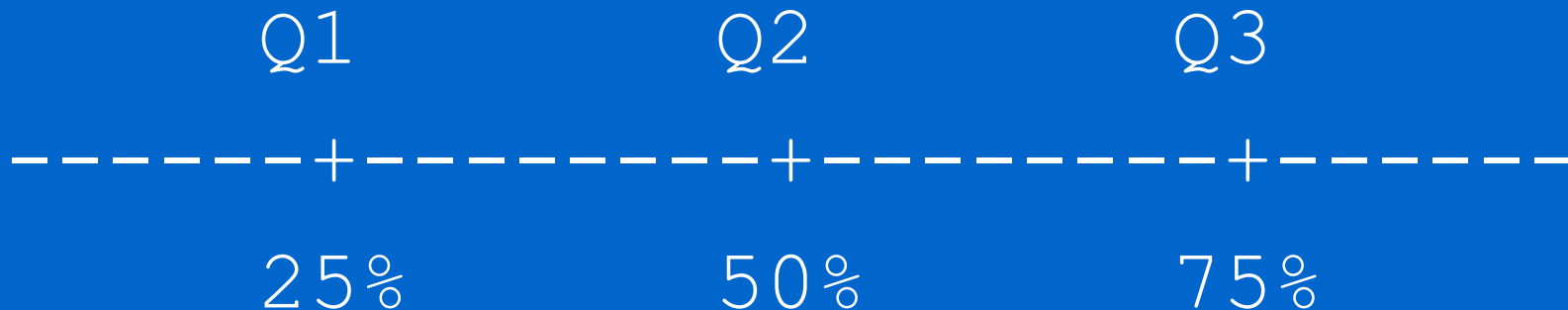
## Cuartiles y percentiles



- Esto da lugar al concepto de percentil.
- Si dividimos la serie en 100 partes iguales, un percentil en particular dejará un porcentaje específico de datos por debajo de él.

•  
•  
•

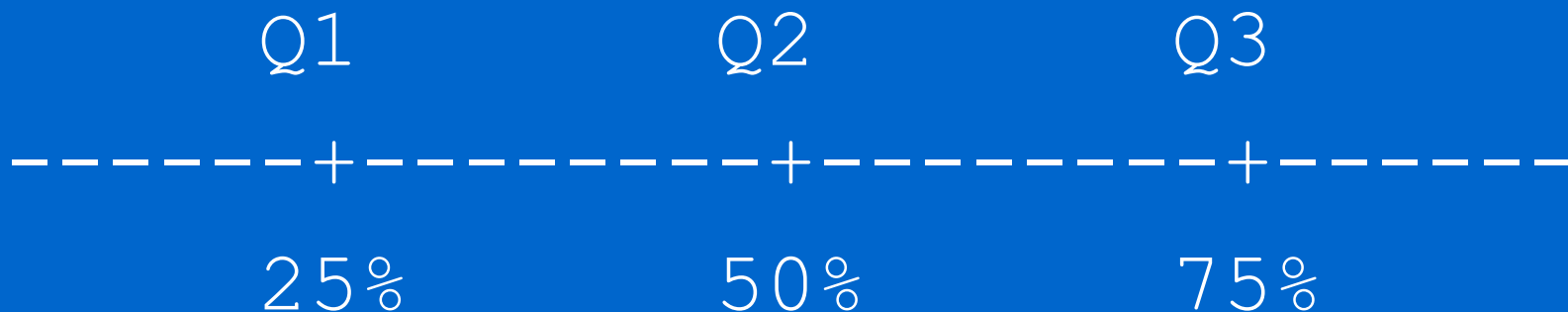
## Cuartiles y percentiles



- De tal forma que el percentil 25 ( $P_{25}$ ) dejará un 25% debajo de él. El percentil 60 ( $P_{60}$ ) dejará un 60% por debajo de él, etc.

•  
•  
•

# Cuartiles y percentiles



- El Q1 corresponde con el percentil 25
- El Q2 corresponde con el percentil 50 y la mediana.
- El Q3 corresponde con el percentil 75

- 
- 
- 

# Rango Intercuartil

A grandes rasgos, la idea es dividir los datos en cuatro grupos iguales y ver que tan separados se encuentran los grupos de los extremos.



