

**Statistics**  
**Ph.D. Preliminary Examination**  
**2003**

(Open book, open notes, but No old Exams and solutions)

Name \_\_\_\_\_

**1. (20 points)** The height of adult male in a particular city is normally distributed with  $\mu = 69.5''$  and standard deviation  $\sigma = 2.65''$ . A sample of  $n = 10$  men are to be measured. Find (as closely as possible) the probabilities that:

- a)  $\bar{X}$  falls with  $\mu \pm 0.75''$
- b)  $s$  falls within  $\sigma \pm 0.1''$

**2. (20 points)** Again the height of adult male,  $Y$ , in a particular area is assumed to be normally distributed, that is,  $Y \sim N(\mu, \sigma^2)$ . A sample of heights was measured by a study group and the sample mean,

$\bar{X}_1$  and sample standard deviation  $s_1$  are computed, the sample size of this group is  $n_1$ . Another study group, not knowing the work of the first group, takes another sample of height data, with sample mean,  $\bar{X}_2$  and sample standard deviation  $s_2$  are computed, the sample size of this group is  $n_2$ . The director of this study decided to combine the results of these two study groups and compute better estimates of  $\mu$  and  $\sigma$ .

- a) Develop the best statistical estimates of  $\mu$  and  $\sigma^2$ . Explain your reasons.
- b) Derive the confidence intervals for the statistical estimates of  $\mu$  and  $\sigma^2$  developed in a)

**3. (20 points)** The following data have been collected to compare the grade achieved by full time students and part time students.

Full time students

Full time students	1	2	3	4	5	6	7	8
GPA	3.3	3.9	2.9	3.3	3.2	3.5	2.8	3.8

Part time students

Part time students	1	2	3	4	5	6	7	8
GPA	3.1	3.8	3.1	3.5	2.7	3.7	3.2	3.9

- a) Construct a 95% confidence interval estimate for the difference in population mean GPAs for full time students and part time students.
- b) Test the null hypothesis that full time students earn grade at least as high as part time students.

**4. (20 points)** An industrial engineering student has collected the following sample data for 10 skilled worker. Productivity is a dependent variable, the years of experience and Sex are independent variables.

Productivity	Years of Experience	Sex
--------------	---------------------	-----

(unit/hr)		
60	2	Male
51	1	Female
53	2	Female
61	3	Female
72	5	Male
74	6	Male
73	6	Male
78	4	Female
75	7	Male
91	7	Female

Design statistical procedures to analyze above data in order to find out how ‘years of experience’ and ‘sex’ will affect ‘productivity, if any. (**design only, do not solve**) You have to give detailed descriptions of your procedures and explanations to justify your procedures.

**5. (20 points)** A petroleum refinery supervisor wishes to evaluate alternative procedures for removal of sulfur compounds from distillates in an intermediate stages of processing. Sulfur is a harmful material, so the more it gets removed, the better. The following data have obtained for pilot runs under each procedure. The value represent the percentage reduction in sulfur content of the raw distillate processed at that stage. Each test run was made with crude from one of the five major regions.

Crude Oil Source	Sulfur removal procedure			
	Evaporator	Precipitator	Slurry Reactor	Tertiary Processing
Alaska	80%	91%	88%	94%
Arabia	30	28	10	40
Louisiana	72	65	47	75
Texas	65	57	62	74
Venezuela	48	53	58	71

The following partial MINITAB output is provided:

Analysis of Variance for %Sulfur, using Adjusted SS for Tests

Source	DF	Seq SS
Crude So	4	7741.3
Procedur	3	837.2
Error	12	658.3
Total	19	9236.8

- Complete above ANOVA analysis, determine if significant differences exist among these 4 sulfur removal procedures
- What is the best procedure? What is the worst?