Universiti Malaysia PAHANG

Engineering • Technology • Creativity

TEST 1

NAME	NUR SYUHAIDAH BINTI ISMAIL
STUDENT ID	CB13006
SECTION	10G
COURSE CODE	BUM2413 APPLIED STATISTICS
DATE	24 MARCH 2014
DURATION	1 HOUR AND 30 MINUTES
SESSION/SEMESTER	SESSION 2013/2014 SEMESTER II

INSTRUCTIONS TO CANDIDATE:

1. Fill in the above particulars clearly.
2. Write your student ID and the question number at the top of every answer sheet.
3. Answer all questions.
4. Write your answers in the spaces provided. All calculations and assumptions must be clearly stated.

TEST REQUIREMENTS:

1. Statistical Tables and Formula
2. Scientific calculator

Question number	FOR EXAMINER USE ONLY
	Mark
Part A	3 /4
1	18 / 19
2	6 / 8
3	11 / 11
4	8 / 8
Total marks	46 /50

DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO DO SO

This test paper consists of **EIGHT (8)** printed pages including front page.

PART A

1. What is the best measures of central tendency for the skewed data?
   - A Mean
   - B Median
   - C Mode
   - D Midrange

2. Nixon Corporation manufactures computer monitors. The following data are numbers of computer monitors produced by the company for duration of 10 days. Based on the measures of central tendency, what is the distribution of these data?

24 25 25 26 27 28 31 33 35 40

24 31 27 25 35 33 26 40 25 28

   - A Symmetric
   - B Left-skewed
   - C Right-skewed
   - D Uniform

mean 29.4
median 27.5
mode = 25

3. Which one is NOT an exploratory data analysis technique?
   - A Histogram
   - B Stem-and-leaf plot
   - C Boxplot
   - D Interquartile range

4. A university registrar has received numerous complaints about the online registration procedure at her university, claiming that the system is slow, confusing and error prone. She wants to estimate the proportion of all students at the university who are dissatisfied with the online registration procedure. Students are listed by their level of seniority; 1st year, 2nd year, 3rd year and 4th year. 1st year and 4th year students are randomly selected, then, all of them are chosen as a sample. Identify the type of sample obtained.

   - A Random samples
   - B Systematic samples
   - C Stratified samples
   - D Cluster samples

PART B

QUESTION 1

Table 1 shows the number and colour of computer keyboards assembled at Factory A and Factory B for a sample of 13 days.

Table 1: Number and colour of computer keyboards at Factory A and Factory B

Day	Factory A		Factory B	
	Number of computer keyboards	Colour	Number of computer keyboards	Colour
1	45	White	51	Green
2	52	Black	53	Purple
3	48	Red	51	White
4	41	Blue	48	Black
5	56	Green	46	Red
6	46	Purple	43	Blue
7	44	White	52	White
8	42	Black	50	Black
9	48	Red	54	Red
10	54	Blue	47	Blue
11	64	Red	39	Green
12	65	Blue	61	Purple
13	31	White	61	Black

(a) Based on Table 1,

(i) what are the variables under study?

- Number of computer keyboards
- colour of computer keyboards

(2 Marks)

(ii) categorise each variable as qualitative or quantitative.

Number of computer keyboards - quantitative
Colour - qualitative

(1 Mark)

(iii) identify the level of measurement for each variable.

Number of computer keyboards - ratio
Colour - nominal

(2 Marks)

(b) Based on the data sets in Table 1, construct a back-to-back stem-and-leaf plot of the number of computer keyboards for the two factories and compare the distributions. **(5 Marks)**

Factory A

1	3	9
8 8 6 5 4 2 1	4	3 6 7 8
6 4 2	5	0 1 1 2 3 4
5 4	6	1 1

Factory B

The shaped distribution for factory A is left skewed

The shape distribution for factory B is left skewed

By comparing the shape distribution, both of two factories A and B are left skewed distribution.

(c) The box-plots of the number of computer keyboards have been constructed in Figure 1.

