Linearization of Nonlinear Equations

- In some cases, it is not easily understood whether the equation is linear or not.

- In these cases, the equations are reorganized to resemble the equation of "**y = a x + b**." The example of this situation is given in Table 2.2.

Tablo 3 Linearization of some nonlinear equations (y = ax + b)

Nonlinear equation	Linear equation	y	x	Slope	Intercept
$y = \frac{a}{x} + b$					
$y = \frac{1}{ax + b}$					
$y = \frac{x}{ax + b}$					
$y = b x^a$					
$y = b a^x$ (exp function, where «x» occurs as an exponent)					

Tablo 3 Linearization of some nonlinear equations

Nonlinear equation	Linear equation	y	x	Slope	Intercept
$y = \frac{a}{x} + b$	$y = a \frac{1}{x} + b$	y	1/x	a	b
$y = \frac{1}{ax + b}$	$\frac{1}{y} = a(x) + b$	1/y	x	a	b
$y = \frac{x}{ax + b}$	$\frac{x}{y} = a(x) + b$	x/y	x	a	b
$y = b x^a$	$\log y = a \log(x) + \log b$	log y	log x	a	log b
$y = b a^x$	$\log y = \log a (x) + \log b$	log y	x	log a	log b

Linear Regression

- Once the experimental data is plotted, some deviation from linearity usually occurs (not all data points are on the straight line

- The question is: «**where the straight line should pass?**"

✓ **1) The answer would be:**

«Freehand method of curve fitting»

✓ **2)** The **correct** answer is:

Run «**linear regression analysis.**»

- **Slope (a)** and **intercept (b)** are calculated from the following statistical equations:

$$\mathbf{a} = \frac{\Sigma\, xy - (\Sigma x\, \Sigma y / n)}{\Sigma\, x^2 - [(\Sigma x)^2 / n)]}$$

$$\mathbf{b} = \frac{\Sigma\, y\, \Sigma\, x^2 - \Sigma x\, \Sigma xy}{n\, (\Sigma\, x^2 - [(\Sigma x)^2 / n)])}$$

- Linear regression analysis has two purposes:

  - To determine the equation of straight line **(regression equation)**.

  - To determine if a relation exists between **dependent (y)** and **independent (x)** variables (calculate «**r**» or «R^2»).

Correlation coefficient (r)

- Shows whether or not there is a relationship between **two dependent (y) variables**.

Example: **phenolic content** vs **antioxidant activity**. (The higher the phenolic content of foods, the higher the antioxidant activity of foods)

$$r = \frac{n \sum xy - \sum x \sum y}{[[n \sum x^2 - (\sum x)^2]\ [n \sum y^2 - (\sum y)^2]]^{1/2}}$$

- Relationship between x and y can be increasing or decreasing.
- “r” is the same sign as slope.
- Values of “r” change from **<u>0</u>** to **<u>1.</u>**
- «r» value should be reported as at least **<u>three digits</u>** (such as, r= 0.975).

- If straight line passes from all data points, then «r» will be equal to 1.

- If data points deviate from straight line, then «r» will be less than 1.

- If «r» value is **over 0.8**, then relationship between «x» and «y» are good;

- If «r» value is **lower than 0.5**, then relationship between «x» and «y» is poor.

Determination coefficient (R^2)

- Shows whether or not there is a relationship between **independent (x)** and **dependent (y)** variables.

Example: **phenolic content** vs **storage time**

(As the storage time increases, the phenolic content of food decreases)

- Best way to test whether experimental data fits regression equation **(y = ax + b)** is to determine R^2. (measurement of straight line)

- If R^2 is equal to 1, this proves that experimental data fits perfectly to regression equation. (all data points are on the straight line)

- The more the experimental data points, the higher the confidence to R^2.

- R^2 is also used to decide **reaction order** for a quality factor in question (or intrest).

