UNIVERSITI SAINS MALAYSIA

First Semester Examination
2015/2016 Academic Session

December 2015 / January 2016

EBB 300/2 – Engineering Statistics
[Statistik Kejuruteraan]

Duration : 2 hours
[Masa : 2 jam]

Please ensure that this examination paper contains TWENTY FOUR printed pages before you begin the examination.
[Sila pastikan bahawa kertas peperiksaan ini mengandungi DUA PULUH EMPAT muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

This paper consists of SIX questions. THREE questions in PART A and THREE questions in PART B.
[Kertas soalan ini mengandungi ENAM soalan. TIGA soalan di BAHAGIAN A dan TIGA soalan di BAHAGIAN B.]

Instruction: Answer **FOUR** questions. Answer **TWO** questions from PART A and **TWO** questions from PART B. If a candidate answers more than five questions only the first five questions answered in the answer script would be examined.
*[**Arahan**: Jawab **EMPAT** soalan. Jawab **DUA** soalan dari BAHAGIAN A dan **DUA** soalan dari BAHAGIAN B. Jika calon menjawab lebih daripada empat soalan hanya empat soalan pertama mengikut susunan dalam skrip jawapan akan diberi markah.]*

The answers to all questions must start on a new page.
[Mulakan jawapan anda untuk semua soalan pada muka surat yang baru.]

You may answer a question either in Bahasa Malaysia or in English.
[Anda dibenarkan menjawab soalan sama ada dalam Bahasa Malaysia atau Bahasa Inggeris.]

In the event of any discrepancies in the examination questions, the English version shall be used.
[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah digunapakai.]

...2/-

PART A / *BAHAGIAN A*

1. [a] Give two examples of

(i) attribute data and
(ii) variable data

Berikan dua contoh bagi

(i) data atribut dan
(ii) data boleh ubah

(20 marks/*markah*)

[b] The following is a series of annual sales (in RM millions) of company XYZ over a ten years period.

Year:	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Sales:	13.0	17.0	19.0	20.0	20.5	20.5	19.0	17.0	15.0	15.0

(i) Determine the median, mean, variance and standard deviation of the sales over this ten years period.

Berikut adalah satu siri jualan tahunan (dalam juta RM) bagi syarikat XYZ dalam tempoh sepanjang 10 tahun.

Tahun:	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*
Penjualan:	*13.0*	*17.0*	*19.0*	*20.0*	*20.5*	*20.5*	*19.0*	*17.0*	*15.0*	*15.0*

(i) Tentukan median, mean, varians dan sisihan piawai bagi jualan sepanjang tempoh 10 tahun ini.

(20 marks/*markah*)

(ii) Sketch a time-series plot.
Lakarkan satu plot siri masa.

(20 marks/*markah*)

(iii) Does there appear to be any significant changes in annual sales over time? Justify your answer.

Adakah sebarang perbezaan ketara dalam jualan tahunan sepanjang tempoh ini. Justifikasikan jawapan anda.

(25 marks/*markah*)

[c] List 3 characteristics of Normal Distribution.
Senaraikan 3 ciri-ciri Taburan Normal.

(15 marks/*markah*)

2. [a] A factory is expected to produce plastic bags with a mean length of 11 cm and standard deviation of 0.02 cm. As a process engineer, you select a random sample of 100 sheet of bags. The mean length of plastic bags is 10.998 cm.

Sebuah kilang dijangka akan menghasilkan beg plastik dengan panjang puratanya 11 sm dan sisihan piawai 0.02 sm. Sebagai seorang jurutera proses, anda memilih 100 helai beg. Panjang purata bagi beg plastik adalah 10.998 sm.

(i) Construct a 90% confident interval estimate for the population mean plastic bag length. Based on your result, is there anything has gone wrong in the production process? Justify your answer.

Binakan jangkaan selang keyakinan 90% untuk mean panjang populasi beg plastik. Berdasarkan keputusan anda, adakah sesuatu yang kurang baik telah berlaku dalam proses pengeluaran? Justifikasikan jawapan anda.

(30 marks/*markah*)

(ii) Construct 95% and 99% confident intervals. Comment on the confident intervals with increasing confident level.

Binakan selang keyakinan 95% dan 99%. Ulaskan selang keyakinan tersebut apabila paras keyakinan meningkat.

(30 marks/*markah*)

[b] (i) Define Type I error and Type II error.
Takrifkan ralat Jenis I dan ralat Jenis II.

(10 marks/*markah*)

(ii) A sample of six replicates of sludge from a wastewater treatment plant was taken. The mean pH of the sludge was 6.68 with a standard deviation of 0.20. Can we conclude that the mean pH of sludge is less than 7.0? (Use $\alpha = 0.01$)

Suatu sampel enapcemar dengan enam replikasi dari sebuah kilang rawatan air buangan telah diambil. pH purata bagi enapcemar tersebut adalah 6.68 dengan sisihan piawai 0.20. Bolehkah kami menyimpulkan bahawa pH purata enapcemar tersebut adalah kurang daripada 7.0? (Guna $\alpha = 0.01$)

(30 marks/*markah*)

3. [a] Discuss two advantages of ANOVA with examples.
Bincangkan dua kebaikan ANOVA dengan contoh-contoh.

