# 13.2 Comparing Two Proportions

## Steps for a CI of $p_A - p_B$

① Name the procedure in context

Two sample z CI for $p_A - p_B$, where $p_A - p_B$ is the difference between the proportion of all ($p_A$) ______ and all ($p_B$) ______ that ______.

② Check Conditions

SI - two independent, random samples from the populations of interest stated in question. I $N \geq 10n$

N - Sampling distributions are ~N

$n_A\hat{p}_A$, $n_A(1-\hat{p}_A)$, $n_B\hat{p}_B$, $n(1-\hat{p}_B)$ are all $\geq 10$

③ Computation

$$(\hat{p}_A - \hat{p}_B) \pm z^* \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$

($z^*$ invNorm)

④ Conclusion

Based on these samples, I am __% confident...

Example 1 As part of the Pew Internet and American Life Project, researchers conducted two surveys in 2012. The first survey asked a random sample of 799 US teens about their use of social media. A second survey asked 2253 adults. In the studies, 80% of teens and 69% of adults used social networking sites.
Construct and interpret a 90% CI for the difference in the proportion of all teens and adults who use social networking sites.

① Two sample CI for $p_A - p_B$, where $p_A - p_B$ is the difference between the proportion of all US teens ($p_A$) and all US adults ($p_B$) that use social networking sites.

② Independent, random samples are stated.

$799(.80) = 639.2 \quad 799(.2) = 159.8$

$2253(.69) = 1554.57 \quad 2253(.31) = 698.43$

All values are $\geq 10$, so the sampling distributions are ~N.

③ $$(.89 - .69) \pm 1.645\sqrt{\frac{(.8)(.2)}{799} + \frac{(.69)(.31)}{2253}} = (.081, .138)$$

④ Based on these samples, I am 90% confident that the difference in the proportion of all U.S. teens and all U.S. adults who use social networking sites (teens - adults) is between .081 and .138.

Calculator check: 2PropZInt

13.2 Cont'd

Steps in a Significance Test for $p_A - p_B$

* Combined sample proportion $\hat{p}_c$

$$\frac{\text{Successes in both samples}}{\text{count in both samples}} = \frac{X_1 + X_2}{n_1 + n_2}$$

1. State the hypothesis: $H_0: p_A - p_B = 0$

   $H_a: p_A - p_B \lesseqgtr 0$
2. Name the procedure in context

   Two sample z-test for $p_A - p_B$ where $p_A - p_B$ is the difference in the proportion of all ($p_A$) ________ and all ($p_B$) ______ that _______

3. Check conditions

   S/I - Two independent, random samples from the populations of interest stated in the question. I - independence; $N \geq 10(n)$

   N - sampling distribution is $\sim N$

   $n_A \hat{p}_c$, $n_A(1-\hat{p}_c)$, $n_B \hat{p}_c$, $n_B(1-\hat{p}_c)$ are all $\geq 10$

4. Computation $z = \dfrac{\hat{p}_A - \hat{p}_B}{\sqrt{\dfrac{\hat{p}_c(1-\hat{p}_c)}{n_A} + \dfrac{\hat{p}_c(1-\hat{p}_c)}{n_B}}}$ find $p =$ ______

5. Conclusion: Because the p-value ______ is $\lessgtr$ ______ $\alpha =$ _____, I (reject/don't reject) ...

Example 2 Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At school 2, an SRS of 150 students included 26 who had not eaten breakfast. Do the data provide convincing statistical evidence that there is a difference in proportion of students at these two schools that don't eat breakfast.

① $H_0: p_1 - p_2 = 0$ $H_a: p_1 - p_2 \neq 0$

② Two sample z-test for $p_1 - p_2$ where $p_1 - p_2$ is the difference in the proportion of all students at School 1 ($p_1$) and all students at School 2 ($p_2$) who do not eat breakfast.

③ It is stated in the question the samples are independent and random. $p_c = \frac{19+26}{80+150} = .196$

$80(.196) = 15.68$ $80(1-.196) = 64.32$

$150(.196) = 29.4$ $150(1-.196) = 120.6$

Since all are $\geq 10$, the sampling distribution is $\sim N$

④ $$z = \frac{\frac{19}{80} - \frac{26}{150}}{\sqrt{\frac{.196(.804)}{80} + \frac{.196(.804)}{150}}} = \frac{.0642}{\sqrt{.0019698 + .00105056}} = 1.168$$

$\boxed{13.2}$ Cont'd

4. $z = 1.168$

$2P(z > 1.168) = 2(.1214)$
$= .243$

p-value = .243

5. Since the p-value = .243 > $\alpha$ = .05, we fail to reject $H_0$. There is not evidence to conclude that the proportions of all students at the two schools who don't eat breakfast are different.

Calculator check: 2Prop z-test