



### Identificación del Objeto de aprendizaje

<b>Fecha</b>	Julio 2013
<b>Asignatura</b>	Estadística y Probabilidad I
<b>Unidad</b>	Unidad 2. Datos Bivariados
<b>Tiempo disponible</b>	10 horas
<b>Aprendizajes</b>	<p>El alumno al final de la unidad II:</p> <ul style="list-style-type: none"> <li>• Establece intuitivamente el tipo de relación entre dos variables.</li> <li>• Construye tablas de contingencia para representar la relación entre dos variables cualitativas.</li> <li>• Interpreta la información que contienen las tablas de contingencia.</li> <li>• Construye diagramas de dispersión para representar gráficamente la relación entre dos variables cuantitativas.</li> <li>• Conoce e interpreta los conceptos de regresión y correlación lineal simple.</li> <li>• Calcula e interpreta los valores estimados de la pendiente y la ordenada al origen de la recta de mínimos cuadrados.</li> <li>• Grafica la recta de regresión.</li> <li>• Calcula e interpreta el coeficiente de correlación lineal simple.</li> <li>• Utiliza la recta de ajuste para predecir valores de alguna de las variables.</li> </ul>
<b>Tema</b>	<ol style="list-style-type: none"> <li>1. Relación entre dos variables.</li> <li>2. Relación entre dos variables cuantitativas.</li> <li>3. Regresión lineal y correlación</li> </ol>
<b>Palabras claves</b>	Relación entre dos variables, tablas de contingencia, variables cualitativas, tablas de contingencia, diagramas de dispersión, gráfica de la relación entre dos variables cuantitativas, interpreta los conceptos de regresión y correlación lineal simple, Calcula e interpreta los valores estimados de la pendiente y la ordenada al origen de la recta de mínimos cuadrados, grafica la recta de regresión, calcula e interpreta el coeficiente de correlación lineal simple y utiliza la recta de ajuste para predecir valores de alguna de las variables.
<b>Autor</b>	Juan José Hernández González

### Objetivo (para el profesor)

<ul style="list-style-type: none"> <li>• El profesor deberá propiciar la socialización del trabajo entre los estudiantes y la discusión tanto en equipo como grupal de las diversas ideas estocásticas. Deberá fomentar las formas de expresión oral y escrita para lograr un mayor dominio de la lengua castellana, lo que redundará en una mejor comprensión y organización de ideas y argumentos.</li> <li>• El profesor deberá diseñar actividades para proporcionar al alumno</li> </ul>	<b>Objetivo (para el profesor)</b>
---	------------------------------------



experiencias de aprendizaje, de preferencia contextualizadas, que le permitan acceder a las fuentes del fenómeno aleatorio, le proporcionen información para ser tratada estadísticamente y lo confronten con las dificultades metodológicas en las etapas de captura, procesamiento, interpretación y predicción. La construcción de muestras y el empleo de instrumentos como cuestionarios y encuestas breves pueden ser de gran utilidad.

- Los contenidos deberán tratarse fundamentalmente a través de la resolución de problemas pertinentes, proporcionando oportunidades para que el estudiante aplique los procedimientos básicos, equilibrando este aspecto con los desarrollos conceptuales. De ser necesario, las precisiones teóricas deberán establecerse cuando los alumnos adquieran la experiencia suficiente para garantizar su comprensión.
- Es necesario mantenerse dentro de los tiempos propuestos para cubrir íntegramente los programas. Es pertinente que el profesor continuamente consulte el programa, en particular el rubro de aprendizajes relevantes, los que deberán ser compartidos cotidianamente con los estudiantes para su reflexión y aplicación.
- Es conveniente insistir que el uso de la calculadora y de la computadora es un apoyo para el aprendizaje significativo de los alumnos, por lo que el profesor debe diseñar actividades de aprendizaje que utilicen esta tecnología.

