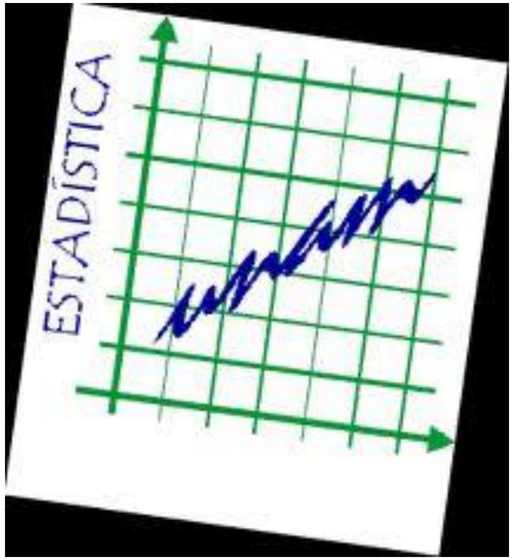
Introducción a la Estadística

Módulo III y IV

Mtro. Carlos Alberto Salgado Treviño
Plantel Texcoco de la Escuela Preparatoria

Índice

1. Coeficiente de Variación
2. Asimetría
3. Covarianza
4. Correlación Lineal
5. Recta de Regresión



TEMA: COEFICIENTE DE VARIACIÓN, COVARIANZA Y RECTA DE REGRESIÓN

PROPÓSITO

Comprender el uso y aplicación de las medidas de dispersión e interpretar el coeficiente de variación, la asimetría desviación media, varianza, desviación estándar y recta de regresión; para el cálculo y apreciación de la dispersión de los datos para su análisis.

COMPETENCIA DISCIPLINAR

4. Argumenta la solución obtenida de un problema, con métodos numéricos, gráficos, analíticos o variacionales, mediante el lenguaje verbal, matemático y el uso de las tecnologías de la información y la comunicación.

TEMA: DATO, VARIABLE Y SU CLASIFICACIÓN

COMPETENCIAS

GENÉRICAS Y ATRIBUTOS

1. Se conoce y valora a sí mismo y aborda problemas y retos teniendo en cuenta los objetivos que persigue.
1.1 Enfrenta las dificultades que se le presentan y es consciente de sus valores, fortalezas y debilidades.

4. Escucha, interpreta y emite mensajes pertinentes en distintos contextos mediante la utilización de medios, códigos y herramientas apropiados.
4.1 Expresa ideas y conceptos mediante representaciones lingüísticas, matemáticas o gráficas.

5. Desarrolla innovaciones y propone soluciones a problemas a partir de métodos establecidos.
5.1 Sigue instrucciones y procedimientos de manera reflexiva, comprendiendo como cada uno de sus pasos contribuye al alcance de un objetivo.
5.2 Ordena información de acuerdo a categorías, jerarquías y relaciones.
5.6 Utiliza las tecnologías de la información y comunicación para procesar e interpretar información.

8. Participa y colabora de manera efectiva en equipos diversos.
8.1 Propone maneras de solucionar un problema o desarrollar un proyecto en equipo, definiendo un curso de acción con pasos específicos. 8.2 Aporta puntos de vista con apertura y considera los de otras personas de manera reflexiva.
8.3 Asume una actitud constructiva, congruente con los conocimientos y habilidades con los que cuenta dentro de distintos equipos de trabajo.

Coeficiente de Variación y Asimetría

El Coeficiente de Variación CV es una medida que permite comparar el grado de dispersión , es decir, es el porcentaje (%) de variación.

$$CV = \frac{\sigma}{\bar{X}}$$

El Coeficiente de Asimetría de Pearson: La Asimetría o Sesgo de una distribución de frecuencia, expresa su inclinación respecto al eje vertical. La Asimetría puede ser:

1.- Positiva	Inclinación a la derecha (la Media es mayor que la Mediana).
2.- Negativa	Inclinación a la izquierda (la Media es menor que la Mediana).
3.- Nula	La Media es igual a la Mediana

1 Coef Pearson = $(\bar{X} - \tilde{X})/\sigma)$ 2 Coef Pearson = $(3\bar{X} - 3\tilde{X})/\sigma)$

LIRC	LSRC	f	fac	Mi	fMi	|Mi-X|	f|Mi-X|	(Mi-X)^2	f(Mi-X)^2
45	55	6	6	50	300	19.4	116.4	376.36	2258.16
55	65	10	16	60	600	9.4	94	88.36	883.6
65	75	19	35	70	1330	0.6	11.4	0.36	6.84
75	85	11	46	80	880	10.6	116.6	112.36	1235.96
85	95	4	50	90	360	20.6	82.4	424.36	1697.44
		50			3470		420.8		6082

Me	69.40	
Mo	70.29	
Md	69.74	
Dm	8.42	
Var	121.64	
Dest	11.03	
CV	0.1587	15.87 %

1 Coef Pearson = $(\bar{X} - \tilde{X})/\sigma)$

2 Coef Pearson = $(3\bar{X} - 3\tilde{X})/\sigma)$

LIRC	LSRC	f	fac	Mi	fMi	|Mi-X|	f|Mi-X|	(Mi-X)^2	f(Mi-X)^2
45	55	6	6	50	300	19.4	116.4	376.36	2258.16
55	65	10	16	60	600	9.4	94	88.36	883.6
65	75	19	35	70	1330	0.6	11.4	0.36	6.84
75	85	11	46	80	880	10.6	116.6	112.36	1235.96
85	95	4	50	90	360	20.6	82.4	424.36	1697.44
		50			3470		420.8		6082

Me	69.40	
Mo	70.29	
Md	69.74	
Dm	8.42	
Var	121.64	
Dest	11.03	
CV	0.1587	15.87 %

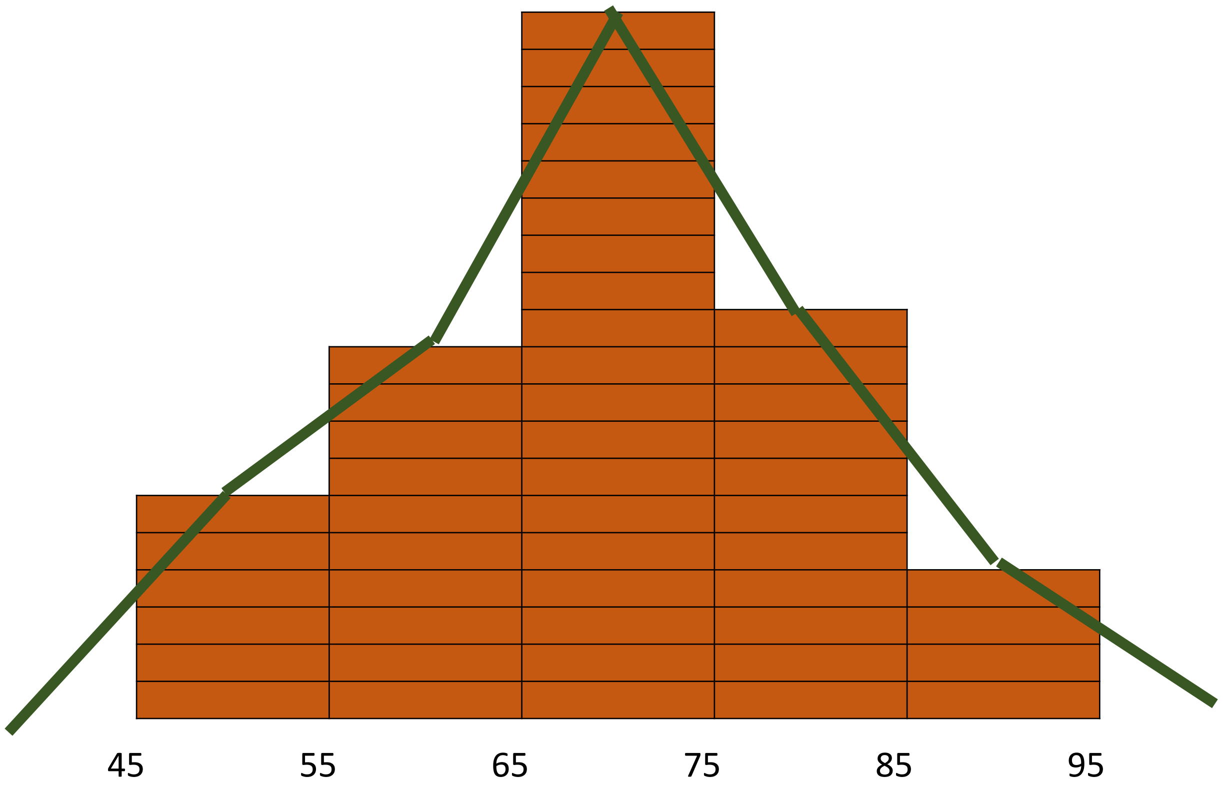
1 Coef Pearson = $(\bar{X} - \tilde{X})/\sigma)$

