Stat 1: Confidence Intervals and Significance Tests - Solutions to Practice Problems

1. Suppose the standard deviation of the heights of Swarthmore women is $\sigma = 2.8$ inches, and that the mean height μ is unknown.

   a) If heights follow a Normal distribution, what proportion of Swarthmore women's heights would be within one inch of μ?

   Given that the standard deviation for individual heights is assumed to be 2.8, a difference of 1 inch corresponds to a difference of $1.0/2.8 = 0.357 \approx 0.4$ standard deviations. From table B we can see that the percent within plus or minus 0.4 standard deviations of the mean is about $65.54 - 34.46 \approx 31\%$.

   b) Imagine you could draw a simple random sample (SRS) of $n = 16$ from the population of Swarthmore women and record the average height $\bar{X}$. Even if the distribution of heights is not Normal, the **central limit theorem** (CLT) implies that the sampling distribution for $\bar{X}$ will be close to Normal. What is the approximate distribution of $\bar{X}$?

   The average value $\bar{X}$ would vary according to a Normal distribution with mean μ (still unknown) and standard deviation $\sigma_{\bar{x}} = 2.8/\sqrt{16} = 0.7$ inches. That is, we expect the average of $n = 16$ values to vary with the same mean, but with a standard deviation that is smaller by a factor of $\sqrt{16} = 4$ than the overall standard deviation of heights.

   c) If you were to generate many averages of $n = 16$ heights, what proportion would be within one inch of μ?

   For averages of $n = 16$, a deviation of 1 inch corresponds to $1/0.7 = 1.428 \approx 1.4$ standard deviations (this is four times larger than the z value of 0.357 in part a). From Table B we can see that the percent within 1.4 standard deviations of the mean is about $91.92 - 8.08 \approx 84\%$.

   d) One sample (possibly representative) of $n = 16$ Swarthmore women reported an average height of $\bar{x} = 65.0$ inches. Treating this as a SRS, construct 68%, 95% and 99% CI's for the mean height of all Swarthmore women.

   For 68% and 95% confidence intervals you can use a margins of error of $z^* = 1$ and $z^* = 2$ standard deviations of $\bar{X}$ (based on the 68/95/99.7 rule). For a 99% interval, you could refer to Table C for the 0.995 quantile (to leave off half of 0.01 in each direction) and find the standard score $z^* = 2.576$, or approximate this by $z^* = 2.6$ using Table A (Percentile= 99.53).

   $$68\% : \quad \bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 65 \pm 1.0(2.8/\sqrt{16}) = 65 \pm 0.7 = (64.3, 65.7).$$

   $$95\% : \quad 65 \pm 2.0(0.7) = 65 \pm 1.4 = (63.6, 66.4) \qquad ((63.63.66.37) \text{ if } z^* = 1.96.)$$

   $$99\% : \quad 65 \pm 2.576(0.7) = 65 \pm 1.4 = (63.2, 66.8).$$

   e) The CDC reports that the mean height of college age women is about $\mu = 64$ inches. Carry out a significance test of whether Swarthmore women differ from the general population.

   The null hypothesis is $H_o : \mu = 64$ and the alternative hypothesis is $H_a : \mu \neq 64$. Under H_o, the sampling distribution for $\bar{X}$ is Normal with mean 64 inches and standard deviation

$2.8/\sqrt{16} = 0.7$ inches (still assuming that $\sigma = 2.8$ inches for heights of Swarthmore women). The evidence for H_a from the sample is the fact that the sample mean is $\bar{x} = 65.0$, which is 1.0 inches larger than the hypothesized mean.

The 2-sided p-value for the test is the probability of getting a value this far from the mean or further, in either direction. It turns out that 0.5 inches is $1.0/0.7 \approx 1.4$ standard deviations for an average of $n = 16$ values . The percentage of Normal observations within 1.4 standard deviations of the mean is $91.92 - 8.08 = 83.84\%$. So the probability of getting a value further from the mean is the 2-sided p-value $= 1 - .8384 = 0.1616$. You could also just note that the probability in either tail is 0.0808 and double this to get a 2-sided p-value of $2(.0808) = 0.1616$.

This is a pretty large p-value, meaning that it wouldn't be surprising to get an average of $n = 16$ heights this far from $\mu = 64$ if this were the true mean. So we do **not** have significant evidence to reject H_o.

2. The mean outcome for one roll of a fair die is $\mu = 3.5$ and the standard deviation is $\sigma = 1.708$.

   a) State hypotheses for testing whether or not a die is fair.

$$H_o: \ \mu = 3.5 \quad \text{vs.} \quad H_a: \ \mu \neq 3.5.$$

   b) If you roll a die $n = 30$ times, what values of $\bar{x}$ would constitute significant evidence at $\alpha = 0.05$ level that the die is biased?

   The standard deviation for the average of $n = 30$ rolls is $\sigma_{\bar{x}} = 1.708/\sqrt{30} \approx 0.3$. To have significance evidence at $\alpha = 0.05$, we must see an average more than 2 (or 1.96, to be more precise) standard deviations from 3.5. This corresponds to average values more than $3.5 + 2(0.3) = 4.1$ or less than $3.5 - 2(0.3) = 2.9$.

   c) A die is rolled 30 times and the average roll is $\bar{x} = 2.67$. Compute the 2-sided p-value for the test. Is the result significant at $\alpha = 0.05$? at $\alpha = 0.01$? What would you conclude?

   We can tell from part b that this result would be significant evidence at $\alpha = 0.05$ (because $2.67 < 2.9$). The p-value for the test is the probability of getting an average roll further from the mean value than $z = (2.67 - 3.5)/0.3 \approx -2.8$ standard deviations. The CLT tells us we may refer to the Normal table, which shows the percent within 2.8 standard deviations is $99.74 - 0.26 = 99.48\%$. The probability of getting a value further away than this is the 2-sided p-value $= 1 - .9948 = 0.0052$. You could also see that 0.26% are below -2.8 and double this to get 0.52%, or a p-value of 0.0052. So this result is significant at $\alpha = 0.01$, and any larger value.

   d) Construct a 99% CI for μ. How many rolls would it take to get a margin of error less than 0.5?

   A 99% CI for μ for this die is

$$2.67 \ \pm \ 2.576(1.708/\sqrt{30}) \ = \ 2.67 \pm 0.80 \ = \ (1.87, 3.47).$$

   Notice that the H_o value $\mu = 3.5$ is not in the 99% CI, which is consistent with there being significant evidence at $\alpha = 0.01$ (p-value ≤ 0.01) that the mean is not $\mu = 3.5$. To get a margin of error less than 0.5, we would need $2.576(1.708/\sqrt{n}) < 0.5$. This would require at least $n = 78$ rolls.