MAT 2375
Midterm exam
Professor Rafal Kulik

Time: 80 minutes
Total: 100 points

17 February 2012

This is a **closed book** exam. Formula sheet and statistical tables are attached. There are **nine** questions worth 100 points. Additionally, there is one bonus question, worth of 10 points.

****

1. (15 points) Assume that $X_1, \ldots, X_n$ is a random sample from a density

$$f(x;\alpha) = \begin{cases} \alpha x^{-(\alpha+1)} & x > 1, \\ 0 & \text{otherwise} \end{cases},$$

with $\alpha > 1$. Find the maximum likelihood estimator for α.

Solution: For the observed data $x_1, \ldots, x_n$, the likelihood function is

$$L(\alpha) = \prod_{i=1}^{n} (\alpha x_i^{-\alpha}) = \alpha^n (x_1 \cdots x_n)^{-\alpha},$$

so that the log-likelihood is

$$\ell(\alpha) = n\log(\alpha) - \alpha \sum_{i=1}^{n} \log(x_i).$$

Taking derivatives with respect to α,

$$\ell'(\alpha) = \frac{n}{\alpha} - \sum_{i=1}^{n} \log(x_i) = 0.$$

Hence, the maximum likelihood estimator for α is

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \log(x_i)}.$$

2. (10 points) Assume that $X_1, \ldots, X_n$ is a random sample from a uniform distribution on $[0, \theta]$, that is

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & x \in [0,\theta], \\ 0 & \text{otherwise} \end{cases}.$$

Find the estimator for θ using method of moments.

Solution: The theoretical mean is $\mu = \theta/2$. Hence, for given data $x_1, \ldots, x_n$, the estimator of θ is obtained by solving

$$\frac{\theta}{2} = \bar{x},$$

that is, the method of moments estimator for θ is

$$\hat{\theta} = 2\bar{x}.$$

3. (10 points) The following data set describes a life length of a randomly selected males:

   ```
   74, 78, 71, 66, 82, 69, 81, 78
   ```

   Find a 95% confidence interval for the population mean μ. Assume that data come from a normal population with an unknown σ^2.

   Solution: The sample mean and the sample standard deviation are, respectively, $\bar{x} = 74.875$ and $s = 5.817154$. Also, $n = 8$, $z_{7,0.025} = 2.36$. The confidence interval has the form

   $$\left(\bar{x} - t_{7,0.025}s/\sqrt{n}, \bar{x} + t_{7,0.025}s/\sqrt{n}\right) = (70.01174, 79.73826).$$

4. (10 points) A particular drug is supposed to reduce a blood pressure. Seven subjects are tested and a systolic blood pressure before (X) and after (Y) the treatment is recorded. The blood pressures $(x_1, y_1), (x_2, y_2), \ldots, (x_7, y_7)$ are as follows:

Subject	x	y
1	122	120
2	121	122
3	118	116
4	116	115
5	124	116
6	130	124
7	126	128

   Find a 95% confidence interval for the difference $\mu_D = \mu_X - \mu_Y$ of two means. Can we conclude that the drug is effective?

   Solution: We compute differences $2, -1, 2, 1, 8, 6, -2$. Their sample mean and the sample standard deviation are, respectively, $\bar{d} = 2.285714$ and $s_d = 3.59232$. Also, $n = 7$, $t_{6,0.025} = 2.45$. The confidence interval has the form

   $$\left(\bar{d} - t_{6,0.025}s_d/\sqrt{n}, \bar{d} + t_{6,0.025}s_d/\sqrt{n}\right) = (-1.036627, 5.608).$$

   Since the confidence interval covers zero, we cannot conclude that the drug is effective.

5. (10 points) In a random sample of 1000 homes in a certain city, it is found that 228 are heated by oil. In a second city, only 123 out of 990 homes are heated by oil.

   - Find a 99% confidence interval for the proportion of homes from the first city that are heated by oil.
   - Find a 95% confidence interval for the difference of proportions.

