UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
(ONLINE)
SEMESTER I
SESSION 2020/2021

COURSE NAME	:	STATISTICS FOR FOOD TECHNOLOGIST
COURSE CODE	:	BWD 11303
PROGRAMME CODE	:	BWD
EXAMINATION DATE	:	JANUARY/FEBRUARY 2021
DURATION	:	3 HOURS 30 MINUTES
INSTRUCTION	:	ANSWER ALL QUESTIONS **OPEN BOOK EXAMINATION**

THIS QUESTION PAPER CONSISTS OF **NINE (9)** PAGES

Q1 (a) Given three coloured jars containing items as follows: Red Jar: Containing 3 bananas and 7 peaches; Green Jar: Containing 4 bananas and 3 peaches; Blue Jar: Containing 3 bananas and 8 peaches. A jar is selected from the three jars and a fruit is taken out randomly where red, green and blue jars are having 40%, 40% and 20% respectively of being selected.

(i) Construct a tree diagram representing this experiment and obtain its sample space.

(4 marks)

(ii) If the selected fruit is banana, find the probability that this banana is taken from the Green Jar.

(8 marks)

(b) The average age of the residents in a city is 70 years, and the standard deviation is 12 years. The distribution of ages is known to be normal distribution. Suppose a group of 20 people is formed to represent all age groups. The average age of this group is 70 years.

(i) Find the sampling distribution of the average age of the residents in a city.

(3 marks)

(ii) Calculate the chance that the average age of a randomly selected group of 20 people from this population is at least 60 years old.

(5 marks)

Q2 (a) A continuous random variable X has probability function as defined below:

$$f(x) = \begin{cases} 3k, & 0 < x < 5 \\ 0, & \text{others} \end{cases}$$

Calculate the mean of X.

(4 marks)

(b) A study was made on the amount of converted sugar in a certain process (Y) at various temperatures of Celsius (X). The data were coded and recorded as in **Table Q2 (b).**

Table Q2 (b)

X	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8
Y	28.4	29.7	32.8	39.1	39.3	48.9	45.9	48.1	52.3

(i) Predict the amount of converted sugar at 60 degrees of Celsius.

(9 marks)

TERBUKA

(ii) Test the $\beta_1 \neq 0$ at the 0.05 level of significance.

(7 marks)

Q3 The average size of a farm in Jemaluang is 130 acres, while the average size of a farm in Pagoh is 110 acres. Assume that the data from both samples with size of 20 give the standard deviation of 12 acres and 14 acres, respectively. Assume that the populations are normally distributed with equal variances.

(a) Find a 95% confidence interval for average size of a farm in Jemaluang.

(6 marks)

(b) Determine the 99% confidence interval for the difference between means acres for farm in Jemaluang and Pagoh.

(7 marks)

(c) Find a 90% confidence interval for the ratio of variance acres for farm in Pagoh and Jemaluang.

(7 marks)

Q4 (a) A recently published study claimed that the food technologist year salary of 20 years' experience in Malaysia is RM240,000. New food technologist at a certain manufacturing guess that the average salary of the experienced food technologist is higher than RM240,000 and so has decided to test the null hypothesis where μ is the average salary of experienced food technologist. A random sample of seven experienced food technologist produced the following salaries (in units of RM1,000):

283, 256, 220, 240, 283, 305, 225

Is the null hypothesis rejected at the 5% level of significance?

(10 marks)

(b) A random sample of 10 hot drinks from Dispenser A had a mean volume of 304 ml and a standard deviation of 4 ml. A random sample of 15 hot drinks from Dispenser B gave corresponding values of 301 ml and 5 ml. The amount dispensed by each machine may be assumed to be normally distributed. Test the hypothesis that there is no difference in the variability of the volume dispensed by the two machines at 5% significance level,.

