Name Spring, 2011

Applied Statistics Comprehensive Examination
Statistical Methods I & II

1. (20 Points) The following table presents the racial makeup for New York City's under-18 population.

Population	White	Black	Hispanic	Other
Under-18	29.2%	28.2%	31.5%	11.1%

The following table presents the racial makeup of New York City's 26,181 police officers.

Population	White	Black	Hispanic	Other
Police Officers	16,965	3,796	5,001	419

Do the police officers reflect the ethnic composition of the city's youth? Test an appropriate hypothesis and state your conclusion.

2. (25 Points) It is hypothesized that 40 percent of the student body at a large university are in favor of a ban on drinking in the dormitories. To test the hypothesis, ten students are randomly selected and asked whether they support the ban. Using as the test statistic the number of students that support the ban, and supposing the hypotheses for the test are $H_o : \pi = 0.40$ versus $H_a : \pi < 0.40$, answer the following questions.

   (a) (15 Points) What is the rejection region for the hypothesis test if the Type I Error rate is to be as close to 0.05 as possible?

   (b) (10 Points) What is the probability of a Type II Error if the true percentage of students that support the ban is 20 percent?

3. (30 Points) A pharmaceutical company tested three formulations of a pain relief medicine for migraine headache sufferers. For the experiment, 15 volunteers were selected and 5 were randomly assigned to each of three drug formulations. The subjects were instructed to take the drug during their next migraine headache episode and to report their pain on a scale of 1 (no pain) to 10 (extreme pain) 30 minutes after taking the drug. The data are displayed in the following table.

Formulation A	Formulation B	Formulation C
4	6	6
5	8	7
4	5	7
3	8	6
2	6	5
Sum = 18	Sum = 33	Sum = 31

Note: the sum of squared errors is 15.20.

(a) (10 Points) Write the mathematical model. State all assumptions and briefly describe how one would verify that the assumptions hold for these data.

(b) (15 Points) Perform the analysis of variance. What do you conclude?

(c) (5 Points) Perform Tukey's W (also known as Tukey's H.S.D.) multiple comparisons procedure to determine which formulation(s) minimize reported pain 30 minutes after taking the drug. The appropriate critical value from the studentized range distribution is 3.773.

4. (25 Points) A nutrition laboratory tests 30 "reduced sodium" hot dogs, finding that the mean sodium content is 310 mg with a standard deviation of 36 mg.

(a) (10 Points) Construct a 95% confidence interval for the true mean sodium content. Be sure to state any assumptions necessary for the confidence interval.

(b) (5 Points) Suppose 60 hot dogs were used instead of 30, and the sample mean and sample standard deviation are the same as given above. If all assumptions for the confidence interval are met, is a 95% confidence interval created based on these 60 hot dogs more, less, or equally likely to contain the true sodium content, compared to a 95% confidence interval based on 30 hot dogs? Explain

(c) (10 Points) Construct a 95% confidence interval for the true standard deviation of the sodium content. Be sure to state any assumptions necessary for the confidence interval.