UNIVERSITI SAINS MALAYSIA

Kolej Legenda

First Semester Examination
Academic Session 2008/2009

November 2008

MAA 161 – Statistics for Science Students

Duration : 3 hours

INSTRUCTIONS TO CANDIDATE:

- Please ensure that this examination paper contains **TEN** questions in **SEVEN** printed pages before you begin the examination.

- Answer ALL questions.

On each page, write only your ***Index Number***.

1. Given the following stem-and-leaf display where the leaf unit is equal to1:

6	3 3 6 7 9
7	0 0 1 1 2 2 4 5
8	4 5 5 5 6
9	0 0 1 4 5
10	4 5

   (a) How many data values are shown on this display?
   (b) List the first four data values.
   (c) Find the first quartile, mean, median and mode.

[10 marks]

2. The mean lifetime of a certain tyre is 30,000 kilometers and the standard deviation is 2,500 kilometers.
   (a) If we assume that the distribution is bell-shaped, what percentage of all such tyres will last more than 22,500 km?
   (b) If we assume nothing about the shape of the distribution, approximately what percentage of all such tyres will last between 22,500 and 37,500km?

[10 marks]

3. One student is selected at random from a group of 200 students known to consist of 140 full-time (80 female and 60 male) students and 60 part-time (40 female and 20 male) students. Events A and C are defined as follows:
   A = the student selected is full-time
   C = the student selected is female
   (a) Are events A and C independent? Justify your answer.
   (b) Find the probability $P(A \text{ or } C)$.
   (c) Find the probability $P(A \mid C)$.

[10 marks]

4. If 2% of the batteries manufactured by a company are defective, find the probability that
   (a) in a case of 20 batteries, there are 3 defective ones.
   (b) in a case of 144 batteries, there are at most 3 defective ones. Use the Poisson approximation

[8 marks]

5. The average cholesterol content of a certain brand of eggs is 215 milligrams and the standard deviation is 15 milligrams. Assume that the variable is normally distributed.
   (a) If a single egg is selected, find the probability the cholesterol content will be greater than 220 milligrams.
   (b) If a sample of 35 eggs is selected, find the probability that the mean of the sample will be larger than 220 milligrams.
   (c) If 10% of the eggs contain less than x milligrams of cholesterol, find the value of x.

[12 marks]

6. A normal population has a mean of 38 and a variance of 16. Samples of size 36 are randomly chosen.

   (a) Describe the distribution of $\overline{x}$, the mean of samples of size 36.
   (b) Find a value of k such that 95% of all such samples will have a mean $\overline{x}$ within the interval $38-k<\overline{x}<38+k$.

[8 marks]

7. A recent study indicated that 40% of the 120 women over age 35 were singles.

   (a) How large a sample must be taken to be 95% confident that the estimate is within 0.10 of the true proportion of women over age 35 who are singles?
   (b) If no estimate of the sample proportion is available, how large should the sample be?

[8 marks]

8. A manufacturer of a new drug claimed that the drug could effectively reduce the diastolic blood pressure. An experiment was designed to estimate the reduction in diastolic blood pressure using a sample of 9 people. The following data shows the diastolic blood pressure readings of these 9 people before and after consuming the drugs for two weeks.

Before	92	110	102	89	108	98	105	111	96
After	91	108	100	92	106	102	103	106	98

Perform an appropriate statistical test to test the claim that the drug could effectively reduce the diastolic blood pressure assuming that the diastolic blood pressure readings are normally distributed. Use $\alpha = 0.05$.

[12 marks]

9. The applicants for a certain course at a private college are given the Mathematics and English tests. They were then grouped into one of three categories, Poor, Average and Good based on their score in the tests. The number of applicants based on these categories is shown the table below.

		English		
		Poor	Average	Good
Mathematics	Poor	12	11	9
	Average	39	56	20
	Good	19	27	7

The administrator of the college wishes to test whether the scores in Mathematics is independent of the scores in English.

   (a) State the hypothesis of the test.
   (b) Determine the critical region and calculate the value of the test statistic. Use $\alpha =$ 0.05.
   (c) State the decision and conclusion.

[12 marks]

10. The following sample data represents the amount of time (in hour) spent on outdoor activities per week for 20 students at certain school.

11.2	12.1	12.5	11.8	10.6
13.2	12.4	10.8	11.6	13.7
9.8	10.8	12.6	12.1	13.0
11.3	12.4	13.1	12.4	11.6

At $\alpha = 0.05$, test the hypothesis that the median amount of time (in hour) spent on outdoor activities per week is 12 hours.