## Índice de navegación del Objeto de aprendizaje

### 1. Introducción. Datos Bivariados

#### 1.1 Relación entre dos variables

Actividad 1: Investigar en qué consiste el método de mínimos cuadrados un método

Actividad 2: Investigar en qué consiste el método de mínimos cuadrados otro método

#### 1.2. Relación entre dos variables cuantitativas

Actividad 3: Explicar el ejemplo 1 con palabras propias

Actividad 4: Explicar el ejemplo 2 con palabras propias

#### 1.3. Regresión lineal y correlación

Actividad 5: Resolver de los problemas propuestos el 1 y el 2, en la hoja de cálculo Excel y en el software de Geogebra.

Actividad 6: Resolver el problema 3



Universidad Nacional  
Autónoma de México

Portal  
académico  
Colegio de Ciencias  
y Humanidades



Actividad 7: Resolver los problemas propuestos del 4 al 10, en la hoja de cálculo Excel y en el software de Geogebra

2. Actividad Final

3. Glosario

4. Referencias

5. Créditos



## DATOS BIVARIADOS

### UNIDAD II

**PROPÓSITO:** El alumno comprenderá la manera en que se establece una relación lineal entre dos variables a partir de tablas, diagramas, regresiones y correlaciones, para describir la naturaleza e intensidad de dicha relación.

#### *Estrategias:*

Presentar una serie de variables pertenecientes al entorno del alumno, para que seleccione parejas de variables, argumentando intuitivamente el grado de la relación entre ellas.

- A partir de problemas, discutir con los alumnos sobre la construcción e interpretación de las tablas de contingencia.
- Plantear problemas donde se haga notar que en un diagrama de dispersión es posible trazar una gran cantidad de rectas, pero que la que se obtiene con el método de mínimos cuadrados es la que mejor se ajusta a la nube de puntos.
- Dado un grupo de datos bivariados, solicitar a los alumnos que:  
Calculen y grafiquen la recta de regresión.
- Hagan predicciones del valor de alguna de las variables utilizando la regla de regresión, como un primer acercamiento a la idea de inferencia.
- Calculen el coeficiente de correlación.
- Discutan en equipo y con el grupo los resultados
- Utilizará la computadora con la hoja de cálculo Excel, Geogebra, para construir tablas de contingencia y gráficas de dispersión de datos bivariados agrupados y no agrupados.



Recta de mínimos cuadrados

Primer método

### Recta de Mínimos Cuadrados

$$y = mx + b$$

$$m = a_1 = \frac{N \sum X * Y - (\sum X) * (\sum Y)}{N \sum X^2 - (\sum X)^2} =$$

$$b = a_0 = \frac{(\sum Y) * (\sum X^2) - (\sum X) * (\sum X * Y)}{N \sum X^2 - (\sum X)^2} =$$

N =

$$\sum_{i=1}^n x_i =$$

$$\sum_{i=1}^n y_i =$$

$$\sum_{i=1}^n x_i * y_i =$$

$$\sum_{i=1}^n x_i^2 =$$

$$\sum_{i=1}^n y_i^2 =$$

$$\left( \sum_{i=1}^n x_i \right)^2 =$$

$$\left( \sum_{i=1}^n y_i \right)^2 =$$

**Karl Pearson** ([Londres 27 de marzo 1857](#) - Londres, [27 de abril](#) de [1936](#))

Coefficiente de Correlación de Karl Pearson

$$r = \frac{N \sum X * Y - (\sum X) * (\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2] * [N * \sum Y^2 - (\sum Y)^2]}}$$

### Recta de Mínimos Cuadrados



Segundo método

## Recta de Mínimos Cuadrados

$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$$

N =

$$\sum_{i=1}^n x_i =$$

$$\sum_{i=1}^n y_i =$$

$$\sum_{i=1}^n x_i^2 =$$

$$\sum_{i=1}^n y_i^2 =$$

$$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2}{N} - (\bar{x})^2 =$$

$$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2}{N} - (\bar{y})^2 =$$

$$\sigma_{xy} = \frac{\sum_{i=1}^n x_i y_i}{N} - (\bar{x} \cdot \bar{y}) =$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} =$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{N} =$$

