Mathematics 376 – Probability and Statistics II
Solutions – Midterm Exam 2
April 19, 2004

I.

A) We have

$$E(Y) = \int_0^\theta y \cdot \frac{3y^2}{\theta^3}\, dy = \frac{3}{\theta^3}\frac{y^4}{4}\Big|_0^\theta = \frac{3\theta}{4}$$

For the method of moments estimator, we set this equal to the sample mean (the first sample moment), and solve for θ:

$$\frac{3\theta}{4} = \overline{Y} \Rightarrow \widehat{\theta}_{MM} = \frac{4\overline{Y}}{3}$$

B) The likelihood function is

$$L(y_1, \ldots, y_n|\theta) = \frac{3(y_1 \cdots y_n)^2}{\theta^{3n}}$$

if $y_i \leq \theta$ for all i, and zero otherwise. This has no critical points (it's always decreasing as a function of θ, for $\theta > 0$) – the same is true for $\ln(L)$. Hence to maximize the likelihood, we want to take θ as *small as possible*, consistent with the requirement that all the y_i must be less than or equal to θ. This means that

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

(the sample maximum).

II.

A) We have:

$$H_0 : p = .4$$
$$H_1 : p > .4$$
$$\text{Test statistic} : Z = \frac{Y/n - .4}{\sqrt{\frac{(.4)(.6)}{n}}}$$
$$RR : \{z : z > z_{.01}\} = \{z : z > 2.33\}$$

We use the value $p = .4$ in the formula for the standard error of the estimator $\widehat{p} = Y/n$ because we are working under the assumption that H_0 is true. See, for instance, the discussion in Example 10.6 on page 470 of the text.

B) Using the given data, our test statistic value is

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

This is not in the rejection region, so we cannot reject the null hypothesis. In other words, this data is *not sufficient* to conclude that p is larger than .4, at the $\alpha = .01$ level of significance.

C) The $\widehat{p}$-value corresponding to the boundary point of the rejection region in A using $n = 3000$ is $.4 + (2.33)\sqrt{\frac{(.4)(.6)}{3000}} = .42084$. We want

$$\begin{aligned}
\beta &= P(z \notin RR | p = .43) \\
&= P(Y/3000 < .42048 | p = .43) \\
&= P\left(z < \frac{.42084 - .43}{\sqrt{\frac{(.43)(.57)}{3000}}}\right) \\
&= P(z < -1.01) \\
&= P(z > 1.01) \doteq .1562
\end{aligned}$$

(Note that we use the value $p = .43$ here in the formula for the standard error of the estimator Y/n because we are now working under the assumption that p has actually increased from .4 to calculate the Type II error probability.)

III.

A) Since the sample size is small ($7 < 30$), we use a one-tail t-test for a single mean here. The null hypothesis is that $\mu = \mu_0 = 172$, the alternative is that $\mu < 172$. We compute $\overline{x} = 167.86$ and $S^2 = \frac{1}{6}\sum_{i=1}^{6}(x_i - \overline{x})^2 = 43.81$. Then our test statistic is

$$t = \frac{\overline{x} - \mu_0}{S/\sqrt{n}} = \frac{167.86 - 172}{\sqrt{43.81/7}} \doteq -1.655$$

The rejection region for the lower-tail t-test is

$$RR = \{t : t < -t_{.1}(6)\} = \{t : t < -1.44\}$$

from the t-table. Since t is in the rejection region, we reject the null hypothesis. In other words, there is sufficient evidence to say that $\mu < 172$ (at the $\alpha = .1$ level of significance).

B) To test $\sigma^2 = 40$ against $\sigma^2 \neq 40$, we want a two-tailed χ^2-test. The test statistic is

$$\chi^2 = 6S^2/\sigma^2 = 6(43.81)/40 = 6.57135$$

The rejection region is

$$RR = \{\chi^2 < \chi^2_{.95}(6)\} \cup \{\chi^2 > \chi^2_{.05}(6)\} = \{\chi^2 < 1.63539\} \cup \{\chi^2 > 12.5916\}$$

Our value 6.57135 is not in the union of these two intervals, so we do not have sufficient evidence to reject the null hypothesis here.

IV.

A) Using the example of a small sample test on a single mean (like III A above), the test statistic

$$\frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

has a t-distribution with $n - 1$ degrees of freedom. For small n ($n < 30$), this is different enough from a standard normal (more probability in the tails) that the z-test would lead to rejection regions that are "too large". For example, with an upper tail test at the $\alpha = .1$ level and $n = 7$, $t_{.1}(6) = 1.44$, but $z_{.1} = 1.28$.

B) The p-value for the two-tailed test would be twice as large as the p-value for the one-tail test. For example, in III A above, the p-value for the one-tail test is $< .1$ and $> .05$ from the book's table. Exact value using Maple is approximately .0745. If we were doing a two-tail t-test, the value 1.655 would be $t_{\alpha/2}(6)$ if $\alpha/2 = .0745$, so $\alpha = 2(.0745) = .1490$ is the p-value of the test.