SG 7.2 AP Statistics Name:

1. According to the 2000 U.S. Census, 80% of Americans over the age of 25 have earned a high school diploma. Suppose we take a random sample of 120 Americans and record the proportion, $\hat{p}$, of individuals in our sample that have a high school diploma.

(a) What are the mean and standard deviation of the sampling distribution of $\hat{p}$?

(b) What is the approximate shape of the sampling distribution? Justify your answer.

(c) Suppose our sample size was 30 instead of 120. Compare the shape, center, and spread of this sampling distribution to the one in parts (a) and (b).

(d) You live in a small town with only 500 residents over the age of 25. What is the largest possible sample you can take from your town and still be able to calculate the standard deviation of sampling distribution of $\hat{p}$ using the method presented in the textbook? Explain.

2. George is a big fan of music from the 1960s, and 22% of the songs on his mp3 player are Beatles songs. Suppose George sets his mp3 player to "shuffle," so that it selects songs randomly (assume the shuffle function permits repetition of songs). During a long drive, George plays 50 randomly-selected songs.

(a) What are the mean and standard deviation of the proportion of the 50 randomly-selected songs that are Beatles songs?

(b) Calculate the probability that more than 30% of the 50 randomly-selected songs are Beatles songs.

7.2 SG Answers

$$\sigma_{\hat{p}} = \frac{1}{n}\sqrt{np(1-p)} = \sqrt{\frac{np(1-p)}{n^2}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.8(0.2)}{120}}$$

1.(a) $\mu_{\hat{p}} = 0.8$ $\sigma_{\hat{p}} = 0.0365$

b. np=(120)(0.8)=96 and n(1-p)=(120)(0.2)=24 both numbers are greater than 10 thus the distribution is approximately normal.

c. $\mu_{\hat{p}}$ would not change, $\sigma_{\hat{p}}$ or the spread, would be larger (0.0756) because the sample size was reduced, and the distribution would be non-Normal since np=(30)(0.8)=24 and n(1-p)=6, which is less than 10.

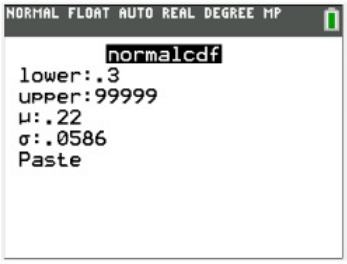
d. The largest sample we can take is 50, otherwise the sample would be more than 10% of the population, and sampling without replacement would require a finite population correction to calculate standard deviation.

2.(a) $\mu_{\hat{p}} = 0.22$ $\sigma_{\hat{p}} = 0.0586$

b.

$$P\left(z > \frac{0.3-0.22}{\sqrt{\frac{(0.22)(0.78)}{50}}}\right) = P\left(z > \frac{0.08}{0.0586}\right) = 1.365$$

Then use z-table = 0.9131 and 1-0.9131 (because >) = 0.0869.



or on your calculator

The probability that more than 30% of the 50 randomly-selcected songs are Beatles songs is 0.0869.

$P(\hat{p} > .3) = P\left(\hat{p} > \frac{3-.22}{.05858}\right)$

$P(z > 1.37) = .9147$ 0.0853

1 - 0.9147 =

or normcdf(.3, 100..., .22, .05858)

This is calc speak. Do NOT write your answer like this on your quiz/test.

You need to say, "Using statistical software, the function normal cdf with the low value of 0.3, the high value as 100, the mean as 0.22, and standard deviation as 0.05858 results in a probabily of 0.0869. This means the probability that more than 30% of the 50 randomly-selcected songs are Beatles songs is 0.0869.

(Don't forget context!!!!)