- R^2 is calculated using the following equation.

$$R^2 = 1 - \frac{\Sigma (Yi - \hat{Y})^2}{\Sigma (Yi - \tilde{Y})^2}$$

where;

Y_i : Experimental data,

$\hat{Y}$: Data calculated from regression equation **(corrected y values),**

$\tilde{Y}$: Mean value of experimental data.

- **Example 2.5:** By using the same experimental data in Example 2.1 for ascorbic acid degradation;

a) Find out regression coefficients (**a** and **b**) and $\mathbf{R^2}$ by “linear regression method.”

a) Draw the straight line using regression equation.

Table 2.1 Ascorbic acid contents of orange juice stored at 30°C

Time (days)	Ascorbic acid conc. ($mg\ L^{-1}$)
2	457
4	305
5	251
6	148

$$a = \frac{\Sigma\, xy - (\Sigma x\, \Sigma y / n)}{\Sigma\, x^2 - [(\Sigma x)^2 / n)]}$$

$$b = \frac{\Sigma\, y\, \Sigma\, x^2 - \Sigma x\, \Sigma xy}{n\, (\Sigma\, x^2 - [(\Sigma x)^2 / n)])}$$

Table 2.4 Regression data

X	X^2	Y	Y^2	XY
2		457		
4		305		
5		251		
6		148		
$\Sigma X=$	$\Sigma X^2=$	$\Sigma Y=$	$\Sigma Y^2=$	$\Sigma XY=$

Table 2.4 Regression data

X	X^2	Y	Y^2	XY
2	4	457	208 849	914
4	16	305	93 025	1220
5	25	251	63 001	1255
6	36	148	21 904	888
$\Sigma X=17$	$\Sigma X^2= 81$	$\Sigma Y=1161$	$\Sigma Y^2=$ 386 779	$\Sigma XY=$ 4277

$\Sigma X=17;$ $\Sigma X^2= 81;$ $\Sigma Y=1161$

$\Sigma Y^2=386\ 779;$ $\Sigma XY= 4277$

$$\mathbf{a} = \frac{\Sigma\, xy \;-\; (\Sigma x\, \Sigma y \,/\, n)}{\Sigma\, x^2 - [(\Sigma x)^2 / n)]}$$

$$\mathbf{b} = \frac{\Sigma\, y\, \Sigma\, x^2 \;-\; \Sigma x\, \Sigma xy}{n\,(\Sigma\, x^2 - [(\Sigma x)^2 / n)])}$$

$a = -75.11 \text{ mg L}^{-1} \text{ day}^{-1}$

$b = 609.49 \text{ mg L}^{-1}$

$$R^2 = 1 - \frac{\Sigma (Yi - \hat{Y})^2}{\Sigma (Yi - \tilde{Y})^2}$$

where;

Y_i : Experimental data,

$\hat{Y}$: Data calculated from regression equation,

$\tilde{Y}$: Mean value of experimental data

Calculation of R^2

Place storage periods (x) in regression equation and calculate the "corrected y" values.

$$y = -75.11\,x + 609.49$$

$$y = -75.11\,x + 609.49$$

$$x = 2 \rightarrow \hat{Y} = 459.27 \; (y=457)$$

$$x = 4 \rightarrow \hat{Y} = 309.05 \; (y=305)$$

$$x = 5 \rightarrow \hat{Y} = 233.94 \; (y=251)$$

$$x = 6 \rightarrow \hat{Y} = 158.83 \; (y=148)$$

If «x» vs «Ŷ» are plotted on arithmetic scaled graph paper, then all data points will be on the straight line.

- Place «corrected y» values in R^2 equation.

