Section R

The Central Limit Theorem for Proportions

Since we can also estimate and draw conclusions about the population proportion, we need to know the sampling distribution of the sample proportion; since the sample proportion will be used to estimate the population proportion.

Suppose that a random sample of size n is obtained from a population in which each individual either does or does not have a certain characteristic. The **sample proportion**, denoted $\hat{p}$, is given by $\hat{p} = \frac{x}{n}$, where x is the number of individuals in the sample with the specified characteristic. The sample proportion, $\hat{p}$, is a statistic that estimates the population proportion, p.

The Sampling Distribution of the Sample Proportion, $\hat{p}$

For a simple random sample of size n with a population proportion p.

- The shape of the sampling distribution of $\hat{p}$ is approximately normal provided $np \geq 10$ and $n(1-p) \geq 10$.
- The mean of the sampling distribution of $\hat{p}$ is $\mu_{\hat{p}} = p$
- The standard deviation of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.

Therefore,

we can say, $\hat{p}$ is normally distributed with parameters $\mu_{\hat{p}}$ and $\sigma_{\hat{p}}$, where $\mu_{\hat{p}} = p$ and $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.

Note: Since the sampling distribution of the sample proportion is approximately normal under certain conditions you can use the normal approximation to find probabilities, therefore you need to convert $\hat{p}$ to a z-score.

Converting $\hat{p}$ to a z-score: $Z = \dfrac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}}$

Examples:

1) The National Center for Educational Statistics reported that 60% of full-time, first-time students at 4-year institutions in 2010 who were seeking a bachelor's or equivalent degree completed a bachelor's or equivalent degree within 6 years at the institution where they began their studies. A random sample of 165 of full-time, first-time students at 4-year institutions is selected.

a) What is the sampling distribution of the proportion of full-time, first time students at 4-year institutions? (i.e. What is the sampling distribution of $\hat{p}$?)

n = 165

p = .6

$np = (165)(.6) = 99 > 10$ ✓

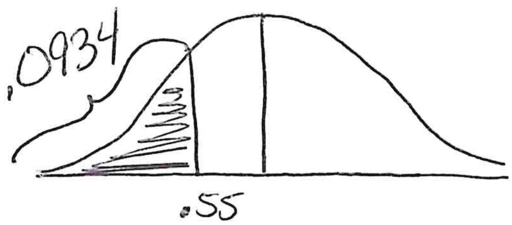
$n(1-p) = (165)(1-.6) = 66 > 10$ ✓

so $\hat{p}$ is normally distributed

$\mu_{\hat{p}} = .6$

$\sigma_{\hat{p}} = \sqrt{\frac{.6(1-.6)}{165}} = .038$

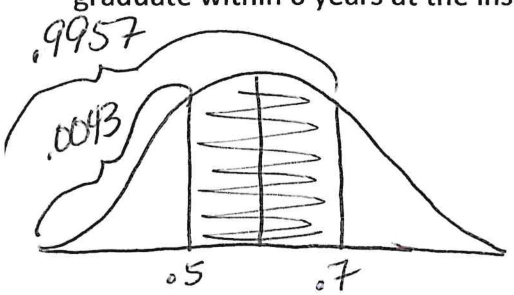
b) Find the probability that less than 55% of full-time, first-time students at 4-year institutions graduate within 6 years at the institution where they began their studies.



$P(\hat{p} < .55) = \underline{.0934}$

$Z = \frac{.55 - .6}{.038} = -1.32 \Rightarrow .0934$

c) Find the probability that between 50% and 70% of full-time, first-time students at 4-year institutions graduate within 6 years at the institution where they began their studies.

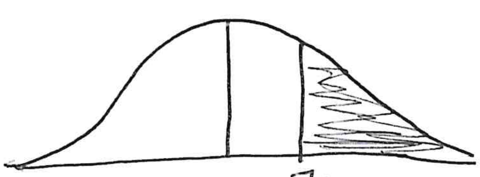


$P(.5 < \hat{p} < .7) = .9957 - .0043 = \underline{.9914}$

$Z = \frac{.5 - .6}{.038} = -2.63 \Rightarrow .0043$

$Z = \frac{.7 - .6}{.038} = 2.63 \Rightarrow .9957$

d) Find the probability that more than 70% of full-time, first-time students at 4-year institutions graduate within 6 years at the institution where they began their studies. Would this be unusual? Explain.



$P(\hat{p} > .7) = 1 - .9957 = \underline{.0043}$

yes, because .0043 < .05

$Z = \frac{.7 - .6}{.038} = 2.63 \Rightarrow .9957$

2) The Bureau of Labor Statistics reported in a recent year that 5% of employed adults in the United States held multiple jobs. A random sample of 350 employed adults is chosen.

a) What is the sampling distribution of the proportion of employed adults in the United States who hold multiple jobs? (i.e. What is the sampling distribution of $\hat{p}$?)

n = 350

p = .05

$np = 350(.05) = 17.5 > 10$ ✓

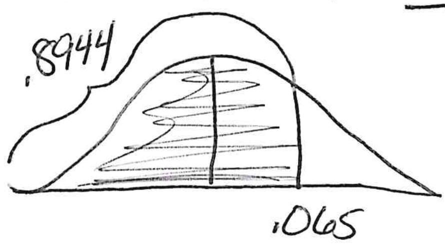
$n(1-p) = 350(1-.05) = 332.5 > 10$ ✓

so $\hat{p}$ is normally distributed

$\mu_{\hat{p}} = .05$

$\sigma_{\hat{p}} = \sqrt{\frac{.05(1-.05)}{350}} = .012$

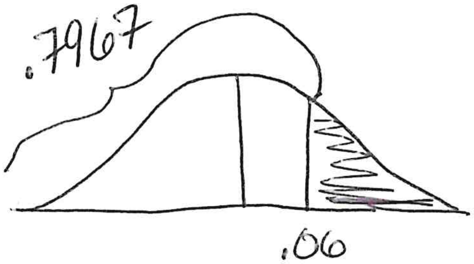
b) Find the probability that less than 6.5% of the individuals in the sample hold multiple jobs.



