Seat No.

King Mongkut's University of Technology Thonburi

Final Examination, Semester 2/2011

Subject : STA 302 Statistics for Engineers Department: Computer Engineering (English Program)

Date : 18 May 2012 Time 1.00-4.00 P.M.

Insturctions 1. There are 6 questions, 13 pages, (include this page) and 90 points

2. The books or documents are not allowed in the exam
3. The calculator is allowed in the exam
4. The statistical table is given and you have to send it back after the exam
5. Some formulae are given (4 pages)
6. Write down your ID on every sheet

Dr. Sukuman Sarikavanij

Tel : 9546

Name...ID..................Department.......................

This exam paper has been considered by the board of department of mathematics.

D. Sukawat

(Dr. Dusadee Sukawat)

Head of Mathematics Department

Department .... ID.....................................

1. The distribution of heights of a certain breed of terrier dogs is normal with a mean height of 72 centimeters and a standard deviation of 10 centimeters, whereas the distribution of heights of a certain breed of poodles is normal with a mean height of 28 centimeters and a standard deviation of 5 centimeters. If we randomly select 61 terriers and 121 poodles, find the probability that

   a.) the sample mean of heights of terriers is more than the sample mean of heights of poodles at least 44.2 centimeters, (6 points)

   b.) $P(S_1^2 / S_2^2 > 5.72)$ (6 points)

Department .... ID......................................

2. Suppose a cross-sectional study is conducted to investigate cardiovascular risk factors among a sample of patients seeking medical care at one of two local hospitals. A total of 200 patients are enrolled.

Family History of CVD	Enrollment Site	
	Hospital 1	Hospital 2
Yes	24	14
No	76	86
Total	100	100

a.) Test if there is a significant difference in proportions of patients with a family history of cardiovascular disease between two hospitals using 1% level of significance. (10 points)

Department .... ID......................................

b.) Construct a 95% confidence interval for the difference in proportions of patients with a family history of CVD between two hospitals. (6 points)

Department .... ID......................................

3. The health department at a major university is interested in whether there is a difference in the mean number of visits to the student health center between college freshmen and sophomores. The following data are collected on random samples of freshmen and sophomores, respectively, over the course of one academic year:

Year in School	Number of Students	Mean Number of Visits to the Student Health Center	Variance in Number of Visits to the Student Health Center
Freshmen	21	3.4	4.5
Sophomores	25	4.5	1.3

Suppose the distribution of the number of visits is normal. Test if there is a significant difference in the mean number of visits to the student health center between university freshmen and sophomores using a 10% level of significance. (18 points)

Department .... ID.......................................

4. Assume that the helium porosity (in percentage) of coal samples taken from any particular seam is normally distributed with true standard deviation 0.75.

   a.) Compute 98% CI for the true average porosity of a certain seam if the average porosity for 20 specimens from the seam was 4.85. (6 points)

   b.) What sample size is necessary to estimate true average porosity to within 0.2 with 99% confidence? (5 points)

Department .... ID......................................

5. In an experiment designed to study the effects of illumination level on task performance, subjects were required to insert a fine-tipped probe into the eyeholes of ten needles in rapid succession both for a low light level with a black background and a higher level with a white background. Each data value is the time (sec) required to complete the task.

	Subject								
	1	2	3	4	5	6	7	8	9
Black	25.85	28.84	32.05	25.74	20.89	41.05	25.01	24.96	27.47
White	18.23	20.84	22.96	19.68	19.50	24.98	16.61	16.07	24.59

Does the data indicate that the higher level of illumination yields a decrease of more than 5 sec in true average task completion time? Test the appropriate hypotheses.

(10 points)

Department .... ID....................................

6. A company would like to predict how the trainees in its salesmanship course will perform. At the beginning of their two-months course, the trainees are given an aptitude test. This is the x-score shown below. Records are kept of the sales records of each salesman and these constitute y-values.

X	18	26	28	34	36	42	48	52	54	60
Y	54	64	54	62	68	70	76	66	76	74

a.) Find the regression line relating performance on the test to sales. (6 points)

b.) What levels of sales would you expect from 3 salesman show scored 40, 50 and 70 on the aptitude test? (3 points)

Department .... ID.......................................

c.) Is the relation between the score and the sales linear? (10 points)

d.) Calculate the coefficient of determination for the regression of the sales on the score. (4 points)

Department .... ID....................................