- 
- 
- 

# Rango Intercuartil



$$Q = Q3 - Q1$$

- Para el cálculo de los cuartiles, se utiliza el método de interpolación lineal.

- 
- 
- 

Estaturas	f
123.5 - 128.5	1
128.5 - 133.5	4
133.5 - 138.5	9
138.5 - 143.5	24
143.5 - 148.5	29
148.5 - 153.5	22
153.5 - 158.5	14
158.5 - 163.5	5
	108

- El primer cuartil, se encuentra en la posición “n/4”

- 
- 
- 

Estaturas	f
123.5 - 128.5	1
128.5 - 133.5	4
133.5 - 138.5	9
138.5 - 143.5	24
143.5 - 148.5	29
148.5 - 153.5	22
153.5 - 158.5	14
158.5 - 163.5	5
	108

- El primer cuartil, se encuentra en la posición “n/4”
- $108 / 4 = 27$

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	
148.5 - 153.5	22	
153.5 - 158.5	14	
158.5 - 163.5	5	
	108	

- El primer cuartil, se encuentra en la posición “n/4”
- $108 / 4 = 27$
- La posición 27 se encuentra en el 4º intervalo de clase

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	
148.5 - 153.5	22	
153.5 - 158.5	14	
158.5 - 163.5	5	
	108	

- Luego dividimos la amplitud (5) entre la frecuencia del intervalo (24).
- Calculamos, cuanto nos falta desde el inicio del intervalo para alcanzar  $n/4$ .

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	
148.5 - 153.5	22	
153.5 - 158.5	14	
158.5 - 163.5	5	
	108	

- Luego dividimos la amplitud (5) entre la frecuencia del intervalo (24).
- Calculamos, cuanto nos falta desde el inicio del intervalo para alcanzar  $n/4$ .

$$27 - 14 = 13$$

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	
148.5 - 153.5	22	
153.5 - 158.5	14	
158.5 - 163.5	5	
	108	

- Averiguamos el valor del segmento hasta  $n/4$ .

$$13(5/24) = 2.71$$

- Lo sumamos al límite inferior del intervalo  
 $138.5 + 2.71 = 141.21$
- $Q_1 = 141.21$

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	67
148.5 - 153.5	22	89
153.5 - 158.5	14	103
158.5 - 163.5	5	108
	108	

- Luego buscamos la posición del 2° cuartil.

$$(n/4) * 2 =$$

$$27 * 2 = 54$$

- Y luego la de Q3

$$(n/4) * 3 = 81$$

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	67
148.5 - 153.5	22	89
153.5 - 158.5	14	103
158.5 - 163.5	5	108
	108	

- Y repetimos el método de interpolación lineal en cada caso.

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	67
148.5 - 153.5	22	89
153.5 - 158.5	14	103
158.5 - 163.5	5	108
	108	

- Y repetimos el método de interpolación lineal en cada caso.
- $Q1 = 141.21$
- $Q2 = 146.26$
- $Q3 = 151.68$

- 
- 
- 

Estaturas	f	fa
123.5 - 128.5	1	1
128.5 - 133.5	4	5
133.5 - 138.5	9	14
138.5 - 143.5	24	38
143.5 - 148.5	29	67
148.5 - 153.5	22	89
153.5 - 158.5	14	103
158.5 - 163.5	5	108
	108	

