

Homework : Linear Algebra for OS

Topic Linear Regression (10-Sep-22)

2024 ภาคฤดูร้อน ผู้สอน: ✓

#D : 65056071

การทำนาย

แบบฝึกหัดครั้งที่ 1

1. กำหนดข้อมูลให้ดังนี้

i	X	Y
1	1	41
2	4	59
3	11	90
4	11	88
5	2	52
6	2	53
7	1	53
8	5	63
9	10	87
10	8	74

1.1 จงเขียนแผนภาพการกระจายระหว่าง X และ Y

1.2 จงหาสมการถดถอยของ X และ Y

1.3 จงหาช่วงความเชื่อมั่น 95% ของพารามิเตอร์ของตัวแบบ (β_0, β_1) และ σ^2

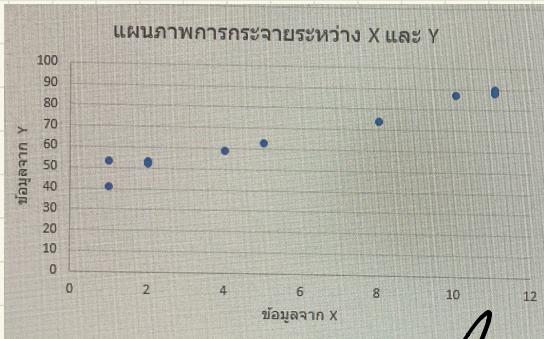
1.4 จงทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ของตัวแบบที่ระดับนัยสำคัญ 0.01 (β_0, β_1)

1.5 จงคำนวณค่าของ R^2

1.1

จุดเดียวที่ไม่ใช่เส้นทางเดียว

แนวคิดใหม่	
X	Y
1	41
2	59
3	90
4	88
5	52
6	53
7	53
8	63
9	87
10	74



ทำนายใน Excel; ผลลัพธ์ Ms. Excel คือการดึงไปที่เครื่อง

Aos
///

1.2) រាយការណ៍នូវលទ្ធផលនៃសម្រាប់
 $x = 10 = y$

Ans Step

- ① ស្ថិត \bar{x} និង \bar{y} នៅលើ $\sum_{i=1}^{10} x_i = \bar{x}$ និង $\sum_{i=1}^{10} y_i = \bar{y}$
- ② $x_i - \bar{x} = Y_i - \bar{Y}$
- ③ $(x_i - \bar{x}) \cdot (Y_i - \bar{Y})$
- ④ $(x_i - \bar{x})^2$

$$\sum_{i=1}^{10} Y_i = \bar{Y}$$

$$\sum_{i=1}^{10} x_i = \bar{x}$$

រាយការណ៍នូវលទ្ធផល (OLS)
 b_1, b_2 នៃគឺ

$$b_1 = \frac{\sum_{i=1}^{10} (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{10} (x_i - \bar{x})^2}$$

$$b_0 = \bar{Y} - b_1 \bar{x}$$

5.1

5.2

$$\hat{Y}_i = b_0 + b_1 x_i$$

* ព័ត៌មាន...
* នៅ step ①-④
 ① ② ③ ④
 ⑤ ⑥

គោលការណ៍សម្រាប់លទ្ធផល		គោលការណ៍សម្រាប់លទ្ធផល				
i	X	Y	$(x_i - \bar{x})$	$(Y_i - \bar{Y})$	$(x_i - \bar{x})(Y_i - \bar{Y})$	$(x_i - \bar{x})^2$
1	1	41	-4.50	-25.00	112.50	20.25
2	4	59	-1.50	-7.00	10.50	2.25
3	11	90	5.50	24.00	132.00	30.25
4	11	88	5.50	22.00	121.00	30.25
5	2	52	-3.50	-14.00	49.00	12.25
6	2	53	-3.50	-13.00	45.50	12.25
7	1	53	-4.50	-13.00	58.50	20.25
8	5	63	-0.50	-3.00	1.50	0.25
9	10	87	4.50	21.00	94.50	20.25
10	8	74	2.50	8.00	20.00	6.25
\bar{x}		\bar{Y}				
5.50	66.00					
			$\sum_{i=1}^{10} (x_i - \bar{x})(Y_i - \bar{Y})$	$\sum_{i=1}^{10} (x_i - \bar{x})^2$		
			645.00	154.50		