(10 marks)

Q5 (a) A scientist wants to study the experiments of the effectiveness of vitamin X on a person's growth. He selected 80 random samples of the 10-year-old children. However, 40 samples of children will be given vitamin X, while 40 samples of children will only be a control test without given vitamin X. Experiments are conducted within 1 year. Explain to him, what is the best way to carry out his experiment.

(6 marks)

(b) A researcher wants to study a diet problem on 16 patients. He wants to know the relationship between the final weight (y) and 11 other independent variables. The independent variables consist of weight, 1st interim weight, 2nd interim weight, 3rd interim weight, triglyceride, 1st interim triglyceride, 2nd interim triglyceride, 3rd interim triglyceride, final triglyceride, age in years and gender. The data are analysed using multiple linear regression and the output of SPSS is shown as in **Appendix A.**

(i) State **TWO (2)** assumptions that need to be fulfilled before we can use the multiple linear regression model. Give your comments of these assumptions with the output in **Appendix A.**

(4 marks)

(ii) Referring to the SPSS output in **Appendix A**, draw a conclusion about this diet problem involving 16 patients.

(10 marks)

– END OF QUESTIONS –

TERBUKA

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2020/2021 PROGRAMME CODE: BWD
COURSE NAME : STATISTICS FOR FOOD TECHNOLOGIST COURSE CODE: BWD 11303

FORMULA

Random variables:

$$\sum_{i=-\infty}^{\infty} P(x_i) = 1,\quad E(X)=\sum_{\forall x} x\cdot P(x),\quad E(X^2)=\sum_{\forall x} x^2\cdot P(x),\quad \int_{-\infty}^{\infty} f(x)\,dx = 1,$$

$$E(X)=\int_{-\infty}^{\infty} x\cdot P(x)\,dx,\quad E(X^2)=\int_{-\infty}^{\infty} x^2\cdot P(x)\,dx,\quad Var(X)=E(X^2)-[E(X)]^2.$$

Special Probability Distributions :

$$P(x=r)={}^nC_r\cdot p^r\cdot q^{n-r},\ r=0,1,\ldots,n,\ X\sim B(n,p),\ P(X=r)=\frac{e^{-\mu}\cdot\mu^r}{r!},\ r=0,1,\ldots,\infty,$$

$$X\sim P_0(\mu),\ Z=\frac{X-\mu}{\sigma},\ Z\sim N(0,1),\ X\sim N(\mu,\sigma^2).$$

Sampling Distributions :

$$\overline{X}\sim N(\mu,\sigma^2/n),\ Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1),\ T=\frac{\overline{x}-\mu}{s/\sqrt{n}},\ \overline{X}_1-\overline{X}_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right).$$

Estimations :

$$n=\left(\frac{Z_{\alpha/2}\cdot\sigma}{E}\right)^2,\quad \overline{x}-Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}<\mu<\overline{x}+Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}},\quad \overline{x}-t_{\alpha/2,v}\sqrt{\frac{s^2}{n}}<\mu<\overline{x}+t_{\alpha/2,v}\sqrt{\frac{s^2}{n}}$$

$$\left(\overline{x}_1-\overline{x}_2\right)-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\left(\overline{x}_1-\overline{x}_2\right)+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},$$

$$\left(\overline{x}_1-\overline{x}_2\right)-Z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<\left(\overline{x}_1-\overline{x}_2\right)+Z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},$$

$$\left(\overline{x}_1-\overline{x}_2\right)-t_{\alpha/2,v}\cdot S_p\sqrt{\frac{2}{n}}<\mu_1-\mu_2<\left(\overline{x}_1-\overline{x}_2\right)+t_{\alpha/2,v}\cdot S_p\sqrt{\frac{2}{n}};v=2n-2$$

$$\left(\overline{x}_1-\overline{x}_2\right)-t_{\alpha/2,v}\cdot S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<\left(\overline{x}_1-\overline{x}_2\right)+t_{\alpha/2,v}\cdot S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

where Pooled estimate of variance, $S_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ with $v=n_1+n_2-2$,

$$\left(\overline{x}_1-\overline{x}_2\right)-t_{\alpha/2,v}\sqrt{\frac{1}{n}\left(s_1^2+s_2^2\right)}<\mu_1-\mu_2<\left(\overline{x}_1-\overline{x}_2\right)+t_{\alpha/2,v}\sqrt{\frac{1}{n}\left(s_1^2+s_2^2\right)}\text{ with } v=2(n-1),$$

$$\left(\overline{x}_1-\overline{x}_2\right)-t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<\left(\overline{x}_1-\overline{x}_2\right)+t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\text{ with } v=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$$