- $Q1 = 141.21$
- $Q2 = 146.26$
- $Q3 = 151.68$

$$Q = Q3 - Q1$$

$$Q = 151.68 - 141.21$$

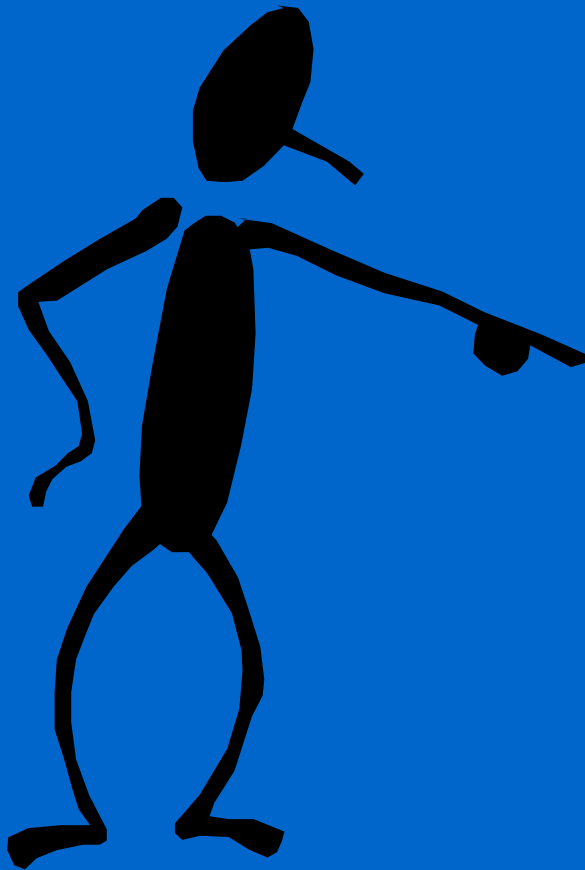
$$Q = 10.47 \text{ cm}$$

- 
- 
- 

## Inconvenientes del Rango Inter cuartil

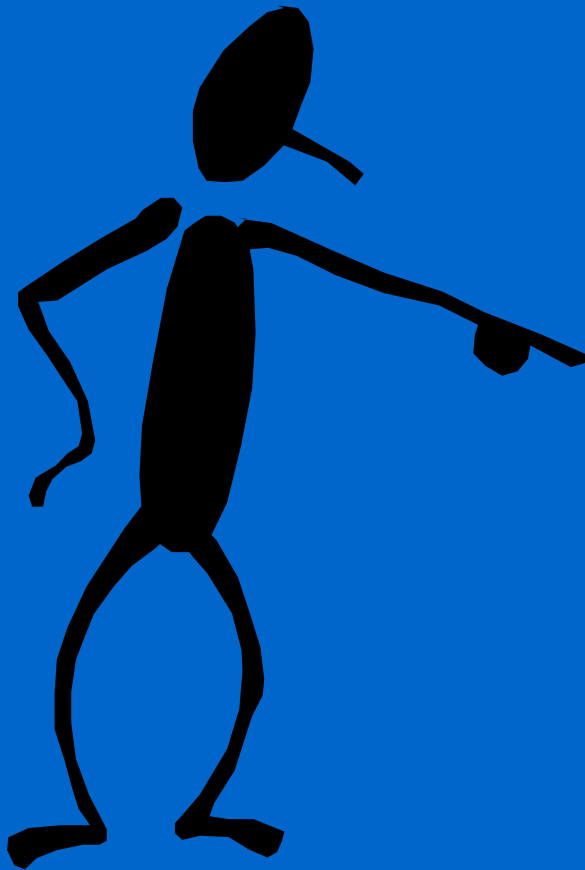
- No es posible, conociendo solo  $Q$ , ubicar con precisión una observación en particular dentro de la distribución.
- Al igual que la mediana (que es  $Q_2$ ), no tiene propiedades que le permitan intervenir en operaciones algebraicas propias de la estadística inferencial.

# Desviación Cuartil



- Al dividir el rango intercuartil (Q) dentro de dos, obtenemos la desviación cuartil.
- $Q/2 = (Q3 - Q1) / 2$

# Desviación Cuartil



- $Q/2 = (Q3 - Q1) / 2$
- Nos indica la dispersión que tienen los datos alrededor de la mediana.

- 
- 
- 

- Estas medidas de dispersión tienen aplicación, sobre todo en la estadística no paramétrica, o cuando el grupo presenta demasiada asimetría.
- Para grupos con comportamiento “normal”, se prefiere el uso de medidas de dispersión con respecto a la media aritmética.

- 
- 
- 

Medidas de Dispersión con respecto a la media aritmética.