   Solution:

   $$\widehat{p} \pm z_{.005}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} = \frac{228}{1000} \pm 2.576\sqrt{\frac{228/1000\,(1-228/1000)}{1000}}$$

   So $[0.194, 0.262]$.

The confidence interval for $p_1 - p_2$ is

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}.$$

Therefore, the 95% confidence interval is ($z_{0.05/2} = 1.96$)

$$\frac{228}{1000} - \frac{123}{990} + 1.96\sqrt{\frac{228/1000(1-228/1000)}{1000} + \frac{123/990(1-123/990)}{990}} = (0.071, 0.137).$$

6. (10 points) It is assumed that the weight of a new born Canadian baby is normally distributed with mean μ and the known variance $\sigma^2 = 588^2$. Based on nine new born babes, it was found that the mean weight at birth was $\bar{X} = 3315g$. Find a 95% confidence interval for μ.

   Solution: The confidence interval is ($z_{0.025} = 1.96$)

$$\left(3315 - 1.96588/\sqrt{9}, 3315 + 1.96588/\sqrt{9}\right) = (2930.84, 3699.16).$$

7. (15 points) Assume that $X_1, \ldots, X_{13}$ and $Y_1, \ldots, Y_7$ are two independent samples from a normal distribution with the same variance $\sigma_X^2 = \sigma_Y^2 = \sigma^2$. Compute

$$\mathbb{P}\left(\frac{S_X^2}{S_Y^2} > 3\right),$$

   where S_X^2 and S_Y^2 are sample variances.

   Solution: Since S_X^2/S_Y^2 has F distribution with $(12, 6)$ degrees of freedom, we ahve

$$\mathbb{P}\left(\frac{S_X^2}{S_Y^2} > 3\right) = 0.05.$$

8. (10 points) A journal announced that 45% of a population is going to vote for a *Great Future* party. It was also announced that an error of the estimation was 4%, and the level of confidence was 95%. What was the sample size used in this study?

   Solution: The sample size for proportion is given by

$$n = \frac{z_{\alpha/2}^2}{4\epsilon^2} = \frac{1.96^2}{40.04^2} = 600.25,$$

   so that $n = 601$.

   Note: The first sentence about 45% is redundant here. This is not a preliminary sample, rather the number obtained from the study of 601 people. Since it may have been confusing, I counted also the following solution

$$n = \frac{z_{\alpha/2}^2 0.45(1-0.45)}{\epsilon^2} = \frac{1.96^2}{40.04^2} = 594.24,$$

   so that $n = 595$. Note also, that point were subtracted if your answer was not the integer.

9. (10 points) In a company, a production time is normally distributed with mean μ and variance $\sigma^2 = 0.25$ hours. We want to construct a confidence interval for μ, using the confidence level 95%, such that the error of estimation is $\varepsilon = 0.05$. What is the required sample size?

   Solution:
$$n = \frac{z_{\alpha/2}^2 \sigma^2}{\epsilon^2} = \frac{1.96^2}{40.04^2} = 384.16,$$
   so that $n = 385$.

10. (Bonus question - 10 points) A manufacturer of rollers wishes to know if a certain enzyme reduces the drying time for the glue that is applied to the back of its rollers. The manufacturer claims that addition of an enzyme decreases drying time. To justify the claim, 22 pieces of the original product and 18 pieces of the modified product were tested. The mean drying times were, respectively, $\bar{X} = 143$ minutes and $\bar{Y} = 129$ minutes. It is further known that the drying time of the original product is normally distributed with standard deviation 32. Also, the modified product is normally distributed with the same variability as the original one. Find the 98% confidence interval for the difference $\mu_X - \mu_Y$, where μ_X is the mean drying time of the original product and μ_Y is the mean drying time of the modified product. Can you conclude that the drying time is reduced for the modified product?

    Solution: Here, $\sigma_X = \sigma_Y = 32$ is known. The confidence interval for the difference is
$$\bar{x} - \bar{y} \pm 2.326\sqrt{\frac{32^2}{22} + \frac{32^2}{18}} = 14 \pm 23.656 = (-9.656, 37.656).$$
    No reduction in the drying time.