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2020/2021 PROGRAMME CODE: BWD
COURSE NAME : STATISTICS FOR FOOD TECHNOLOGIST COURSE CODE: BWD 11303

FORMULA

$$\frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2,v}} < \sigma^2 < \frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2,v}} \quad \text{with } v = n - 1,$$

$$\frac{s_1^2}{s_2^2}\cdot\frac{1}{f_{\alpha/2}(v_1,v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\cdot f_{\alpha/2}(v_2,v_1) \quad \text{with } v_1 = n_1 - 1 \text{ and } v_2 = n_2 - 1$$

Hypothesis Testing :

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{with } v = n_1 + n_2 - 2,$$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n}\left(s_1^2 + s_2^2\right)}}, \quad T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{with}$$

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}\ ; \quad S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}\ ; \quad \chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$F = \frac{s_1^2}{s_2^2}, \quad \text{with } \frac{1}{f_{\alpha/2}(v_2,v_1)} \quad \text{and} \quad f_{\alpha/2}(v_1,v_2)$$

Simple Linear Regressions :

$$S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}, \quad S_{xx} = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}, \quad S_{yy} = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}, \quad \bar{x} = \frac{\sum x}{n}, \quad \bar{y} = \frac{\sum y}{n},$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x, \quad r = \frac{S_{xy}}{\sqrt{S_{xx}\ S_{yy}}}, \quad SSE = S_{yy} - \hat{\beta}_1 S_{xy}, \quad MSE = \frac{SSE}{n-2},$$

$$T = \frac{\hat{\beta}_1 - \beta_1^*}{\sqrt{\frac{MSE}{S_{xx}}}} \sim t_{n-2}, \quad T = \frac{\hat{\beta}_0 - \beta_0^*}{\sqrt{MSE\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim t_{n-2}$$

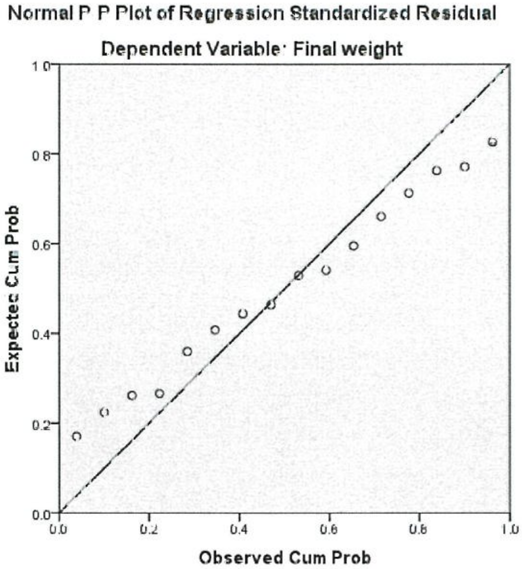
TERBUKA

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2020/2021 PROGRAMME CODE: BWD
COURSE NAME : STATISTICS FOR FOOD TECHNOLOGIST COURSE CODE: BWD 11303

