

Measure the General Trend

قياس الاتجاه العام

2) Least Squares Method

طريقة المربعات الصغرى

The least squares method is one of the most accurate methods for estimating the equation of the general trend line. Assuming that the general trend line is a straight line, its equation is written as follows:

$$y_i = a + bt_i + \varepsilon_i$$

Where:

y_i : represents the value of the observation with the number (i).

تمثل قيمة المشاهدة

t_i : represents the sequence of the year with the number (i).

تمثل تسلسل السنة

ε_i : represents the random error with the number (i).

تمثل الخطأ العشوائي

a, b: are constants that represent the parameters of the general trend line equation.

a, b عبارة عن ثوابت تمثل معلمات معادلة الاتجاه العام الخطي

Using the least squares method that makes the sum of squares of the error ($\sum_{i=1}^n \varepsilon_i^2$) as small as possible and after performing partial differentiation for (a and b), we obtain the estimated values of the parameters of the general trend line (\hat{b} and \hat{a}) as follows:

$$\varepsilon_i = y_i - a - bt_i$$

$$\varepsilon_i^2 = (y_i - a - bt_i)^2$$

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - a - bt_i)^2 \quad \dots (1)$$

Taking the derivative (partial differential) of equation (1) with respect to (a) and setting the derivative equal to zero, we get:

بأخذ المشتقة (التفاضل الجزئي) للمعادلة (1) بالنسبة الى (a) ومساوات المشتقة بالصفر نحصل على:

$$\frac{\partial \sum_{i=1}^n \varepsilon_i^2}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bt_i)$$

$$0 = -2 \sum_{i=1}^n (y_i - a - bt_i)$$

Multiply both sides of the above equation by (-1/2)

$$0 = \sum_{i=1}^n (y_i - a - bt_i)$$

$$0 = \sum_{i=1}^n y_i - na - b \sum_{i=1}^n t_i$$

$$\sum_{i=1}^n y_i = na + b \sum_{i=1}^n t_i \quad \dots (2)$$

$$na = \sum_{i=1}^n y_i - b \sum_{i=1}^n t_i$$

$$\hat{a} = \frac{\sum_{i=1}^n y_i}{n} - b \frac{\sum_{i=1}^n t_i}{n} \quad \dots (3)$$

By taking the derivative of equation (1) with respect to (b) and also setting the derivative equal to zero, we get:

$$\frac{\partial \sum_{i=1}^n \varepsilon_i^2}{\partial b} = -2 \sum_{i=1}^n (y_i - a - bt_i)t_i$$

$$0 = -2 \sum_{i=1}^n (y_i - a - bt_i)t_i$$

Multiply both sides of the above equation by (-1/2)

$$0 = \sum_{i=1}^n (y_i - a - bt_i)t_i$$

$$0 = \sum_{i=1}^n y_i t_i - a \sum_{i=1}^n t_i - b \sum_{i=1}^n t_i^2 \quad \dots (4)$$

By substituting the value of (\hat{a}) in the above equation, we get the value of \hat{b}

$$\sum_{i=1}^n y_i t_i = \left(\frac{\sum_{i=1}^n y_i}{n} - b \frac{\sum_{i=1}^n t_i}{n} \right) \sum_{i=1}^n t_i + b \sum_{i=1}^n t_i^2$$

$$\sum_{i=1}^n y_i t_i = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n t_i}{n} - b \frac{(\sum_{i=1}^n t_i)^2}{n} + b \sum_{i=1}^n t_i^2$$

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

Multiply both sides of the above equation by (n)

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

$$\hat{b} = \frac{n \sum_{i=1}^n y_i t_i - \sum_{i=1}^n y_i \sum_{i=1}^n t_i}{n \sum_{i=1}^n t_i^2 - \left(\sum_{i=1}^n t_i \right)^2} \dots (5)$$

By substituting the estimated values of the two parameters (\hat{b} and \hat{a}) in the general trend line equation, we obtain the estimated (predictive) general trend line equation as follows:

$$\hat{y} = \hat{a} + \hat{b}t$$

ملاحظة:

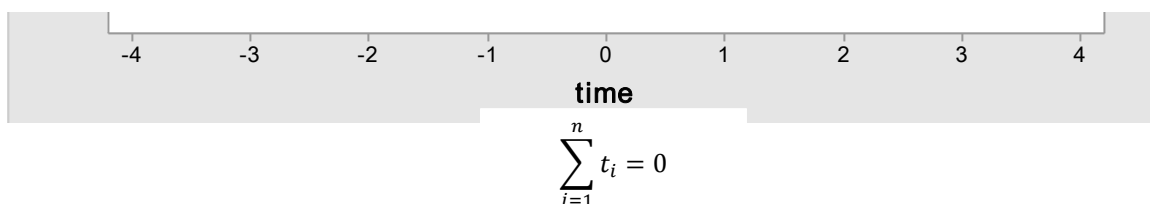
يتم استخدام المعادلتين (3) و (5) لتقدير المعلمتين \hat{a} و \hat{b} في حالة اعطاء السنوات تسلسلات ($t=1,2,3,4,5,\dots$) بغض النظر عن عدد سنوات السلسلة الزمنية. إذا كانت قيمة (\hat{b}) اشارتها موجبة هذا يعني ان قيمة الظاهرة (\hat{y}) في تزايد مستمر أما إذا كانت قيمة (\hat{b}) اشارتها سالبة فان قيمة (\hat{y}) في تناقص.

The calculations in finding the values of (\hat{b} and \hat{a}) can be simplified by moving the origin point with respect to time to the middle values. Then, to determine the origin point, there are two cases:

1) **If the number of values is odd:**

إذا كان عدد القيم فردي

The time value takes zero at the middle value, so the time periods following it are (1,2,3,...) and the periods preceding it are (-1,-2,-3,...) and thus the sum of the values is equal to zero, i.e. ($\sum_{i=1}^n t_i = 0$) as shown in the following figure:



Using the previous natural equations to determine the value of $(\hat{b}$ and $\hat{a})$ using the least squares method, we substitute in equation (2) for $\sum_{i=1}^n t_i = 0$, and we get the value of \hat{a} :

$$\sum_{i=1}^n y_i = na + b \sum_{i=1}^n t_i$$

$$na = \sum_{i=1}^n y_i - b(0) \rightarrow na = \sum_{i=1}^n y_i$$

$$\hat{a} = \frac{\sum_{i=1}^n y_i}{n} \quad \dots (6)$$

We substitute in equation (4) for $\sum_{i=1}^n t_i = 0$ and we get the value of \hat{b} :