$$\sigma_x =$$

$$\sigma_y =$$

Coefficiente de Correlación de Karl Pearson

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} =$$

## Recta de Mínimos Cuadrados



## Recta de Mínimos Cuadrados

$$\bar{x} = \frac{\sum_{i=1}^n x_i * f_i}{N} =$$

$$y = \quad + \quad * (x - \quad)$$

Dame la x

*Estimación*

$$\bar{y} = \frac{\sum_{i=1}^n y_i * f_i}{N} =$$

$$y = \underline{16,0}$$

$$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2 * f_i}{N} - (\bar{x})^2 =$$

$$\sigma_x =$$

$$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2 * f_i}{N} - (\bar{y})^2 =$$

$$\sigma_y =$$

$$\sigma_{xy} = \frac{\sum_{i=1}^n x_i * y_i * f_i}{N} - (\bar{x} * \bar{y}) =$$

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} =$$

$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$$



# Datos No Agrupados

$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i * y_i$

$\sum_{i=1}^n x_i =$	$\sum_{i=1}^n y_i =$	$\sum_{i=1}^n x_i^2 =$	$\sum_{i=1}^n y_i^2 =$	$\sum_{i=1}^n x_i * y_i =$
----------------------	----------------------	------------------------	------------------------	----------------------------

$N =$		
$\sum_{i=1}^n x_i =$	$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} =$	
$\sum_{i=1}^n y_i =$	$\bar{y} = \frac{\sum_{i=1}^n y_i}{N} =$	
$\sum_{i=1}^n x_i * y_i =$		
$\sum_{i=1}^n x_i^2 =$	$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2}{N} - (\bar{x})^2 =$	$\sigma_x =$
$\sum_{i=1}^n y_i^2 =$	$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2}{N} - (\bar{y})^2 =$	$\sigma_y =$

Recta de Minimos Cuadrados

$y = mx + b$

$$m = a_1 = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{[n \sum x_i^2 - (\sum x_i)^2] [n \sum y_i^2 - (\sum y_i)^2]}$$

$$m = a_1 = \frac{N \sum X * Y - (\sum X) * (\sum Y)}{N \sum X^2 - (\sum X)^2} =$$

$$b = a_0 = \frac{(\sum Y) * (\sum X^2) - (\sum X) * (\sum X * Y)}{N \sum X^2 - (\sum X)^2} =$$

$$\sigma_{xy} = \frac{\sum_{i=1}^n x_i y_i}{N} - (\bar{x} * \bar{y}) =$$

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} =$$



$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$$

## Datos Agrupados

$x_i$	$y_i$	$f_i$	$x_i * f_i$	$y_i * f_i$	$f_i * x_i^2$	$f_i * y_i^2$	$f_i * x_i * y_i$
2	7	1	2	7	4	49	14
2	7	1	2	7	4	49	14
2	7	1	2	7	4	49	14
2	7	1	2	7	4	49	14
$\sum x_i = 2$	$\sum y_i = 7$	$\sum f_i = 1$	$\sum x_i * f_i = 2$	$\sum y_i * f_i = 7$	$\sum f_i * x_i^2 = 4$	$\sum f_i * y_i^2 = 49$	$\sum f_i * x_i * y_i = 14$

### Recta de Minimos Cuadrados

$$\bar{x} = \frac{\sum_{i=1}^n x_i * f_i}{N} = \quad y = \quad + \frac{\quad}{\quad} * (x - \quad)$$

Dame la x

Estimación

$$y = \underline{16.0}$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i * f_i}{N} =$$

$$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2 * f_i}{N} - (\bar{x})^2 =$$

$$\sigma_x =$$

$$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2 * f_i}{N} - (\bar{y})^2 =$$

$$\sigma_y =$$

$$\sigma_{xy} = \frac{\sum_{i=1}^n x_i * y_i * f_i}{N} - (\bar{x} * \bar{y}) =$$

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} =$$

$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$$



Comentarios y Conclusiones:

Comentario:

Conclusión:

Ejemplo 1:

Una compañía de seguros considera que el número de vehículos ( $y$ ) que circulan por una determinada autopista a más de 120 km/h, puede ponerse en función del número de accidentes ( $x$ ) que ocurren en ella. Durante 5 días obtuvo los siguientes resultados:

<b>Accidentes (<math>x_i</math>)</b>	<b>5</b>	<b>7</b>	<b>2</b>	<b>1</b>	<b>9</b>
<b>Número de vehículos (<math>y_i</math>)</b>	15	18	10	8	20

- Calcula el coeficiente de correlación lineal.
- Si ayer se produjeron 6 accidentes, ¿cuántos vehículos podemos suponer que circulaban por la autopista a más de 120 km / h?
- ¿Es buena la predicción?

Construimos una tabla, teniendo en cuenta que la frecuencia absoluta es uno. Debemos conocer la media aritmética de las dos variables, las varianzas, las desviaciones típicas y la covarianza.

		Media aritmética		Varianza		Covarianza
	$f_i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i * y_i$
	1	5	15	25	225	75
	1	7	18	49	324	126
	1	2	10	4	100	20
	1	1	8	1	64	8
	1	9	20	81	400	180
$\Sigma$	N=5	24	71	160	1113	409

## Medias aritméticas

$$\bar{x} = \frac{\sum x_i}{N} = \frac{24}{5} = 4,8 \quad \bar{x} = 4,8 \quad \bar{y} = \frac{\sum y_i}{N} = \frac{71}{5} = 14,2 \quad \bar{y} = 14,2$$

## Varianzas y desviaciones típicas

$$\sigma_x^2 = \frac{\sum (x_i)^2}{N} - (\bar{x})^2 = \frac{160}{5} - (4,8)^2 = 8,96 \quad \sigma_x = \sqrt{8,96} = 2,993$$

$$\sigma_y^2 = \frac{\sum (y_i)^2}{N} - (\bar{y})^2 = \frac{1113}{5} - (14,2)^2 = 20,96 \quad \sigma_y = \sqrt{20,96} = 4,578$$

$$\text{Covarianza } \sigma_{xy} \Rightarrow \sigma_{xy} = \frac{\sum x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{409}{5} - (4,8 \cdot 14,2) = 13,64$$

$$\text{a) Correlación lineal de Pearson } r \Rightarrow r = \frac{\sigma_{xy}}{\sigma_x \times \sigma_y} = \frac{13,64}{2,993 \cdot 4,578} = 0,995 \Rightarrow r = 0,995$$

Comentarios:

La covarianza es positiva, correlación directa. Al aumentar la velocidad aumentará el número de accidentes.

El valor de r está muy próximo a 1, la estimación realizada estará muy cerca del valor real.  
Dependencia funcional fuerte.

$$\text{b) Recta de regresión de } y \text{ sobre } x \Rightarrow y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x})$$

$$y = 14,2 + \frac{13,64}{8,96} (x - 4,8) \Rightarrow y = 14,2 + 1,52(x - 4,8) \Rightarrow y = 1,52x + 6,9$$

Para  $x = 6$  accidentes el número de vehículos estimado es:  $y = 1,52 \cdot 6 + 6,9 = 16$

Podemos suponer que ayer circulaban 16 vehículos a más de 120 km/h

c) La predicción hecha es buena, el coeficiente de correlación está muy próximo a uno.



Solucion en la hoja de calculo Excel:

Datos Agrupados							
$x_i$	$y_i$	$f_i$	$x_i * f_i$	$y_i * f_i$	$f_i * x_i^2$	$f_i * y_i^2$	$f_i * x_i * y_i$
5	15	1	5	15	25	225	75
7	18	1	7	18	49	324	126
2	10	1	2	10	4	100	20
1	8	1	1	8	1	64	8
9	20	1	9	20	81	400	180
$\sum_{i=1}^n x_i =$	$\sum_{i=1}^n y_i =$	$\sum_{i=1}^n f_i =$	$\sum_{i=1}^n x_i * f_i =$	$\sum_{i=1}^n y_i * f_i =$	$\sum_{i=1}^n f_i * x_i^2 =$	$\sum_{i=1}^n f_i * y_i^2 =$	$\sum_{i=1}^n f_i * x_i * y_i =$
24	71	5	24	71	160	1113	409