Appendix A

First Analysis



Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999[a]	.999	.996	2.145

a. Predictors: (Constant), 2nd interim weight, 3rd interim triglyceride, 2nd interim triglyceride, Triglyceride, Final triglyceride, Age in years, 1st interim triglyceride, Gender, Weight, 1st interim weight

b. Dependent Variable: Final weight

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16818.434	10	1681.843	365.557	.000[b]
	Residual	23.004	5	4.601		
	Total	16841.438	15			

a. Dependent Variable: Final weight

b. Predictors: (Constant), 2nd interim weight, 3rd interim triglyceride, 2nd interim triglyceride, Triglyceride, Final triglyceride, Age in years, 1st interim triglyceride, Gender, Weight, 1st interim weight

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2020/2021
COURSE NAME : STATISTICS FOR FOOD TECHNOLOGIST

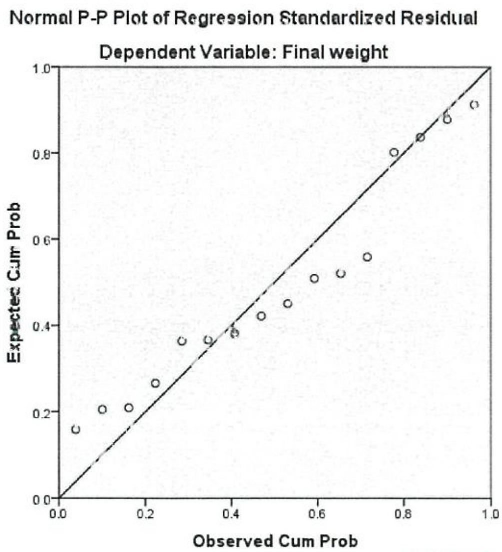
PROGRAMME CODE: BWD
COURSE CODE: BWD 11303

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	22.783	18.242		1.249	.267		
	Age in years	-.111	.172	-.023	-.644	.548	.217	4.616
	Gender	-7.596	3.760	-.116	-2.020	.099	.083	12.102
	Triglyceride	-.038	.033	-.033	-1.142	.305	.337	2.967
	1st interim triglyceride	.067	.058	.050	1.158	.299	.147	6.814
	2nd interim triglyceride	.043	.067	.028	.648	.545	.145	6.920
	3rd interim triglyceride	-.038	.043	-.038	-.899	.410	.153	6.552
	Final triglyceride	.004	.034	.004	.124	.906	.313	3.195
	Weight	-.521	.450	-.520	-1.157	.299	.001	739.371
	1st interim weight	-.436	.534	-.438	-.816	.452	.001	1052.990
	2nd interim weight	1.860	.678	1.860	2.745	.041	.001	1679.940

a. Dependent Variable: Final weight

Second Analysis



TERBUKA

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2020/2021 PROGRAMME CODE: BWD
COURSE NAME : STATISTICS FOR FOOD TECHNOLOGIST COURSE CODE: BWD 11303

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999[a]	.998	.995	2.270

a. Predictors: (Constant), 3rd interim weight, 3rd interim triglyceride, 2nd interim triglyceride, Triglyceride, Final triglyceride, Age in years, 1st interim triglyceride

b. Dependent Variable: Final weight

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16800.217	7	2400.031	465.789	.000[b]
	Residual	41.221	8	5.153		
	Total	16841.438	15			

a. Dependent Variable: Final weight

b. Predictors: (Constant), 3rd interim weight, 3rd interim triglyceride, 2nd interim triglyceride, Triglyceride, Final triglyceride, Age in years, 1st interim triglyceride

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-1.519	11.568		-.131	.899		
	Age in years	-.273	.154	-.056	-1.773	.114	.302	3.309
	Triglyceride	.001	.024	.001	.055	.957	.689	1.450
	1st interim triglyceride	.077	.052	.058	1.491	.174	.202	4.949
	2nd interim triglyceride	.088	.065	.057	1.342	.216	.169	5.924
	3rd interim triglyceride	-.066	.041	-.065	-1.617	.144	.189	5.290
	Final triglyceride	.023	.033	.021	.705	.501	.359	2.785
	3rd interim weight	.994	.022	1.006	44.646	.000	.603	1.659

a. Dependent Variable: Final weight

TERBUKA