$$0 = \sum_{i=1}^n y_i t_i - a \sum_{i=1}^n t_i - b \sum_{i=1}^n t_i^2$$

$$0 = \sum_{i=1}^n y_i t_i - a(0) - b \sum_{i=1}^n t_i^2$$

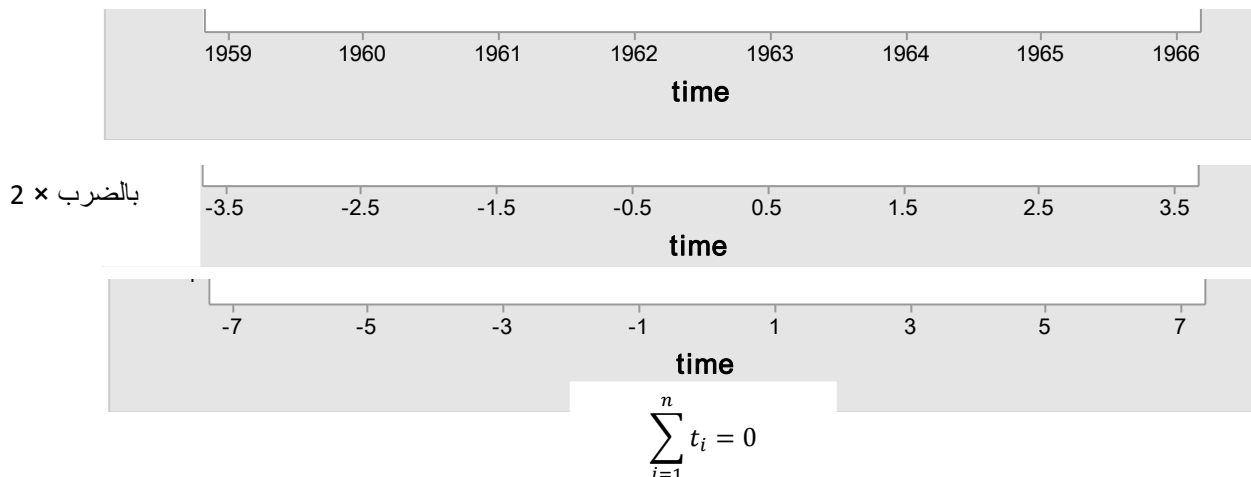
$$\sum_{i=1}^n y_i t_i = b \sum_{i=1}^n t_i^2$$

$$\hat{b} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i^2} \quad \dots (7)$$

2) If the number of values is even:

إذا كان عدد القيم زوجي

In this case, we take zero between the two middle values. If we consider the unit of time to be half the period, then the time values in the first half of the series will be (1,3,5,7,...) and in the upper half of the series (-1,-3,-5,-7,...) Thus, $\sum_{i=1}^n t_i = 0$ will be as shown in the following figure:



Example: Find the equation of the general trend line by least squares for the following time series when $\sum_{i=1}^n t_i = 0$ and forecast the value of production for the year 1990.

Years	Production	t	t*y	t ²
1980	10.0	-4	-40.0	16
1981	12.7	-3	-38.1	9
1982	12.4	-2	-24.8	4
83	11.9	-1	-11.9	1
84	12.5	0	0.0	0
85	13.0	1	13.0	1
86	14.9	2	29.8	4
87	16.5	3	49.5	9
88	18.7	4	74.8	16
Sum	$\sum_{i=1}^n y_i = 122.6$	$\sum_{i=1}^n t_i = 0$	$\sum_{i=1}^n y_i t_i = 52.3$	$\sum_{i=1}^n t_i^2 = 60$

$$\hat{y} = \hat{a} + \hat{b}t$$

$$\hat{a} = \frac{\sum_{i=1}^n y_i}{n} = \frac{122.6}{9} = 13.62$$

$$\hat{b} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i^2} = \frac{52.3}{60} = 0.87$$

$$\hat{y}_i = 13.62 + 0.87t$$

We replace the value of t in the equation above with the number (6) because the sequence of the year 1990 in the years is the number (6).

$$\hat{y} = 13.62 + 0.87(6) = 18.84$$

Example (2): The following data represents Iraq's imports of reinforcing steel in thousands of tons for the period 1991-2000.

Required:

- 1) Estimate the general trend line equation using the least squares method when $\sum_{i=1}^n t_i = 0$.
- 2) Estimate the general trend line equation using the series half-means method. (Homework)

Years	Imports	t	t*y	t ²
1991	40	-9	-360	81
92	35	-7	-245	49
93	65	-5	-325	25
94	50	-3	-150	9
95	60	-1	-60	1
96	70	1	70	1
97	80	3	240	9
98	75	5	375	25
99	85	7	595	49
2000	90	9	810	81
Sum	$\sum_{i=1}^n y_i = 650$	$\sum_{i=1}^n t_i = 0$	$\sum_{i=1}^n y_i t_i = 950$	$\sum_{i=1}^n t_i^2 = 330$

$$\hat{y} = \hat{a} + \hat{b}t$$

$$\hat{a} = \frac{\sum_{i=1}^n y_i}{n} = \frac{650}{10} = 65$$

$$\hat{b} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i^2} = \frac{950}{330} = 2.88$$

$$\hat{y}_i = 65 + 2.88t$$