Mathematics 376 – Mathematical Statistics
Solutions for Midterm Exam 2
April 26, 2012

I. Let $Y_1, \ldots, Y_n$ be a random sample from a population with density function

$$f(y|\theta) = \begin{cases} \frac{4y^3}{\theta^4} & \text{if } 0 \le y \le \theta \\ 0 & \text{otherwise.} \end{cases}$$

A) (10) Determine the method of moments estimator for θ.

Solution: Computing $E(Y)$ from this density, we find:

$$E(Y) = \int_0^\theta y \cdot \frac{4y^3}{\theta^4}\, dy = \frac{4\theta}{5}.$$

For the method of moments, we set this equal to $\overline{Y}$, the first sample moment, and solve for θ:

$$\widehat{\theta}_{MOM} = \frac{5\overline{Y}}{4}.$$

B) (10) Determine the maximum likelihood estimator for θ.
Solution: The likelihood function in this case is

$$L(y_1, \ldots, y_n \mid \theta) = \frac{4^n (\prod_{i=1}^n y_i)^3}{\theta^{4n}} \tag{1}$$

As a function of θ, neither this nor

$$\ln(L) = n\ln(4) + 3\sum_{i=1}^n \ln(y_i) - 4n\ln(\theta)$$

has a critical point. Both are strictly *decreasing* as functions of θ. Therefore, we want to choose θ as small as possible, consistent with the sample values, in order to maximize L. This occurs when $\theta = Y_{(n)}$, the sample maximum. Therefore

$$\widehat{\theta}_{ML} = Y_{(n)}.$$

C) (5) Is $U = \sum_i \ln(Y_i)$ a sufficient statistic for θ? Why or why not?
Solution: The answer is *yes* by the factorization criterion. From (1) above we have

$$L(y_1, \ldots, y_n \mid \theta) = \frac{4^n e^{3U}}{\theta^{4n}} = g(U, \theta)$$

We can take the other factor $h(y_1, \ldots, y_n) = 1$.

II. Assume that the hardness of iron bars produced by one plant is normally distributed with mean μ and variance σ^2. A random sample of $n = 7$ bars from the output of the plant is measured to determine whether the average hardness meets the desired figure $\mu_0 = 172$.

A) (20) The following measurements:

$$167, 174, 179, 164, 163, 160, 168$$

were obtained. Test $H_0 : \mu = 172$ versus the alternative $H_a : \mu \neq 172$ using an appropriate test with $\alpha = .01$.

Solution: Since $n < 30$, we must use a t-test here. We have $\overline{Y} \doteq 167.86$ and $S^2 \doteq 43.81$. We have

$$t = \frac{167.82 - 172}{\sqrt{\frac{43.81}{7}}} \doteq -1.66$$

The rejection region for the test is determined from the t-table. The form of the alternative hypothesis indicates we want

$$RR = \{t \mid |t| > t_{.005}\} = \{t \mid |t| > 3.707\}$$

(for 6 df). There is not sufficient evidence to reject H_0 here.

B) (10) How is your answer from A related to the 99% confidence interval for μ that we learned to compute in the first section of the course?

Solution: The 99% confidence interval would be

$$167.82 \pm 3.707\sqrt{\frac{43.81}{7}} = 167.82 \pm 9.27 = (158.58, 177.13).$$

The value $\mu_0 = 172$ is one of the numbers in this interval, so it is a "believable" value for μ based on the data. Alternatively, the rejection region for the test is the *complement* of the interval

$$172 \pm 3.707\sqrt{\frac{43.81}{7}} = 172 \pm 9.27$$

Since the sample $\overline{Y}$ is in this interval we do not reject H_0.

III. In this problem, let S represent the proportion of women who agree that men are basically selfish and egotistical.

A) (15) Suppose that surveys done 20 years ago showed that S was 0.40 at that time, but a new survey finds that $Y = 1260$ out of a random sample of $n = 3000$ women said that men are basically selfish and egotistical. Is there sufficient evidence to say the proportion S has increased? Give the null and alternative hypotheses for a suitable test, carry out the analysis, estimate the attained significance level (p-value) and state your conclusion succinctly and clearly.

Solution: The wording of the question is supposed to suggest using $H_0 : S = .4$ and $H_a : S > .4$. Since $n = 3000$ is large, we use a Z-test:

$$Z = \frac{\frac{1260}{3000} - .4}{\sqrt{\frac{(.4)(.6)}{3000}}} = 2.236.$$

For the upper tail test, we would reject H_0 for any α with $z_\alpha \leq 2.236$. From the standard normal table, this says the p-value is about .0125. There is relatively strong evidence to reject H_0.