Sesgo 1	-0.03083

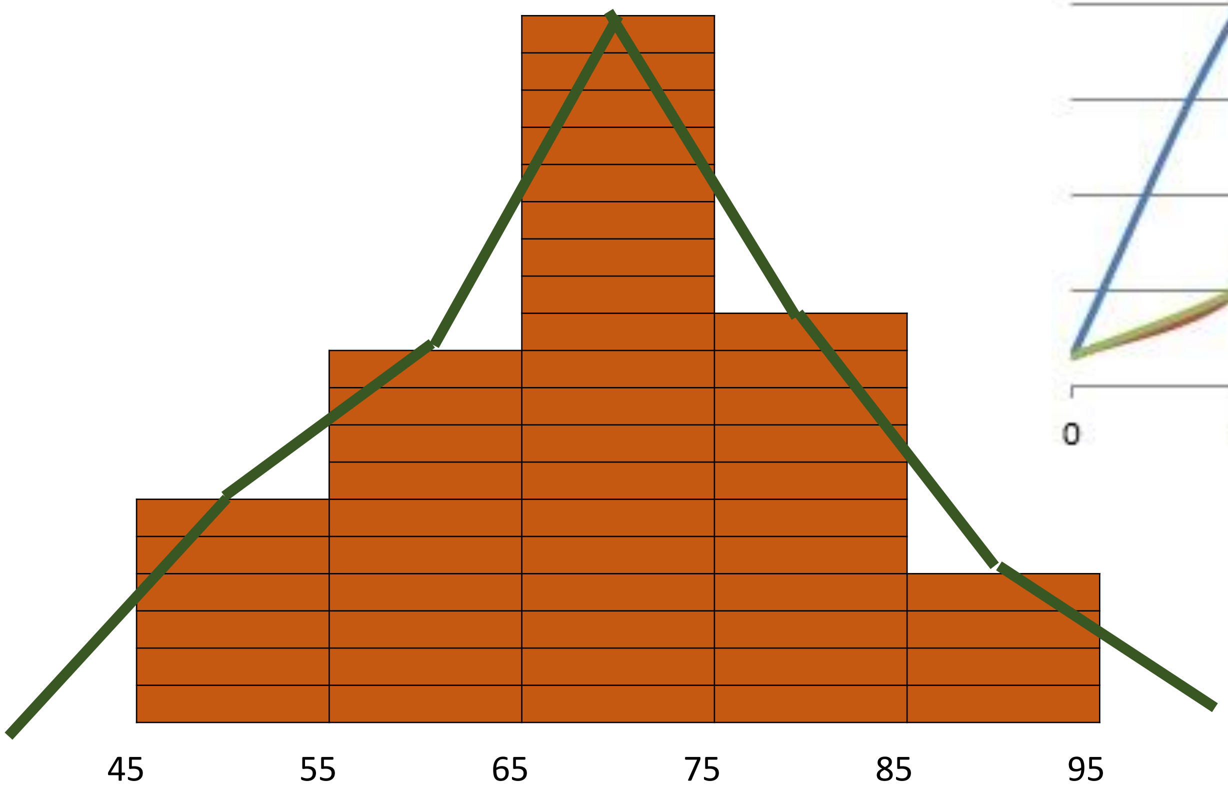
2 Coef Pearson = $(3\bar{X} - 3\tilde{X})/\sigma)$

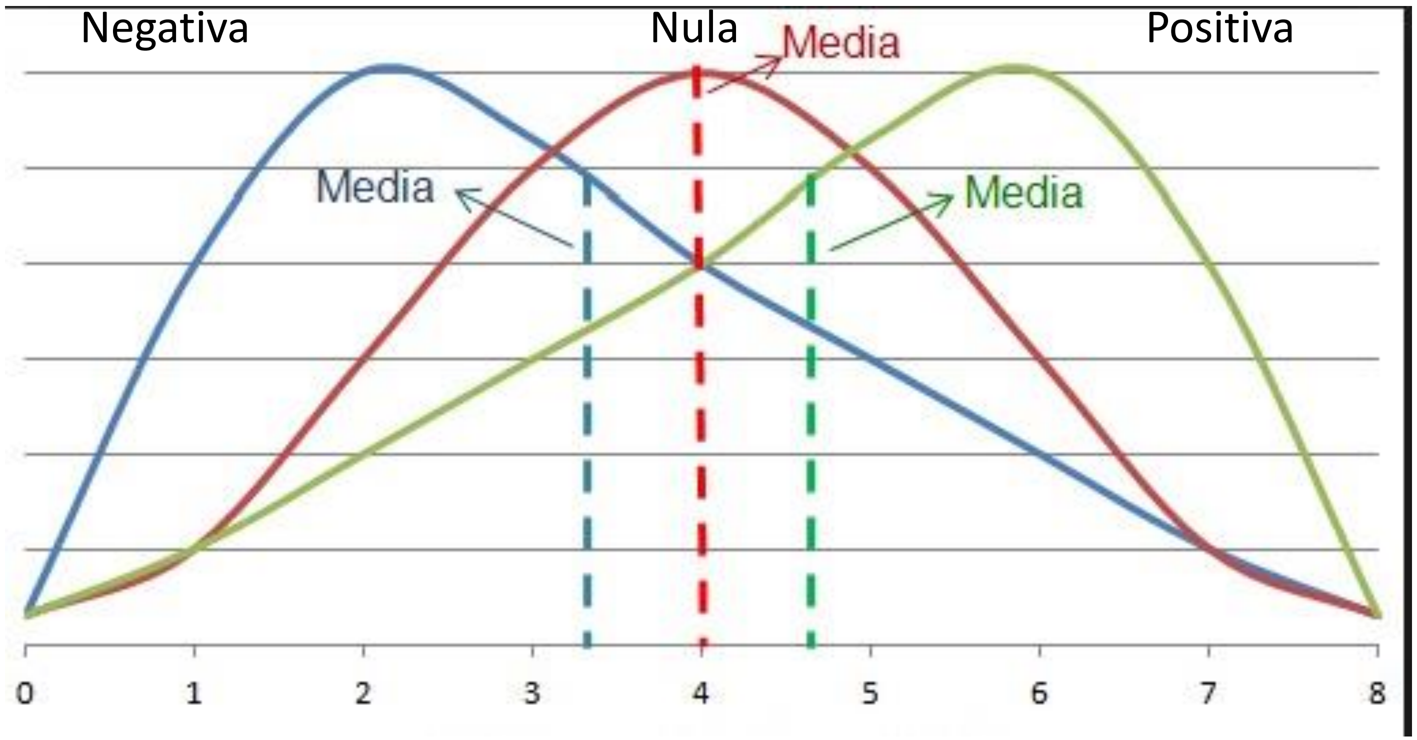
Sesgo 2	-0.09248

Sesgo 1	**-0.03083**
Sesgo 2	**-0.09248**
CV	**15.87 %**



Sesgo 1	-0.03083
Sesgo 2	-0.09248
CV	15.87 %





Positiva Inclinación a la derecha (Media > Mediana

Negativa Inclinación a la izquierda (Media < Mediana).

Nula Media = Mediana

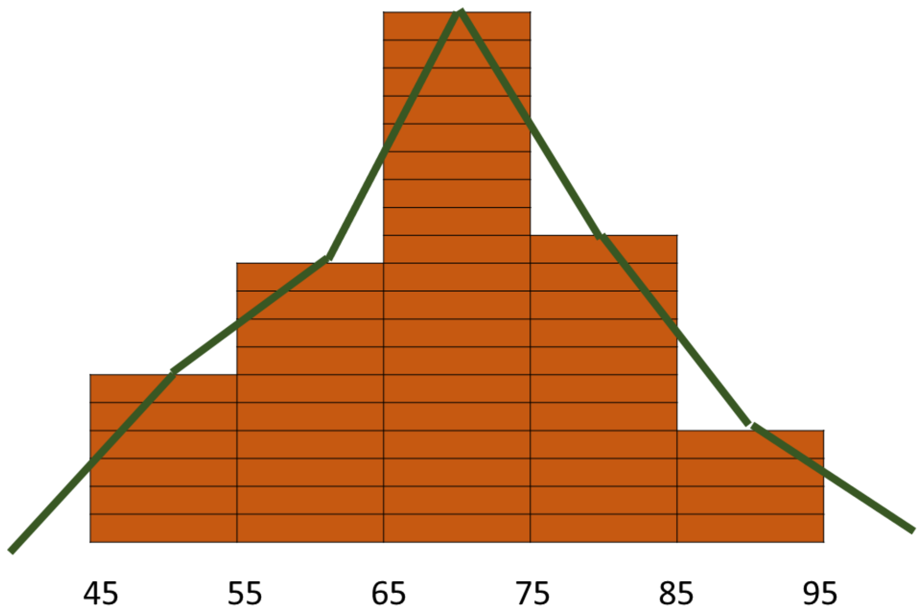
Medidas de Posición

Cuartiles

$$Qi = L + (\frac{i*N-fac}{f})c$$

L = Limite Inferior Real de Clase del Cuartil a calcular
i = Cuartil a calcular
N = Total de datos en la tabla
fac = Frecuencia acumulada de la clase
f = frecuencia de la clase
c = Tamaño del intervalo

LIRC	LSRC	f	fac	Mi	fMi	|Mi-X|	f|Mi-X|	(Mi-X)^2	f(Mi-X)^2
45	55	6	6	50	300	19.4	116.4	376.36	2258.16
55	65	10	16	60	600	9.4	94	88.36	883.6
65	75	19	35	70	1330	0.6	11.4	0.36	6.84
75	85	11	46	80	880	10.6	116.6	112.36	1235.96
85	95	4	50	90	360	20.6	82.4	424.36	1697.44
		50			3470		420.8		6082



