

TUGAS 1 APG

Nama : Zahra Nurul Fatimah
NIM : G14190070

1. Suatu penelitian dibuat untuk menguji kenormalan ganda antara rata-rata waktu belajar per hari (jam) dengan perolehan nilai kalkulus. Penelitian dilakukan terhadap 15 orang dengan hasil sebagai berikut.

	Rata-rata waktu belajar	Nilai kalkulus
1	3,2	89
2	1,8	75
3	2,6	92
4	4	90
5	1,2	70
6	1,6	75
7	2,4	82
8	2,8	83
9	2,9	90
10	3	85
11	3,3	100
12	3,4	80
13	3,5	98
14	3,6	89
15	3,8	84

Buatlah qq-plot dari data tersebut (sertakan langkah pengerjaannya)!

Jawab :

Cara membuat qq-plot

1. Nilai vector rata-rata (\bar{X})

no	rata-rata waktu belajar	nilai kalkulus
1	3.2	89
2	1.8	75
3	2.6	92
4	4	90
5	1.2	70
6	1.6	75
7	2.4	82
8	2.8	83

9	2.9	90
10	3	85
11	3.3	100
12	3.4	80
13	3.5	98
14	3.6	89
15	3.8	84
Xbar	2.873333333	85.46666667

$$\bar{X}_i = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Untuk peubah rata-rata waktu belajar :

$$\bar{X}_i = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{3.2 + 1.8 + \dots + 3.8}{15} = 2.87$$

Untuk peubah nilai kalkulus :

$$\bar{X}_i = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{89 + 75 + \dots + 84}{15} = 85.47$$

2. Nilai matriks varians-kovarians : S

Transpose Xi-Xbar	
0.326667	3.533333
-1.07333	-10.4667
-0.27333	6.533333
1.126667	4.533333
-1.67333	-15.4667
-1.27333	-10.4667
-0.47333	-3.46667
-0.07333	-2.46667
0.026667	4.533333
0.126667	-0.46667
0.426667	14.53333
0.526667	-5.46667
0.626667	12.53333
0.726667	3.533333
0.926667	-1.46667

Matriks kovarian	
0.679238	4.941905
4.941905	70.40952

Invers matriks varians-covariance	
3.008644	-0.211171
-0.211171	0.029024

Matriks Kovarian :

$$\text{Var}(X_1) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 0.679$$

$$\text{Var}(X_2) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 70.40$$

$$\text{Cov}(X_1, X_2) = \frac{\sum_{i=1}^n (x_{i,1} - \bar{x}_1)(x_{i,2} - \bar{x}_2)}{n-1} = 4.94$$

3. Nilai jarak mahalanobis $d_i^2 = (x_i - \bar{x})'S^{-1}(x_i - \bar{x})$; $i = 1, 2, \dots, 15$ dan diurutkan

d_i^2
0.08
0.12
0.20
0.33
0.55
0.87
1.90
2.22
2.26
2.42
2.43
2.92
3.22
4.06
4.44

4. Nilai $p_i = \frac{i-\frac{1}{2}}{n}$, $i = 1, 2, \dots, 15$

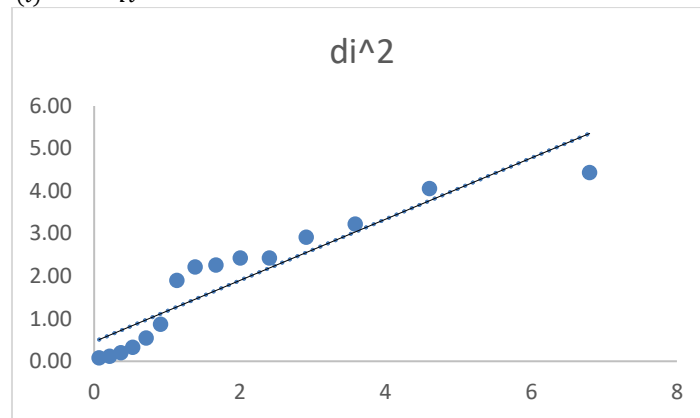
d_i^2	i	pi
0.08	1	0.033333
0.12	2	0.1
0.20	3	0.166667
0.33	4	0.233333
0.55	5	0.3
0.87	6	0.366667
1.90	7	0.433333
2.22	8	0.5
2.26	9	0.566667
2.42	10	0.633333
2.43	11	0.7
2.92	12	0.766667
3.22	13	0.833333
4.06	14	0.9
4.44	15	0.966667

