**STAT 401 - Section 0201**

**FINAL EXAM** - December 17, 2009

**Instructions:** Show all work related to your solution. Credit may be deducted for numerical answers unsupported by valid reasoning or calculations. You may use calculators as needed.

**[25=15+20+10] 1.** Let $(x_1, \ldots, x_n)$ be a sample from a population with probability density function (pdf)

$$f(x;\theta) = \frac{3}{2}\theta x^2,\ -1 \le x \le 0,\ = \frac{3(2-\theta)}{2}x^2,\ 0 \le x \le 1$$

and $f(x;\theta) = 0$ for $x \notin [-1,1]$. Here $\theta$, $0 < \theta < 1$ is a parameter.
(i) Write the likelihood function (mind that the pdf is given by different expressions for positive and negative $x$'s).
(ii) Find the maximum likelihood estimator $\hat{\theta}$ of $\theta$ and show that it is unbiased.
(iii) Find the variance of $\hat{\theta}$.

**[25+10] 2.** The sample mean $\bar{x}$ and standard deviation (sd) $s$ calculated from a (small) random sample of size $n = 17$ drawn from a normal population with unknown mean $\mu$ and sd $\sigma$ are 31.8 and 2.9, respectively.
(i) Calculate a 90% confidence interval for $\mu$.
(ii) Calculate a 90% confidence interval for $\sigma$.

**[15+30] 3.** A group of equally skilled students was given two problems. The time (in min) it takes a student to solve the first problem is a normally distributed random variable with mean $\mu_1$ and sd $\sigma$. For the second problem the parameters are $\mu_2$ and $\sigma$. Assume that the students solved the problems independently. Here are two (small) samples of size $m = 4$ for the first problem and $n = 5$ for the second:

11.4; 9.5; 12.3; 14.0

10.8; 12.3; 13.5; 15.7; 14.2.

(i) State the null hypothesis and the alternative to prove or disprove the claim that the second problem is more difficult than the first.
(ii) Perform the test using the level $\alpha = 0.05$.

**[20+20] 4.** Four samples, each of size 7 from normal populations with means $\mu_1$, $\mu_2$, $\mu_3$, $\mu_4$ and common variance $\sigma^2$ resulted in

$$\bar{x}_{1.} = 4.5, \bar{x}_{2.} = 5.1, \bar{x}_{3.} = 3.4, \bar{x}_{4.} = 3.0; s_1^2 = 0.98, s_2^2 = 1.16, s_3^2 = 1.10, s_4^2 = 0.88.$$

(i) Do the data contradict the hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$? Use the significance level $\alpha = 0.05$.
(ii) If $H_0$ is accepted, stop here. Otherwise, perform multiple comparisons with $\alpha = 0.05$.

**[15+15] 5.** A sample of $n = 18$ pairs $(x_i,\ y_i)$ of random variables resulted in the following summary statistics:

$$\sum x_i = 1950, \sum x_i^2 = 211,450, \sum y_i = 48.0, \sum y_i^2 = 230.6, \sum x_i y_i = 5230.9.$$

(i) If the simple regression model were it to the data, calculate the equation of the least squares line you would use to predict $Y$ from $X$.
(ii) Calculate the sample correlation coefficient.

**[20] 6.** Assume that the response Y is related to the strength $x$ by

$$Y(x) = 2.5 + 1.2x + \epsilon$$

where the error $\epsilon$ has a normal distribution with mean zero and sd $\sigma = 0.3$. Calculate the probability that the average of independent measurements $Y(2)$, $Y(3)$, $Y(4)$ is between 6.0 and 6.3

**[10+10+15] 7.** Two coins are tossed together and the total number of heads in each toss (i. e., 0, 1 or 2) is recorded.
(i) State the null hypothesis that one coin is fair (and the other has an unknown probability $\theta$ of falling head).
(ii) If in $n$ tosses 0, 1 and 2 occured $n_0$, $n_1$ and $n_2$ times, respectively, calculate the maximum likelihood estimator of $\theta$.
(iii) For $n = 50$, $n_0 = 10$, $n_1 = 22$, $n_2 = 18$ use the chi-square statistic to test the null hypothesis at $\alpha = 0.05$ level.