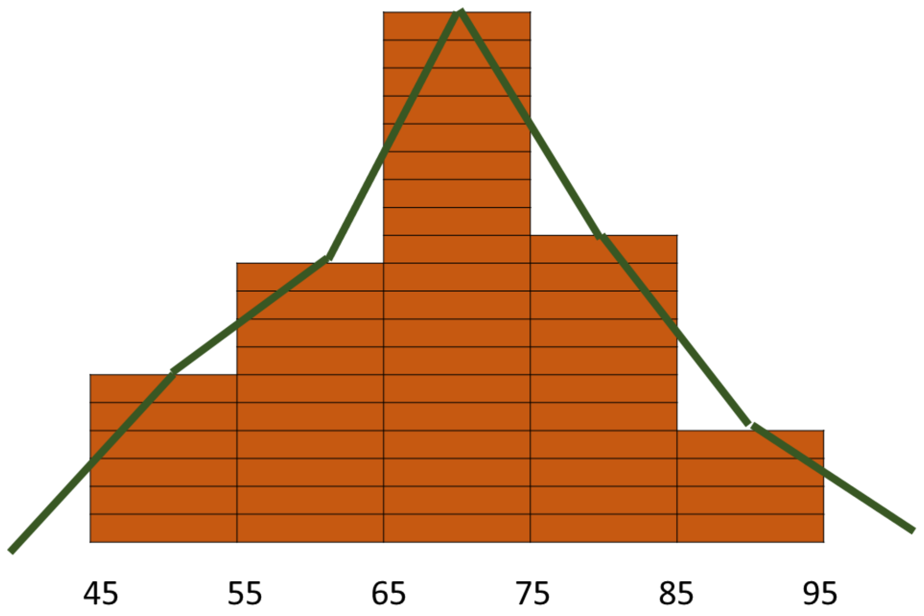
Medidas de Posición

Cuartiles

L = *Limite Inferior Real de Clase del Cuartil a calcular*
i = Cuartil a calcular
N = Total de datos en la tabla
fac = Frecuencia acumulada de la clase
f = frecuencia de la clase
c = Tamaño del intervalo

$$Qi = L + (\frac{i*N-fac}{f})c$$

LIRC	LSRC	f	fac	Mi	fMi	|Mi-X|	f|Mi-X|	(Mi-X)^2	f(Mi-X)^2
45	55	6	6	50	300	19.4	116.4	376.36	2258.16
55	65	10	16	60	600	9.4	94	88.36	883.6
65	75	19	35	70	1330	0.6	11.4	0.36	6.84
75	85	11	46	80	880	10.6	116.6	112.36	1235.96
85	95	4	50	90	360	20.6	82.4	424.36	1697.44
		50			3470		420.8		6082