5. Nilai $q_{i,p}(pi) = X_p^2(\frac{n-i+\frac{1}{2}}{n})$

chi square	d_i^2
0.067803103	0.08
0.210721031	0.12
0.364643114	0.20
0.531406331	0.33
0.713349888	0.55
0.913516805	0.87
1.135968075	1.90
1.386294361	2.22
1.672496048	2.26
2.006604218	2.42

2.407945609	2.43
2.910574465	2.92
3.583518938	3.22
4.605170186	4.06
6.802394763	4.44

6. Scatter plot antara $d_{(i)}^2$ dan q_i



Berdasarkan hasil scatterplot di atas, terlihat plot cenderung membentuk suatu garis lurus. Untuk lebih jelas apakah ada hubungan antara $d_{(i)}^2$ dan q_i , maka dapat dilakukan uji korelasi antar kedua variabel tersebut.

Hasil uji korelasi antara nilai $d_{(i)}^2$ dan q_i

```
Pearson's product-moment correlation
data: data$rata.rata.waktu.belajar and data$nilai.kalkulus
t = 3.6833, df = 13, p-value = 0.002757
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3191895 0.8981042
sample estimates:
cor
0.7146076
```

Berdasarkan hasil di atas, didapat nilai **P-value = 0.002 < 5%**, maka dapat disimpulkan bahwa sudah cukup bukti untuk menyatakan terdapat hubungan antara $d_{(i)}^2$ dan q_i pada taraf nyata 5%.

Hasil uji normalitas ganda dengan Mardia's skewness dan kurtosis

```
$multivariateNormality
      Test      Statistic      p value Result
1 Mardia Skewness 6.25452339476786 0.180929552248116 YES
2 Mardia kurtosis -0.847032706723608 0.396976891405281 YES
3 MVN <NA> <NA> YES

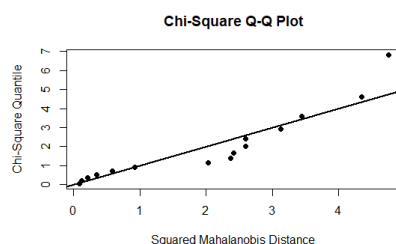
$univariateNormality
      Test      Variable Statistic p value Normality
1 Anderson-Darling rata.rata.waktu.belajar 0.3620 0.3956 YES
2 Anderson-Darling nilai.kalkulus 0.1992 0.8584 YES

$descriptives
      n      Mean Std.Dev Median Min Max 25th 75th Skew
rata.rata.waktu.belajar 15 2.873333 0.824159 3 1.2 4 2.5 3.45 -0.59508956
nilai.kalkulus 15 85.466667 8.391038 85 70.0 100 81.0 90.00 -0.07047411

      kurtosis
rata.rata.waktu.belajar -0.8678301
nilai.kalkulus -0.9213440
```

Berdasarkan hasil di atas, dapat disimpulkan bahwa sudah cukup bukti untuk menyatakan bahwa data menyebar normal ganda pada taraf nyata 5%.

Plot qq-plot dengan software R



2. Berikut merupakan data antara total produksi kacang tanah dalam kg (X1) dengan total produksi jeruk dalam kg (X2). Kedua peubah diasumsikan menyebar normal ganda.

	Produksi kacang (kg)	Produksi jeruk (kg)
1	1500	1790
2	2300	2390
3	1870	2150
4	1680	2400
5	1720	1500
6	3960	4130
7	2340	2690
8	1980	2530
9	3840	4060
10	2880	3160

Tentukanlah:

- Vektor rata-rata dan matriks kovariannya!
- Ujilah apakah vektor rata-rata populasi $\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 2350 \\ 2700 \end{bmatrix}$ pada taraf nyata 5%
- Buatlah selang kepercayaan simultan 95% bagi μ

Jawab :

- Vektor rata-rata dan matriks kovariannya!