[10 marks]

APPENDIX

Confidence Interval

$$\overline{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad \overline{X} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

$$\overline{d} \pm t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n_d}} \qquad b \pm t_{\frac{\alpha}{2}}\, s_b$$

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\left(\hat{p}_x - \hat{p}_y\right) \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$$

$$(\overline{X} - \overline{Y}) \pm Z_{\alpha/2} \cdot \sqrt{\frac{{\sigma_x}^2}{n_x} + \frac{{\sigma_y}^2}{n_y}}$$

$$(\overline{X} - \overline{Y}) \pm t_{\alpha/2} \cdot \sqrt{\frac{{s_x}^2}{n_x} + \frac{{s_y}^2}{n_y}}$$

$$(\overline{X} - \overline{Y}) \pm t_{\alpha/2}\, \mathrm{S_p} \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

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

$$\left(\frac{s}{1 + \frac{Z_{\frac{\alpha}{2}}}{\sqrt{2n}}}, \frac{s}{1 - \frac{Z_{\frac{\alpha}{2}}}{\sqrt{2n}}} \right)$$

$$\left(\frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2},(v_2,v_1)}, \frac{s_1^2}{s_2^2} F_{\frac{\alpha}{2},(v_2,v_1)} \right)$$

Test Statistic

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

$$T = \frac{\overline{X} - \mu}{s / \sqrt{n}}$$

$$T = \frac{\overline{d} - \mu_d}{s_d / \sqrt{n_d}}$$

$$T = \frac{b - \beta_1}{s_b}$$

$$T = r\sqrt{\frac{n-2}{1-r^2}}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$Z = \frac{s - \sigma}{\sigma / \sqrt{2n}}$$

$$Z = \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{{\sigma_x}^2}{n_x} + \frac{{\sigma_y}^2}{n_y}}}$$

$$T = \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sqrt{{S_p}^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$$

$${S_p}^2 = \frac{(n_x - 1){s_x}^2 + (n_y - 1){s_y}^2}{n_x + n_y - 2}$$

$$F = \frac{s_x^2}{s_y^2}$$

$$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\hat{p}(1-\hat{p}) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$$

$$T = \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

$$dk = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y} \right)^2}{\frac{\left(\frac{s_x^2}{n_x} \right)^2}{n_x - 1} + \frac{\left(\frac{s_y^2}{n_y} \right)^2}{n_y - 1}}$$

$$\chi^2 = \sum \frac{(O - E)^2}{E}, \quad E = np$$

Table A5 Table of Critical T Values for Wilcoxon's Signed-Ranks and Matched-Pairs Signed-Ranks Test

	One-tailed level of significance			
	.05	.025	.01	.005
	Two-tailed level of significance			
	.10	.05	.02	.01
n				
5	0	–	–	–
6	2	0	–	–
7	3	2	0	–
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15
16	35	29	23	19
17	41	34	27	23
18	47	40	32	27
19	53	46	37	32
20	60	52	43	37
21	67	58	49	42
22	75	65	55	48
23	83	73	62	54
24	91	81	69	61
25	100	89	76	68
26	110	98	84	75
27	119	107	92	83
28	130	116	101	91
29	140	126	110	100
30	151	137	120	109
31	163	147	130	118
32	175	159	140	128
33	187	170	151	138
34	200	182	162	148
35	213	195	173	159
36	227	208	185	171
37	241	221	198	182
38	256	235	211	194
39	271	249	224	207
40	286	264	238	220
41	302	279	252	233
42	319	294	266	247
43	336	310	281	261
44	353	327	296	276
45	371	343	312	291
46	389	361	328	307
47	407	378	345	322
48	426	396	362	339
49	446	415	379	355
50	466	434	397	373

Table J Critical Values for the Sign Test

Reject the null hypothesis if the smaller number of positive or negative signs is less than or equal to the value in the table.

n	One-tailed, $\alpha = 0.005$ / Two-tailed, $\alpha = 0.01$	$\alpha = 0.01$ / $\alpha = 0.02$	$\alpha = 0.025$ / $\alpha = 0.05$	$\alpha = 0.05$ / $\alpha = 0.10$
8	0	0	0	1
9	0	0	1	1
10	0	0	1	1
11	0	1	1	2
12	1	1	2	2
13	1	1	2	3
14	1	2	3	3
15	2	2	3	3
16	2	2	3	4
17	2	3	4	4
18	3	3	4	5
19	3	4	4	5
20	3	4	5	5
21	4	4	5	6
22	4	5	5	6
23	4	5	6	7
24	5	5	6	7
25	5	6	6	7

Note: Table J is for one-tailed or two-tailed tests. The term n represents the total number of positive and negative signs. The test value is the number of less frequent signs.

Source: From *Journal of American Statistical Association,* vol. 41 (1946), pp. 557–66. W. J. Dixon and A. M. Mood.