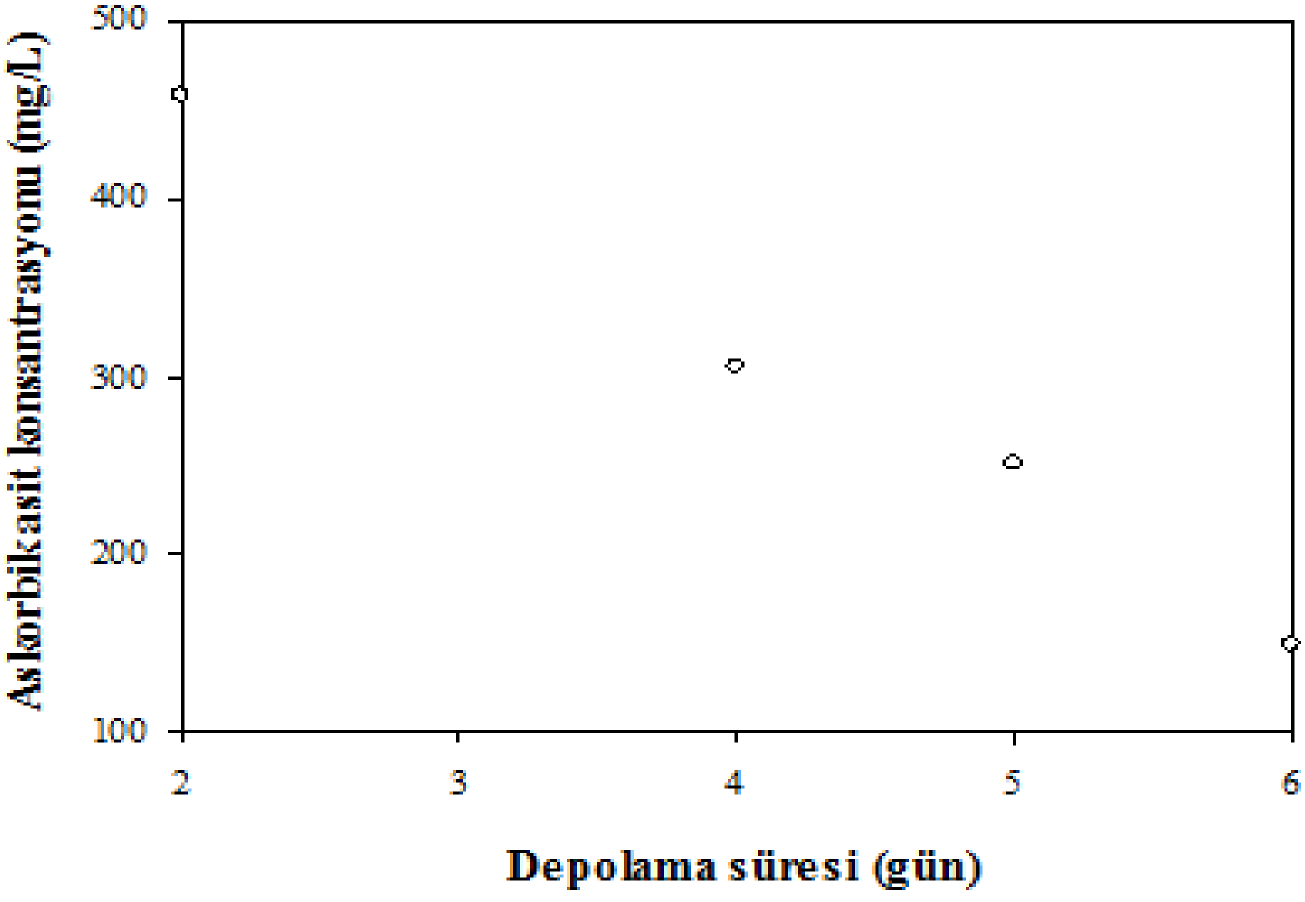
$$R^2 = 1 - \frac{[(457 - 459.27)^2 + (305 - 309.05)^2 + (251 - 233.94)^2 + (148 - 158.83)^2]}{[(457 - 290.25)^2 + (305 - 290.25)^2 + (251 - 290.25)^2 + (148 - 290.25)^2]}$$

$$R^2 = 1 - \frac{429.8879}{49798.7275}$$

$$R^2 = 0.9914$$

- ***Interpretation:*** R^2 value of 0.9914 (very close to 1) indicates experimental data was **perfectly** fitted to straight line.