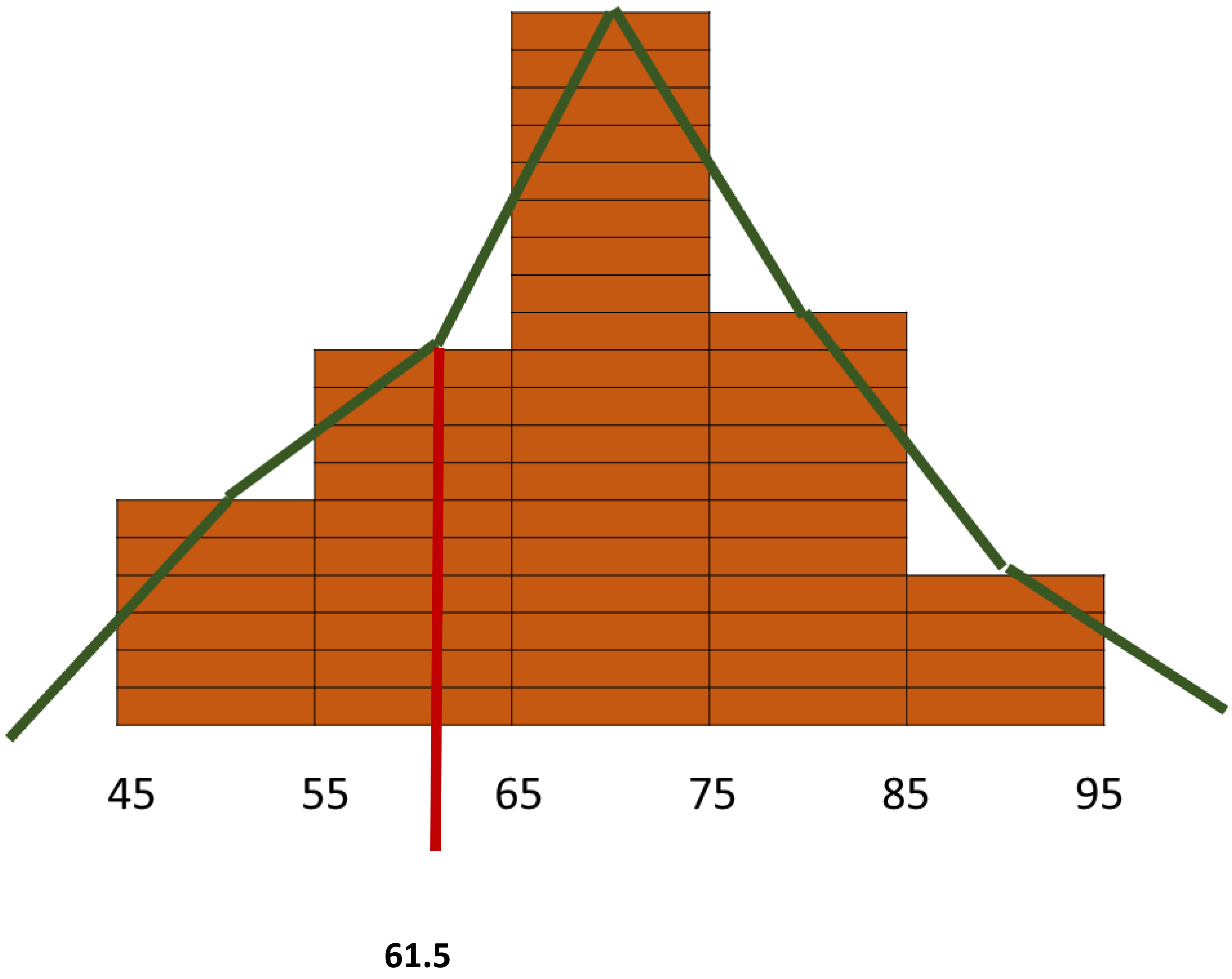
$$Q1 = L + (\frac{1*N-fac}{f})c$$

$$Q2 = L + (\frac{2*N-fac}{f})c$$

$$Q3 = L + (\frac{3*N-fac}{f})c$$

LIRC	LSRC	f	fac
45	55	6	6
55	65	10	16
65	75	19	35
75	85	11	46
85	95	4	50
		50	

$$Q1 = L + \left(\frac{1*N-fac}{f}\right)C$$

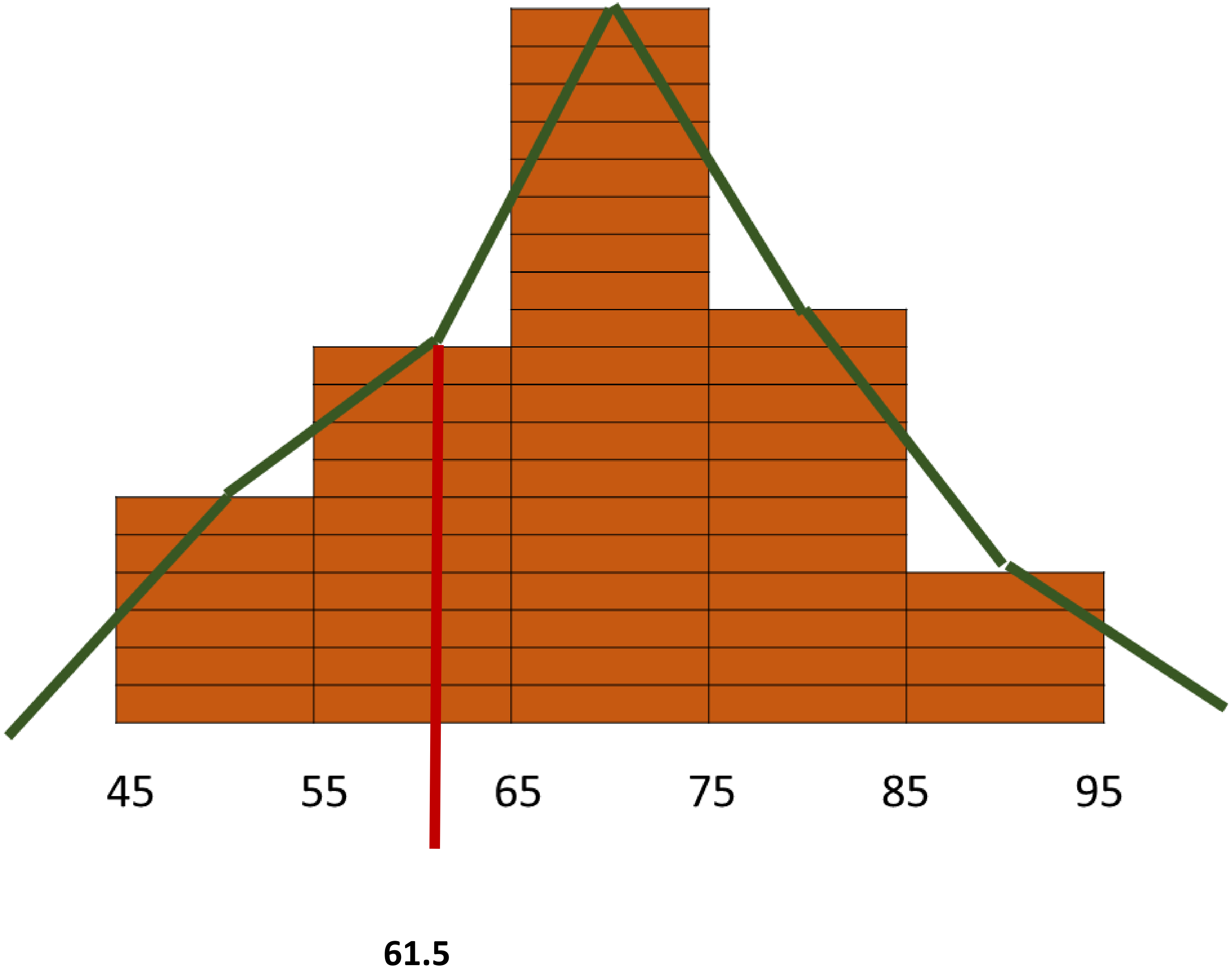


LIRC	LSRC	f	fac
45	55	6	6
55	65	10	16
65	75	19	35
75	85	11	46
85	95	4	50
		50	

$$Q1 = L + \left(\frac{1*N - fac}{f}\right)c$$

*1.- (i * N)4* *(1*50)/4* *12.5*

$$Q1 = 55 + \left(\frac{12.5-6}{10}\right)10 = 61.5$$



LIRC	LSRC	f	fac
45	55	6	6
55	65	10	16
65	75	19	35
75	85	11	46
85	95	4	50
		50	

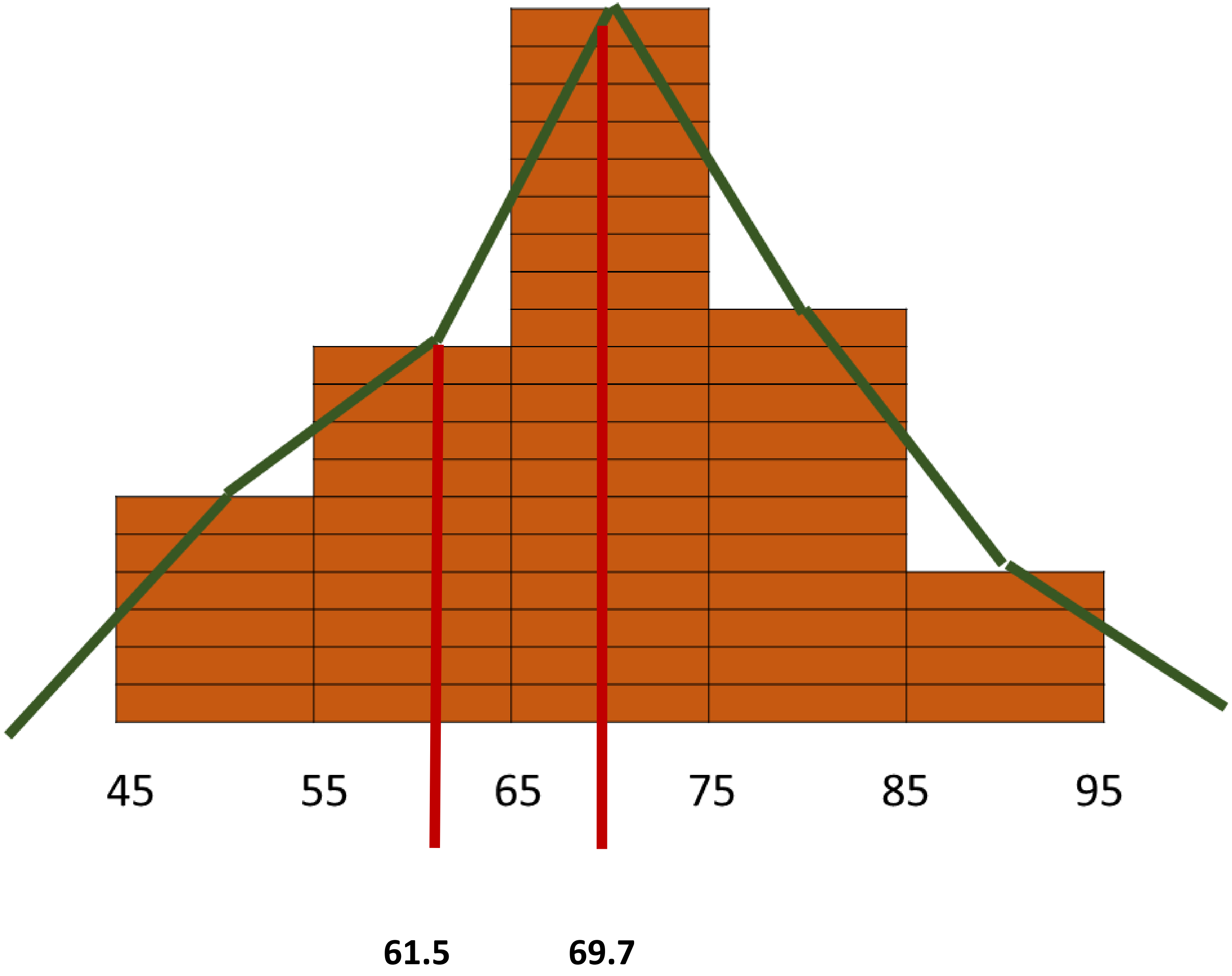
$$Q1 = L + (\frac{1*N-fac}{f})\text{C}$$

*1.- (i * N)4* *(1*50)/4* *12.5*

$$Q1 = 55 + (\frac{12.5-6}{10})10=61.5$$

*2.- (i * N)4* *(2*50)/4* *25*

$$Q2 = 65 + (\frac{25-16}{19})10=69.7$$



LIRC	LSRC	f	fac
45	55	6	6
55	65	10	16
65	75	19	35
75	85	11	46
85	95	4	50
		50	

$$Q1 = L + \left(\frac{1*N - fac}{f}\right)c$$

*1.- (i * N)4* *(1*50)/4* *12.5*

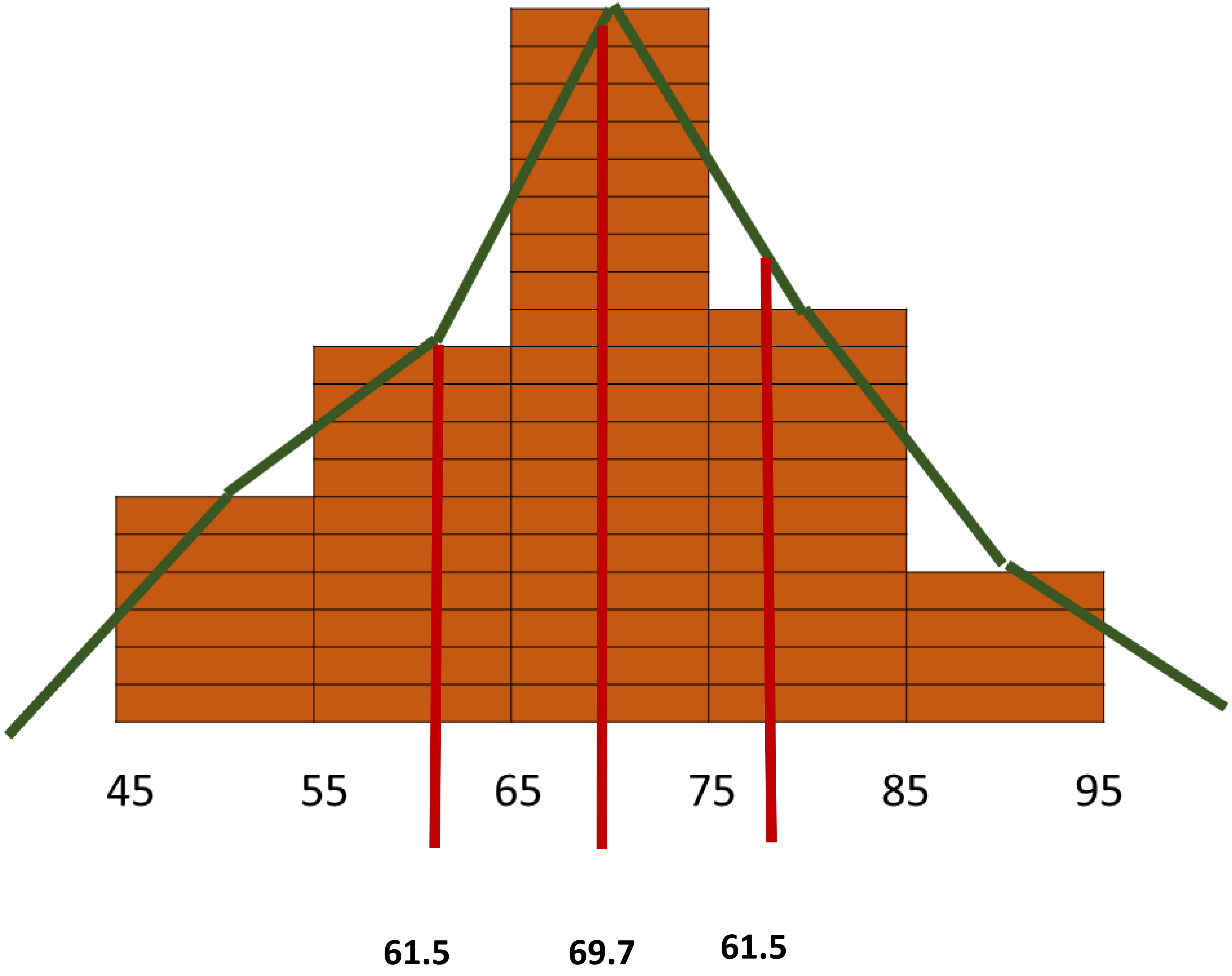
$$Q1 = 55 + \left(\frac{12.5-6}{10}\right)10=61.5$$