① ②

b1	4.1748
b0	43.0388

នូវលទ្ធផល b_0, b_1 នៃគឺ ⑥

នៅពេលនៃគឺ នូវលទ្ធផល \hat{Y}_i

$$\hat{Y}_i = 4.1748x + 43.0388$$

Ans 2

1.3) នូវរាយការណ៍ទីតាំង 95% នៃរាយការណ៍ទីតាំង (β_0, β_1) នៅ = σ^2

Sel ចិត្តស្ថាទៀត បញ្ជាក់ថាគារពិនិត្យសម្រាប់លទ្ធផល = 1

$(1 - \alpha) \cdot 100$ នៃរាយការណ៍ទីតាំង β_1 នៅ

$b_k \pm \frac{t_{\alpha/2}}{2} se(b_k)$ 1.3.1

នៅលើ $k = 0 \text{ និង } 1$

$se(b_k) = s$

$s = \sqrt{\sum_i (x_i - \bar{x})^2}$

នៅលើ 1.3.1 នៃរាយការណ៍ទីតាំង 95%

$(1 - \alpha) \cdot 100 = 95$

$\alpha = \frac{1}{2}$ $\alpha = 0.05$

$\therefore 15700 = 1700 \times 95\%$ នៃរាយការណ៍ទីតាំង β_0 នៅ = $b_0 \pm \frac{0.05}{2} se(b_0)$

នៅលើ 1.2) ឱ្យរាយការណ៍ទីតាំង $b_0 = 43,0388$

$\therefore 15700 = 1700 \times 43,0388 \pm \frac{0.05}{2}$

នៅលើ 1.3.2) $se(b_0) = \sqrt{var(b_0)}$

$var(b_0) = \frac{s^2 \times \sum_{i=1}^n x_i^2}{n \times \sum_{i=1}^n (x_i - \bar{x})^2}$

$s^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$

$n = 10$ ជាមួយ sample 3

10 = 9.1748, $se(b_1) \approx 0.2688$

$$m \quad se(b_1) = \sqrt{var(b_1)} \quad \text{--- (4)}$$

$$var(b_1) = \frac{s^2}{\sum_{i=1}^{n-2} (x_i - \bar{x})^2} \quad \text{--- (5)}$$

97% (1) - (5)

กันน้ำดื่มน้ำ

	x_i	y_i	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$\hat{y}_i = b_0 + b_1 x_i$	$e_i = y_i - \hat{y}_i$	e_i^2	x_i^2		
1	1	41	-4.50	-25.00		112.00	20.25	47.21	-6.21	38.01	1.00
2	4	59	-1.50	-7.00		10.50	2.25	59.74	-0.74	0.54	16.00
3	11	90	5.50	24.00		132.00	30.25	88.96	1.04	1.08	121.00
4	10	88	5.50	22.00		121.00	30.25	88.96	-0.96	0.92	121.00
5	2	62	-3.50	-14.00		49.00	12.25	51.39	0.61	0.37	4.00
6	2	53	-3.50	-13.00		45.50	12.25	51.39	1.61	2.00	4.00
7	1	53	-4.50	-13.00		58.50	20.25	47.21	-5.79	33.48	1.00
8	5	63	-0.50	-3.00		1.50	0.25	63.91	-0.91	0.83	25.00
9	10	87	4.50	21.00		94.50	20.25	84.79	2.21	4.90	100.00
10	9	74	2.50	8.00		20.00	6.25	76.44	-2.44	5.94	64.00
\bar{x}		\bar{y}				$\sum_{i=1}^{10} (x_i - \bar{x})(y_i - \bar{y})$	$\sum_{i=1}^{10} (x_i - \bar{x})^2$	average	$\sum_{i=1}^{10} e_i^2$	$\sum_{i=1}^{10} x_i^2$	
		5.50	66.00			165.00	154.50	66.00	89.28	457.00	



10 = 10 (5) $t_{0.05/2}$ 10 กะบารีตอ 7 $b_1 = 4.1748$

$$\therefore 10 = \sqrt{\frac{1}{n-2}} t_{0.05/2} \quad \frac{1}{n-2} = \frac{1}{10-2} = \frac{1}{8}$$

$$\therefore 10 = \sqrt{\frac{1}{8}} t_{0.025} \quad n-2 = 8$$

→ 10 = $\sqrt{1/8} t_{0.025}$

$t_{0.025} = 2.12$

$$n \rightarrow v = 8$$

n	10 = COUNT(B7:B16)
s^2	11.160 = K18/(V5-2)
$var(b_0)$	3.3011 = (L18 * V6) / (V5 * H18)
$se(b_0)$	1.8169 = SQRT(V7)
$var(b_1)$	0.07 = V6/H18
$se(b_1)$	0.2688 = SQRT(V10)

