UNIVERSITY OF RUHUNA – FACULTY OF MEDICINE

ALLIED HEALTH SCIENCES DEGREE PROGRAMME

SECOND BPHARM PART I EXAMINATION - JULY 2017

PH 2143 BIOSTATISTICS II

DATE: 07/07 /2017 INDEX NO: ....

Three (03) HOURS

INSTRUCTIONS

- No paper should be removed from the examination hall.
- Create a folder on the desktop with your index number as the name of the folder and save all your work in that folder.
- Answer all Questions.

1.

(a) The amounts of decrease in percent body fat for eight participants in an exercise program over three weeks are given as follows:

1.8, 10.6, -1.2, 8.9, 12.1, -2.1, 6.2, 4.8.

(i) Construct the boxplot of the data. Based on the graph, how would you describe the shape of the distribution? (*10 marks*)

(ii) Construct the normal Q-Q plot. Would you reject normality based on this plot? How would you describe the shape? Does it agree with your answer to part (i)? (*10 marks*)

(iii) Out of the mean and median, which one would best describe the center of the distribution? Justify your answer? (*10 marks*)

(iv) Give your estimate of the center. (*10 marks*)

(v) How would you measure the variability in the distribution? What is your estimate? (*10 marks*)

(b) A clinical trial is conducted at the gynecology unit of a major hospital to determine the effectiveness of drug A in preventing premature birth. In the trial, 30 pregnant women are to be studied, 15 in a treatment group to receive drug A and 15 in a control group to receive a placebo. The patients are to take a fixed dose of each drug on a one-time-only basis between the 24th and 28th weeks of pregnancy. The patients are assigned to groups based on computer-generated random numbers, where for every two patients eligible for the study, one is assigned randomly to the treatment group and the other to the control group. Birthweights for each group are as follows:

Baby weight (kg)	
Treatment Group	Control Group
3.13	2.91
3.45	3.04
3.31	2.45
3.45	3.72
3.09	2.41
3.27	3.00
3.63	2.63
2.55	2.59
2.63	2.82
3.31	3.22
3.72	3.18
3.13	3.13
3.09	2.54
2.59	1.91
3.90	3.09

(i) Construct the side-by-side boxplots for the treatment group and the control group. Are the parent distributions symmetric or skewed? **(10 marks)**

(ii) Complete the tests of normality and state your conclusion. **(10 marks)**

(iii) Obtain the sample variance for each group. Do you think the population variances are homogeneous? Explain. **(10 marks)**

(iv) Based on your results in part (i), (ii) & (iii), use the most appropriate test to evaluate the drug compound. Give the value of test statistic and p-value. State your conclusion. **(20 marks)**

2. The data in the following table represent the number of hours of relief provided by four different brands of headache tablets administered to 24 subjects experiencing fevers of 38°C or more.

Tablet			
A	B	C	D
5.2	9.1	2.4	7.1
4.7	7.1	3.4	6.6
8.1	8.2	4.1	9.3
6.2	6.0	1.0	4.2
3.0	9.1	4.0	5.6
5.8	7.2	3.2	6.5

(a) State the null and alternative hypotheses for a one-way analysis of variance test for comparing four brands of headache tablets. **(10 marks)**

(b) Construct side by side boxplots. Does it appear that there is a difference among the four tablets? *(10 marks)*

(c) Construct normal Q-Q plots for each tablet. Does the normality assumption seem to be satisfied? *(10 marks)*

(d) Obtain the sample standard deviation for each brand of tablet. Do you think the population standard deviations are homogeneous? Explain. *(10 marks)*

(e) Do you think that the assumptions for the F test in one-way analysis of variance are satisfied? Explain. *(10 marks)*

(f) Obtain the analysis of variance output for the above data. *(10 marks)*

(g) Test the hypothesis that there is no difference among the four tablets. Be sure to state your conclusions. *(20 marks)*

(h) Determine whether a multiple comparison test is necessary. If it is, complete the test and indicate what differences exist among the four tablets. *(20 marks)*

3.