*2.- (i * N)4* *(2*50)/4* *25*

$$Q2 = 65 + \left(\frac{25-16}{19}\right)10=69.7$$

*3.- (i * N)4* *(3*50)/4* *37.5*

$$Q3 = 75 + \left(\frac{37.5-35}{11}\right)10=77.3$$



lirc	lsrc	f	Mi	fac	fMi	|Mi-X|	f|Mi-X|	(Mi-X)^2	f(Mi-X)^2
50	60	8	55	8	440	24.93	199.44	621.50	4972.04
60	70	10	65	18	650	14.93	149.3	222.90	2229.05
70	80	16	75	34	1200	4.93	78.88	24.30	388.88
80	90	16	85	50	1360	5.07	81.12	25.70	411.28
90	100	10	95	60	950	15.07	150.7	227.10	2271.05
100	110	5	105	65	525	25.07	125.35	628.50	3142.52
110	120	2	115	67	230	35.07	70.14	1229.90	2459.81
		67			5355		854.93		15874.63

Media	79.93
Moda	80.00
Mediana	79.69
Dm	12.76
Var	236.93
DesSand	15.39
CV	19.26
Sesgo 1	0.01559
Sesgo 2	0.04678

lirc	lsrc	f	Mi	
		8		
		10		
70		16	75	
		16		
		10		
		5		
		2		

Mi=LIRC+LSRC/2

Mi=75

LIRC=70

Q1	68.75	16.75
Q2	85.75	33.50
Q3	90.25	50.25

Gráfica

Covarianza

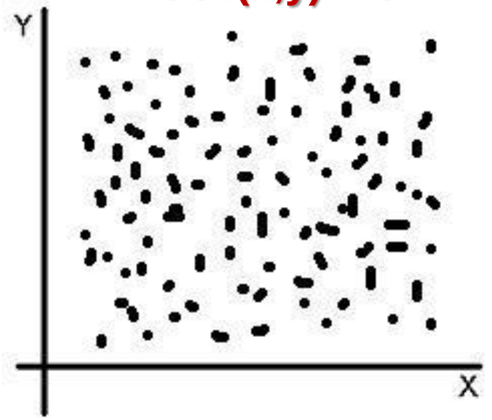
La **covarianza** de una variable bidimensional es la media aritmética de los productos de las desviaciones de cada una de las variables respecto a sus medias respectivas, mide la relación lineal entre dos variables (Datos Bivariados).

La **covarianza** se representa por $\mathbf{s_{xy}}$ o $\boldsymbol{\sigma_{xy}}$.(Cov_{xy})

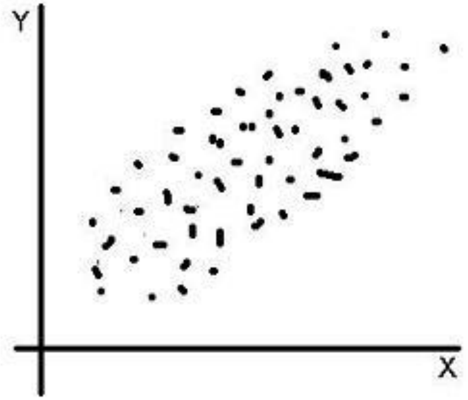
$$Cov_{xy} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{N}$$

- σ_{xy} es **positiva** si los valores altos de X están asociados a los valores altos de Y y viceversa.
- σ_{xy} es **negativa** si los valores altos de X están asociados a los valores bajos de Y y viceversa.
- Si X e Y son variables aleatorias independientes cov(x,y) = 0 .

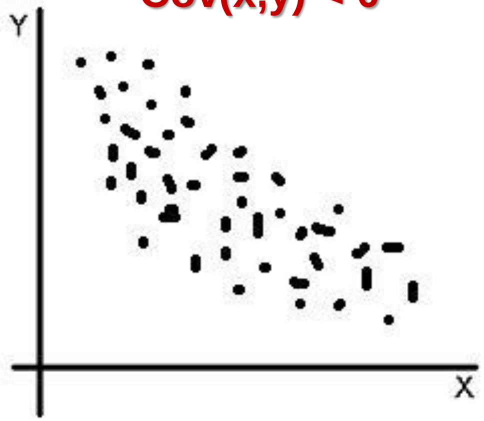
Cov(x,y) = 0



Cov(x,y) > 0



Cov(x,y) < 0



Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i - \bar{X})$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})\ (Y_i - \bar{Y})$
2	**1**			
3	**3**			
4	**2**			
4	**4**			
5	**4**			
6	**4**			
6	**6**			
7	**4**			
7	**6**			
8	**7**			
10	**9**			
10	**10**			

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\ (Y_i-\bar{Y})$
2	**1**			
3	**3**			
4	**2**			
4	**4**			
5	**4**			
6	**4**			
6	**6**			
7	**4**			
7	**6**			
8	**7**			
10	**9**			
10	**10**			
72	**60**			

$$\bar{X} = \frac{\sum X_i}{N} \qquad \bar{X} = \frac{72}{12} = 6$$

$$\bar{Y} = \frac{\sum Y_i}{N} \qquad \bar{Y} = \frac{60}{12} = 5$$

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i$-$\bar{X})$	$(Y_i$-$\bar{Y})$	$(X_i$-$\bar{X})$ $(Y_i$-$\bar{Y})$
2	**1**	-4	-4	
3	**3**	-3	-2	
4	**2**	-2	-3	
4	**4**	-2	-1	
5	**4**	-1	-1	
6	**4**	0	-1	
6	**6**	0	1	
7	**4**	1	-1	
7	**6**	1	1	
8	**7**	2	2	
10	**9**	4	4	
10	**10**	4	5	
72	**60**			

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\ (Y_i-\bar{Y})$
2	**1**	-4	-4	16
3	**3**	-3	-2	6
4	**2**	-2	-3	6
4	**4**	-2	-1	2
5	**4**	-1	-1	1
6	**4**	0	-1	0
6	**6**	0	1	0
7	**4**	1	-1	-1
7	**6**	1	1	1
8	**7**	2	2	4
10	**9**	4	4	16
10	**10**	4	5	20
72	**60**			

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i\text{-}\bar{X})$	$(Y_i\text{-}\bar{Y})$	$(X_i\text{-}\bar{X})\ (Y_i\text{-}\bar{Y})$
2	**1**	-4	-4	16
3	**3**	-3	-2	6
4	**2**	-2	-3	6
4	**4**	-2	-1	2
5	**4**	-1	-1	1
6	**4**	0	-1	0
6	**6**	0	1	0
7	**4**	1	-1	-1
7	**6**	1	1	1
8	**7**	2	2	4
10	**9**	4	4	16
10	**10**	4	5	20
72	**60**			**71**

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i - \bar{X})$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})\ (Y_i - \bar{Y})$
2	**1**	-4	-4	16
3	**3**	-3	-2	6
4	**2**	-2	-3	6
4	**4**	-2	-1	2
5	**4**	-1	-1	1
6	**4**	0	-1	0
6	**6**	0	1	0
7	**4**	1	-1	-1
7	**6**	1	1	1
8	**7**	2	2	4
10	**9**	4	4	16
10	**10**	4	5	20
72	**60**			**71**

$$Cov_{xy} = \frac{\sum (X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = 71/12 \quad = \quad \mathbf{5.92}$$

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	**2**	**3**	**4**	**4**	**5**	**6**	**6**	**7**	**7**	**8**	**10**	**10**
Física	1	3	2	4	4	4	6	4	6	7	9	10

X_i	Y_i	$(X_i$-$\bar{X})$	$(Y_i$-$\bar{Y})$	$(X_i$-$\bar{X})$ $(Y_i$-$\bar{Y})$
2	**1**	-4	-4	16
3	**3**	-3	-2	6
4	**2**	-2	-3	6
4	**4**	-2	-1	2
5	**4**	-1	-1	1
6	**4**	0	-1	0
6	**6**	0	1	0
7	**4**	1	-1	-1
7	**6**	1	1	1
8	**7**	2	2	4
10	**9**	4	4	16
10	**10**	4	5	20
72	**60**			**71**

$$\bar{X} = \frac{\sum X_i}{N} \qquad \bar{X}=\frac{72}{12} = 6$$

$$\bar{Y} = \frac{\sum Y_i}{N} \qquad \bar{Y}=\frac{60}{12} = 5$$

Centro de gravedad

Es el punto donde se encuentran las medias de los datos en una distribución bidimensional.

El centro de gravedad (CG) es uno de los puntos por donde debe pasar la Recta de Regresión.

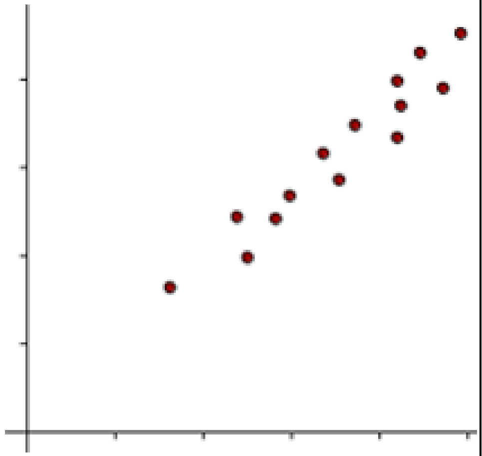
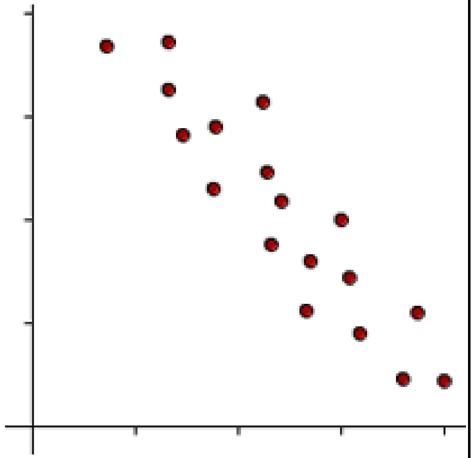
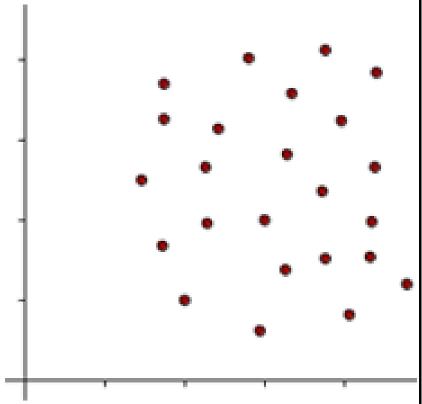
$$CG = (\bar{X}_i \bar{Y}_i) \qquad CG = (6,5)$$