**Recta de Mínimos Cuadrados**

$$\bar{x} = \frac{\sum_{i=1}^n x_i * f_i}{N} = \frac{24}{5} = 4,80$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i * f_i}{N} = \frac{71}{5} = 14,2$$

$$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2 * f_i}{N} - (\bar{x})^2 = \frac{160}{5} - (4,80)^2 = 8,96 \quad \sigma_x = 2,993326$$

$$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2 * f_i}{N} - (\bar{y})^2 = \frac{1113}{5} - (14,2)^2 = 20,96 \quad \sigma_y = 4,578209$$

$$\sigma_{xy} = \frac{\sum_{i=1}^n x_i * y_i * f_i}{N} - (\bar{x} * \bar{y}) = \frac{409}{5} - (4,80 * 14,2) = 13,64$$

$$r = \frac{\sigma_{xy}}{\sigma_x * \sigma_y} = 0,99532$$

Dame la x **6**

**Estimación**

$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$$

$$y = 14,2 + \frac{13,64}{8,96} * (6 - 4,80)$$

$$y = 16,0$$

[prediccion estadistica \(Accidentes vs Vehiculos\).xlsx](#)

[accidentes de automovil.ggb](#)



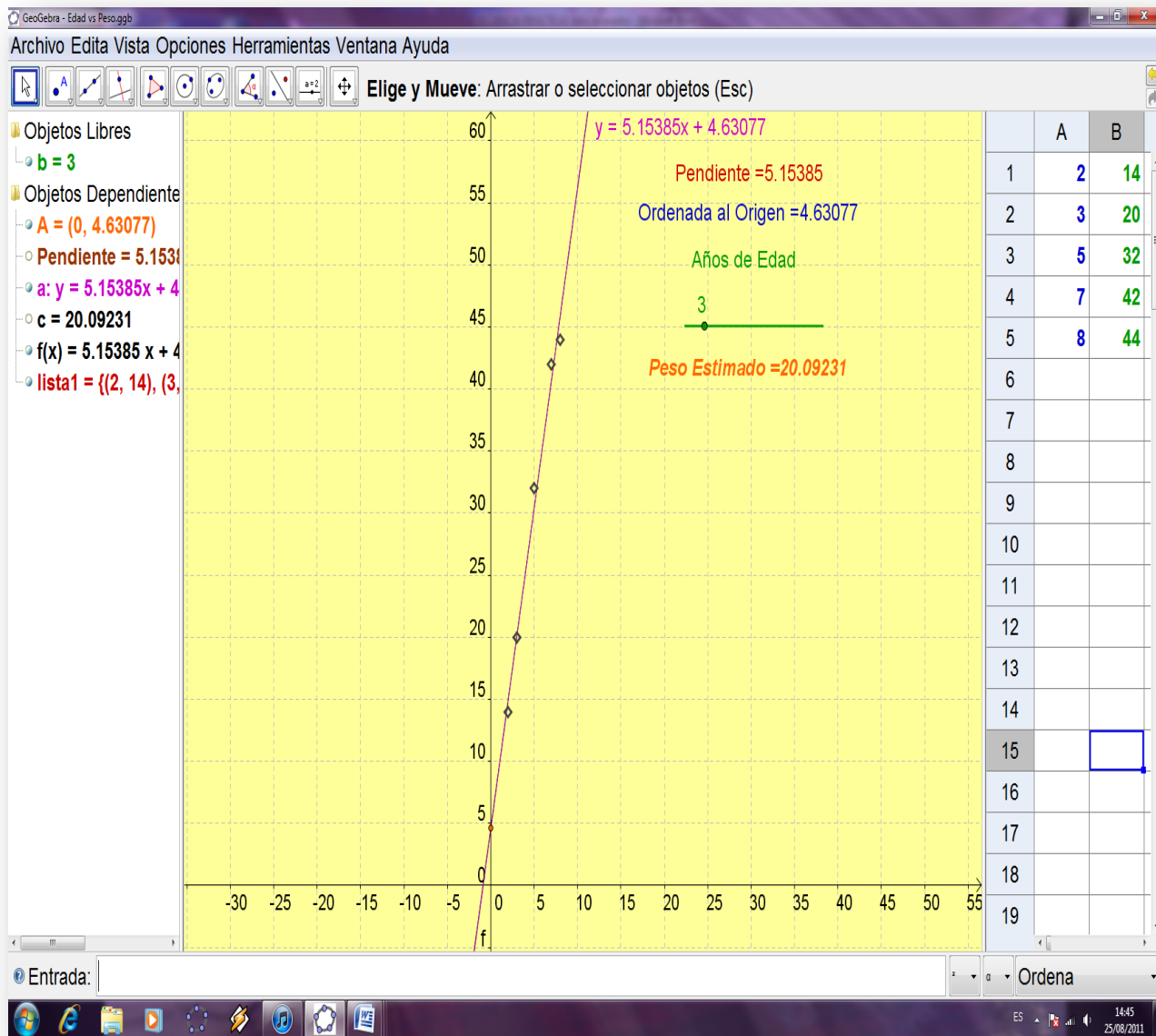
### Ejemplo 2:

Problema.- Cinco niños de 2, 3, 5, 7 y 8 años de edad pesan, respectivamente 14, 20, 32, 42 y 44 kilos.

Hallar la ecuación de la recta de regresión de la edad sobre el peso.

¿Cuál sería el peso aproximado de un niño de seis años?

Solución en Geogebra:





Solución en la Hoja de Cálculo Excel: (Primer método)

Solución en la Hoja de Cálculo Excel: (Segundo método)

Datos Agrupados							
$x_i$	$y_i$	$f_i$	$x_i * f_i$	$y_i * f_i$	$f_i * x_i^2$	$f_i * y_i^2$	$f_i * x_i * y_i$
2	14	1	2	14	4	196	28
3	20	1	3	20	9	400	60
5	32	1	5	32	25	1024	160
7	42	1	7	42	49	1764	294
8	44	1	8	44	64	1936	352
$\sum_{i=1}^n x_i = 25$	$\sum_{i=1}^n y_i = 152$	$\sum_{i=1}^n f_i = 5$	$\sum_{i=1}^n x_i * f_i = 25$	$\sum_{i=1}^n y_i * f_i = 152$	$\sum_{i=1}^n f_i * x_i^2 = 151$	$\sum_{i=1}^n f_i * y_i^2 = 5320$	$\sum_{i=1}^n f_i * x_i * y_i = 894$

Recta de Mínimos Cuadrados			
$N = 5$	$\bar{x} = \frac{\sum_{i=1}^n x_i * f_i}{N} = 5,00$	$y = 30,4 + \frac{26,8}{5,2} * (x - 5,00)$	
		<b>Dame la x</b> 2	
		<b>Estimación</b>	
	$\bar{y} = \frac{\sum_{i=1}^n y_i * f_i}{N} = 30,4$	$y = 14,94$	
	$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2 * f_i}{N} - (\bar{x})^2 = 5,2$	$\sigma_x = 2,280351$	$m = 5,15385$
			$b = 4,6308$
	$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2 * f_i}{N} - (\bar{y})^2 = 139,84$	$\sigma_y = 11,8254$	
	$\sigma_{xy} = \frac{\sum_{i=1}^n x_i * y_i * f_i}{N} - (\bar{x} * \bar{y}) = 26,8$		
	$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = 0,99384$		
	$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$		






## Datos No Agrupados

$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i * y_i$					
2	14	4	196	28					
3	20	9	400	60					
5	32	25	1024	160					
7	42	49	1764	294					
8	44	64	1936	352					
$\sum_{i=1}^n x_i =$	25	$\sum_{i=1}^n y_i =$	152	$\sum_{i=1}^n x_i^2 =$	151	$\sum_{i=1}^n y_i^2 =$	5320	$\sum_{i=1}^n x_i * y_i =$	894