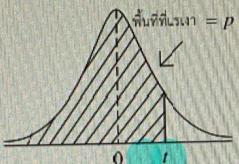
{ (6)

$se(b_0)$	1.8169
$se(b_1)$	0.2688

14

ตารางแสดงค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที่

$$P(T \leq t) = p \text{ เมื่อ } T \sim t(v)$$



v	p										
	0.60	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999	0.9995	
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657	318.309	636.619	
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.599	
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924	
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610	
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869	
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959	
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408	
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041	
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781	
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587	
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437	
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318	
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221	
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140	
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073	
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015	
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898	3.646	3.965	

$$\therefore +_{0.025, 8} = 2.306$$

พื้นที่

จุดที่ $\frac{\alpha}{2}$ นั้น $\alpha/2 = 0.025$ หมายความว่า $P = 0.975$

จุดที่ $1 - \alpha/2$ นั้น $1 - \alpha/2 = 0.975$ หมายความว่า $P = 0.975$

$$\therefore \text{area} + = 1 - P$$

$$\therefore 0.025 = 1 - P \quad \text{ก็ได้}, P = 0.975$$

$\therefore \text{ค่าเฉลี่ย} \quad (1.3.1)$

$$\text{ค่าเฉลี่ย} = \bar{x}_0 \pm \frac{s}{\sqrt{n}} (\text{S.E.})$$

$$\bar{x}_0 = 43.0388 \pm 2.306 (1.8169)$$

$95\% \text{ CI} \bar{x}_0 \pm 1.96 \text{ S.E.}$

$$(38.8491, 47.2286)$$

∴ β_1 සඳහා $1.3.1 \chi^2$ ප්‍රතිඵලිය ඇත
∴ β_1 සඳහා 95% C.I. නො ඇත

$$b_1 \pm + \frac{\alpha}{2} (\text{sech } b_1)$$

$\text{mean} + \frac{\alpha}{2}, \text{sech } b_1 \text{ නො } = b_1$
නො ඇත

$$= 4.1748 \pm 2.306 (0.2688)$$

∴ $95\% \text{ CI}$ නො β_1 දහුවාකේන

$$(3.5550, 4.7945) \text{ Area}$$

$\text{Area} = 1 - \alpha$

$100(1 - \alpha)$ මෙයින්ගේ අංකය =

χ^2_{obs}

$$(n-2) \frac{s^2}{\chi^2_{(\alpha/2)}} \leq \sigma^2 \leq (n-2) \frac{s^2}{\chi^2_{(\alpha/2)}}$$

$$\text{නො } \alpha = 0.05$$

$$1 - \frac{\alpha}{2} = 1 - \frac{0.05}{2} = 0.975$$

∴ $\text{Area} = 0.975$

$$\rho = 0.025 \quad v = 8$$

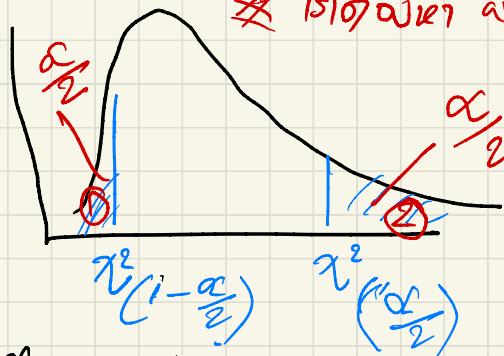
$$\text{නො } \rho = 0.975 \quad v = 8$$

$\text{Area} = 0.975$

$\text{Area} = 0.975$

* 1510 ඉටු ඇත ① නිය

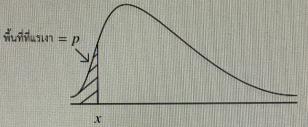
②



6

ตารางที่ 4 ตารางแสดงค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบคิลล์แคร์

$$P(X \leq x) = p \text{ เมื่อ } X \sim \chi^2(\nu)$$



ตารางที่ 4.1 $p = 0.0005, 0.001, 0.005, 0.01, 0.025, 0.05, 0.10, 0.20, 0.30$ และ 0.40