(20 marks/*markah*)

[b] As a production engineer of a synthetic fiber factory, you are investigating the tensile strength of a new synthetic fiber. The strength of fiber is usually affected by the weight % of cotton used in the blend of raw materials. You conducted a randomized experiment with five levels of cotton content and replicated the experiment five times. The data are shown in Table 1.

Sebagai seorang jurutera pengeluaran dalam sebuah kilang pembuatan gentian buatan, anda mengkaji kekuatan tegasan bagi sejenis gentian buatan yang baru. Kekuatan gentian ini biasanya dipengaruhi dengan % berat kapas yang digunakan dalam pencampuran bahan-bahan mentah. Anda menjalankan eksperimen secara rawak untuk lima paras kandangan kapas dan mereplikasikan eksperimen sebanyak lima kali. Data adalah ditunjukkan dalam Jadual 1.

Table 1: Effect of cotton weight % on the strength of fiber

Jadual 1: Kesan % berat kapas ke atas kekuatan gentian

Cotton weight % *% berat kapas*	**Observations** *Pemerhatian*				
15	7	7	15	11	9
20	12	17	12	18	18
25	14	19	19	18	18
30	19	25	22	19	23
35	7	10	11	15	11

(i) State the null hypothesis and alternate hypothesis of this experiment.

Nyatakan hipotesis null dan hipotesis alternatif bagi eksperimen ini.

(10 marks/*markah*)

(ii) The normal probability plot and test for equal variance are shown in Figure 1 and Figure 2, respectively. Do the assumptions for the model still valid? Explain your answer.

Plot keberangkalian normal dan pengujian sama varians ditunjukkan dalam Rajah 1 dan Rajah 2. Adakah anggapan model ini masih sah? Terangkan jawapan anda.

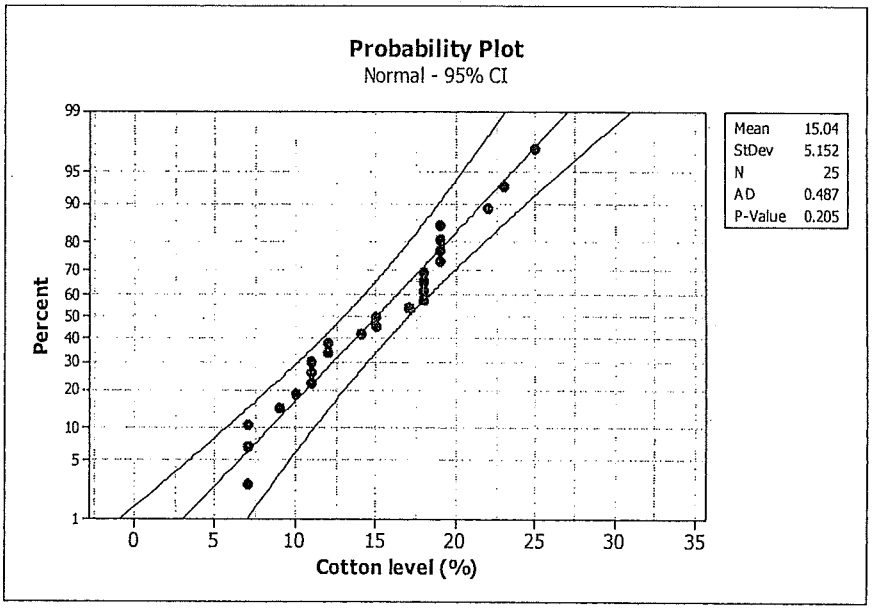

Figure 1: Normal probability plot

Rajah 1: Plot keberangkalian normal

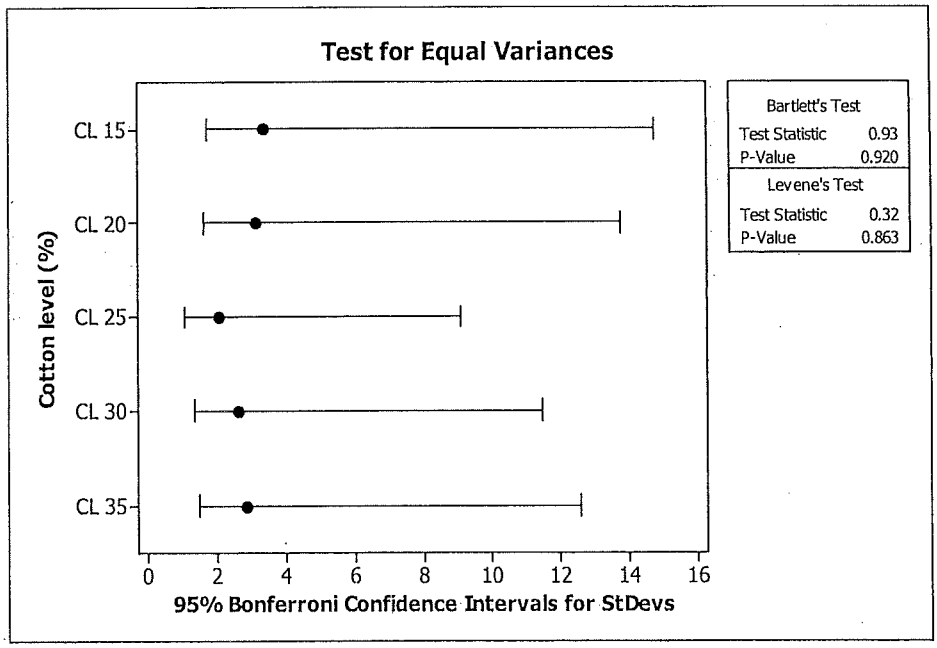

Figure 2: Test for equal variance

Rajah 2: Pengujian sama varians

(20 marks/*markah*)

(iii) Table 2 shows the ANOVA table of the experiment. Based on the ANOVA table, do the cotton weight (%) have the same effect on the strength of fiber? Why?