### prediccion estadistica (Edad vs Peso).xlsx

$\sum_{i=1}^n x_i =$	$N =$	5	$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} =$	5	
$\sum_{i=1}^n y_i =$		152	$\bar{y} = \frac{\sum_{i=1}^n y_i}{N} =$	30,4	
$\sum_{i=1}^n x_i * y_i =$		894	$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2}{N} - (\bar{x})^2 =$	5,2	$\sigma_x =$ 2,28035
$\sum_{i=1}^n x_i^2 =$		151	$\sigma_y^2 = \frac{\sum_{i=1}^n y_i^2}{N} - (\bar{y})^2 =$	139,84	$\sigma_y =$ 11,8254
$\sum_{i=1}^n y_i^2 =$		5320	$\sigma_{xy} = \frac{\sum_{i=1}^n x_i y_i}{N} - (\bar{x} \cdot \bar{y}) =$	26,8	
<b>Recta de Mínimos Cuadrados</b>					
$y = mx + b$					
$r = \frac{N \sum X * Y - (\sum X) * (\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2] * [N * \sum Y^2 - (\sum Y)^2]}}$			0,99384 $r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} =$ 0,99384		
$m = a_1 = \frac{N \sum X * Y - (\sum X) * (\sum Y)}{N \sum X^2 - (\sum X)^2} =$			5,15385		
$b = a_0 = \frac{(\sum Y) * (\sum X^2) - (\sum X) * (\sum X * Y)}{N \sum X^2 - (\sum X)^2} =$			4,63077		
					
$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} * (x - \bar{x})$					
<b>Dame la x</b> 6					
<b>Estimación</b>					
<b>y = 35,55</b>					



**Problemas Propuestos:**

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

<b>Nº de clientes (X)</b>	<b>8</b>	<b>7</b>	<b>6</b>	<b>4</b>	<b>2</b>	<b>1</b>
<b>Distancia (Y)</b>	<b>15</b>	<b>19</b>	<b>25</b>	<b>23</b>	<b>34</b>	<b>40</b>

a).- Calcular el coeficiente de correlación lineal.

b).- Si el centro comercial se sitúa a 2 km, ¿cuántos clientes puede esperar?

c).- Si desea recibir a 500 clientes, ¿a qué distancia del núcleo de población debe situarse?

2.- Las notas obtenidas por cinco alumnos en Matemáticas y Química son:

<b>Matemáticas</b>	<b>6</b>	<b>4</b>	<b>8</b>	<b>5</b>	<b>3.5</b>
<b>Química</b>	<b>6.5</b>	<b>4.5</b>	<b>7</b>	<b>5</b>	<b>4</b>

Determinar las **rectas de regresión** y calcular la nota esperada en Química para un alumno que tiene 7.5 en Matemáticas.

3.- Un conjunto de datos bidimensionales (X, Y) tiene **coeficiente de correlación**  $r = -0.9$ , siendo las medias de las distribuciones marginales  $\bar{x} = 1$ ,  $\bar{y} = 2$ .

Se sabe que una de las cuatro ecuaciones siguientes corresponde a la **recta de regresión** de Y sobre X:

$$y = -x + 2; 3x - y = 1; 2x + y = 4; y = x + 1$$

Seleccionar **razonadamente** esta recta.



4.- Las estaturas (centímetros) y pesos (kilogramos) de 10 jugadores de baloncesto de un equipo son:

Estatura (X)	186	189	190	192	193	193	198	201	203	205
Pesos (Y)	85	85	86	90	87	91	93	103	100	101

Calcular:

- La recta de regresión de Y sobre X.
- El coeficiente de correlación.
- El peso estimado de un jugador que mide 208 cm.

5.- A partir de los siguientes datos referentes a horas trabajadas en un taller (X), y a unidades producidas (Y), determinar la **recta de regresión** de Y sobre X, el **coeficiente de correlación lineal** e interpretarlo.