B) (15) Suppose that the current value of S is actually $S = .43$ and the test in part A was carried out to yield $\alpha = 0.05$. How large would n have to be in a sample like the one from part A to ensure that β (the probability of a Type II error) is .1 or less?

Solution: In terms of $\frac{Y}{n}$, the rejection region for the test with $\alpha = 0.05$ (and a general sample size n) is

$$\left\{ \frac{Y}{n} > .4 + 1.645\sqrt{\frac{(.4)(.6)}{n}} \right\}$$

Recall that β is the chance of *not rejecting* H_0 when H_0 is false. If $S = .43$, we start from

$$P\left(\frac{Y}{n} \leq .4 + 1.645\sqrt{\frac{(.4)(.6)}{n}} \right)$$

and "restandardize" using $S = .43$:

$$\begin{aligned}
\beta &= P\left(\frac{Y}{n} \leq .4 + 1.645\sqrt{\frac{(.4)(.6)}{n}} \right) \\
&= P\left(\frac{\frac{Y}{n} - .43}{\sqrt{\frac{(.43)(.57)}{n}}} \leq \frac{-.03 + 1.645\sqrt{\frac{(.4)(.6)}{n}}}{\sqrt{\frac{(.43)(.57)}{n}}} \right) \\
&= P\left(Z \leq \frac{-.03 + 1.645\sqrt{\frac{(.4)(.6)}{n}}}{\sqrt{\frac{(.43)(.57)}{n}}} \right)
\end{aligned}$$

To get this to be .1, we want

$$\frac{-.03 + 1.645\sqrt{\frac{(.4)(.6)}{n}}}{\sqrt{\frac{(.43)(.57)}{n}}} \doteq -1.28$$

so

$$\sqrt{n} = \frac{1.645\sqrt{(.4)(.6)} + 1.28\sqrt{(.43)(.57)}}{.03} \doteq 47.98$$

so $n \geq 2303$ will do. (In other words, the test with $n = 3000$ will already have β this small(!).)

IV. (15) Let (x_i, y_i, z_i), $i = 1, \ldots, n$ be a collection of data points. Using the matrix formulation, derive the normal equations for the least squares estimators for the coefficients $\beta_0, \beta_1, \beta_2$ in the model $Z = \beta_0 + \beta_1 x + \beta_2 y + \varepsilon$ fitting the data.

Solution: The normal equations are $X^t X\beta = X^t Z$ where

$$X = \begin{pmatrix} 1 & x_1 & y_1 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{pmatrix} \quad \text{and} \quad Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

Hence the normal equations can be written in matrix form as

$$\begin{pmatrix} n & \sum_i x_i & \sum_i y_i \\ \sum_i x_i & \sum_i x_i^2 & \sum_i x_i y_i \\ \sum_i y_i & \sum_i x_i y_i & \sum_i y_i^2 \end{pmatrix} \begin{pmatrix} \widehat{\beta_0} \\ \widehat{\beta_1} \\ \widehat{\beta_2} \end{pmatrix} = \begin{pmatrix} \sum_i z_i \\ \sum_i x_i z_i \\ \sum_i y_i z_i \end{pmatrix}$$

Extra Credit "Thought Questions" – Answer in 2 or 3 complete sentences:

A) (5) Exactly why is it the case that a Z-test is not appropriate when the sample size is small and population variance is not known exactly? Explain using the case of a test on a single mean.

Solution: Solution: The reason is because of the distribution of the *test statistic* we can compute from the data. Under the assumption that the data comes from a normal population, we would estimate μ_0 by $\overline{Y}$ and σ by S. So we would use

$$t = \frac{\overline{Y} - \mu_0}{\frac{S}{\sqrt{n}}}$$

to carry out this sort of test. If the S was the population SD, σ, then (under $H_0 : \mu = \mu_0$) this would have a standard normal distribution and the Z-test would be appropriate. However, with the sample SD (S) estimating σ, this statistic actually has a t-distribution with $n-1$ df, so rejection regions set up using the standard normal would not give enough area in upper and lower tails to account for the variability of t.

B) (5) For the basic small-sample test for equality of means, we need to assume the two samples came from normal populations with equal variances. How would you determine whether applying that test was appropriate?

Solution: Whether there is a drastic deviation from normality can be detected with normal qq-plots or the Shapiro-Wilk test. Detecting nonequality of variances can be done with an F-test. The test statistic S_1^2/S_2^2 has an F-distribution with $n_1 - 1$ degrees of freedom in the numerator and $n_2 - 1$ degrees of freedom in the denominator under the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$.