Jadual 2 menunjukkan jadual ANOVA bagi eksperimen tersebut. Berdasarkan jadual ANOVA ini, adakah % berat kapas mempunyai kesan yang sama pada kekuatan gentian? Kenapa?

One-way ANOVA: 15, 20, 25, 30, 35

Source	DF	SS	MS	F	P
Factor	4	475.76	118.94	14.76	0.000
Error	20	161.20	8.06		
Total	24	636.96			

S = 2.839 R-Sq = 74.69% R-Sq(adj) = 69.63%

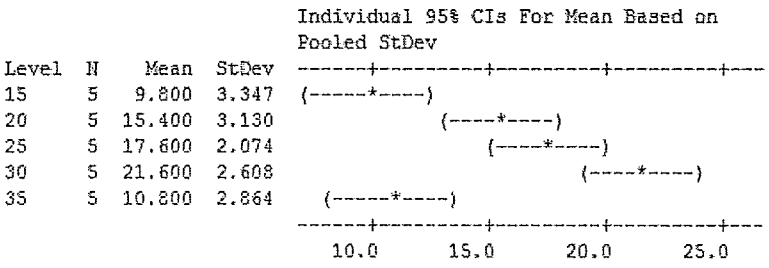

Table 2: ANOVA table

Jadual 2: Jadual ANOVA

(20 marks/*markah*)

(iv) Analyze the residuals from this experiment in Figure 3. Are the basic ANOVA assumptions satisfied?

Buat analisis pada residual daripada eksperimen dalam Rajah 3. Adakah anggapan ANOVA asas telah dipenuhi?

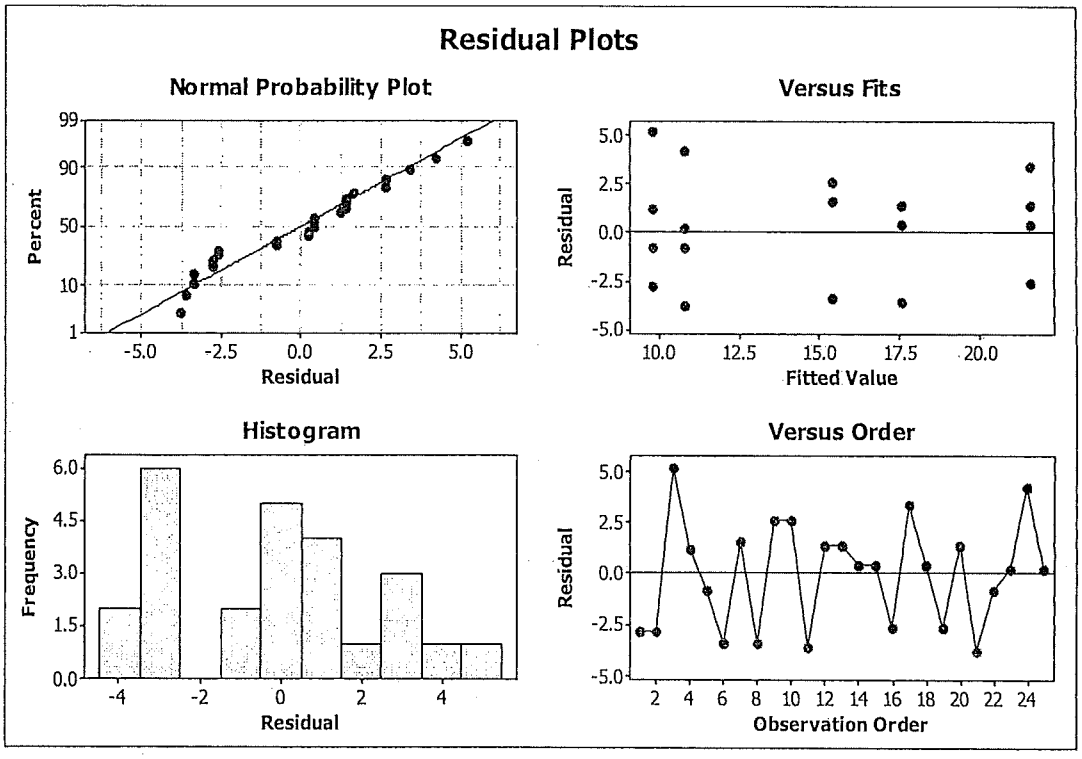

Figure 3: Residual plots

Rajah 3: Plot residual

(20 marks/*markah*)

(v) If you wish to produce fiber with the highest strength, which cotton weight % would you select?

Jika anda ingin menghasilkan gentian dengan kekuatan yang tertinggi, % berat kapas manakah yang akan anda pilih?