Formula

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$ when σ^2 known

$$Z=\frac{\bar{X}-\mu}{S/\sqrt{n}}$$ when σ^2 unknown, $n\geq 30$

$$T=\frac{\bar{X}-\mu}{S/\sqrt{n}},\upsilon=n-1$$ when σ^2 unknown, $n<30$

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{(\sigma_1^2/n_1)+(\sigma_2^2/n_2)}}$$ when σ_1^2,σ_2^2 known

$$\chi^2=\frac{(n-1)S^2}{\sigma^2},\upsilon=n-1$$

$$F=\frac{S_1^2\cdot\sigma_2^2}{S_2^2\cdot\sigma_1^2},\upsilon_1=n_1-1,\upsilon_2=n_2-1$$

$$\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$ when σ^2 known

$$\bar{x}-z_{\alpha/2}\frac{s}{\sqrt{n}}<\mu<\bar{x}+z_{\alpha/2}\frac{s}{\sqrt{n}}$$ when σ^2 unknown, $n\geq 30$

$$\bar{x}-t_{\alpha/2}\frac{s}{\sqrt{n}}<\mu<\bar{x}+t_{\alpha/2}\frac{s}{\sqrt{n}},\upsilon=n-1$$ when σ^2 unknown, $n<30$

$$(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$ when σ_1^2,σ_2^2 known

$$(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$ when σ_1^2,σ_2^2 unknown, $n_1,n_2\geq 30$

$$(\bar{x}_1-\bar{x}_2)-t_{\alpha/2}\cdot s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+t_{\alpha/2}\cdot s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},$$

$$\upsilon=n_1+n_2-2,\ s_p=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$$ when σ_1^2,σ_2^2 unknown, $\sigma_1^2=\sigma_2^2$, $n_1,n_2<30$

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

$$\upsilon=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}}$$ when σ_1^2,σ_2^2 unknown, $\sigma_1^2\neq\sigma_2^2$, $n_1,n_2<30$

$$\bar{d}-t_{\alpha/2}\frac{s_d}{\sqrt{n}}<\mu_D<\bar{d}+t_{\alpha/2}\frac{s_d}{\sqrt{n}},\upsilon=n-1$$

$$\hat{p}-z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}<p<\hat{p}+z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Department .... ID....................................

$$(\hat{p}_1-\hat{p}_2)-z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}<p_1-p_2<(\hat{p}_1-\hat{p}_2)+z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}}<\sigma^2<\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}},\upsilon=n-1$$

$$\frac{s_1^2}{s_2^2}\frac{1}{f_{\alpha/2}(\upsilon_1,\upsilon_2)}<\frac{\sigma_1^2}{\sigma_2^2}<\frac{s_1^2}{s_2^2}f_{\alpha/2}(\upsilon_2,\upsilon_1),\ \upsilon_1=n_1-1,\upsilon_2=n_2-1$$

$$n=\left(\frac{z_{\alpha/2}\cdot\sigma}{e}\right)^2,\ n=\frac{z^2_{\alpha/2}\hat{p}\hat{q}}{e^2},\ n=\frac{z^2_{\alpha/2}}{4e^2}$$

H_0	Test Statistic	H_1	Critical region
1.1. $\mu=\mu_0$	σ^2 known $Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$z>z_\alpha$ $z<-z_\alpha$ $z<-z_{\frac{\alpha}{2}}$ and $z>z_{\frac{\alpha}{2}}$
1.2. $\mu=\mu_0$	σ^2 unknown, $n\geq 30$ $Z=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$z>z_\alpha$ $z<-z_\alpha$ $z<-z_{\frac{\alpha}{2}}$ and $z>z_{\frac{\alpha}{2}}$
1.3. $\mu=\mu_0$	σ^2 unknown, $n<30$ $T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$, $\upsilon=n-1$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$t>t_\alpha$ $t<-t_\alpha$ $t<-t_{\frac{\alpha}{2}}$ and $t>t_{\frac{\alpha}{2}}$
2.1. $\mu_1-\mu_2=d_0$	σ_1^2,σ_2^2 known $Z=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{\sqrt{(\sigma_1^2/n_1)+(\sigma_2^2/n_2)}}$	$\mu_1-\mu_2>d_0$ $\mu_1-\mu_2<d_0$ $\mu_1-\mu_2\neq d_0$	$z>z_\alpha$ $z<-z_\alpha$ $z<-z_{\frac{\alpha}{2}}$ and $z>z_{\frac{\alpha}{2}}$
2.2. $\mu_1-\mu_2=d_0$	σ_1^2,σ_2^2 unknown, $n_1,n_2\geq 30$ $Z=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{\sqrt{(S_1^2/n_1)+(S_2^2/n_2)}}$	$\mu_1-\mu_2>d_0$ $\mu_1-\mu_2<d_0$ $\mu_1-\mu_2\neq d_0$	$z>z_\alpha$ $z<-z_\alpha$ $z<-z_{\frac{\alpha}{2}}$ and $z>z_{\frac{\alpha}{2}}$
2.3. $\mu_1-\mu_2=d_0$	σ_1^2,σ_2^2 unknown, $\sigma_1^2=\sigma_2^2$, $n_1,n_2<30$ $T=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{S_p\sqrt{(1/n_1)+(1/n_2)}}$ $\upsilon=n_1+n_2-2$ $S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$	$\mu_1-\mu_2>d_0$ $\mu_1-\mu_2<d_0$ $\mu_1-\mu_2\neq d_0$	$t>t_\alpha$ $t<-t_\alpha$ $t<-t_{\frac{\alpha}{2}}$ and $t>t_{\frac{\alpha}{2}}$

Department .... ID......................................