Correlación Lineal

Coeficiente de Correlación Lineal

La correlación trata de establecer la relación o dependencia que existe entre las dos variables que intervienen en una distribución bidimensional o datos bivariados. Indica que tan estrecha es la relación entre dos variables. Se expresa mediante la letra **r**

Tipos de correlación

1º Correlación directa: Se da cuando al aumentar una de las variables la otra aumenta. La recta correspondiente a la nube de puntos de la distribución es una recta creciente.	**2º Correlación inversa:** Se da cuando al aumentar una de las variables la otra disminuye. La recta correspondiente a la nube de puntos de la distribución es una recta decreciente.	**3º Correlación nula:** Se da cuando no hay dependencia de ningún tipo entre las variables. En este caso se dice que las variables no tienen correlación y la nube de puntos tiene una forma dispersa o redondeada.
		

$$r = \frac{Cov_{xy}}{S_x S_y} \quad \Leftrightarrow \quad \boldsymbol{r = \frac{Cov_{xy}}{\sigma_x \sigma_y}}$$

Grado de correlación

Indica la proximidad que hay entre los puntos de la nube de puntos. Se pueden dar tres tipos:

Propiedades del coeficiente de correlación lineal

1. El signo del **coeficiente de correlación** es el mismo que el de la **covarianza**.
   - ✓ Si la covarianza es positiva, la correlación es directa.
   - ✓ Si la covarianza es negativa, la correlación es inversa.
   - ✓ Si la covarianza es nula, no existe correlación.
2. El **coeficiente de correlación lineal** es un número real comprendido entre menos −1 y 1.
3. Si el **coeficiente de correlación lineal** toma valores cercanos a −1 la correlación es **fuerte e inversa**, y será tanto más fuerte cuanto más se aproxime r a −1.
4. Si el **coeficiente de correlación lineal** toma valores cercanos a 1 la correlación es **fuerte y directa**, y será tanto más fuerte cuanto más se aproxime r a 1.
5. Si el **coeficiente de correlación lineal** toma valores cercanos a 0, la correlación es **débil**.
6. Si r = 1 ó −1, los puntos de la nube están sobre la recta, entre ambas variables hay **dependencia funcional**.

X_i	Y_i	$(X_i$-$\bar{X})$	$(Y_i$-$\bar{Y})$	$(X_i$-$\bar{X})$ $(Y_i$-$\bar{Y})$
5	15			
7	18			
2	10			
1	8			
9	20			
$\bar{X}$=4.8	$\bar{Y}$=14.2			

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{N}$$

X_i	Y_i	$(X_i\text{-}\bar{X})$	$(Y_i\text{-}\bar{Y})$	$(X_i\text{-}\bar{X})\ (Y_i\text{-}\bar{Y})$
5	15	0.2	0.8	
7	18	2.2	3.8	
2	10	-2.8	-4.2	
1	8	-3.8	-6.2	
9	20	4.2	5.8	
$\bar{X}$=4.8	$\bar{Y}$=14.2			

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\ (Y_i-\bar{Y})$
5	15	0.2	0.8	0.16
7	18	2.2	3.8	8.36
2	10	-2.8	-4.2	11.76
1	8	-3.8	-6.2	23.56
9	20	4.2	5.8	24.36
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

X_i	Y_i	$(X_i\text{-}\bar{X})$	$(Y_i\text{-}\bar{Y})$	$(X_i\text{-}\bar{X})\ (Y_i\text{-}\bar{Y})$
5	15	0.2	0.8	0.16
7	18	2.2	3.8	8.36
2	10	-2.8	-4.2	11.76
1	8	-3.8	-6.2	23.56
9	20	4.2	5.8	24.36
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

X_i	Y_i	$(X_i$-$\bar{X})$	$(Y_i$-$\bar{Y})$	$(X_i$-$\bar{X})$ $(Y_i$-$\bar{Y})$
5	15	0.2	0.8	0.16
7	18	2.2	3.8	8.36
2	10	-2.8	-4.2	11.76
1	8	-3.8	-6.2	23.56
9	20	4.2	5.8	24.36
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\,(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
5	15	0.2	0.8	0.16		
7	18	2.2	3.8	8.36		
2	10	-2.8	-4.2	11.76		
1	8	-3.8	-6.2	23.56		
9	20	4.2	5.8	24.36		
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2		

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

3. Coeficiente de Correlación

$$\boldsymbol{r = \frac{Cov_{xy}}{\sigma_x \sigma_y}}$$

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\,(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
5	15	0.2	0.8	0.16		
7	18	2.2	3.8	8.36		
2	10	-2.8	-4.2	11.76		
1	8	-3.8	-6.2	23.56		
9	20	4.2	5.8	24.36		
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2		

Calcular la Desviación Estándar de *Xi y Yi*

$$\sigma_x = \sqrt{\frac{\sum(Xi - X)^2}{N}} =$$

$$\sigma_y = \sqrt{\frac{\sum(Yi - Y)^2}{N}} =$$

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

3. Coeficiente de Correlación

$$\boldsymbol{r = \frac{Cov_{xy}}{\sigma_x \sigma_y}}$$

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\,(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
5	15	0.2	0.8	0.16	0.04	0.64
7	18	2.2	3.8	8.36	4.84	14.44
2	10	-2.8	-4.2	11.76	7.84	17.64
1	8	-3.8	-6.2	23.56	14.44	38.44
9	20	4.2	5.8	24.36	17.64	33.64
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2	44.8	104.8

Calcular la Desviación Estándar de *Xi y Yi*

$$\sigma_x = \sqrt{\frac{\sum(Xi - X)^2}{N}} =$$

$$\sigma_y = \sqrt{\frac{\sum(Yi - Y)^2}{N}} =$$

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

3. Coeficiente de Correlación

$$\boldsymbol{r = \frac{Cov_{xy}}{\sigma_x \sigma_y}}$$

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\,(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
5	15	0.2	0.8	0.16	0.04	0.64
7	18	2.2	3.8	8.36	4.84	14.44
2	10	-2.8	-4.2	11.76	7.84	17.64
1	8	-3.8	-6.2	23.56	14.44	38.44
9	20	4.2	5.8	24.36	17.64	33.64
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2	44.8	104.8

Calcular la Desviación Estándar de *Xi y Yi*

$$\sigma_x = \sqrt{\frac{\sum(Xi - X)^2}{N}} = \sqrt{\frac{44.8}{5}} = 2.993$$

$$\sigma_y = \sqrt{\frac{\sum(Yi - Y)^2}{N}} = \sqrt{\frac{104.8}{5}} = 4.578$$

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

3. Coeficiente de Correlación

$$\boldsymbol{r = \frac{Cov_{xy}}{\sigma_x \sigma_y}}$$

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\ (Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
5	15	0.2	0.8	0.16	0.04	0.64
7	18	2.2	3.8	8.36	4.84	14.44
2	10	-2.8	-4.2	11.76	7.84	17.64
1	8	-3.8	-6.2	23.56	14.44	38.44
9	20	4.2	5.8	24.36	17.64	33.64
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2	44.8	104.8

Calcular la Desviación Estándar de *Xi y Yi*

$$\sigma_x = \sqrt{\frac{\sum(Xi - X)^2}{N}} = \sqrt{\frac{44.8}{5}} = 2.993$$

$$\sigma_y = \sqrt{\frac{\sum(Yi - Y)^2}{N}} = \sqrt{\frac{104.8}{5}} = 4.578$$

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})\,(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