(10 marks/*markah*)

PART B / *BAHAGIAN B*

4. [a] An article in *Industrial Quality Control* (1956,pp 5-8) describes an experiment to investigate the effect of two factors (glass type and phosphor type) on the brightness of television tube. The response variable measured is the current (in micro amps) necessary to obtain a specified brightness level. The data are as follow:

Artikel dalam 'Industrial Quality Control' (1956, ms 5-8) menerangkan eksperimen untuk mengkaji kesan dua faktor (jenis kaca dan jenis fosfor) kepada kecerahan tiub televisyen. Pembolehubah balas diukur adalah arus (dalam amp mikro) perlu untuk mendapatkan tahap kecerahan yang ditetapkan. Data adalah seperti berikut:

Table 3 : Two Factors (Glass Type and Phosphor Type)

Jadual 3 : Dua Faktor (Jenis Kaca dan Jenis Fosfor)

Glass Type	Phosphor Type		
	1	2	3
1	280	300	290
	290	310	285
	285	295	290
2	230	260	220
	235	240	225
	240	235	230

Table 4 : Analysis of Variance for Brightness of Television Tube

Jadual 4 : Analisis Varian untuk Kecerahan Tiub Televisyen

Source	DF	SS	MS	F	$F_{d1,d2,0.05}$	Conclusion
glasstyp		14450.0	?	?	?	?
phostype		933.3	?	?	?	?
glasstyp*phostype		133.3	?	?	?	?
Error		633.3	?	?		
Total		16150.0				
S = 7.265		R-Sq = 96.08%			R-Sq(adj) = 94.44%	

Assume the significant level is α=0.05. Answer the following questions:

Andaikan aras keertian adalah $\alpha = 0.05$. Jawab soalan-soalan berikut:

(i) State the hypotheses of interest in this experiment.

Nyatakan hipotesis kepentingan dalam eksperimen ini.

(10 marks/*markah*)

(ii) Based on the above given data, construct the table and complete the empty cells.

Berdasarkan data yang diberikan di atas, bina jadual dan lengkapkan sel-sel kosong.

(30 marks/*markah*)

(iii) Test the hypotheses in part (i) and draw your conclusions using the analysis of variance.

Uji hipotesis di bahagian (i) dan buat kesimpulan anda dengan menggunakan analisis varians.

(10 marks/*markah*)

4. [b] From experiment 4(a), you are required to summit a model report to your supervisor. Your supervisor requested you to provide an accurate summary of the overall new finding using the ANOVA from 4(a) , Figure 4 and Figure 5.

Dari eksperimen, 4 (a), anda dikehendaki menyerah satu laporan model kepada penyelia anda. Penyelia anda meminta anda untuk menyediakan satu ringkasan yang tepat mengenai penemuan baru secara keseluruhan menggunakan ANOVA dari 4 (a), Rajah 4 dan Rajah 5.

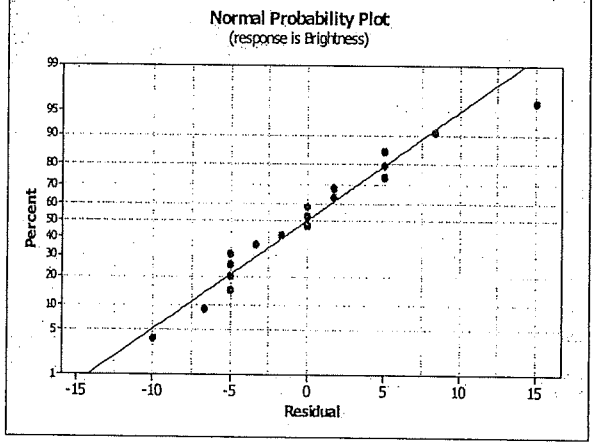

Figure 4 : Normal Probability Plot

Rajah 4 : Kebarangkalian Normal Plot

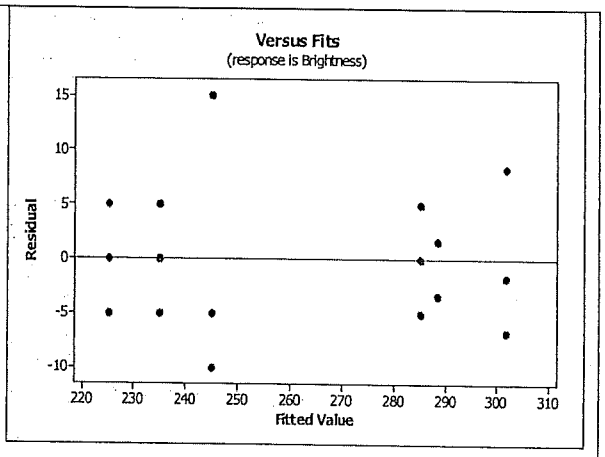

Figure 5 : Residual vs Fitted value

Rajah 5 : Residual vs Nilai Padanan

(i) Analyze the normality and residuals from this experiment using the Figure 4 and Figure 5.

Buat analisis normal dan sisa (residual) daripada eksperimen ini menggunakan Rajah 4 dan Rajah 5.