Figure 2.5 Plotting original experimental data



Plotting the regression line

- To plot the regression line, we do not need to calculate all «corrected y values.»

- Two «corrected y» values (generally the "highest" and "lowest") are calculated by placing the corresponding «x» values in regression equation.

$$y = -75.11\,(x) + 609.49$$

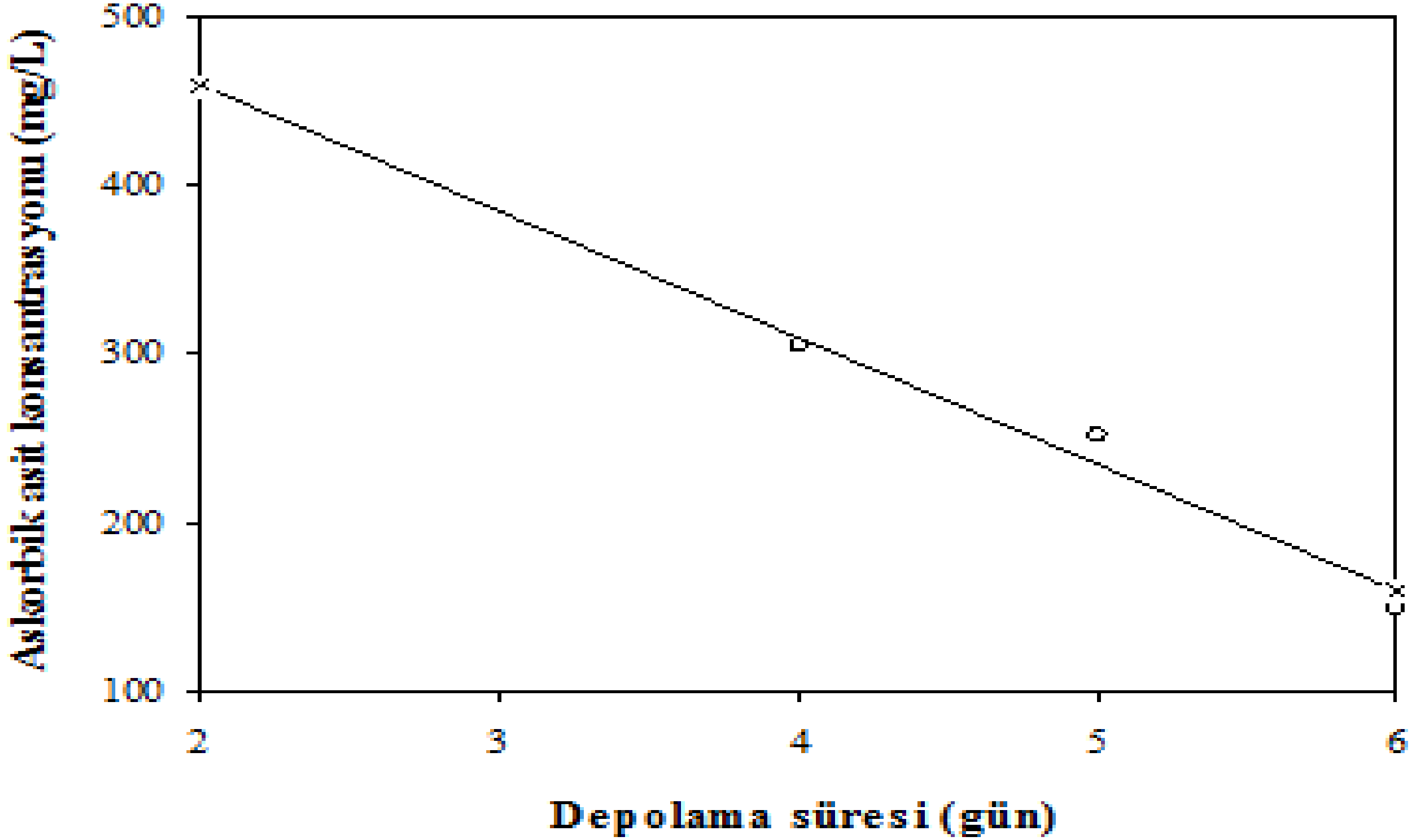
$$\mathbf{x = 2 \;\rightarrow\; y_{corrected} = 459.3}$$