- Una de las propiedades importantes de la media aritmética, es que la sumatoria de las desviaciones de los datos con respecto a si misma es cero.

$$\sum (x_i - \bar{x}) = 0$$

- 
- 
- 

## Medidas de Dispersión con respecto a la media aritmética.



- En un principio, ...
- “La mejor manera de mostrar la dispersión de los datos, es calcular cuanto se desvían en promedio los datos de la media”
- Es decir, “cuanto es el promedio de desviación”

- 
- 
- 

## Medidas de Dispersión con respecto a la media aritmética.



- Sin embargo al querer calcular el promedio de desviaciones se encontró que:

$$\text{Promedio} = \frac{\sum (x_i - \bar{x})}{n}$$

$$\text{Promedio} = 0$$

- 
- 
- 

## Desviación media

- Para poder calcular la desviación media, se eliminan los signos de las diferencias de los datos con respecto a la media, aplicando el valor absoluto.

$$\text{D.M.} = \frac{\sum |x_i - \bar{x}|}{n}$$

- 
- 
- 

## Desviación media

$$D.M. = \frac{\sum |x_i - \bar{x}|}{n}$$

- Es una medida objetiva, cuanto mayor sea su valor mayor será la dispersión de los datos.

- 
- 
- 

## Desviación media para datos agrupados

- Para el cálculo en datos agrupados, asumimos al igual que para la media, que los datos se encuentran en el centro del intervalo, y tienen todos, el valor de la marca de clase

•  
•  
•

## Desviación media para datos agrupados

$$\text{D.M.} = \frac{\sum |x_i - \bar{x}|}{n}$$

- Así, sustituimos  $x_i$  por la marca de clase y la multiplicamos por la frecuencia.

•  
•  
•

## Desviación media para datos agrupados

$$\text{D.M.} = \frac{\sum |x_i - \bar{x}|}{n}$$

$$\text{D.M.} = \frac{\sum f |x - \bar{x}|}{n}$$

- Así, sustituimos  $x_i$  por la marca de clase y la multiplicamos por la frecuencia.

- 
- 
- 

## Desviación media para datos agrupados

$$\text{D.M.} = \frac{\sum |x_i - \bar{x}|}{n}$$

- Series simples

---

$$\text{D.M.} = \frac{\sum f |x - \bar{x}|}{n}$$

- Datos agrupados

- 
- 
- 

# Ejemplo

Int	x	f
27-29	28	1
30-32	31	10
33-35	34	14
36-38	37	33
39-41	40	14
42-44	43	7
45-47	46	3
Totales		82

$$DM. = \frac{\sum f|x-\bar{x}|}{n}$$

- 
- 
- 

# Ejemplo

Int	x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
27-29	28	1	28	9	9
30-32	31	10	310	6	60
33-35	34	14	476	3	42
36-38	37	33	1221	0	0
39-41	40	14	560	3	42
42-44	43	7	301	6	42
45-47	46	3	138	9	27
Totales		82	3034		222