(15 marks/*markah*)

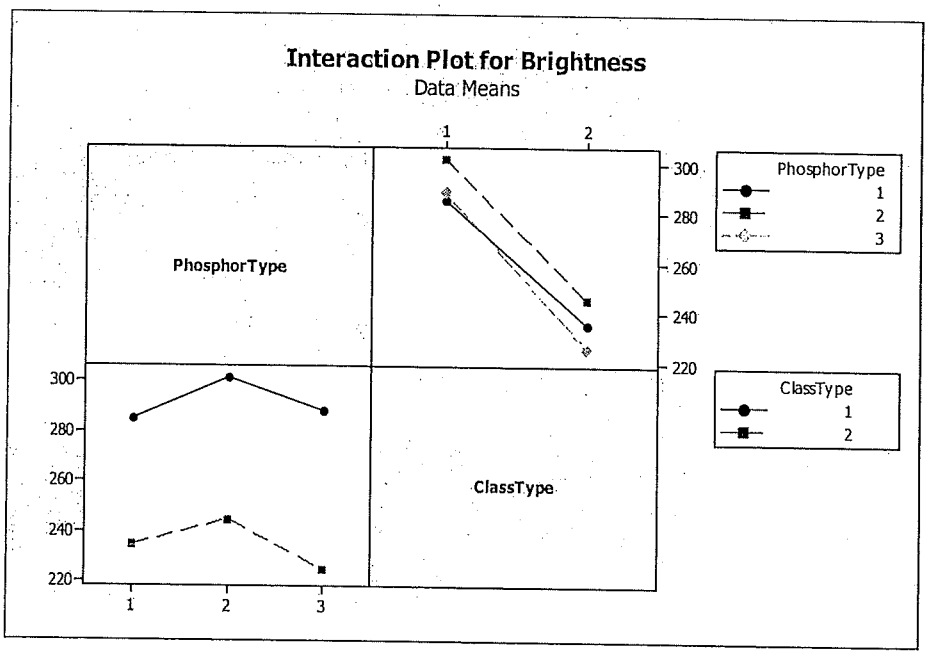

Figure 6 : Interaction Plot For Brightness And Grouping Information On Factors Using Fisher Method

Rajah 6 : Plot Interaksi Untuk Kecerahan Dan Pengumpulan Maklumat Mengenai Faktor-Faktor Menggunakan Kaedah Fisher

Grouping Information on Phosphor Type Using Fisher Method

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	------+---------+---------+---------+---
1	6	260.00	27.75	(----------*----------)
2	6	273.33	32.51	(----------*----------)
3	6	256.67	34.88	(----------*----------)

------+---------+---------+---------+---

240 260 280 300

Grouping Information on ClassType Using Fisher Method

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	------+---------+---------+---------+---
1	9	291.67	9.01	(---*--)
2	9	235.00	11.46	(---*--)

------+---------+---------+---------+---

240 260 280 300

...15/-

(ii) Provide your simple analysis for Brightness and Grouping Information on factors using the Interaction Plot and Fisher Method in Figure 6.

Sediakan satu analisis mudah anda untuk kecerahan dan Pengumpulan Maklumat mengenai faktor-faktor menggunakan Plot Interaksi dan Kaedah Fisher dalam Rajah 6.

(15 marks/*markah*)

(iii) Write a conclusion on the success of this project.
Tulis satu kesimpulan untuk kejayaan projek ini.

(20 marks/*markah*)

5. [a] Refer to Figure 7 and Table 5, consider an investigation on the effect of the concentration of the reactant and the amount of the catalyst on the conversion (yield) in a catalytic chemical process which was produced by Mr. Hakimi a final year students in PPKBSM. The objective was to determine if adjustments to either of these two factors would increase the yield.

Rujuk kepada Rajah 7 dan Jadual 5 pertimbangkan suatu siasatan terhadap kesan kepekatan bahan tindak balas dan jumlah pemangkin kepada penukaran (hasil) dalam proses kimia pemangkin yang dihasilkan oleh En. Hakimi pelajar tahun akhir PPKBSM. Objektifnya adalah untuk menentukan sama ada pelarasan kedua-dua faktor ini akan meningkatkan hasil.