ν	p									
	0.0005	0.001	0.005	0.01	0.025	0.05	0.10	0.20	0.30	0.40
1	0.000	0.000	0.000	0.000	0.001	0.004	0.016	0.064	0.148	0.275
2	0.001	0.002	0.010	0.020	0.051	0.103	0.211	0.446	0.713	1.022
3	0.015	0.024	0.072	0.115	0.216	0.352	0.584	1.005	1.424	1.869
4	0.064	0.091	0.207	0.297	0.484	0.711	1.064	1.649	2.195	2.753
5	0.158	0.210	0.412	0.554	0.831	1.145	1.610	2.343	3.000	3.655
6	0.299	0.381	0.676	0.872	1.237	1.635	2.204	3.070	3.828	4.570
7	0.485	0.598	0.989	1.239	1.700	2.167	2.833	3.822	4.671	5.493
8	0.710	0.857	1.344	1.646	2.180	2.733	3.490	4.594	5.527	6.423
9	0.972	1.152	1.735	2.088	2.700	3.525	4.168	5.380	6.393	7.357
10	1.265	1.479	2.156	2.558	3.247	3.940	4.865	6.179	7.267	8.295

ตารางที่ 4.2 $p = 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.975, 0.990, 0.995, 0.999$ และ 0.9995

ν	p									
	0.50	0.60	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999
1	0.455	0.708	1.074	1.642	2.706	3.841	5.024	6.635	7.879	10.
2	1.386	1.833	2.408	3.219	4.605	5.991	7.378	9.210	10.597	13.
3	2.366	2.946	3.665	4.642	6.251	7.815	9.349	11.345	12.838	16.
4	3.357	4.045	4.878	5.989	7.779	9.488	11.143	13.277	14.860	18
5	4.351	5.132	6.064	7.289	9.236	11.070	12.833	15.086	16.750	20
6	5.348	6.211	7.231	8.558	10.645	12.592	14.449	16.812	18.548	22
7	6.346	7.283	8.383	9.803	12.017	14.067	16.183	18.475	20.278	24
8	7.344	8.351	9.524	11.030	13.362	15.507	17.535	20.090	21.955	26
9	8.343	9.414	10.656	12.242	14.684	16.919	19.166	21.666	23.589	27
10	9.342	10.173	11.781	13.442	15.987	18.307	20.483	23.209	25.188	29

$$\because \alpha = 0.05 \quad \therefore \chi^2_{\alpha/2} = 17.535$$

$$\text{ดัง } \chi^2_{(\frac{\alpha}{2})} = \chi^2_{0.025} = 17.535$$

$$\chi^2_{(1-\frac{\alpha}{2})} = \chi^2_{0.975} = 2.180$$

$$\therefore \text{สูตรความน่าจะเป็น} = 0.95 \text{ ถ้า } G^2$$

คือ $G^2 \geq 17.535$ (กรณีที่ $G^2 \geq 2.180$)

$$P\left(G^2 \geq \chi^2_{\alpha/2}\right) = 1 - \alpha \quad (8)$$

$$\begin{aligned} & \text{กรณี } G^2 \geq 17.535 \quad \chi^2_{\frac{\alpha}{2}} = \chi^2_{0.025} = 17.535 \\ & \alpha = 0.05 \\ & \chi^2 = 3993.595 \quad \chi^2_{(1-\frac{\alpha}{2})} = \chi^2_{0.975} = 2.180 \end{aligned}$$

$$\begin{aligned} & \text{ถ้า } G^2 \geq 17.535 \quad \text{เมื่อ } \alpha = 0.05 \\ & P\left(8 \times \frac{11.160}{17.535} \leq G^2 \leq 8 \times \frac{11.160}{2.180}\right) = 0.95 \\ & P(5.09 \leq G^2 \leq 40.95) \end{aligned}$$

หาก $G^2 \geq 17.535$ \Rightarrow $\alpha = 0.05$ \Rightarrow $G^2 \geq 2.180$ \Rightarrow $\alpha = 0.95$

∴ $\hat{y} = 6.2 + 0.95x$
 95% IC (5.09, 40.95)

1.4) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$ (0.01 (β_0, β_1))

$$H_1: \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

(5) H_0, H_1

Intercept $\rightarrow \beta_0$ | Slope $\rightarrow \beta_1$

$$H_0: \beta_0 = 0$$

$$H_1: \beta_0 \neq 0$$

$$\alpha = 0.01$$

$$1.4.1) \text{Test Statistic } \beta_0 = 0$$

$$t_c = \frac{\hat{\beta}_0 - 0}{\text{se}(\hat{\beta}_0)}$$

$$-t_{\frac{\alpha}{2}, (n-2)} \leq t_c \leq t_{\frac{\alpha}{2}, (n-2)}$$

$$t_c < -t_{\frac{\alpha}{2}, (n-2)}$$

$$t_c > t_{\frac{\alpha}{2}, (n-2)}$$

$$t_c < -t_{\frac{\alpha}{2}, (n-2)}$$

$$t_c > t_{\frac{\alpha}{2}, (n-2)}$$

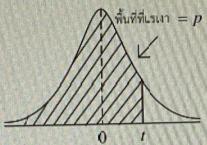
$$t_c = 43.0388 - 0$$

$$t_c = 43.0388$$

18.