$$DM. = \frac{\sum f|x-\bar{x}|}{n}$$

- 
- 
- 

# Ejemplo

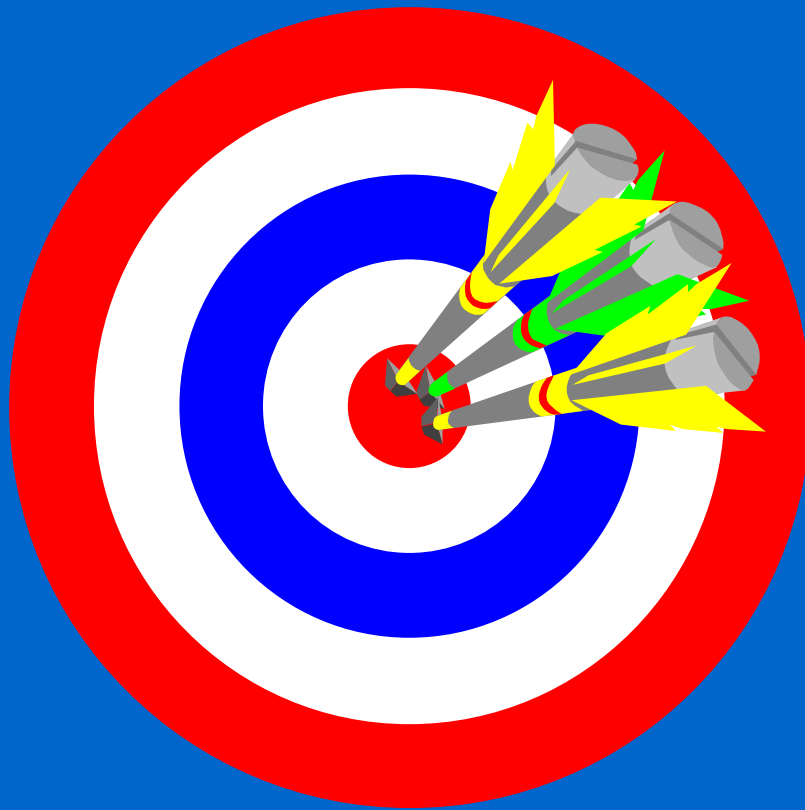
Int	x	f	fx	$ x-\bar{x} $	f $ x-\bar{x} $
27-29	28	1	28	9	9
30-32	31	10	310	6	60
33-35	34	14	476	3	42
36-38	37	33	1221	0	0
39-41	40	14	560	3	42
42-44	43	7	301	6	42
45-47	46	3	138	9	27
Totales		82	3034		222

$$DM. = \frac{\sum f|x-\bar{x}|}{n}$$

$$DM = 222 / 82$$

$$DM = 2.71$$

# Ejercicio



- Dos personas compiten para ser seleccionadas a un torneo de dardos.
- Se numeran los círculos de la diana, para registrar los resultado de cada uno de los competidores.

- 
- 
- 

# Ejercicio

A

x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
5	1	5	2	2
4	8	32	1	8
3	14	42	0	0
2	5	10	1	5
1	1	1	2	2
		29	90	17

B

x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
5	4	20	2	8
4	9	36	1	9
3	7	21	0	0
2	5	10	1	5
1	4	4	2	8
		29	91	30

- 
- 
- 

# Ejercicio

A

x	f	fx	$ x-\bar{x} $	f $ x-\bar{x} $
5	1	5	2	2
4	8	32	1	8
3	14	42	0	0
2	5	10	1	5
1	1	1	2	2
	29	90		17

- $\bar{X} = 3.10$

- $DM = 0.59$

- 
- 
- 

# Ejercicio

A

x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
5	1	5	2	2
4	8	32	1	8
3	14	42	0	0
2	5	10	1	5
1	1	1	2	2
		29	90	17

B

x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
5	4	20	2	8
4	9	36	1	9
3	7	21	0	0
2	5	10	1	5
1	4	4	2	8
		29	91	30

- 
- 
- 

# Ejercicio

- $\bar{X} = 3.14$

- $DM = 1.03$

B

x	f	fx	$ x - \bar{x} $	$f x - \bar{x} $
5	4	20	2	8
4	9	36	1	9
3	7	21	0	0
2	5	10	1	5
1	4	4	2	8
		91	30	

- 
- 
- 

# Ejercicio

A

x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
5	1	5	2	2
4	8	32	1	8
3	14	42	0	0
2	5	10	1	5
1	1	1	2	2
		29	90	17

B

x	f	fx	$ x-\bar{x} $	$f x-\bar{x} $
5	4	20	2	8
4	9	36	1	9
3	7	21	0	0
2	5	10	1	5
1	4	4	2	8
		29	91	30

- 
- 
- 

# Ejercicio

A

- $\bar{X} = 3.10$
- $DM = 0.59$

B

x	f	fx	$ x - \bar{x} $	f $ x - \bar{x} $
5	4	20	2	8
4	9	36	1	9
3	7	21	0	0
2	5	10	1	5
1	4	4	2	8
		91		
		29	30	