H_0	Test Statistic	H_1	Critical region
2.4. $\mu_1 - \mu_2 = d_0$	σ_1^2, σ_2^2 unknown, $\sigma_1^2 \neq \sigma_2^2$, $n_1, n_2 < 30$ $T = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{(S_1^2/n_1) + (S_2^2/n_2)}}$ $\upsilon = \dfrac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\dfrac{\left(S_1^2/n_1\right)^2}{n_1 - 1} + \dfrac{\left(S_2^2/n_2\right)^2}{n_2 - 1}}$	$\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 \neq d_0$	$t > t_\alpha$ $t < -t_\alpha$ $t < -t_{\frac{\alpha}{2}}$ and $t > t_{\frac{\alpha}{2}}$
2.5. $\mu_D = d_0$	Pair Observation, $n < 30$ $T = \dfrac{\bar{D} - d_0}{S_D/\sqrt{n}}, \upsilon = n - 1$	$\mu_D > d_0$ $\mu_D < d_0$ $\mu_D \neq d_0$	$t > t_\alpha$ $t < -t_\alpha$ $t < -t_{\frac{\alpha}{2}}$ and $t > t_{\frac{\alpha}{2}}$
3.1. $p = p_0$	$n \geq 30$ $Z = \dfrac{X - np_0}{\sqrt{np_0 q_0}}$	$p > p_0$ $p < p_0$ $p \neq p_0$	$z > z_\alpha$ $z < -z_\alpha$ $z < -z_{\frac{\alpha}{2}}$ and $z > z_{\frac{\alpha}{2}}$
3.2. $p = p_0$	$n < 30$ $X \sim b(x; n, p_0)$	$p > p_0$ $p < p_0$ $p \neq p_0$	$X \geq x$ $X \leq x$ $X \leq x$ if $x < np_0$ or $X \geq x$ if $x > np_0$
4.1. $p_1 - p_2 = 0$	$n_1, n_2 \geq 30$ $Z = \dfrac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}\hat{Q}(1/n_1 + 1/n_2)}}$ $\hat{P}_1 = \dfrac{X_1}{n_1}, \hat{P}_2 = \dfrac{X_2}{n_2}$ $\hat{P} = \dfrac{X_1 + X_2}{n_1 + n_2}$	$p_1 - p_2 > 0$ $p_1 - p_2 < 0$ $p_1 - p_2 \neq 0$	$z > z_\alpha$ $z < -z_\alpha$ $z < -z_{\frac{\alpha}{2}}$ and $z > z_{\frac{\alpha}{2}}$
4.2. $p_1 - p_2 = d_0$ and $d_0 \neq 0$	$n_1, n_2 \geq 30$ $Z = \dfrac{(\hat{P}_1 - \hat{P}_2) - d_0}{\sqrt{(\hat{P}_1\hat{Q}_1/n_1) + (\hat{P}_2\hat{Q}_2/n_2)}}$	$p_1 - p_2 > d_0$ $p_1 - p_2 < d_0$ $p_1 - p_2 \neq d_0$	$z > z_\alpha$ $z < -z_\alpha$ $z < -z_{\frac{\alpha}{2}}$ and $z > z_{\frac{\alpha}{2}}$
5. $\sigma^2 = \sigma_0^2$	$\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ $\upsilon = n - 1$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 > \chi_\alpha^2$ $\chi^2 < \chi_{1-\alpha}^2$ $\chi^2 < \chi_{1-\frac{\alpha}{2}}^2$ and $\chi^2 > \chi_{\frac{\alpha}{2}}^2$
6. $\sigma_1^2 = \sigma_2^2$	$F = \dfrac{S_1^2}{S_2^2}$ $\upsilon_1 = n_1 - 1, \upsilon_2 = n_2 - 1$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$f > f_\alpha$ $f < f_{1-\alpha}$ $f < f_{1-\frac{\alpha}{2}}$ and $f > f_{\frac{\alpha}{2}}$

Department .... ID.....................................

$$\hat{y} = a + bx$$

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$a = \overline{y} - b\overline{x} \; .$$

$$SST = \sum y_i^2 - n\overline{y}^2, \quad SSR = b\left(\sum x_i y_i - n\overline{x}\,\overline{y}\right)$$

$$R^2 = \frac{SSR}{SST}$$