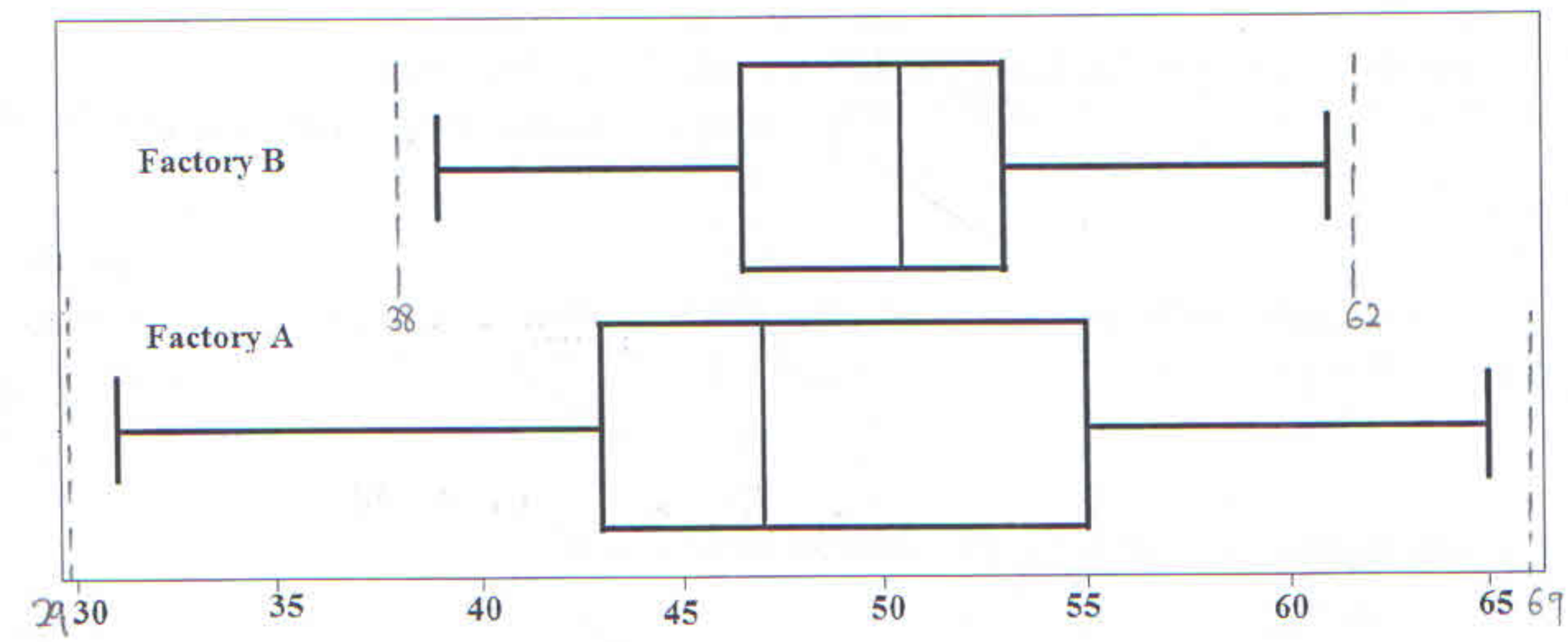

Figure 1: Box-plots of the number of computer keyboards for Factory A and Factory B

(i) Based on the box-plots in Figure 1, show that there are no outliers exist for both factories. Given that, Factory A, $Q_1 = 44$, $Q_2 = 48$ and Factory B, $Q_2 = 51$, $Q_3 = 53$.

(6 Marks)

Find outlier for factory A

$Q_3 = \frac{n_3(13)}{4} = n_{9.75} = n_{10} = 54$

$IQR = Q_3 - Q_1$
$= 54 - 44 = 10$

$Q_1 - 1.5(IQR) = 44 - 1.5(10)$
$= 29$

$Q_3 + 1.5(IQR) = 54 + 1.5(10)$
$= 69$

∴ Factory A, there is no outlier

Find outlier for factory B

$Q_1 = \frac{n_1(13)}{4} = n_{3.25} = n_4 = 47$

$IQR = Q_3 - Q_1$
$= 53 - 47 = 6$

$Q_1 - 1.5(IQR) = 47 - 1.5(6)$
$= 38$

$Q_3 + 1.5(IQR) = 53 + 1.5(6)$
$= 62$

∴ Factory B, there is no outlier

(ii) By referring to Figure 1, compare the number of computer keyboards assembled at the two factories in terms of average and variability.

(3 Marks)

- By comparing the shape distribution, Factory A is right skewed while Factory B is left skewed
- By comparing the average, Factory B has a higher median than Factory A
- By comparing the variability, Factory A has a higher variable than Factory B

QUESTION 2

Catalyst A and catalyst B in a batch of chemical process were being compared for their effect on the process reaction with the standard deviation of 4 and 5, respectively. A sample of 12 batches was prepared using catalyst A gave an average yield of 85 while 10 batches from catalyst B gave an average yield of 81.