- 
- 
- 

## Ejercicio

A

- $\bar{X} = 3.10$

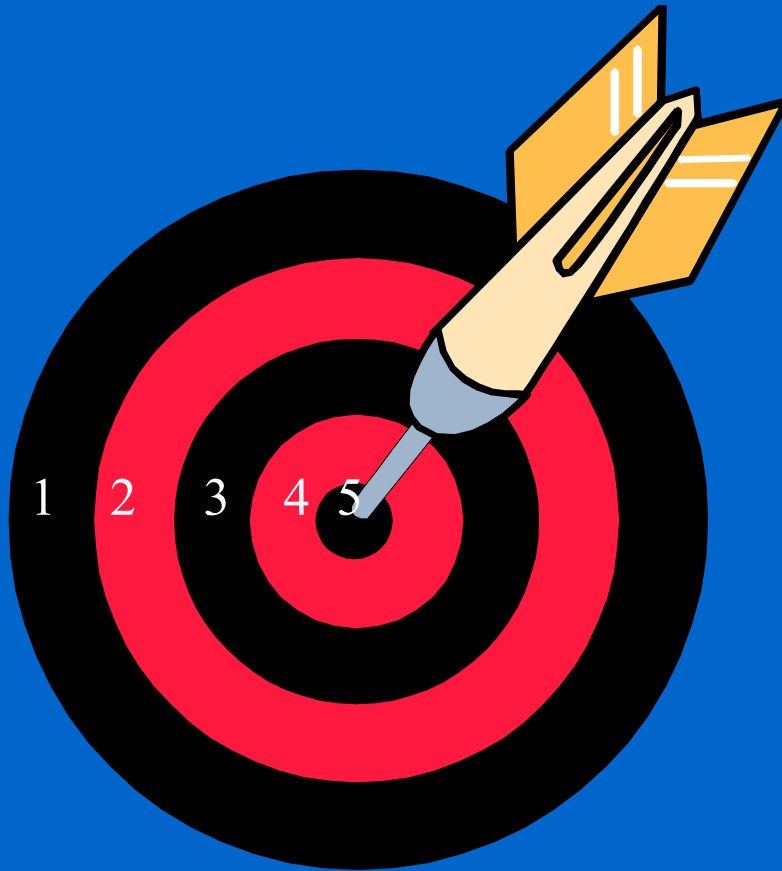
- $DM = 0.59$

B

- $\bar{X} = 3.14$

- $DM = 1.03$

# Ejercicio



- $\bar{X} = 3.10$   
A

- $DM = 0.59$

---

- $\bar{X} = 3.14$   
B

- $DM = 1.03$



# Varianza



- Algunas personas, (matemáticos y estadísticos), no estaban conformes con emplear el “valor absoluto” en el cálculo de la desviación media, por su “falta de fundamento”.





# Varianza



- Buscando la forma de eliminar los signos de los resultados de  $(x - \bar{x})$  se consideró utilizar el valor cuadrado.
- Dando como resultado que  $(x - \bar{x})^2$ , siempre tiene un resultado positivo.



- 
- 
- 

# La Varianza

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

- De tal forma que la varianza es:  
El promedio del cuadrado de las desviaciones de los datos con respecto a su media aritmética.

- 
- 
- 

## La varianza

- Cuando se calcula para una muestra, la varianza tiende a dar un resultado que resulta más pequeño que el verdadero parámetro.
- Es decir que es un estimador sesgado.

- 
- 
- 

## La varianza

- Como el error de estimación se multiplica al cuadrado  $\{(x - \mu)^2\}$ , es necesario hacer una compensación para que cuando se calcule el estadístico, éste se parezca al parámetro.
- Esta compensación se hace substituyendo la “N”, en la fórmula, por los grados de libertad “n-1”.

