

## **.Cátedra No 4 UNIDAD II. (continuación)**

Otra clasificación de las regresiones la podemos ver al contar en número de variables. Los casos anteriores de:

$$y = 2x + 1 \quad y \quad y = 10 - 2x$$

Observamos que tienen 2 variables, Una Independiente y otra dependiente. Las Regresiones Múltiple tienen más de dos variables, o bien la variable independiente tiene

### **2.6 Métodos de ajuste de la curva que describe un diagrama de dispersión.**

La línea de estimación es aquella que describe la tendencia que siguen los puntos de un diagrama de dispersión en un estudio de regresión.

Existen varias formas para determinar una línea de estimación, pero se distinguen 2 metodologías principalmente.

- Método Gráfico libre
- Método de la ecuación de una recta: Método de los mínimos Cuadrados.

#### **2.6.1 Método gráfico libre.**

En este método se procura trazar una línea por donde el investigador visualiza que va la tendencia de los puntos del diagrama de dispersión en un estudio de regresión.

Tiene en contra, de que por su carácter subjetivo, la línea determinada va estar supeditada al criterio, o visualización del investigador.

Así en un estudio de regresión, por el método gráfico libre se podrán tener tantas líneas de estimación como investigadores opinen al respecto.

Por este carácter subjetivo este método es poco empleado y nosotros nunca lo solicitaremos en una práctica o prueba.

#### **2.6.2 Método de la ecuación de una recta: Método de los mínimos Cuadrados.**

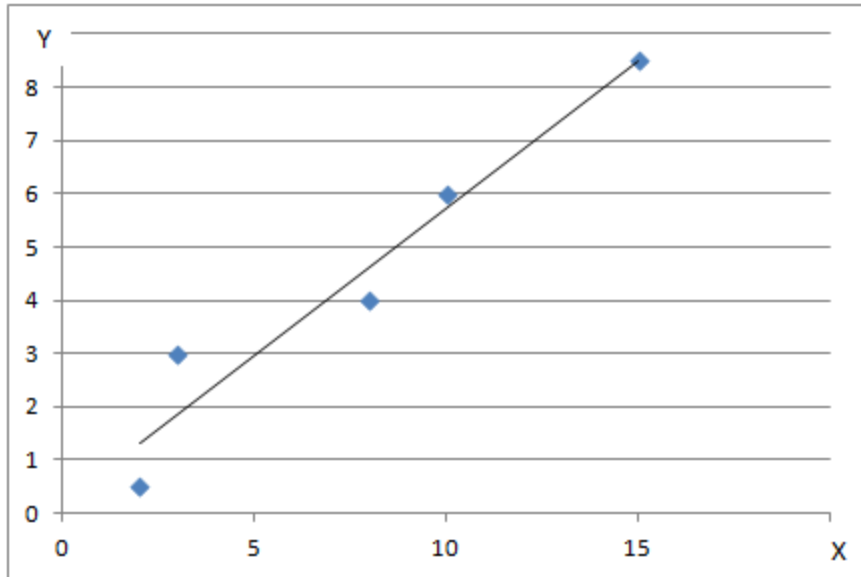
En esta metodología, se trata de obtener una ecuación que represente la línea de estimación en un estudio de regresión. Existen varias formas de obtener esta ecuación, pero el método más empleado por su grado de precisión y facilidad de cálculo es el método de los mínimos cuadrados.

En el método de los mínimos cuadrados el procedimiento consiste en minimizar la suma de los cuadrados de las distancias verticales entre los valores de los datos y los de la regresión estimada, es decir, minimizar la suma de los residuos al cuadrado, teniendo como residuo la diferencia de los datos observados y los valores del modelo (línea).

Así para el siguiente conjunto de puntos

<b>x</b>	2	3	8	10	15
<b>y</b>	0.5	3	4	6	8.5

Tendremos el siguiente diagrama de dispersion con su línea de estimación:



Observaremos los errores entre los puntos reales, y los puntos observados sobre la línea serán:

<b>y</b>	<b><math>\hat{y}</math></b>	<b><math>y - \hat{y}</math></b>
0.5	1.5	-1
3	1.8	1.2
4	4.6	-0.6
6	5.8	0.2
8.5	8.5	0

-0.2 Residuo

Dado un conjunto de  $n$  parejas de datos  $(x,y)$ , se debe encontrar la ecuación de la recta

$$y = a + bx$$

que pasa lo mas cerca posible de los puntos.

El método del ajuste por mínimos cuadrados o regresión lineal, permite obtener los valores de las constantes **a** y **b**, siendo

a = valor constante (numero) que representa el punto de intercepción de la línea de estimación con el eje de la

b = valor constante (numero) que representa la pendiente de la línea de estimación.

x=Variable independiente

$\hat{y}$ = Variable dependiente estimada.

Los valores de las constantes b y a se calculan por:

$$b = \frac{\sum xy - n (\bar{x})^2 (\bar{y})^2}{\sum x^2 - n (\bar{x})^2}$$

$$a = \bar{y} - b \bar{x}$$

En la proxima clase se observara una capitulacion con estos conceptos, y tendremos el 1er ejemplo aplicado.