Horas (X)	80	79	83	84	78	60	82	85	79	84	80	62
Producción (Y)	300	302	315	330	300	250	300	340	315	330	310	240

6.- Se ha solicitado a un grupo de 50 individuos información sobre el número de horas que dedican diariamente a dormir y ver la televisión. La clasificación de las respuestas ha permitido elaborar la siguiente tabla:

Nº de horas dormidas (X)	6	7	8	9	10
Nº de horas de televisión (Y)	4	3	3	2	1
Frecuencias absolutas ( $f_i$ )	3	16	20	10	1

- Calcular el coeficiente de correlación.
- Determinar la ecuación de la recta de regresión de Y sobre X.



c).- Si una persona duerme ocho horas y media, ¿cuánto cabe esperar que vea la televisión?

7.- El número de horas dedicadas al estudio de una asignatura y la calificación obtenida en el examen correspondiente, de ocho personas es:

<b>Horas (X)</b>	<b>20</b>	<b>16</b>	<b>34</b>	<b>23</b>	<b>27</b>	<b>32</b>	<b>18</b>	<b>22</b>
<b>Calificación (Y)</b>	<b>6.5</b>	<b>6</b>	<b>8.5</b>	<b>7</b>	<b>9</b>	<b>9.5</b>	<b>7.5</b>	<b>8</b>

Hallar la:

a).- Recta de regresión de Y sobre X.

b).- Calificación estimada para una persona que hubiese estudiado 28 horas.

8.- En la tabla siguiente se indica la edad (en años) y la conducta agresiva (medida en una escala de cero a 10) de 10 niños.

Edad	<b>6</b>	<b>6</b>	<b>6.7</b>	<b>7</b>	<b>7.4</b>	<b>7.9</b>	<b>8</b>	<b>8.2</b>	<b>8.5</b>	<b>8.9</b>
Conducta Agresiva	<b>9</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>7</b>	<b>4</b>	<b>2</b>	<b>3</b>	<b>3</b>	<b>1</b>

a).- Obtener la recta de regresión de la conducta agresiva en función de la edad.

b).- A partir de dicha recta, obtener el valor de la conducta agresiva que correspondería a un niño de 7.2 años.

9.- Las calificaciones de 40 alumnos en psicología evolutiva y en estadística han sido las de la tabla adjunta.

<b>Psicología (x)</b>	3	4	5	6	6	7	7	8	10
<b>Estadística (y)</b>	2	5	5	6	7	6	7	9	10
<b>Nº de alumnos (fi)</b>	4	6	12	4	5	4	2	1	2

a).- Obtener la ecuación de la recta de regresión de calificaciones de estadística respecto de las calificaciones de psicología.



b).- ¿Cuál será la nota esperada en estadística para un alumno que obtuvo un 4,5 en psicología?

10.-Las notas obtenidas por 10 alumnos en Matemáticas y en Música son:

Matemáticas	6	4	8	5	3,5	7	5	10	5	4
Música	6,5	4,5	7	5	4	8	7	10	6	5

- Calcula la covarianza y el coeficiente de correlación.
- ¿Existe correlación entre las dos variables?
- ¿Cuál será la nota esperada en Música para un alumno que hubiese obtenido un 8,3 en Matemáticas?

#### BIBLIOGRAFÍA BÁSICA:

Spiegel, M.	Estadística	Mc Graw Hill, 1970
Spiegel, M.	Probabilidad y Estadística	Mc Graw Hill, 1975
Stevenson, W.	Probabilidad y Estadística para Administración y Economía	Harla, 1994
Christensen, H.	Estadística paso a paso	Trillas, 1997
Daniel, W.	Estadística Aplicada a las Ciencias Sociales y a la Educación	Mc Graw Hill, 1998
Mendenhall, W.	Estadística para Administración y Economía	Iberoamérica, 1978
Johnson, R.	Estadística Elemental	Iberoamérica, 1990
Willoughby, S.	Probabilidad y Estadística	PCSA, 1993
Wonnacott, T.	Fundamentos de Estadística para Administración y Economía	Limusa, 1989

[prediccion estadistica \(Jaime Ramirez Sanchez\).xlsx](#)  
[Curso Profesores regresion y correlacion](#)