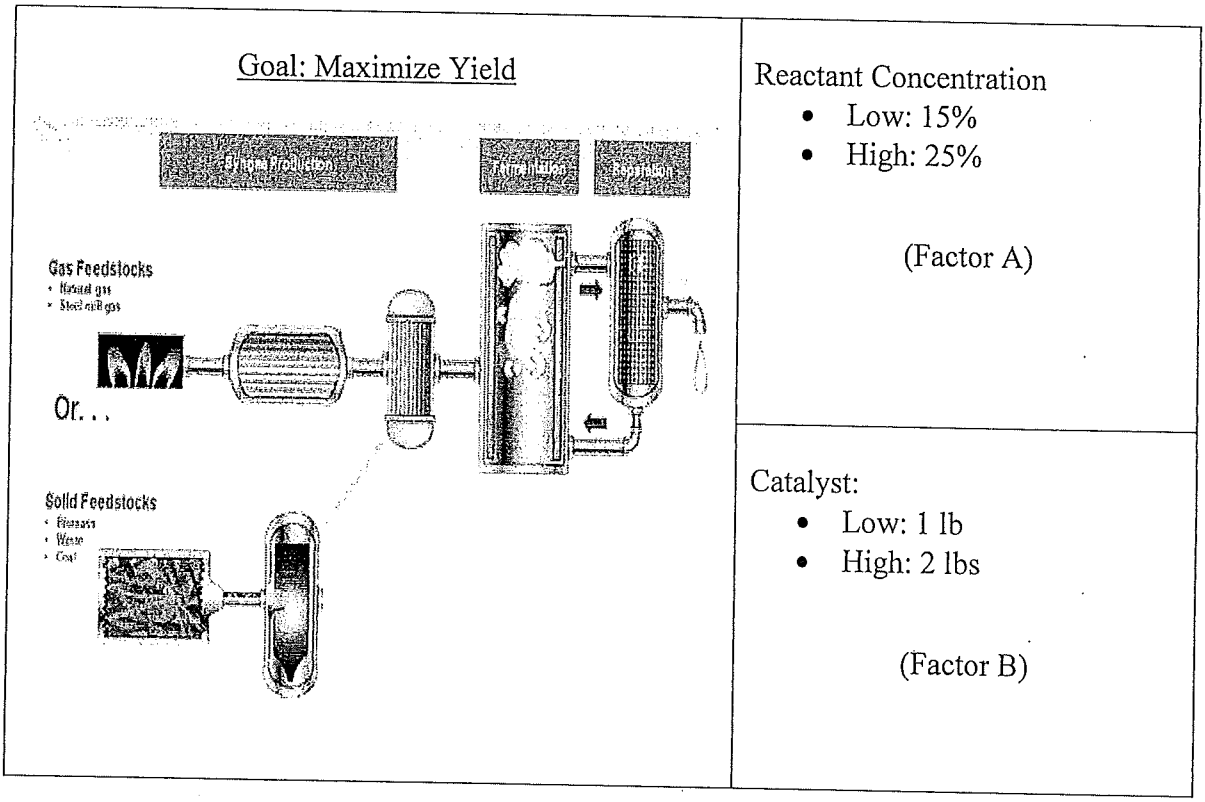

Figure 7 : Reactant Concentration(A) vs Catalyst (B)

Rajah 7 : Bahan tindakbalas Kepekatan (A) vs Catalyst (B)

Source	**DF**	**SS**	**MS**	**F**	**P**
Reactant Concentration	1	208.333	208.333	53.19	0.000
Catalyst	**1**	75.000	75.000	19.15	0.002
Interaction	**1**	8.333	8.333	2.13	0.183
Error	8	31.333	3.917		
Total	11	323.000			
S = 1.979		**R-Sq = 90.30%**		**R-Sq(adj) = 86.66%**	

Table 5 : Two-way ANOVA: Yield versus Reactant Concentration, Catalyst in Minitab

Jadual 5 : Dua Hala ANOVA: Hasil Berbanding Kepekatan Bahan Tindakbalas, Catalyst Di Minitab

(i) Explain the value S = 1.979 in Table 5 to relate the effectiveness of model response.

Terangkan nilai S = 1.979 dalam Jadual 5 untuk mengaitkan keberkesanan tindak balas model.

(10 marks/*markah*)

(ii) Compare the R-sq and R-sq (adj) values and provide your explanation on this model.

Bandingkan nilai R-sq dan R-sq (adj) dan berikan penjelasan anda untuk model ini.

(15 marks/*markah*)

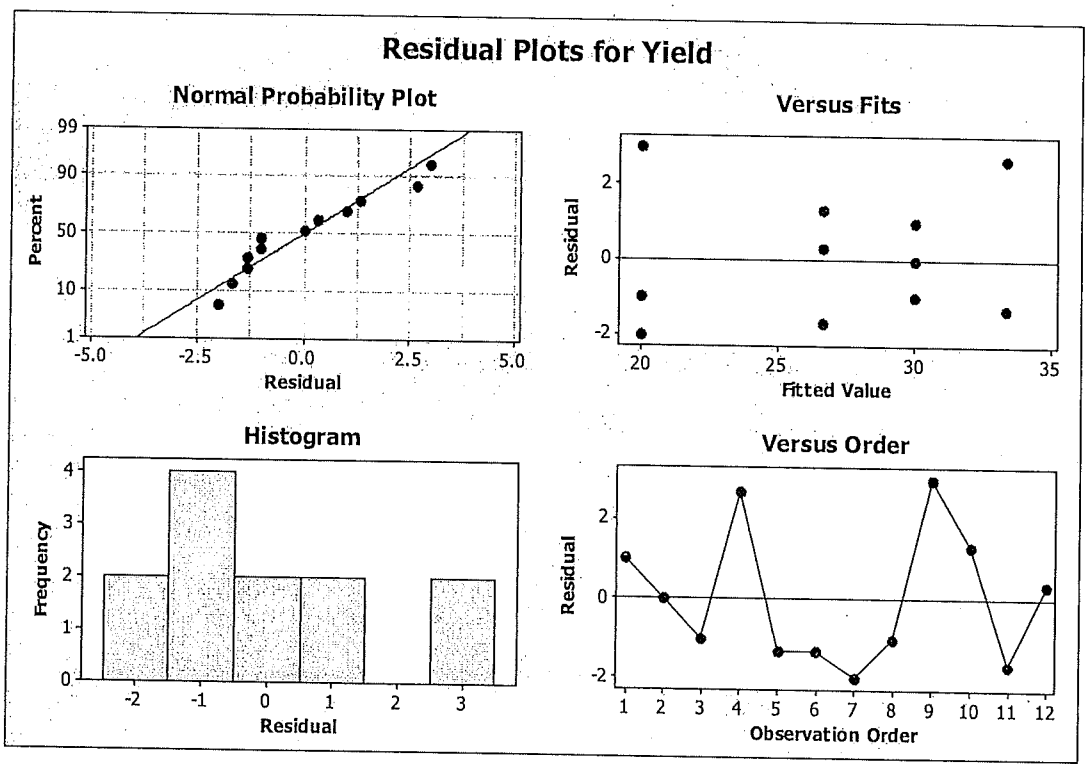

Figure 8 : Residual Plots For Yield

Rajah 8 : Plot Residual Untuk Hasil

(iii) Use the residual plots for Yield in Figure 8 and provide your analysis of model for each graphs.

