STATISTICS 101 - Homework 6 Answers

Due Wednesday, March 26, 2008

This assignment is worth a total of 75 points.

1. (31 pts) The maker of M&M's says on its website that 14% of Milk Chocolate M&M's are yellow. Suppose that M&Ms are packaged at random. We wish to examine the sample proportion of yellow Milk Chocolate M&M's, $\hat{p}$, in various sized bags.

   (a) (20 pts) For each of the different sized bags, give the mean and standard deviation of the sampling distribution of $\hat{p}$. Also comment on whether or not the success/failure condition is met for the sampling distribution to be approximately normal.

   For all of these parts $p = 0.14$ and the sample size n is the number of M&M's in the bag.

   i. (5 pts) Fun size bags containing 25 Milk Chocolate M&M's.
   The mean of the sampling distribution will be $p = 0.14$.
   The standard deviation of the sampling distribution will be $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.14(0.86)}{25}} = 0.0694$.
   The success/failure condition is not met ($np = 25(0.14) = 3.5$ and $n(1-p) = 21.5$) and so the sampling distribution is not approximately normal.

   ii. (5 pts) Small size bags containing 50 Milk Chocolate M&M's.
   The mean of the sampling distribution will be $p = 0.14$.
   The standard deviation of the sampling distribution will be $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.14(0.86)}{50}} = 0.0491$.
   The success/failure condition is not met ($np = 50(0.14) = 7$ and $n(1-p) = 43$) and so the sampling distribution is not approximately normal.

   iii. (5 pts) Medium size bags containing 100 Milk Chocolate M&M's.
   The mean of the sampling distribution will be $p = 0.14$.
   The standard deviation of the sampling distribution will be $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.14(0.86)}{100}} = 0.0347$.
   The success/failure condition is met ($np = 100(0.14) = 14$ and $n(1-p) = 86$) and so the sampling distribution is approximately normal.

   iv. (5 pts) Extra Large size bags containing 500 Milk Chocolate M&M's.
   The mean of the sampling distribution will be $p = 0.14$.
   The standard deviation of the sampling distribution will be $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.14(0.86)}{500}} = 0.0155$.
   The success/failure condition is met ($np = 500(0.14) = 70$ and $n(1-p) = 430$) and so the sampling distribution is approximately normal.

   (b) (6 pts; 2 pts for each) For the extra large bags containing 500 Milk Chocolate M&M's, use the 68-95-99.7 Rule to describe how the sample proportion of yellow Milk Chocolate M&M's might vary from bag to bag.

   The sampling distribution has an approximately normal distribution with mean 0.14 and standard deviation 0.0155. This means that

   Approx. 68% of all samples of size 500 will have a $\hat{p}$ value between $0.14 \pm 0.0155 = (0.1245, 0.1555)$
   Approx. 95% of all samples of size 500 will have a $\hat{p}$ value between $0.14 \pm 2 * 0.0155 = (0.1090, 0.1710)$
   Approx. 99.7% of all samples of size 500 will have a $\hat{p}$ value between $0.14 \pm 3 * 0.0155 = (0.0935, 0.1865)$

   (c) (5 pts) In an extra large bag of 500 Milk Chocolate M&M's there are 90 yellow. In this an unusually large proportion of yellow? Explain your answer.

   90 yellow out of 500 total makes a $\hat{p} = 90/500 = 0.18$. Under the sampling distribution from part (b), we would expect to have approximately 95% of all samples of size 500 to have a $\hat{p}$ value between 0.1090 and 0.1710 and approximately 99.7% of all samples of size 500 to have a $\hat{p}$ value between 0.0935 and 0.1865. Since 0.18 is between the upper bounds for 95% and 99.7%, this seems to be an unusually large number of yellow M&Ms.

Students could also answer this question by calculating a probability. If they choose to do this, the probability calculation is:

$$\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.18-0.14}{\sqrt{\frac{0.14(0.86)}{500}}} = \frac{0.04}{0.0155} = 2.5806$$

The probability of being greater than this value is found using the normal z-table and is the table value for z = 2.58 subtracted from 1. This value is $1 - 0.9951 = 0.0049$.

Again, this seems to be an unusually large number of yellow M&Ms.

2. (12 pts; 2 pts for the success/failure condition, 1 pt for the normal, 1 pt for the mean, 2 pts for the s.d., and 6 pts for the 3 intervals.) It is believed that 42% of all college students in the United States engage in binge drinking (5 or more drinks at a sitting for men, 4 or more for women). Consider a random sample of 100 college students. Verify that the success/failure condition is met. Use the 68-95-99.7 Rule to describe the sampling distribution model for the sample proportion of students who engage in binge drinking.