$$\mathbf{x = 6 \;\rightarrow\; y_{corrected} = 158.8}$$

- Place the calculated new coordinates (**<u>2; 459</u>**) and (**<u>6; 159</u>**) on graph.

- Connect these two data points.

Figure 2.6 Regression line



In graphical presentations

- ✓ Only **original experimental data points** are shown.
- ✓ Data points calculated from regression equation (corrected y values) are never shown in graph (not in the exam!!!!).

Example 2.6 (homework): Ascorbic acid degradation was studied in orange juice stored at various temperatures. Periodically, samples were drawn from storage and samples were analyzed for aa. AA contents of samples are presented in Table 2.6.

a) Calculate regression coefficients (a and b) and R^2 by linear regression.
b) Plot the regression lines.

- Ascorbic acid degradation was studied in orange juice stored at various temperatures.
    - What was the food product?
    - What food constituent was studied in this research?
    - Was this study «processing» or «storage»?
    - How can you do this experiment?
    - What was the purpose of this study?

Temperature (°C)	Storage time (h)	AA content (mg mL^{-1})
23	20	0.948
	40	0.476
	60	0.004
35	5	1.029
	10	0.758
	20	0.261
45	0	1.200
	5	0.655
	10	0.109

Temperature (°C)	Storage time (h)	AA content (mg mL^{-1})
23	**20**	**0.948**
	40	0.476
	60	0.004
35	5	1.029
	10	**0.758**
	20	**0.261**
45	0	1.200
	5	0.655
	10	**0.109**

- How many graphs are you going to draw? Justify your answer.

Solution

To calculate slopes, intercepts and determination coefficients at 23^{o}, 35^{o} and $45^{o}C$, we first need to calculate the values for the regression equations.

Table 2.7 Data for regression analysis for the aa degradation in orange juice stored at 23°C

X	**X^2**	**Y**	**Y^2**	**XY**
20		0.948		
40		0.476		
60		0.004		
$\Sigma X=$	$\Sigma X^2=$	$\Sigma Y=$	$\Sigma Y^2=$	$\Sigma XY=$