Gunakan plot residual untuk Hasil dalam Rajah 8 dan berikan analisis anda untuk setiap graf bagi model ini.

(25 marks/*markah*)

[b] Use Figure 7 on chemical process for further regression analysis.

Gunakan proses kimia Rajah 7 untuk melanjutkan analisis regresi.

Table 6 : Regression analysis

Jadual 6 : Analisis Regresi

Predictor	Coef	SE Coef	T	P
Constant	32.500	2.642	12.30	0.000
Reactant Concentration	-8.333	1.212	-6.88	0.000
Catalyst	5.000	1.212	4.13	0.003

S = 2.09938 R-Sq = 87.7% R-Sq(adj) = 85.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	283.33	141.67	32.14	0.000
Residual Error	9	39.67	4.41		
Total	11	323.00			

Source	DF	Seq SS
Reactant Concentration	1	208.33
Catalyst	1	75.00

(i) State the general linear model which can fit for this experiment.

Nyatakan model linear am yang boleh dipadankan untuk eksperimen ini.

(10 marks/*markah*)

(ii) Use the corresponding ANOVA table to construct polynomial regression (if in the coded variables form).

Gunakan jadual ANOVA yang sepadan untuk membentuk regresi polinomial (jika dalam bentuk pembolehubah berkod).

(10 marks/*markah*)

(iii) State the orthogonally coded (or the conversion to coded to uncoded) factors for the above Reactant concentration (A) and Catalyst (B)?

Nyatakan faktor-faktor yang dikodkan dalam bentuk ortagon (atau penukaran kepada yang dikodkan kepada tidak dikodkan) untuk kepekatan Reaktan (A) dan Pemangkin (B) di atas.

(10 marks/*markah*)

(iv) Let say Mr. Hakimi is trying to predict the chemical reaction for Reactant concentration (A) =20 % and Catalyst (B) = 1.5 lbs using general linear model which was produced by Regression Analysis. Do you think Mr. Hakimi able to get maximized Yield for his experiment? Justify your answer.

Katakan En. Hakimi cuba meramalkan tindak balas kimia untuk kepekatan Bahan tindakbalas (A) = 20% dan Catalyst (B) = 1.5lbs menggunakan model linear am yang telah dihasilkan oleh Analisis Regresi. Adakah anda fikir En. Hakimi mendapat hasil yang maksimum untuk percubaan beliau? Jelaskan jawapan anda.

(20 marks/*markah*)

6. [a] An article in Oikos: A journal of Ecology entitled "Regulation of Root Vole Population dynamics by food supply and Predation: A Two-Factor Experiment" (2005,Vol.109,pp 387-395) investigated on how food supply interacts with predation in the regulation of root vole (Microtus oeconomus Pallas) population dynamics. A replicated two-factor field experiment manipulating both food supply and predation condition for root voles was conducted. Four treatments were conducted for different combination of factor A (predator-access) and factor B (food-supplemented).

Artikel dalam Oikos: Jurnal Ekologi bertajuk "Regulation of Root Vole Population dynamics by food supply and Predation: A Two-Factor Experiment" (2005, Vol.109, ms 387-395) meyiasat bagaimana bekalan makanan berinteraksi dengan pemangsa dalam kawalseliaan akar vole (Microtus oeconomus Pallas) dinamik populasi. Replikasi eksperimen dua faktor yang memanipulasi bidang kedua-dua bekalan makanan dan keadaan pemangsa untuk akar voles telah dijalankan. Empat rawatan telah dijalankan untuk kombinasi yang berbeza faktor A (pemangsa- akses) dan faktor B (makanan ditambah).

Table 7 : A Two-Factor Experiment For Regulation Of Root Vole

Jadual 7 : Satu Eksperimen Dua-Faktor Untuk kawalseliaan Akar Vole

	Coded Food Supply (A)	Coded Predation (B)	(Root vole) Replication			(Root vole) Total
			1	2	3	
(1)	–	–	65.439	89.089	172.339	326.867
a	+	–	88.589	114.059	200.979	403.627
b	–	+	40.799	47.959	74.439	163.197
ab	+	+	56.949	97.079	78.759	232.787

...22/-

Table 8 : Analysis of variance table for Root Vole

Jadual 8 : Analisis Jadual Varians Bagi Akar Vole

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	11113.89	3	3704.63	2.02	0.19
A-Food_Supply	1784.86	1	1784.86	0.97	0.353
B-Predation	9324.75	1	9324.75	5.08	0.0542
AB	4.28	1	4.28	2.33E-03	0.9627
Pure Error	14686.21	8	1835.78		
Cor Total	25800.1	11			

(i) Infer an appropriate general statistical model for this experiment?

