Exam 1 PS217, Fall 1996

1. Below is a sample. For the data from this sample, compute the mean, median, mode, variance and standard deviation. [10 pts]

1
4
6
5
3
5
2
5
3

ΣX = 34

ΣX^2 = 150

$$\bar{X} = \frac{\sum X}{n} = \frac{34}{9} = 3.78$$

$$SS = \sum X^2 - \frac{(\sum X)^2}{n} = 150 - \frac{1156}{9} = 21.56$$

$$s^2 = \frac{SS}{n-1} = \frac{21.56}{8} = 2.695$$

$$s = \sqrt{s^2} = \sqrt{2.695} = 1.64$$

The median would be 4 (4 scores above and 4 scores below). The mode would be 5 (most frequently occurring score). Thus, the distribution would be a bit negatively skewed.

2. Test the hypothesis that the sample you see in the first question was drawn from a normal population with μ = 5 and σ = 2. [10 pts]

Because σ is known, the appropriate statistic would be a z-score.

H_0: μ = 5

H_1: μ ≠ 5

Decision Rule: If $|z_{Obt}| \geq 1.96$, reject H_0.

$$\sigma_{\bar{X}} = \sigma / \sqrt{n} = 2 / \sqrt{9} = .67$$

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{3.78 - 5}{.67} = -1.82$$

Because $|z_{Obt}| < 1.96$, we would retain H_0. Thus, it is plausible that the sample was drawn from a population with $\mu = 5$. OTOH, you should realize that you may be making a Type II error, because with only 9 scores in your sample, you may have insufficient power.

3. You all remember the OJ Simpson trial, right? OK, tell me in words what a Type I and a Type II error would be for that trial, which tested the H_0 that Simpson was not guilty. In this context, which error would be more serious? Why? In typical psychological research, we set $\alpha = .05$. What level do you think α is set to for most criminal trials in this country? Why? [10 pts]

A Type I Error would be concluding that OJ was guilty, when in fact he was innocent. A Type II Error would be concluding that OJ was innocent, when in fact he was guilty. Because we generally think that it's a very serious error to falsely imprison (or execute) an innocent person, a Type I Error would be thought of as more serious. On that basis, courtroom decisions are likely to have a probability of Type I Error set at a "probability" that is much smaller than .05 (one would hope .001 or so), because we would not want to tolerate falsely convicting 5 people out of 100 on trial.

4. The average age for registered voters in the county is $\mu = 39.7$ years with $\sigma = 11.8$. The distribution of ages is approximately normal. During a recent jury trial in the county courthouse, a statistician noted that the average age for the 12 jurors was $\overline{X} = 50.4$ years. [10 pts]
a. How likely is it to obtain a jury this old or older by chance?

$$\sigma_{\overline{X}} = \sigma/\sqrt{n} = 11.8/\sqrt{12} = 3.41$$

$$z = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{50.4 - 39.7}{3.41} = 3.14$$

The probability of a z of 3.14 or greater is .0008.

b. Is it reasonable to conclude that this jury is not a random sample of registered voters?
Yes, it is extremely unlikely to draw a sample of 12 from a population with $\mu = 39.7$ and have the sample mean turn out to be 50.4,. (Of course, it is possible to get such a sample from that population, just very unlikely! So, you could be making a Type I Error.)

5. As we saw in class, gestation periods are normally distributed with $\mu = 268$ days and $\sigma = 16$ days. Suppose that 16 women became pregnant. How likely is it that the mean gestation period for this sample would fall between 276 and 284 days? [5 pts]

$$\sigma_{\overline{X}} = \sigma/\sqrt{n} = 16/\sqrt{16} = 4$$

$$z = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{276 - 268}{4} = 2$$

$$z = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{284 - 268}{4} = 4$$

The proportion of scores greater than or equal to z=2 would be .0228. The proportion of scores greater than or equal to z=4 would be .00003. Thus, the proportion between z=2 and z=4 would be .02277, or roughly 2.3%.

6. A normal distribution has a mean of 120 and a standard deviation of 20. For this distribution, [5 pts]

a. What score separates the top 40% (highest scores) from the rest?

.40 in Column C yields a z = .25, which would mean a score = 125.

b. Scores between 60 and 100 make up what percentage of the distribution?

$$z = \frac{X - \mu}{\sigma} = \frac{60 - 120}{20} = -3$$
$$z = \frac{X - \mu}{\sigma} = \frac{100 - 120}{20} = -1$$

The proportion of scores less than z = -1 would be .1587. The proportion of scores less than z = -3 would be .0013. Thus, the proportion of scores between 60 and 100 would be .1587 - .0013 = .1574.

c. What range of scores would form the middle 60% of this distribution?

The middle 60% of the distribution means that you're eliminating the extreme 20% of the scores in each end, so you'd look up .20 in Column C. That yields z = 0.84, so –0.84 and +0.84 would cut off the lower and upper 20% of the distribution (leaving the middle 60%). Thus, the values that would cut off the middle 60% of the distribution would be 103.2 and 136.8.