Sampling Distribution of $\hat{p}$-hat

1. In a survey, 600 mothers and fathers were asked about the importance of sports for boys and girls. Of the parents interviewed, 70% said the genders are equal and should have equal opportunities to participate in sports.

   1. What are the mean, standard deviation, and shape of the distribution of the sample proportion $\hat{p}$ of parents who say the genders are equal and should have equal opportunities? Be sure to justify your answer for the shape of the distribution. Use $n = 600$.

 $$n = 600, \; p = 0.7$$

      $$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = 0.0187$$

      The distribution is normal since both mean and standard deviation $n(1-p)$ are more than 10. Therefore, the distribution is normal, centered around 0.7 with a standard deviation of 0.0187.

   2. Using the normal approximation without the continuity correction, sketch the probability distribution curve for the distribution of $\hat{p}$. Shade equal areas on both sides of the mean to show an area the represents a probability of .95, and label the upper and lower bounds of the shaded area as values of p-hat. Show your calculations for the upper and lower bounds.

      $$0.7 - (2 \cdot 0.0187) = 0.663$$

      $$0.7 + (2 \cdot 0.0187) = 0.737$$
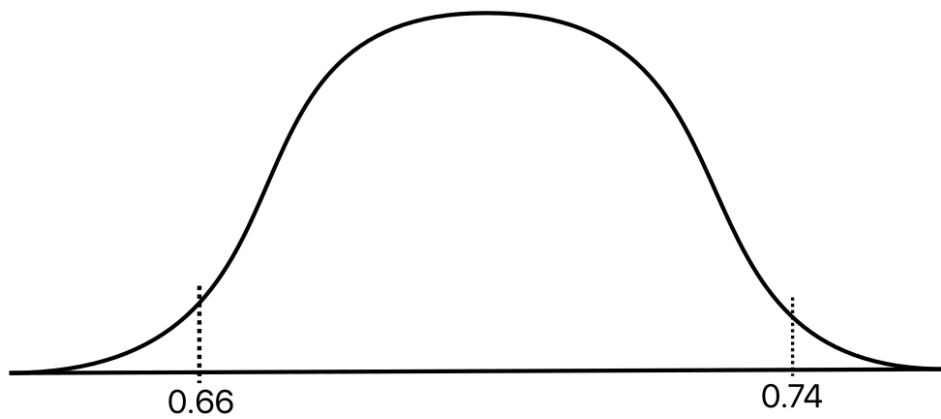
3. Considering the sketch in part B, the shaded area shows a .95 probability of what happening? In other words, what does the probability of 0.95 represent?

   The 0.95 represents the chance that the sample proportion will fall in between 0.65 and 0.74.

4. Using the normal approximation, what's the probability a randomly drawn sample of parents of size 600 will have a sample proportion between 67% and 73%? Draw a sketch of the probability curve, shade the area representing the probability you're finding, and label the z-scores that represent the upper and lower bounds of the probability you're finding. Don't use the continuity correction.

   The z-scores of the bounds are -1.604 and 1.604. Using our calculator with the command

   ```
   normcdf(0.67,0.73,0.7,0.0187)
   ```

   we find that the probability is $\boxed{0.8913}$.
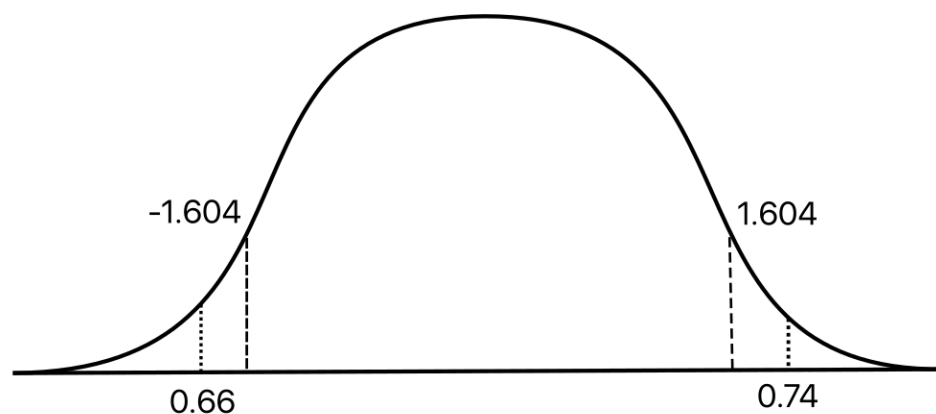
5. Now, use the exact binomial calculation to find the probability of getting between, but not including, 67% and 73% of the respondents in a sample of 600 who say the genders are equal and should have equal opportunities. To use the exact binomial, you'll need to convert the proportions to counts by multiplying each proportion by 600.