3. Coeficiente de Correlación

$$\boldsymbol{r = \frac{Cov_{xy}}{\sigma_x \sigma_y}}$$

$$\boldsymbol{r} = \frac{13.64}{(2.993*4.578)} = 0.9953$$

Diagrama de Dispersión

Diagrama de Dispersión

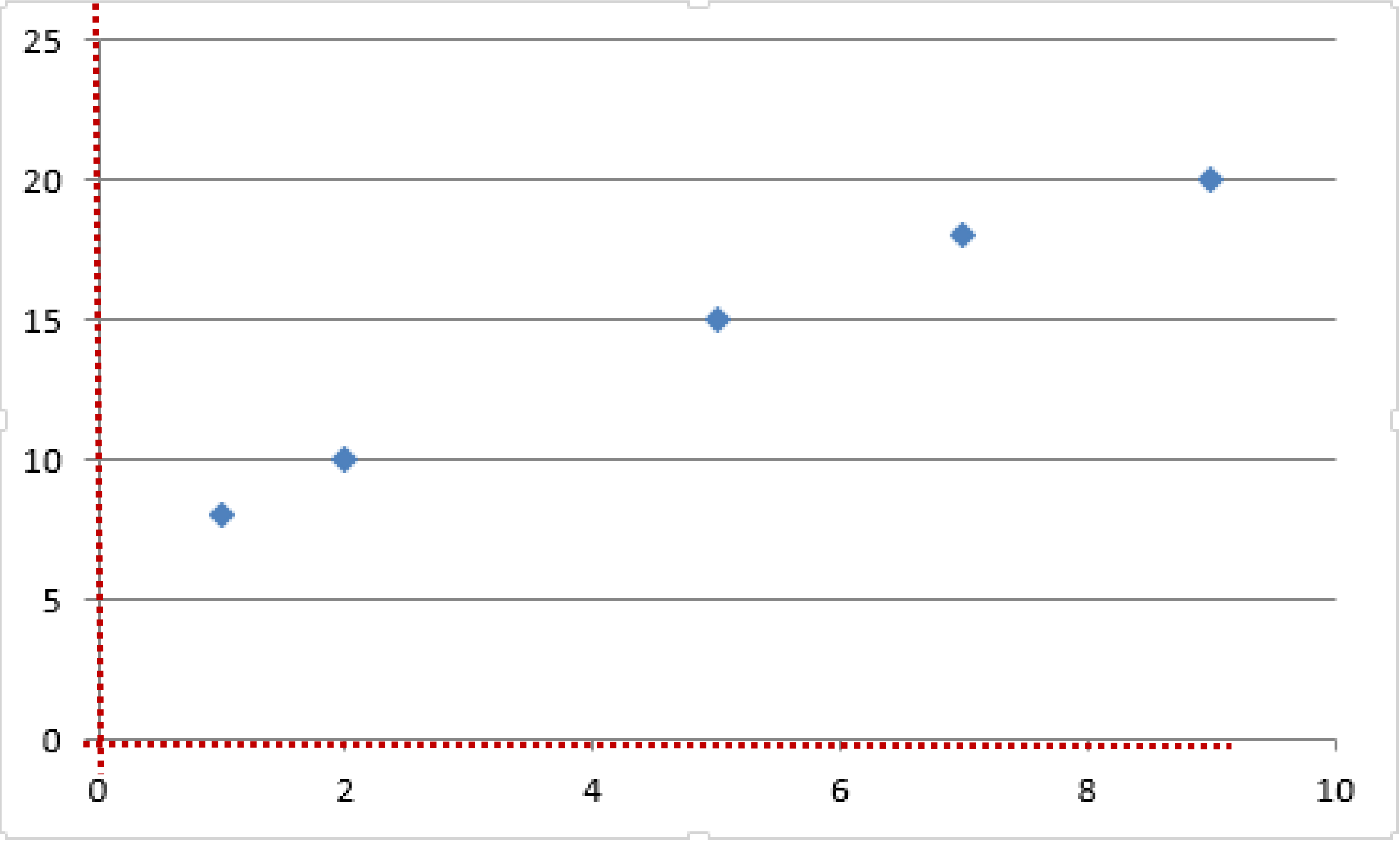
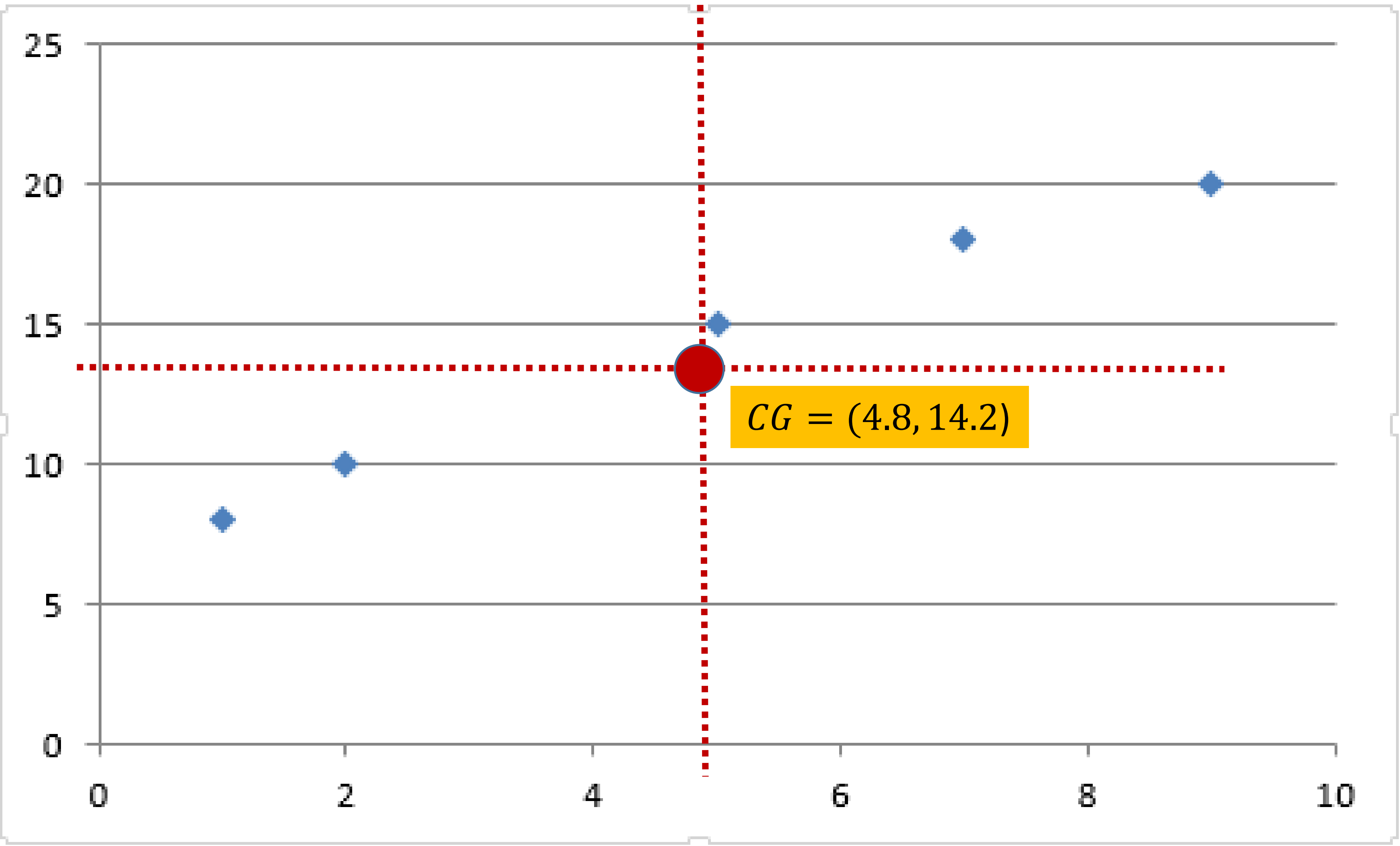


Diagrama de Dispersión



Recta de Regresión

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
5	15	0.2	0.8	0.16	0.04	0.64
7	18	2.2	3.8	8.36	4.84	14.44
2	10	-2.8	-4.2	11.76	7.84	17.64
1	8	-3.8	-6.2	23.56	14.44	38.44
9	20	4.2	5.8	24.36	17.64	33.64
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2	44.8	104.8

1. Covarianza

$$Cov_{xy} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{N}$$

$$Cov_{xy} = \frac{68.2}{5} = 13.64$$

2. Centro de Gravedad

$$CG = (\bar{X}_i \bar{Y}_i)$$

$$CG = (4.8, 14.2)$$

3. Coeficiente de Correlación

$$\boldsymbol{r} = \frac{\boldsymbol{Cov_{xy}}}{\boldsymbol{\sigma_x \sigma_y}}$$

$$\boldsymbol{r} = \frac{13.64}{(2.993*4.578)} = 0.9953$$

4. Recta de Regresión

y = mx + b

$$m = \frac{Cov_{xy}}{\sigma_x^2}$$

$$b = \bar{Y} - m\bar{X}$$

Recta de regresión

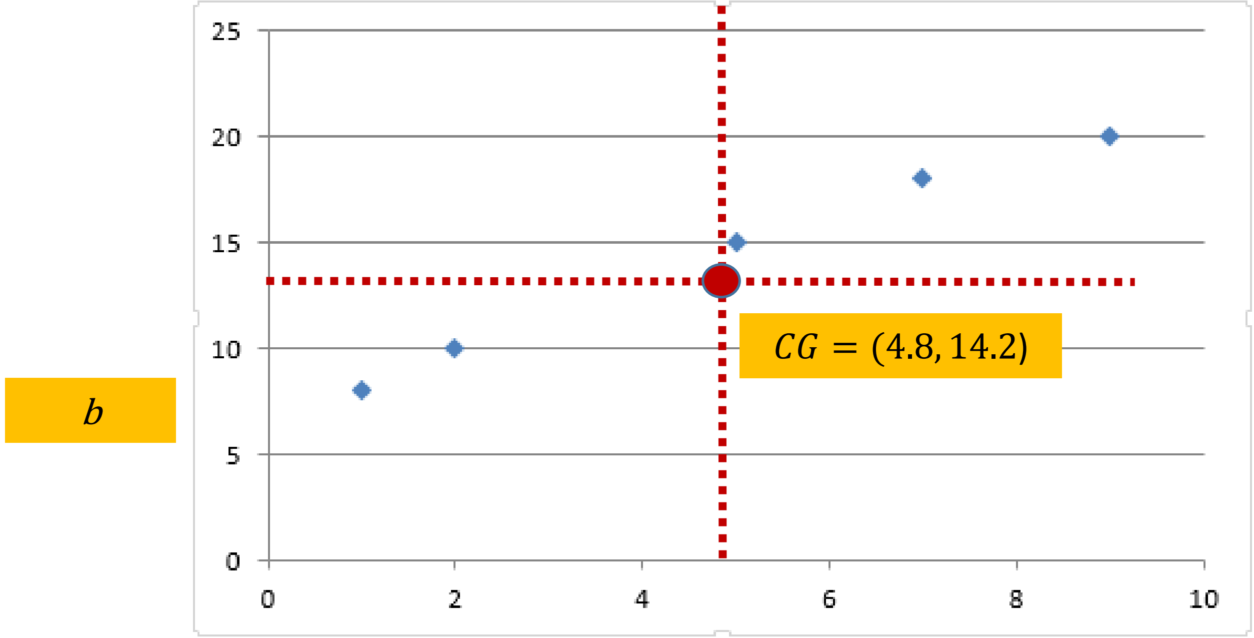
La Ecuación de una Recta se determina de acuerdo a las condiciones que cumple el conjunto de puntos que la forman. Cuando una recta se conoce la pendiente (ángulo de inclinación) y su ordenada de origen Ecuación:

$$\mathbf{y = mx + b}$$

Donde

m = pendiente de la recta $m = \dfrac{Cov_{xy}}{\sigma_x^2}$

b = punto donde se intersecta con el eje *y* $b = \bar{Y} - m\bar{X}$



X_i	Y_i	$(X_i\text{-}\bar{X})$	$(Y_i\text{-}\bar{Y})$	$(X_i\text{-}\bar{X})\ (Y_i\text{-}\bar{Y})$	$(X_i\text{-}\bar{X})^2$
5	15	0.2	0.8	0.16	0.04
7	18	2.2	3.8	8.36	4.84
2	10	-2.8	-4.2	11.76	7.84
1	8	-3.8	-6.2	23.56	14.44
9	20	4.2	5.8	24.36	17.64
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2	44.8