Tulis model statistik umum yang sesuai untuk eksperimen ini?

(10 marks/*markah*)

(ii) Construct the coded linear regression model using orthogonal design.

Bina satu model regresi linear berkod menggunakan reka bentuk ortogon.

(25 marks/*markah*)

(iii) Use the ANOVA table and coded linear regression model to draw your conclusions for the experiment.

Gunakan jadual ANOVA dan model regresi linear berkod untuk membuat kesimpulan anda untuk eksperimen.

(15 marks/*markah*)

[b] There were two-factors the time (second) and temperature (degree celsius) being used to produce yield and molecular weight as multiple responses.

Terdapat dua faktor masa (kedua) dan suhu (darjah celsius) yang digunakan untuk menghasilkan hasil dan berat molekul sebagai respon berbilang.

$$Yield = 79.94 + 0.99x_1 + 0.52x_2 - 1.38x_1^2 - 1.00x_2^2$$

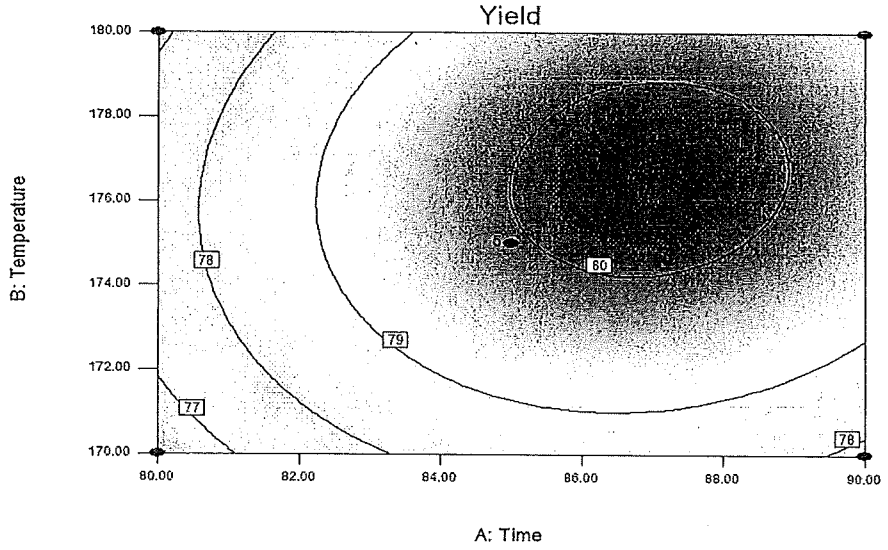

Figure 9 : 2D Contour Plot for Yield

Rajah 9 : 2D Kontur Plot untuk Yield

(i) Based on your observation in Figure 9, and overall results of regression on coded model, perform concise analysis on this experiment.

Berdasarkan pemerhatian anda dalam Rajah 9, dan keputusan keseluruhan regresi model berkod, cuba lakukan analisis ringkas mengenai eksperimen ini.

(25 marks/*markah*)

Molecular weight $= 3423.08 + 120.25x_1 + 177.35x_2$

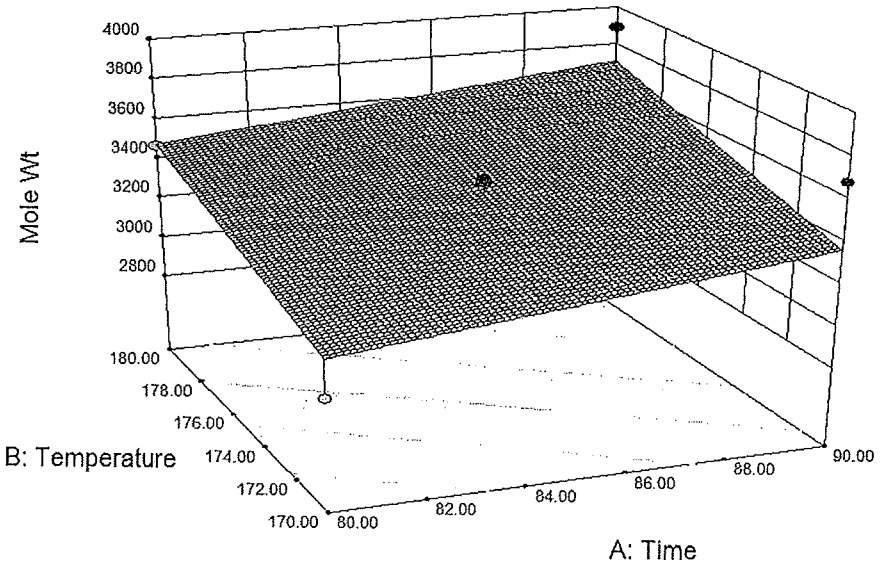

Figure 10 : 3D Surface For Molecular Weight

Rajah 10 : Permukaan 3D Untuk Berat Molekul

(ii) By using Figure 10 we can identify the local optimum results. Provide simple solution as a discussion of the molecular weight response for this experiment.

Dengan menggunakan Rajah 10 kita boleh mengenal pasti hasil optimum tempatan. Sediakan satu penyelesaian ringkas sebagai perbincangan untuk respon berat molekul dalam eksperimen ini.

(25 marks/*markah*)

- oooOooo -