(a) Two drugs, amantadine (A) and rimantadine (R), are being studied for use in combatting the influenza virus. A single 100-milligram does is administered orally to healthy adults. The variable studied is T_{max}, the time in minutes required to reach maximum plasma concentration. The following data are obtained:

$T_{max}(A)$				$T_{max}(R)$		
105	205	203		230	236	290
126	300	134		253	246	240
120	200	142		256	320	283
119	123	108		227	271	516
133	180			264	280	

(i) Construct the side-by-side boxplots for the $T_{max}(A)$ group and the $T_{max}(R)$ group. Are the parent distributions symmetric or skewed? *(10 marks)*

(ii) Complete the tests of normality and state your conclusion. *(10 marks)*

(iii) Based on your results in part (i) & (ii), use the most appropriate test to compare the center of the two distributions. Give the value of test statistic and p-value. State your conclusion. *(30 marks)*

(b) A survey of children under 15 years of age residing in the city area were classified according to ethnic group and hemoglobin level. The results were as follows:

Ethnic group	Hemoglobin level (g/100 ml)			
	10.0 or greater	9.0-9.9	<9.0	Total
A	80	100	20	200
B	99	190	96	385
C	70	30	10	110
Total	249	320	126	695

Do these data provide sufficient evidence to indicate, at the 0.05 level of significance, that the two variables are related? Answer the question by carrying out the following steps:

    (i) State the null and alternative hypotheses of the test.   **(10 marks)**
    (ii) Obtain the output of the statistical test for the hypothesis in Prat (i).

                                                       (10 marks)

    (iii) Write down the value of the Pearson Chi-Square test statistic.

                                                       (10 marks)

    (iv) Write down the p-value.

                                                       (10 marks)

    (v) Write down your conclusion.

                                                       (10 marks)

4. The following data were obtained in a study of the relationship between the weight (x) and chest size (y) of infants at birth.

Weight (kg)	Chest Size (cm)
2.75	29.5
2.15	26.3
4.41	32.2
5.52	36.5
3.21	27.2
4.32	27.7
2.31	28.3
4.30	30.3
3.71	28.7

(a) Calculate the sample correlation, r.   **(10 marks)**
(b) Test the null hypothesis that $\rho = 0$ against the alternative that $\rho > 0$ at the 0.01 level of significance.   **(10 marks)**
(c) Obtain the least square estimates of β_0 and β_1 for the model $y = \beta_0 + \beta_1 x + \varepsilon$.

                                                       (10 marks)
(d) Interpret the meaning of β_1 in this problem.   **(20 marks)**
(e) Test the hypothesis that $\beta_1 = 0$. Write down the value of test statistic and p-value. Do the results of this test indicate that a linear trend is significant?   **(20 marks)**

(f) What percentage of the variation in infant chest sizes (y) is explained by the linear regression model? **(15 marks)**

(g) Predict the chest size of the infant when the infant weight is 3.5kg at the birth.

(15 marks)

5.

(a) Briefly explain the following designs:
    (i) Completely randomized design, **(15 marks)**
    (ii) Randomized block design **(15 marks)**

(b) Write down the mathematical model and the relevant null hypotheses for the fixed effects two-factor model. **(25 marks)**

(c) An experiment was conducted to study the effects of temperature and type of oven on the life of a particular component. Four types of ovens and 3 temperature levels were used in the experiment. Twenty-four pieces were assigned randomly, two to each combination of treatments, and the following results were recorded.

Temperature (^{0}F)	Oven			
	O1	O2	O3	O4
500	227	214	225	260
	221	259	236	229
550	187	181	232	246
	208	179	198	273
600	174	198	178	206
	202	194	213	219

Using a 0.05 level of significance, test the hypothesis that

    (a) different temperatures have no effect on the life of the component; **(15 marks)**
    (b) different ovens have no effect on the life of the component; **(15 marks)**
    (c) the type of oven and temperature do not interact. **(15 marks)**

@@@@@@@@@@@@@@