No	produksi kacang (kg)	Produksi jeruk (kg)
1	1500	1790
2	2300	2390
3	1870	2150
4	1680	2400
5	1720	1500
6	3960	4130
7	2340	2690
8	1980	2530
9	3840	4060
10	2880	3160
Rataan	2407	2680

$$\bar{X}_i = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Untuk peubah produksi kacang (kg) :

$$\bar{X}_i = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1500 + 2300 + \dots + 2880}{10} = 2407$$

Untuk peubah produksi jeruk (kg):

$$\bar{X}_i = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1790 + 2390 + \dots + 3160}{10} = 2680$$

Matriks kovarian	
777645.6	739655.5556
739655.6	765311.1111

$$\text{Var}(X_1) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 777645.6$$

$$\text{Var}(X_2) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 765311.1111$$

$$\text{Cov}(X_1, X_2) = \frac{\sum_{i=1}^n (x_{i,1} - \bar{x}_1)(x_{i,2} - \bar{x}_2)}{n-1} = 739655.6$$

b. Ujilah apakah vector rata-rata populasi $[\mu_1] = [2350]$ pada taraf nyata 5%

Hipotesis :

$$H_0 : [\mu_1 \quad \mu_2]' = [2350 \quad 2700]$$

$$H_1 : [\mu_1 \quad \mu_2]' \neq [2350 \quad 2700]$$

Statistik Uji :

miu nol	N	S-1		Xbar	Xbar-miu nol	(Xbar-miu nol)'	
2350	10	1.59272E-05	-1.53933E-05	2407	57	57	-20
2700		-1.53933E-05	1.61839E-05	2680	-20		

$$T^2 = n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) = \mathbf{0.933179}$$

$$c^2 = \frac{(n-1)p}{(n-p)} F_{(p,n-p)} \alpha = \frac{(10-1)^2}{(10-2)} F_{(p,n-p)} \alpha = \frac{18}{8} (4.45897) = \mathbf{10.03268}$$

Kesimpulan :

$T^2 < C_2$, **Tak tolak H_0** . maka belum cukup bukti untuk menyatakan bahwa minimal ada salah satu dari rata-rata produksi kacang dan jeruk yang memiliki nilai rata-rata tidak sama dengan 2350 kg atau 2700 kg pada taraf nyata 5%.

c. Buatlah selang kepercayaan simultan 95% bagi μ

Untuk peubah X1 (produksi kacang)

$$\bar{x}_i - \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha)} \frac{\sqrt{S_{ii}}}{\sqrt{n}} \leq \mu_i \leq \bar{x}_i + \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha)} \frac{\sqrt{S_{ii}}}{\sqrt{n}}$$

$$2407 - \sqrt{\frac{2*(10-1)}{8}} 4.4589 \frac{\sqrt{777645.6}}{\sqrt{10}} \leq \mu_i \leq 2407 + \sqrt{\frac{2*(10-1)}{8}} 4.4589 \frac{\sqrt{777645.6}}{\sqrt{10}}$$

$$1523.718 \leq \mu_i \leq 3283.249$$

Pada tingkat kepercayaan 95% total produksi kacang tersebut adalah antara 1523.718 hingga 3283.249

Untuk peubah X2 (produksi Jeruk)

$$\bar{x}_i - \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha) \frac{\sqrt{S_{ii}}}{\sqrt{n}}} \leq \mu_i \leq \bar{x}_i + \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha) \frac{\sqrt{S_{ii}}}{\sqrt{n}}}$$

$$2680 - \sqrt{\frac{2*(10-1)}{8} 4.4589 \frac{\sqrt{765311.1111}}{\sqrt{10}}} \leq \mu_i \leq 2680 + \sqrt{\frac{2*(10-1)}{8} 4.4589 \frac{\sqrt{765311.1111}}{\sqrt{10}}}$$

$$1803.751 \leq \mu_i \leq 3556.249$$

Pada tingkat kepercayaan 95% nilai produksi jeruk tersebut adalah antara 1803.751 hingga 3556.249

3. Diketahui suatu penyakit yang dapat didiagnosis dengan gejala demam, tekanan darah rendah, dan nyeri tubuh. Sebuah penelitian menguji obat untuk penyakit tersebut dengan memberi obat tsb dan placebo (obat palsu) kepada 35 pasien. Dari 35 pasien, 20 pasien diberi obat baru dan 15 orang diberi placebo. Lakukan pengujian apakah obat tersebut memberikan hasil yang berbeda? (alpha=5%)