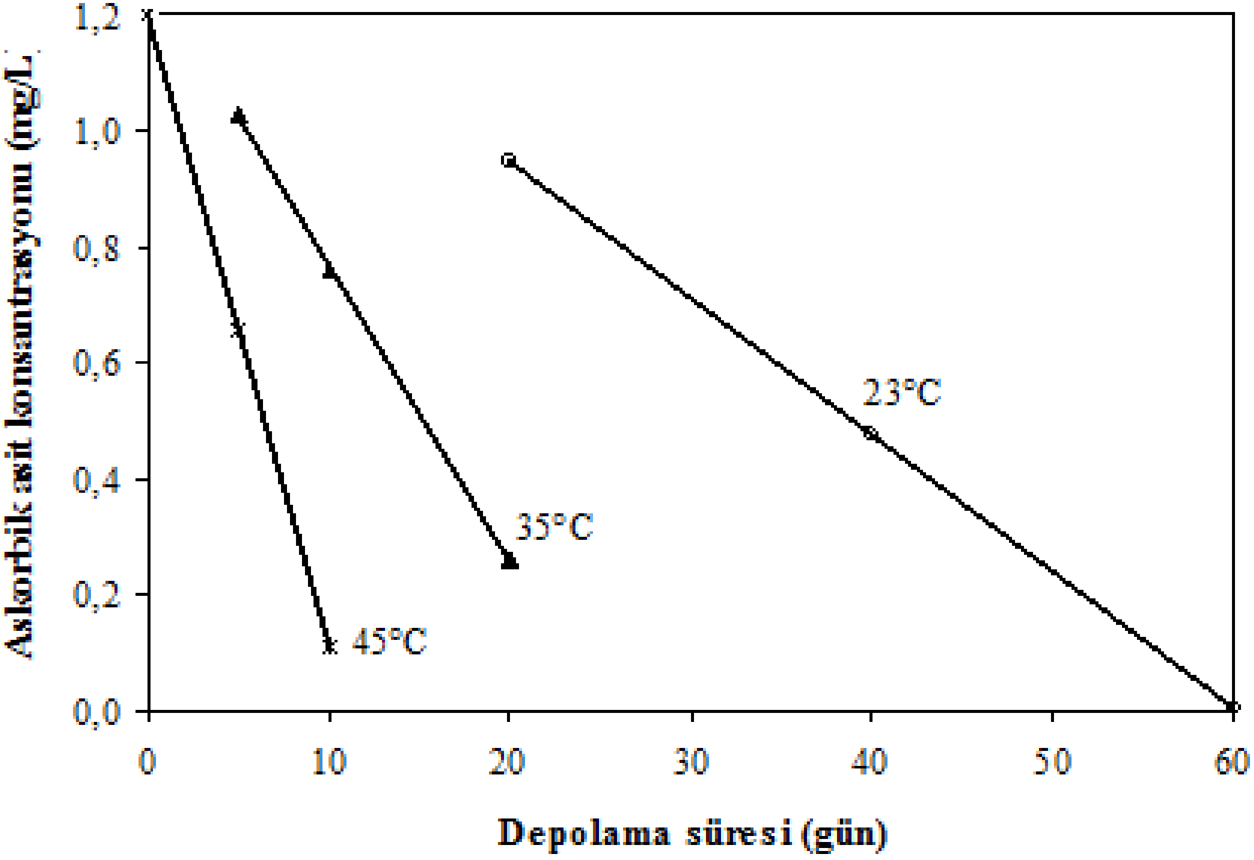
Table 2.7 Data for regression analysis for the aa degradation in orange juice stored at 23°C

X	X^2	Y	Y^2	XY
20	400	0.948	0.898704	18.96
40	1600	0.476	0.226576	19.04
60	3600	0.004	0.000016	0.24
$\Sigma X=$ 120	$\Sigma X^2=$ 5600	$\Sigma Y=$ 1.428	$\Sigma Y^2=$ 1.125296	$\Sigma XY=$ 38.24

Table 2.10 Regression coefficients for the aa degradation in orange juice stored at various temperatures.

Temperature (°C)	Slope (a) (mg mL^{-1} h^{-1})	Intercept (b) (mg mL^{-1})	R^2
23	–0.0236	1.4200	1
35	–0.0510	1.2775	0.9995
45	–0.1091	1.2002	1

Figure 2.7 Regression lines for aa degradation in orange juice stored at various temperatures.



Example 2.7

Formation of brown color as a result of non-enzymatic browning reactions was measured in dried apricots desulfrized at 50°C. Brown color was spectrophotometrically determined by measuring absorbance values at 420 nm (A_{420}). Experimental data is presented in Table 2.11.

- Formation of brown color as a result of non-enzymatic browning reactions was measured in dried apricots desulfrized at 50°C.

  - ✓ What was the food product?
  - ✓ What food constituent was studied in this research?
  - ✓ Was this study «processing» or «storage»?