$$0.67 \cdot 600 = 402, \ 0.73 \cdot 600 = 438$$

```
binomcdf(600,0.7,437) - binomcdf(600,0.7,402) = 0.88
```

6. Now try it again, but this time find the probability of getting at least 67% but no more than 73%. Use the exact binomial calculation.

```
binomcdf(600,0.7,438) - binomcdf(600,0.7,402) = 0.901
```

2. Random samples of size n = 500 were selected from a binomial population with p = .1.

1. Is it appropriate in this case to use the normal distribution to approximate the distribution of $\hat{p}$? What are the mean, standard deviation, and shape of the distribution?

   It is appropriate, since both np and $n(1-p)$ are greater than 10.

   The mean is the same as the mean of the population. $\boxed{0.1}$

   $$\sigma_{\hat{p}} - \sqrt{\frac{0.09}{500}} = 0.013$$

   The shape is approximately normal with a mean of 0.1 and a standard deviation of 0.013.

2. Using the normal approximation without the continuity correction, find the probability that $\hat{p} < .12$.

```
normcdf(∞,0.12,0.1,0.0134) = 0.93
```

3. Using the normal approximation, find $P(\hat{p} < .12)$. This time use the continuity correction.

```
normcdf(∞,60.5,50,6.71) = 0.9412
```

4. Using the exact binomial calculation, find $P(\hat{p} < .12)$. Compare your answer with the answers you got for parts B and C. Why are they different?

```
binomcdf(500,0.1,60) - binomcdf(500,0.1,0) = 0.938
```

The answers are different because the first two ways we found the probability of $\hat{p} < 12$ were approximations.

3. One of the ways Americans relieve stress is to reward themselves with sweets. Suppose a study claims 52% of Americans admit to overeating sweets when stressed. Suppose also that the 52% figure is correct for the population and that random samples of size $n = 100$ Americans are selected.

1. Does the distribution of $\hat{p}$ have an approximately normal distribution? If so, what are its mean and standard deviation?

It does have an approximately normal distribution, since $np, n(1-p) > 10$.

$$\mu_{\hat{p}} = p = 0.52$$

$$\sigma_{\hat{p}} = \sqrt{\frac{0.2496}{100}} = 0.05$$

2. Using the normal approximation of p ˆ without the continuity correction, what's the probability of getting a sample ($n = 100$) with $\hat{p}$ greater than .6?

```
normcdf(0.6,∞,0.52,0.05) = 0.055
```

3. Using the normal approximation of the binomial distribution with the continuity correction, what's the probability of getting a sample ($n = 100$) with $\hat{p}$ greater than .6?

```
normcdf(60.5,∞,52,4.996) = 0.044
```

4. Using the exact binomial calculation, what's the probability of getting sample ($n = 100$) with $\hat{p}$ greater than .6?

```
binomcdf(100,0.52,100)-binomcdf(100,0.52,61)
```

4. In 1996 there was a battle in the courts and in the marketplace
   between Intel and Digital Equipment Corp. about the technology
   behind Intel's Pentium microprocessing chip. Digital accused
   Intel of willful infringement on Digital's patents. Although
   Digital's Alpha microprocessor chip was the fastest on the
   market at the time, its speed fell victim to Intel's marketing
   clout. That same year, Intel shipped 76% of the microprocessor
   market. Suppose a random sample of $n = 1,000$ personal computer
   (PC) sales is monitored and the type of microprocessor installed
   is recorded. Let $\hat{p}$ be the proportion of personal computers in
   the sample with a Pentium microprocessor.

   1. What are the mean, standard deviation, and shape of the
      distribution of $\hat{p}$?

      $$\mu_{\hat{p}} = 0.76$$

      $$\text{SD} = \sqrt{\frac{0.1824}{1000}} = 0.0135$$

      It is a normal distribution centered around 0.76 with
      standard deviation of 0.0135.

   2. Using the normal approximation without the continuity
      correction, what's the probability you'd draw a random
      sample of 1,000 PCs with a proportion of Pentium chips
      exceeding 80%?

      ```
      normcdf(0.8,∞,0.76,0.0135) = 0.0015
      ```

   3. Looking at the answer you got for part B, if you got a
      simple random sample with $\hat{p} > .8$, would you conclude the
      population proportion may be higher than .76? Why?

      Yes, i would, since a 0.1% chance is really really low. This is
      also the number without the continuity correction.