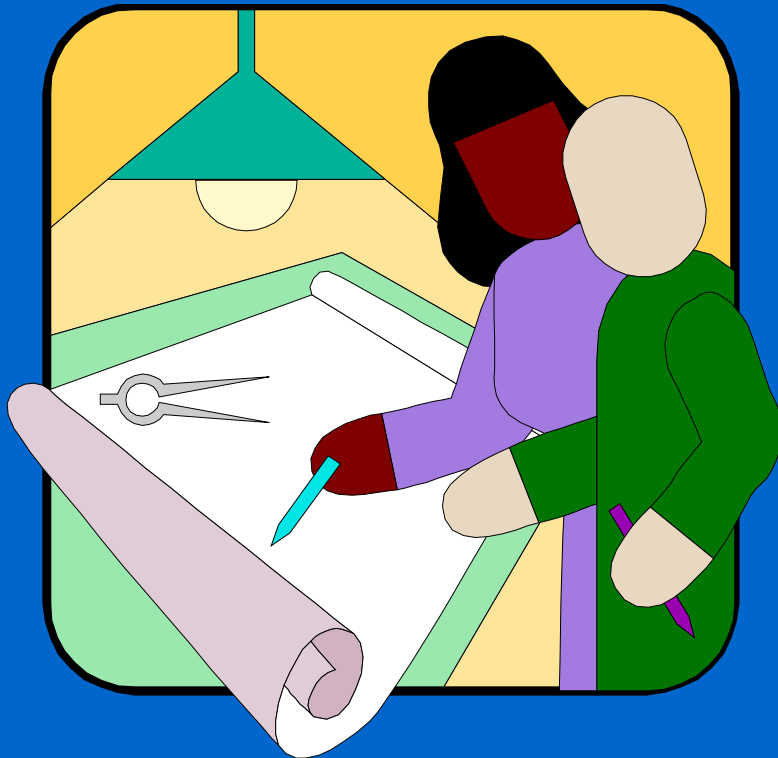
- 
- 
- 

# La Varianza

- $\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$
- Varianza para población (parámetro)

- 
- $s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$
  - Varianza para muestra (estadístico)

# La varianza para datos agrupados



- Cuando trabajamos con datos agrupados, procedemos al igual que con la desviación media.
- Substituimos “ $x_i$ ” por la marca de clase “ $x$ ” y multiplicamos por la frecuencia.

- 
- 
- 

# Varianza para datos agrupados

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

- Varianza para población (parámetro)

---

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

- Varianza para muestra (estadístico)

# Desviación estándar



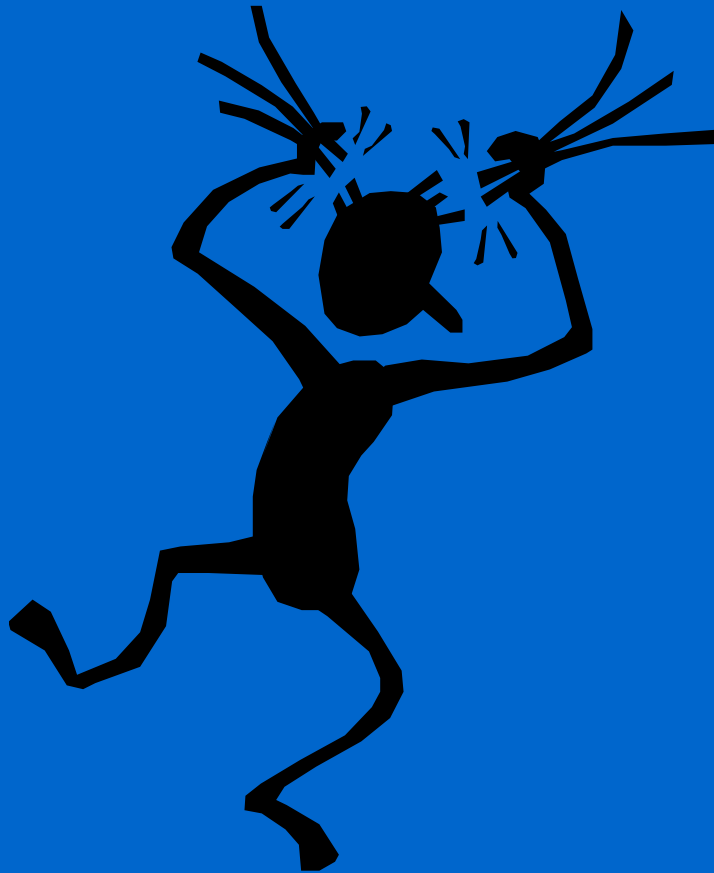
- La varianza responde a las necesidades de los matemáticos, y a mayor varianza mayor dispersión.
- Parece que todo está solucionado, ... Pero la varianza tiene un gran problema.