(a) Construct a 90% confidence interval for the difference between the population means yield of the two catalysts and give the interpretation of your answer.

(5 Marks)

Catalyst A

$\bar{x} = 85$

$\sigma = 4$

$n = 12$

Catalyst B

$\bar{x} = 81$

$\sigma = 5$

$n = 10$

$Z_{0.1/2} = 1.6449$

A 90% CI for population between means

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= (85-81) \pm (1.6449) \sqrt{\frac{(4)^2}{12} + \frac{(5)^2}{12}}$$

$$= 4 \pm 3.0405$$

$$= (0.9595, 7.0405)$$

We are 90% confident that are population means yield of two catalysts is between interval (0.9595, 7.0405).

(b) How many batches of catalyst B should be used in the study if we are 99% confident that the error of the mean will lie within ±0.3?

(3 Marks)

$Z_{0.01/2} = 2.5758$

$$n = \left[\frac{Z_{\alpha/2}\,\sigma}{E}\right]^2$$

$$= \left[\frac{(2.5758)(5)}{0.3}\right]^2$$

$$= 1842.99 \approx 1843 \text{ batches}$$

QUESTION 3

Two different brands of paint, Paint K and Paint L were considered. Table 2 shows the drying times (in hour) for two different brands of paint.

Table 2: Drying times (in hour) for two different brands of paint

Paint K	3.5	2.7	3.9	4.2	3.6	2.7
Paint L	4.7	3.9	4.5	5.5	4.0	5.3

(a) Construct a 98% confidence interval on the population variance of drying times for Paint K..

(7 Marks)

$\chi^2_{\alpha/2,v} = \chi^2_{0.02/2,\,6-1} = 15.0863$

$S = 0.6186$

$n = 6$

$\chi^2_{1-\alpha/2,v} = \chi^2_{1-0.02/2,\,6-1} = 0.5543$

A 98% CI for population variance

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}} , \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}} \right)$$

$$= \left[\frac{(6-1)(0.6186)^2}{15.0863} , \frac{(6-1)(0.6186)^2}{0.5543} \right]$$

$$= (0.1268 , 3.4518)$$

(b) Construct a 99% confidence interval on the ratio of two population variances for drying times of the two brands.

(4 Marks)

Paint K

$S = 0.6186$

$n = 6$

Paint L

$S = 0.6565$

$n = 6$

$f_{\alpha/2,v_1,v_2} = f_{0.01/2,5,5} = 14.9396$

$f_{\alpha/2,v_2,v_1} = f_{0.01/2,5,5} = 14.9396$

A 99% CI for ratio of two population Variance.

$$\left(\frac{S_1^2}{S_2^2} f_{1-\alpha/2,v_2,v_1} , \frac{S_1^2}{S_2^2} f_{\alpha/2,v_2,v_1} \right)$$

$$= \left[\frac{(0.6186)^2}{(0.6565)^2} \cdot \frac{1}{14.9396} , \frac{(0.6186)^2}{(0.6565)^2} \cdot 14.9396 \right]$$

$$= (0.0594 , 13.2649)$$

QUESTION 4

The cure rate for the standard treatment of a disease is 45%. Dr. Amani has introduced her new treatment which she claims is much better. She tested the new treatment on 50 patients with the disease and cured 25 of them.

(a) Find the sampling distribution of the cure rate of a disease. **(3 Marks)**

$\mu_P = P = 0.45$
$n = 50$

$P \sim N\left(0.45, \frac{0.45(1-0.45)}{50}\right)$

$= P \sim N(0.45, 0.0050)$

(b) Construct a 95% confidence interval on the cure rate of a disease for the new treatment. **(5 Marks)**

$p = \frac{25}{50} = 0.5$ $\quad n = 50$

$Z_{0.05/2} = 1.9600$

$\left(p - z\alpha/2\sqrt{\frac{p(1-p)}{n}}, p + z\alpha/2\sqrt{\frac{p(1-p)}{n}}\right)$

$= \left[0.5 - (1.9600)\sqrt{\frac{0.5(1-0.5)}{50}}, 0.5 + (1.9600)\sqrt{\frac{0.5(1-0.5)}{50}}\right]$

$= (0.3614, 0.6386)$