



UNIVERSITY OF RUHUNA
DEPARTMENT OF MATHEMATICS

BACHELOR OF SCIENCE (SPECIAL) DEGREE (LEVEL II)
MATHEMATICS
MST 4053: STATISTICS LABORATORY

Assignment No: 10

Semester I, 2010

.../04/2010

1. Consider the following samples.

Group A	10	12	14	23	25	26	27	27	31	33	36	37	38	40	42	43	64	65
Group B	8	12	22	24	24	26	30	33	35	35	38	39	41	44	45	50	52	
Group C	22	24	26	35	38	39	39	40	43	45	46	53	55	57	60	62		

- Obtain the stem-and-leaf plot and the descriptive statistics.
- Judging from the stem-and-leaf plots, do you think it is reasonable to assume that the samples are from normally distributed populations?
- Judging from the standard deviations of the three groups, do you think the population standard deviations are homogeneous?
- Examine the interquartile ranges of the three groups. Do their values support the homogeneous standard deviation assumption?
- From the standard deviations, calculate the mean square for error.
- Examine the means of the three groups. Do you think that there is evidence to reject the hypothesis that the population means are equal?
- Can the grand mean be calculated by averaging the means of the three groups given in the descriptive statistics? Explain.

2. A psychologist studying the effect of drug and electroshock therapy on a subject's ability to solve simple tasks. The number of tasks completed in a 10-minute period is recorded for the subjects, who have been randomly assigned to four treatment groups; drug with electroshock, drug without electroshock, no drug with electroshock, and no drug and no electroshock.

Drug w/shock	3	1	2	1	4	3	2	4	2	1	3	4	5	3	4	3	
Drug wo/shock	2	3	4	0	2	4	2	4	4	5	6	2	3	5	6	2	3
No Drug w/shock	1	2	2	0	1	2	3	1	0	3	1	3	2	0	1	2	
No Drug w/no shock	5	3	2	3	4	6	5	4	6	3	4	5	6	4	2	4	

- a) State null and alternative hypothesis to determine whether the drug and electroshocks therapies have an effect on ability to solve simple tasks.
- b) From the appearance of the box plots, do you think that the null hypothesis should be rejected?
- c) Judging from the box plot, do you think the population standard deviations are homogeneous?
- d) Is there evidence in the box plots that would suggest that the populations are not normally distributed?
- e) Do you think that the overall conditions for the ordinary F -test are met, or should we look for an alternative test?

3. The following error scores were obtained for four group of experimental animals running a maze under different experimental conditions.

Condition A	16	12	15	13	15	14	15	14
Condition B	20	18	22	17	21	19	18	18
Condition C	9	11	14	15	8	10	11	10
Condition D	15	14	18	20	16	17	17	16

- a) State the null and alternative hypothesis for an ANOVA.
- b) Based on the normal probability plots, do you think the normality assumption for an analysis of variance is met?
- c) Based on the box plots, does it appear that the population standard deviations are somewhat homogeneous.
- d) Do the box plot indicate that the null hypothesis should be rejected? Explain your answer.

4. To compare the efficiency of the pit crews of major NASCAR teams, the durations of pit stops were measured for three top crews. These are the times (in seconds) for 12 randomly selected pit stops:

Team A	25	22	18	30	24	15	40	23	10	20	45	25
Team B	25	30	24	26	22	15	32	46	20	28	35	25
Team C	30	35	32	26	37	43	36	40	35	25	55	33

- a) Complete side-by-side box plots and comment on the shapes of the distributions. Are the tails of the distributions unusually long? What about symmetry?
- b) Construct normal probability plots of the three samples. Do you detect any significant departures from normality?
- c) To test the equality of the centers of the distributions, should an ordinary F -test or the Kruskal-Wallis test to be conducted?

5. The delay times (in minutes) were recorded on 20 flights selected randomly from each of four major air carriers. Use data set *Carriers*.

- a) Construct normal probability plots and box plots of the four samples.
- b) Based on the results of part (a), should the analysis of the data be an ordinary F -test or Kruskal-Wallis test?
- c) Using the test you recommended in part (b), perform an analysis to determine whether a significant difference exists between the delay times for the four air carriers.

6. A study was conducted to compare the hostility levels of high school students in rural, suburban, and urban areas. A psychological test, the Hostility Level Test (HLT), was used to measure the degree of hostility. Fifteen students were randomly selected from each type of school and given the HLT. Summarize the data using side-by-side box plot and stem-and-leaf plots. Based on these results select appropriate test. Is there evidence to indicate that the hostility levels of students differ in the three school environments? Use data set *Hostile*.

7. Three engineering universities wished to compare the salaries of their graduates 10 years after graduation. Seventeen graduates were randomly and independently selected from each of the three universities. Summarize the data using side-by-side box plot and stem-and-leaf plots. Based on these results choose appropriate test. Is there statistical evidence to indicate that the salaries of the graduates of the three engineering schools differ? Use data set *Engineer*.

8. A home heating contractor sells three types of oil heaters. To compare the heating units, these efficiency ratings were obtained on samples of each type of heater. Use data set *Heating*.

- a) Construct side-by-side box plots and normal probability plots of the three samples. Is normality a reasonable assumption in each case? Are there many outliers?
 - b) Should the ordinary F -test or Kruskal-Wallis test be used to compare the three distributions?
 - c) Complete the test recommended in part (b) to determine whether there is statistical evidence of a difference in the mean efficiency ratings of the three types of heating units. Be sure to state a conclusion and justify your choice of statistical test.
-

9. The prospective owners of a soon-to-be-built motor lodge wished to evaluate three prospective locations for the business. On randomly selected occasions, they recorded the numbers of vehicles that passed prospective sites in 1-hour periods of time. Use data set *Lodge*.
- a) Construct normal probability plots of the three samples. Is normality a reasonable assumption in each case?
 - b) Construct box plots of the three samples. Does the variability in the three distributions appear the same?
 - c) Should the ordinary F -test or Kruskal-Wallis test be used to compare the three distributions?
 - d) Complete the test recommend in part (c) to determine whether there is a significant difference in the amount of traffic that passes the three points.