  - ✓ How can you do this experiment?
  - ✓ What was the purpose of this study? (**SO_2 and browning**)

- Calculate regression coefficients (slope and intercept) and «R^2» by linear regression.
- Plot the experimental data on arithmetic graph paper and draw the regression line.
- Calculate the brown color formation after **65 h** desulfurization at 50°C.
- Calculte the brown color formation after **180 h** desulfurization at 50°C.
- **Analyse your results for 65 h or 180 h, which of your results is more reliable?**

Table 2.11 The concentration of brown pigments formed in dried apricots after desulfirization at 50°C

Time (h)	**A_{420} g^{-1} dried weight**
0	0.144
24	0.161
48	0.171
72	0.196
96	0.213
120	0.233
144	0.255

Table 2.12 Data for regression analysis for the formation of brown colored pigments in dried apricots after desulphurization at 50°C

X	X^2	Y	Y^2	XY
0		0.144		
24		0.161		
48		0.171		
72		0.196		
96		0.213		
120		0.233		
144		0.255		
ΣX=???	ΣX^2=???	ΣY=???	ΣY^2=???	ΣXY=???

Table 2.12 Data for regression analysis for the formation of brown colored pigments in dried apricots after desulphurization at 50°C

X	X^2	Y	Y^2	XY
0	0	0.144	0.021	0
24	576	0.161	0.026	3.864
48	2304	0.171	0.029	8.208
72	5184	0.196	0.038	14.112
96	9216	0.213	0.045	20.448
120	14400	0.233	0.054	27.960
144	20736	0.255	0.065	36.720
$\Sigma X=504$	$\Sigma X^2=52416$	$\Sigma Y=1.373$	$\Sigma Y^2=0.279$	$\Sigma XY=111.312$

Calculation of slope

$$a = \frac{111.312 - (504)(1.373) / 7}{52416 - [(504)^2 / 7]} = 0.00077\ A_{420}\ g^{-1}\ h^{-1}$$

Calculation of y-intercept

$$b = \frac{(1.373)\,(52416) - (504)\,(111.312)}{7\,[52416 - (504)^2 / 7]} = 0.1405\ A_{420}\ g^{-1}$$

Equation for brown color formation

$$y = 0.00077\,x + 0.1405$$

Calculation of the "corrected y" values

$$y = 0.00077\,(0) + 0.1405 \rightarrow y_1 = 0.141$$

$$(24) \rightarrow y_2 = 0.159$$

$$(48) \rightarrow y_3 = 0.177$$

$$(72) \rightarrow y_4 = 0.196$$

$$(96) \rightarrow y_5 = 0.214$$

$$(120) \rightarrow y_6 = 0.233$$

$$(144) \rightarrow y_7 = 0.251$$

Calculation of R^2

$$R^2 = 1 - \frac{[(0.144 - 0.141)^2 + (0.161 - 0.159)^2 + (0.171 - 0.177)^2 + (0.196 - 0.196)^2 + (0.213 - 0.214)^2 + (0.233 - 0.233) + (0.255 - 0.251)]}{[(0.144 - 0.196)^2 + (0.161 - 0.196)^2 + (0.171 - 0.196)^2 + (0.196 - 0.196)^2 + (0.213 - 0.196)^2 + (0.233 - 0.196) + (0.255 - 0.196)]}$$

$R^2 = 1 - 0.004596$

$R^2 = 0.993$

Drawing the regression line

- Mark the experimental values in an arithmetic scaled graph paper.

- To draw the regression line:
  - Place the lowest and the highest experimental "x" values in regression equation and calculate the "corrected y" values.

$$y = 0.00077\,(0) + 0.1405 \rightarrow y_{corrected} = 0.141$$
$$(144) \rightarrow y_{corrected} = 0.251$$

- Mark the “x” and “corrected y” values in graph paper.

- Connect the two points in graph to draw the straight line (**<u>regression line</u>**).

Figure 2.8 Regression line for formation of brown pigment during desulfurization of dried apricots at 50°C