   The success/failure condition is met. ($np = 100(0.42) = 42$ and $n(1-p) = 58$). The sampling distribution is therefore approximately normal with mean $p = 0.42$ and standard deviation $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.42(0.58)}{100}} = 0.0494$. Using the 68-95-99.7 rule for normal distributions we would expect

   Approx. 68% of all samples of size 100 will have a $\hat{p}$ value between $0.42 \pm 0.0494 = (0.3706, 0.4694)$
   Approx. 95% of all samples of size 100 will have a $\hat{p}$ value between $0.42 \pm 2 * 0.0494 = (0.3212, 0.5188)$
   Approx. 99.7% of all samples of size 100 will have a $\hat{p}$ value between $0.42 \pm 3 * 0.0494 = (0.2718, 0.5682)$

3. (14 pts) In 2002, 22.5% of all adults (18 years or older) in the United States were current smokers.

   (a) (2 pts) For a random sample of 1000 U.S. adults, is the 10% condition met? Explain your answer.
   As long as the population size is 10000 adults or more, the 10% condition is met. This is true so the condition is met.

   (b) (2 pts) Is the success/failure condition met? Explain your answer.
   Yes, $np = 1000(0.225) = 225$ and $n(1-p) = 775$.

   (c) (10 pts; 1 pt for the normal, 1 pt for the mean, 2 pts for the s.d. and 6 pts for the 3 intervals) Use the 68-95-99.7 Rule to describe the sampling distribution model for the proportion of current smokers in a random sample of 1000 adults in the United States in 2002.
   Since the two conditions are met, the sampling distribution of the sample proportion statistic is approximately normal with mean $p = 0.225$ and standard deviation $\sqrt{\frac{0.225(0.775)}{1000}} = 0.0132$. Using the 68-95-99.7 rule for normal distributions we would expect
   Approx. 68% of all samples of size 1000 will have a $\hat{p}$ value between $0.225 \pm 0.0132 = (0.2118, 0.2382)$
   Approx. 95% of all samples of size 1000 will have a $\hat{p}$ value between $0.225 \pm 2 * 0.0132 = (0.1986, 0.2514)$
   Approx. 99.7% of all samples of size 1000 will have a $\hat{p}$ value between $0.225 \pm 3 * 0.0132 = (0.1854, 0.2646)$

4. (13 pts; 4 pts for verifying conditions, 4 pts for the sampling distribution, 5 pts for the probability) A seed corn distributor advertises a germination rate of 90% for its corn. What is the probability that out of 400 randomly selected corn seeds from the distributor, fewer than 350 will germinate? As a part of your answer, verify that the appropriate conditions for computing this probability are met.

   We must first show the success/failure condition and the 10% condition are met. For the success/failure condition, $np = 400 * 0.90 = 360$ and $n(1-p) = 40$. So this condition is met. For the 10% condition to be met, we must assume that the population of seeds produced by this corn distributor is more than 4000. This seems a very reasonable assumption and so the 10% condition is met.

Since both conditions are met, the sampling distribution of the sample proportion statistic will be approximately normal with mean $p = 0.9$ and standard deviation $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.9(0.1)}{400}} = 0.015$. Having fewer than 350 seeds germinate is equivalent to having a sample proportion $\hat{p} = 350/400 = 0.875$ or less.

Standardizing the value 0.875 gives $z = \frac{0.875-0.9}{0.015} = -1.67$. The probability of being less than this value on the z-table is 0.0475.

So chances are pretty small (around 0.0475 or 4.75%) that you would get 350 seeds or less to germinate if the true germination rate of the corn was 90%.

5. (5 pts) In Lab #7, you took a sample of size 5 and a sample of size 10 from a population of 250 students. In this lab, the proportion of students in the population with blue eyes was 0.312. Suppose you now have a very large population (so large so that meeting the 10% condition is not a problem) with the same proportion of blue eyes in the population (0.312). What is the smallest sample size you can take from this very large population for the sampling distribution of the sample proportion of students with blue eyes to be approximately normal? Explain your answer.

   Since the 10% condition is met, we need to find a value of n that will satisfy the success/failure condition for $p = 0.312$. As long as $np \geq 10$, the other condition $n(1-p) \geq 10$ will hold. So we just need to find the value of n that will satisfy $np \geq 10$. Since $p = 0.312$, $n \geq 10/0.312 = 32.05$. n is a sample size and must be a whole number, so $n = 33$.