Note: Asumsi lingkungan penelitiannya heterogen

NO	Populasi 1		
	Demam	Tekanan darah	Nyeri
1	38.4	73	13
2	36.8	85	10
3	40	58	20
4	39.8	80	26
5	38.5	68	29
6	40.1	52	22
7	39.4	67	16
8	38.5	77	18
9	38.1	65	22
10	40	78	10
11	39.8	57	13
12	39.4	48	16
13	37.2	75	18
14	37.8	70	8
15	38.5	53	26
16	40.4	100	16
17	37.4	63	28
18	39.1	79	20
19	39.9	75	18
20	37.1	80	14

No	Populasi 2		
	Demam	Tekanan darah	Nyeri
1	40.9	56	21
2	40.6	64	24
3	37.3	75	15
4	41.6	57	19
5	38.9	78	20
6	39.6	92	22
7	40.3	78	24
8	39.6	56	21
9	38.1	48	19
10	39.5	63	23
11	40.3	64	27
12	39.5	67	14
13	38.5	52	14
14	37.9	59	16
15	40.1	79	21

Jawab

Hipotesis

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Statistik Uji

NO	Populasi 1		
	Demam	Tekanan darah	Nyeri
1	38.4	73	13
2	36.8	85	10
3	40	58	20
4	39.8	80	26
5	38.5	68	29
6	40.1	52	22
7	39.4	67	16
8	38.5	77	18
9	38.1	65	22
10	40	78	10
11	39.8	57	13
12	39.4	48	16
13	37.2	75	18
14	37.8	70	8
15	38.5	53	26
16	40.4	100	16
17	37.4	63	28
18	39.1	79	20
19	39.9	75	18
20	37.1	80	14
rataan	38.81	70.15	18.15

No	Populasi 2		
	Demam	Tekanan darah	Nyeri
1	40.9	56	21
2	40.6	64	24
3	37.3	75	15
4	41.6	57	19
5	38.9	78	20
6	39.6	92	22
7	40.3	78	24
8	39.6	56	21
9	38.1	48	19
10	39.5	63	23
11	40.3	64	27
12	39.5	67	14
13	38.5	52	14
14	37.9	59	16
15	40.1	79	21
Rataan	39.51333	65.86667	20

$$n_1 = 20$$

$$n_2 = 15$$

$$\bar{x}_1 - \bar{x}_2 = \begin{bmatrix} -0.7033 \\ 4.28333 \\ -1.85 \end{bmatrix} \quad (\bar{x}_1 - \bar{x}_2)' = [-0.7033 \quad 4.28333 \quad -1.85]$$

$$S_1 = \begin{bmatrix} 1.288316 & -1.54368 & 0.435263 \\ -1.54368 & 161.8184 & -22.6553 \\ 0.435263 & -22.6553 & 36.55526 \end{bmatrix} \quad S_2 = \begin{bmatrix} 1.425524 & 0.409048 & 2.728571 \\ 0.409048 & 148.6952 & 10.71429 \\ 2.728571 & 10.71429 & 15.14286 \end{bmatrix}$$

$$\left(\frac{s_1}{n_1} + \frac{s_2}{n_2} \right) = \begin{bmatrix} 0.159451 & -0.04991 & 0.203668 \\ -0.04991 & 18.00394 & -0.41848 \\ 0.203668 & -0.41848 & 2.837287 \end{bmatrix}$$

$$\left(\frac{s_1}{n_1} + \frac{s_2}{n_2} \right)^{-1} = \begin{bmatrix} 6.905654 & 0.00765 & -0.49458 \\ 0.00765 & 0.055743 & 0.007673 \\ -0.49458 & 0.007673 & 0.389083 \end{bmatrix}$$

$$T^2 = (\bar{x}_1 - \bar{x}_2)' \left(\frac{s_1}{n_1} + \frac{s_2}{n_2} \right)^{-1} (\bar{x}_1 - \bar{x}_2) = \mathbf{4.31568}$$

$$X^2_{(a,p)} = X^2_{(0.05,3)} = \mathbf{7.81473}$$

Kesimpulan: $T^2 < X^2_{(0.05,3)}$, **tak tolak H_0** . maka belum cukup bukti untuk menyatakan bahwa pengujian obat tersebut memberikan hasil yang berbeda pada taraf nyata 5%.