# Desviación estándar



- La varianza al utilizar el cuadrado de la diferencia, eleva al cuadrado las dimensionales.
- Lo que nos da una dispersión en una dimensional distinta a la de la tendencia central.

# Desviación estándar



- Así, si medimos alturas en metros (m.), la varianza nos dará dispersión en metros cuadrados (m<sup>2</sup>).

- 
- 
- 

# Desviación Estándar

- Para resolver este problema, se le aplica la raíz cuadrada a la varianza, y obtenemos como resultado:

La desviación estándar

- $\sigma = \sqrt{\sigma^2}$

- $\sigma = \sqrt{\frac{\Sigma f(x - \mu)^2}{N}}$

- 
- 
- 

# La desviación estándar

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

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

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

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

- 
- 
- 



Inter	x	f	fx	$ x-\bar{x} $	f $ x-\bar{x} $	$(x-\bar{x})^2$	f $(x-\bar{x})^2$
27-29	28	1	28	9	9	81	81
30-32	31	10	310	6	60	36	360
33-35	34	14	476	3	42	9	126
36-38	37	33	1221	0	0	0	0
39-41	40	14	560	3	42	9	126
42-44	43	7	301	6	42	36	252
45-47	46	3	138	9	27	81	243
Totales		82	3034		222		1188



- 
- 
- 
- 
- 
- 
- 
-

- 
- 
- 

## Dispersión relativa

- Cuando las variables que se estudian en una investigación no poseen unidades de medida iguales, no es posible compararlas utilizando las medidas de dispersión absoluta, que se han revisado anteriormente.

# Dispersión relativa



- Por ejemplo si comparamos dos grupos cuyos pesos están medidos uno en lbs. y el otro en Kg.

# Dispersión relativa



- La media y la desviación estándar de estos datos no nos darán información útil, ya que “la regla” que se utilizó para medir cada grupo fue diferente.

•  
•  
•

## Dispersión relativa

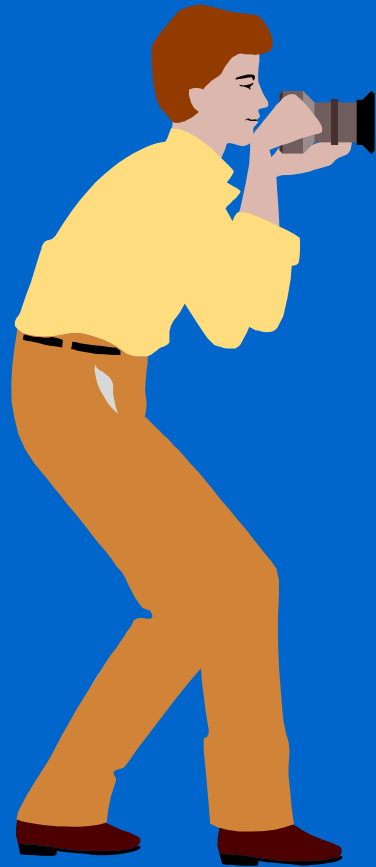
- Para poder comparar grupos de datos expresados en medidas diferentes, aplicamos la dispersión relativa que es adimensional.
- La desviación relativa no es mas que dividir la dispersión absoluta dentro de la media aritmética.



- La dispersión relativa es aplicable a cualquier medida de dispersión absoluta.

- Dispersión relativa de la desviación media

$$DM / \mu \quad \text{ó} \quad DM / \bar{x}$$



- La dispersión relativa es aplicable a cualquier medida de dispersión absoluta.

- Dispersión relativa de la varianza

$$s^2 / \bar{x} \quad \text{ó} \quad \sigma^2 / \mu$$



- Cuando se aplica a la desviación estándar, tiene nombre propio.
- Coeficiente de variación

$$\text{C.V.} = s / \bar{x}$$

$$\text{C.V.} = \sigma / \mu$$



- Dudas o preguntas ?

