Chapter 12 Inference For Proportions

12.1 Inference for a Population Proportion – Day 1

<u>OBJ</u>: You will answer questions about the proportion of some outcome in a population.

p - unknown proportion of a population
$\hat{p}$ - Estimate of parameter p (Sample prop.)

$$\hat{p} = \frac{\text{count of successes}}{\text{count of OBS.}} = \frac{x}{n}$$

Example 12.3

12.1 – 12.3

Sampling Distribution of $\hat{p}$ (Section 9.2):

$$\left.\text{Mean} = p \quad \text{Std. Dev} = \sqrt{\frac{p(1-p)}{n}}\right\} \text{ if } \begin{cases} \text{Pop} \geq 10n \\ np \geq 10,\ n(1-p) \geq 10 \\ \text{Dist. of } \hat{p} \text{ is} \approx \text{Normal} \end{cases}$$

$$\text{Standardize } \hat{p}: \left. z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \right\} N(0,1)$$

But we don't Know the value of p!!

$\therefore$ We can't check in $np \geq 10$ or $n(1-p) \geq 10$
So....

* To Test $H_0: p = p_0$ That the unknown p has some value p_0, Replace p with p_0.

$$\boxed{z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}}$$

<u>Conditions:</u>
SRS
Pop $\geq 10n$
$np_0 \geq 10$
$n(1-p_0) \geq 10$

Think: $\dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

$\Rightarrow$

* In a conf. Int. for p, we don't have a value to substitute (See p_o Above). $\hat{p}$ will be close to p in large samples. So replace p with $\hat{p}$ to determine np and $n(1-p)$. Replace the STD. Dev. by the Standard Error of $\hat{p}$.

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

$$CI: \boxed{\hat{p} \pm Z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

Conditions:

SRS

Pop $\geq 10n$

$n\hat{p} \geq 10$

$n(1-\hat{p}) \geq 10$

Think: $\bar{x} \pm Z^* \frac{\sigma}{\sqrt{n}}$

12.4

HW: 12.5, Examples (12.7 + 12.8) Tomorrow (To pg 696

12.1 Inference For A Pop. Prop. - Day 2

OBJ - You will find the sample size required for a desired margin of error.

But Before we Begin

12.7: Done w/ TI-84 Techniques!

Now - Margin of Error! A.K.A - What Sample Size is needed?

The margin of error in a C.I. for p is: $m = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Looking for n

m and z^* are Known

$\hat{p}$ is **not** Known - we will call our guess p^*.

How do we get p^*?

1) Use p^* based on a study or past experience w/ statistics. (It will be stated in the problem)

-OR-

2) Use $p^* = .5$ m is largest when $p^* = .5$, so we have a Conservative guess. If $\hat{p} \neq .5$, we get a margin of error smaller than planned.

$$z^* \sqrt{\frac{p^*(1-p^*)}{n}} \le m$$

General Rule:

If you expect $\hat{p}$ to be:

$.3 \le \hat{p} \le .7$, use $p^* = .5$

$\hat{p} < .3$ OR $\hat{p} > .7$, use a better guess from a pilot study

12.11

HW: 12.9, 12.18, $\underbrace{12.21}_{\text{use calc.}}$

Quiz Tomorrow!!

12.2 Comparing Two Proportions- Day 1

OBJ: You will use the conf. int. for comparing two proportions

Example 12.2 - pg 685

Notation:

Pop	Pop. Prop.	Sample Size	Sample Prop.
1	p_1	n_1	$\hat{p}_1$
2	p_2	n_2	$\hat{p}_2$

$\bar{x}$ Estimates μ

$\bar{x}_1 - \bar{x}_2$ Estimates $\mu_1 - \mu_2$

$\hat{p}$ Estimates p

$\hat{p}_1 - \hat{p}_2$ Estimates $p_1 - p_2$

STD. DEV of $P_1 - P_2 = \sqrt{\text{Variance}} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

SO...

STD. ERROR of $\hat{p}_1 - \hat{p}_2 = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

∴ Conf. Int. For Comparing Two Prop.:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Conditions:

Pop $\geq 10n_1$	Pop $\geq 10n_2$
$n_1\hat{p}_1 \geq 5$	$n_2\hat{p}_2 \geq 5$
$n_1(1-\hat{p}_1) \geq 5$	$n_2(1-\hat{p}_2) \geq 5$

2 Indep SRS.

12.24

HW: 12.22

12.2 Comparing Two Prop. - Day 2

<u>OBJ</u>: You will determine if there is a sig. Diff. in the prop. of two groups.

- To Test $H_0: p_1 = p_2$, Standardize $\hat{p}_1 - \hat{p}_2$ to get a z statistic.
- If H_0 is true, all the obs. in both samples really come from a single pop.
- So instead of estimating the SE $\hat{p}_1 - \hat{p}_2$ we pool the two samples and use the overall sample prop. to estimate the single pop. parameter p. (Pooled Sample Prop. $\hat{p}$).

Pooled Sample Prop $\hat{p} = \hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$

<u>Now</u> we use $\hat{p}$ in place of $\hat{p}_1$ and $\hat{p}_2$ For SE $\hat{p}_1 - \hat{p}_2$.

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \Rightarrow \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}}$$

<u>Conditions</u>: 2 Indep. SRS

$n_1\hat{p} \geq 5$	$n_2\hat{p} \geq 5$
$n_1(1-\hat{p}) \geq 5$	$n_2(1-\hat{p}) \geq 5$
Pop $\geq 10n_1$	Pop $\geq 10n_2$

12.25

<u>HW</u>: 12.26 (a, b)