$P(\hat{p} < .065) = .8944$

$Z = \frac{.065 - .05}{.012} = 1.25 \Rightarrow .8944$

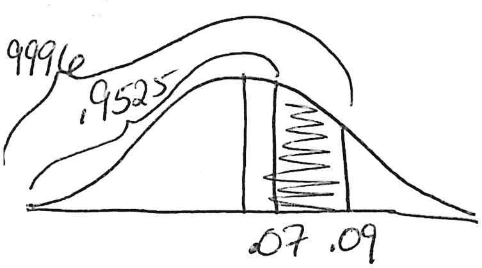
c) Find the probability that more than 6% of the individuals in the sample hold multiple jobs.



$P(\hat{p} > .06) = 1 - .7967 = .2033$

$Z = \frac{.06 - .05}{.012} = .83 \Rightarrow .7967$

d) Find the probability that the proportion of individuals in the sample who hold multiple jobs is between 0.07 and 0.09.

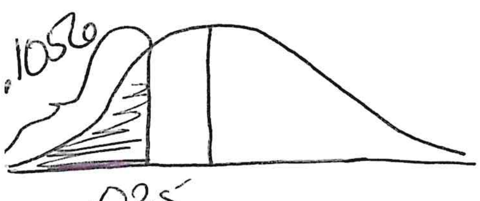


$P(.07 < \hat{p} < .09) = .9996 - .9525$
$= .0471$

$Z = \frac{.07 - .05}{.012} = 1.67 \Rightarrow .9525$

$Z = \frac{.09 - .05}{.012} = 3.33 \Rightarrow .9996$

e) Would it be unusual if less than 3.5% of the individuals in the sample who held multiple jobs?



$P(\hat{p} < .035) = .1056$

No, because $.1056 > .05$

$Z = \frac{.035 - .05}{.012} = -1.25 \Rightarrow .1056$

3) A Pew Research report indicated that 73% of teenagers aged 12 – 17 have access to smartphones. A random sample of 150 teenagers is drawn.

a) What is the sampling distribution of the proportion of teenagers aged 12 – 17 have access to smartphones? (i.e. What is the sampling distribution of $\hat{p}$?)

n = 150

p = .73

$np = 150(.73) = 109.5 > 10$ ✓

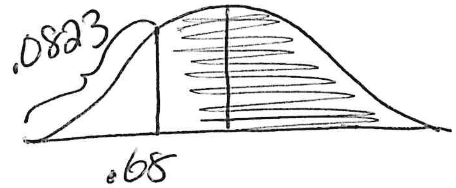
$n(1-p) = 150(1-.73) = 40.5 > 10$ ✓

So $\hat{p}$ is normally distributed

$\mu_{\hat{p}} = .73$

$\sigma_{\hat{p}} = \sqrt{\frac{.73(1-.73)}{150}} = .036$

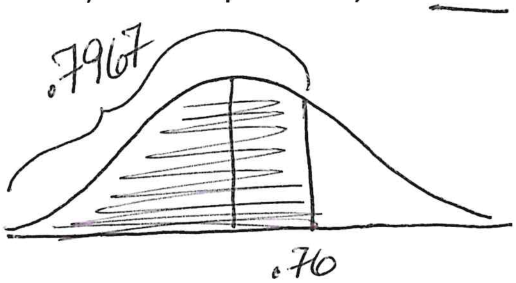
b) Find the probability that more than 68% of the sampled teenagers have access to smartphones.



$P(\hat{p} > .68) = 1 - .0823 = \underline{.9177}$

$Z = \frac{.68 - .73}{.036} = -1.39 \Rightarrow .0823$

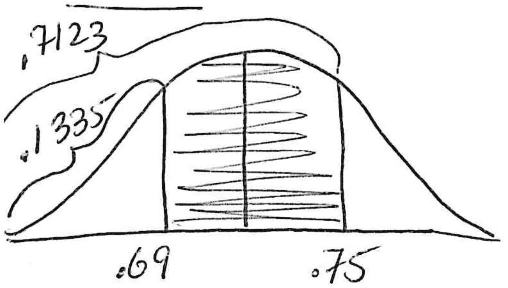
c) Find the probability that less than 76% of the sampled teenagers have access to smartphones.



$P(\hat{p} < .76) = \underline{.7967}$

$Z = \frac{.76 - .73}{.036} = .83 \Rightarrow .7967$

d) Find the probability that the proportion of the sampled teenagers who have access to smartphones is between 0.69 and 0.75.

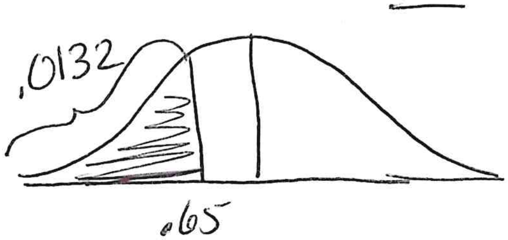


$P(.69 < \hat{p} < .75) = .7123 - .1335 = \underline{.5788}$

$Z = \frac{.69 - .73}{.036} = -1.11 \Rightarrow .1335$

$Z = \frac{.75 - .73}{.036} = .56 \Rightarrow .7123$

e) Would it be unusual if less than 65% of the sampled teenagers have access to smartphones?



$P(\hat{p} < .65) = \underline{.0132}$

$Z = \frac{.65 - .73}{.036} = -2.22 \Rightarrow .0132$