Calcular la Varianza de *Xi*

$$\sigma_x^2 = \frac{\sum(Xi - X)^2}{N}$$

Calcular la Pendiente de *Xi*

$$m = \frac{Cov_{xy}}{\sigma_x^2}$$

Calcular punto de intersección en el eje *Yi*

$$b = \bar{Y} - m\bar{X}$$

y = mx + b

X_i	Y_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\ (Y_i-\bar{Y})$	$(X_i-\bar{X})^2$
5	15	0.2	0.8	0.16	0.04
7	18	2.2	3.8	8.36	4.84
2	10	-2.8	-4.2	11.76	7.84
1	8	-3.8	-6.2	23.56	14.44
9	20	4.2	5.8	24.36	17.64
$\bar{X}$=4.8	$\bar{Y}$=14.2			68.2	44.8

Calcular la Varianza de *Xi*

$$\sigma_x^2 = \frac{\sum(Xi - X)^2}{N}$$

$$\sigma_x^2 = \frac{44.8}{5} = 8.96$$

Calcular la Pendiente de *Xi*

$$m = \frac{Cov_{xy}}{\sigma_x^2}$$

$$m = \frac{13.64}{8.96} = 1.5223$$

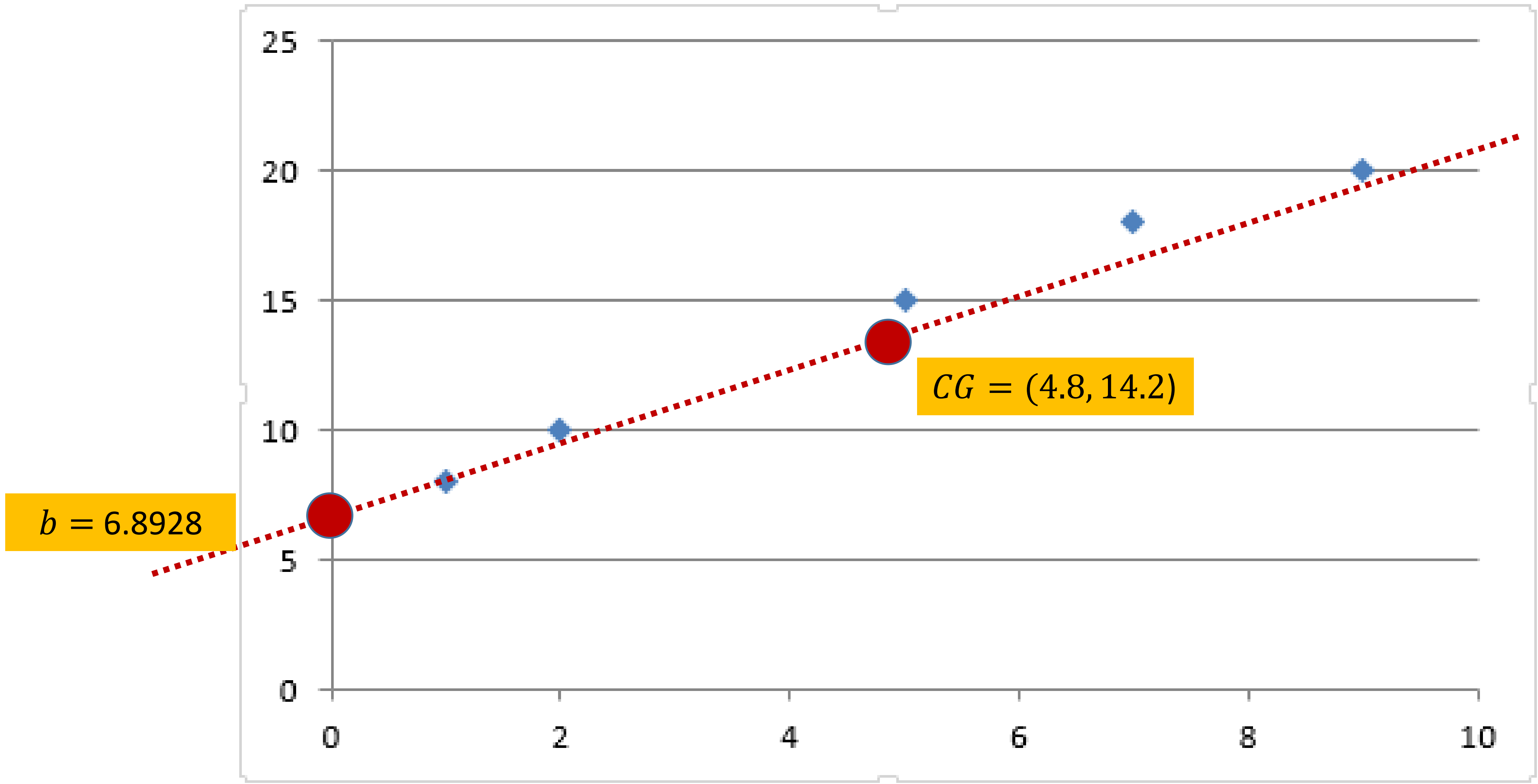
Calcular punto de intersección en el eje *Yi*

$$b = \bar{Y} - m\bar{X}$$

b= 14.2 – 1.5223(4.8) = 6.8928

y = mx + b → **y = 1.5223x + 6.8928**

Diagrama de Dispersión



Recta de Regresión

Series de Tiempo

$$y = 1.5223x + 6.8928$$

X_i	Y_i
5	15
7	18
2	10
1	8
9	20

Xi	**Edad**	**5**	**7**	**2**	**1**	**9**	**6**	**8**
Yi	**Peso**	**15**	**18**	**10**	**8**	**20**		

Con base en la Ecuación de la Recta, calcular:

1.- Cuál sería el peso aproximado de un niño de 6 años

2.- Cuál sería el peso aproximado de un niño de 8 años

Recta de Regresión

Series de Tiempo

y = 1.5223x + 6.8928

X_i	Y_i
5	15
7	18
2	10
1	8
9	20

Xi	**Edad**	**5**	**7**	**2**	**1**	**9**	**6**	**8**
Yi	**Peso**	**15**	**18**	**10**	**8**	**20**	**16.02**	**19.07**

Con base en la Ecuación de la Recta, calcular:

1.- Cuál sería el peso aproximado de un niño de 6 años

y = 1.5223(6) + 6.8928 **y = 16.02**

2.- Cuál sería el peso aproximado de un niño de 8 años

y = 1.5223(8) + 6.8928 **y = 19.07**

En las series de tiempo, las observaciones son tomadas a lo largo del tiempo. Una serie de tiempo es el resultado de observar los valores de una variable X a lo largo del tiempo Y.

Ante una serie de tiempo, lo primero que se debe hacer es representar su grafico de secuencia; esto es, representar gráficamente cada observación xt frente al instante t en que se observa, y luego unir con segmentos cada uno de los T puntos. El grafico de secuencia nos permitirá observar cómo evoluciona la serie a lo largo del tiempo; específicamente, podremos ver las principales características de la serie de tiempo:

Un centro comercial sabe en función de la distancia, en kilómetros, a la que se sitúe de un núcleo de población, acuden los clientes, en cientos, que figuran en la tabla:

- ✓ Calcular el coeficiente de correlación lineal
- ✓ Grafique su ejercicio (Diagrama de Dispersión)
- ✓ Con la Recta de Regresión obtenida, calcule lo siguiente: Si desea recibir a 5 clientes, ¿a qué distancia del núcleo de población debe situarse?

Nº de Clientes (X)	**Distancia (Y)**
8	15
7	19
6	25
4	23
2	34
1	40

EJERCICIO FINAL

En una competencia, dos atletas de salto de longitud registran los siguientes datos:

Intentos	1	2	3	4	5	6	7
Competidor 1	2.03	1.98	2.07	2.13	2.09	2.11	2.16
Competidor 2	1.99	2.07	2.04	2.11	2.14	2.14	2.16

Con base en los resultados de su análisis, responda lo siguiente:

1. ¿Cuál es el promedio de los dos atletas en sus intentos y cuál es que más alto?
2. Con base en su gráfica, ¿Cómo es la Correlación?
3. Obtenga la ecuación de la Recta
4. Pronostique ¿cuál es el sería la altura en el intento 8, 9 y 10
5. A cuál atleta le apostaría según el rendimiento

Atleta 1 ______ Atleta 2 _______

Por qué: ____