ตารางแสดงค่าความน่าจะเป็นของจั่วและปริสุที่เมื่อการแยกแจ้งความน่าจะจะเป็นแบบที่

$$P(T \leq t) = p \text{ เมื่อ } T \sim t(v)$$



v	0.60	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999	0.9995
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781

$$\text{ท.ก. } \alpha = 0.01$$

$$\frac{\alpha}{2} = 0.005$$

$$\text{จากตารางท.ก. } + \frac{1}{n} [p = 0.995] \quad n(v) = 8$$

$$\text{ต้องการ } \beta_1 = 0 \rightarrow + 0.005 = 3.355$$

ดังนั้น จึงต้องหา t_c ที่ $P(T \geq t_c) = 0.005$

$$\therefore \frac{\alpha}{2}, (n-2) \rightarrow + 0.005, 8 = 3.355 \quad (11)$$

$$\therefore - \frac{\alpha}{2}, n-2 \rightarrow - 0.005, 8 = -3.355 \quad (12)$$

104.2 กิจกรรมบันทึก

$$t_c = \frac{b_1 - 0}{\sec b_1} \approx n-2 \quad (13)$$

$$t_c \approx 1.4 \cdot 1 \quad (13)$$

$$t_c \approx 1.4 \cdot 1 \quad (13)$$

$$t_c = \frac{4.1748 - 3.355}{0.2688}$$

$$\therefore t_c = 15.533 \quad (14)$$

$$t_c \approx 17 \quad (14)$$

1.4.1 for $\beta_0 = 0$;

$$-3.355 \leq 23.688 \leq 3.355 \rightarrow \text{False} \rightarrow \text{True}$$

$23.688 < -3.355 \rightarrow 0$ $23.688 > 3.355$

$\therefore \text{ที่ } t_c = 23.688 \text{ นั้น } \beta_0 \neq 0 \text{ นะ}$

~~ข้อที่ 6 วิเคราะห์ค่าคงที่และตัวแปรอิสระ~~ $\beta_0 \neq 0$ นะ

1.4.2 for $\beta_1 = 0$

$$-3.355 \leq 15.533 \leq 3.355 \rightarrow \text{False}$$

$15.533 < -3.355 \rightarrow 0$ $15.533 > 3.355$

$\therefore \text{ที่ } t_c = 15.533 \text{ นั้น } \beta_1 \neq 0 \text{ นะ}$

~~ข้อที่ 6 วิเคราะห์ค่าคงที่และตัวแปรอิสระ~~ $\beta_1 \neq 0$ นะ

ผลลัพธ์ของ $(\hat{Y}_i - \bar{Y})^2$ แล้ว
 $(Y_i - \bar{Y})^2$ หารด้วย r^2
 ดูที่ห้อง

Index	Y_i	ที่ 1.3			ที่ 1.5		
		$(Y_i - \bar{Y})$	$\hat{Y}_i = b_0 + b_1 X_i$	$e_i = Y_i - \hat{Y}_i$	e_i^2	x_i^2	$(\bar{Y} - \hat{Y})^2$
41	-25.00	47.21	-6.21	38.61	1.00	352.93	625.00
59	-7.00	59.74	-0.74	0.54	16.00	35.21	49.00
90	24.00	88.96	1.04	1.08	121.00	527.22	576.00
88	22.00	88.56	-0.96	0.92	121.00	527.22	484.00
52	-18.00	51.39	0.61	0.37	4.00	213.50	196.00
53	-13.00	47.21	5.79	33.48	1.00	352.93	169.00
53	-13.00	53.91	-0.91	0.83	25.00	4.36	9.00
63	-3.00	84.79	2.23	4.90	100.00	352.93	441.00
87	21.00	76.44	5.94	64.00	108.93	64.00	64.00
74	8.00	average	2.44	5.94	16.00	213.50	169.00
	66.00		66.00	89.28	457.00	2892.72	2782.00

หาผลลัพธ์ของ r^2
 แล้ว $\frac{1}{15}$

ผลลัพธ์ $\frac{1}{15}$

$$R^2 = \frac{1}{15} = 0.9679$$

6.5 หาค่า R^2

ที่ 1.5

$$R^2 = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2}{\sum_i (Y_i - \bar{Y})^2} = \frac{1}{15}$$

ที่ 1.5
 16