Comparing 2 proportions

BPS chapter 21

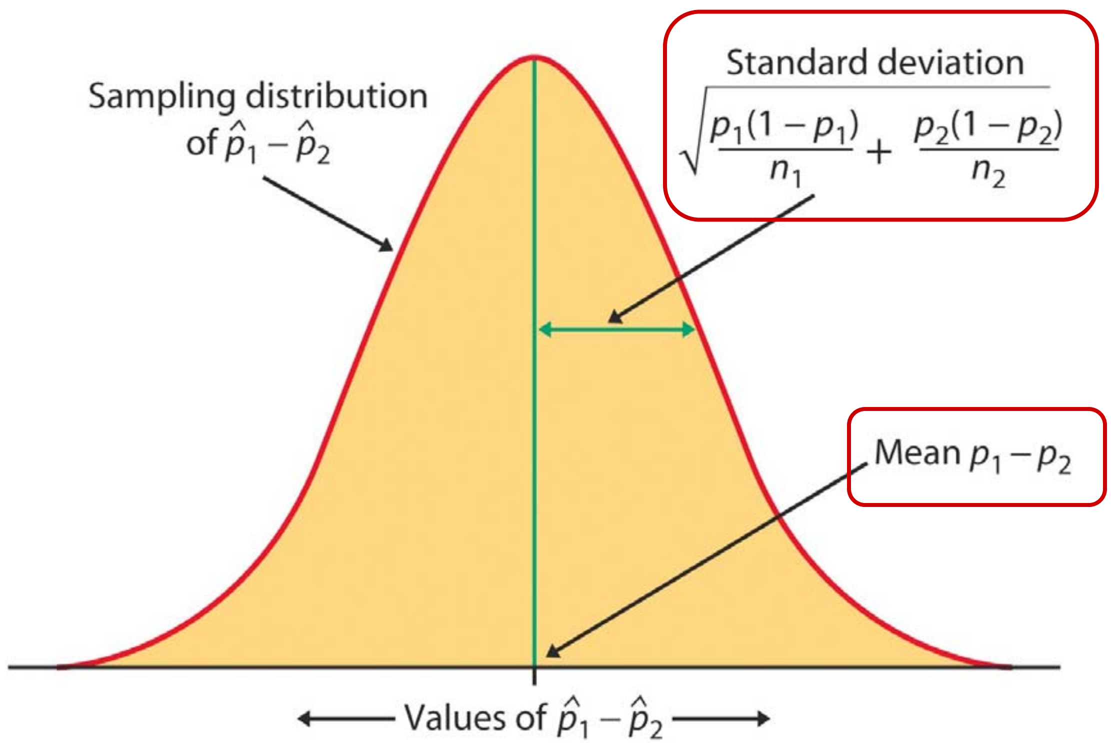
Objectives (BPS chapter 21)

Comparing two proportions

- The sampling distribution of a difference between proportions
- Large Sample confidence intervals for comparing two proportions
- Using technology
- Accurate confidence intervals for comparing two proportions
- Significance tests for comparing proportions

Comparing two independent samples

We often need to compare two treatments used on **independent** samples. We can compute the difference between the two sample proportions and compare it to the corresponding, approximately normal sampling distribution for ($\hat{p}_1 - \hat{p}_2$):



Large-sample CI for two proportions

For two independent SRSs of sizes n_1 and n_2 with sample proportion of successes $\hat{p}_1$ and $\hat{p}_2$ respectively, an **approximate level *C* confidence interval** for $p_1 - p_2$ is:

$$(\hat{p}_1 - \hat{p}_2) \pm m, \quad m \text{ is the margin of error}$$

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

C is the area under the standard normal curve between −z and z*.*

Use this method only when the populations are at least 10 times larger than the samples and the number of successes and the number of failures are each at least 10 in each sample.

Cholesterol and heart attacks

How much does the cholesterol-lowering drug Gemfibrozil help reduce the risk of heart attack? We compare the incidence of heart attack over a 5-year period for two random samples of middle-aged men taking either the drug or a placebo.

	H. attack	n	$\hat{p}$
Drug	56	2051	2.73%
Placebo	84	2030	4.14%

Standard error of the difference $p_1 - p_2$:

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

$$SE = \sqrt{\frac{0.0273(0.9727)}{2051} + \frac{0.0414(0.9586)}{2030}} = 0.00764$$

The confidence interval is $(\hat{p}_1 - \hat{p}_2) \pm z^* SE$

So the 90% CI is (0.0414−0.0273) ± 1.645*0.00746 = 0.0141 ± 0.0125

We are 90% confident that the percentage of middle-aged men who suffer a heart attack is 0.16% to 2.7% lower when taking the cholesterol-lowering drug.

Using technology

Here are some examples of output for a different set of data.

Texas Instruments TI-83 or TI-84

```
2-PropZInt
 (.05912,.11399)
 p̂1=.437638704
 p̂2=.3510840624
 n1=2253
 n2=2629
```

CrunchIt!

Two sample Proportion with summary

95% confidence interval results:

p_1 : proportion of successes for population 1

p_2 : proportion of successes for population 2

$p_1 - p_2$: difference in proportions

Difference	Count1	Total1	Count2	Total2	Sample Diff.	Std. Err.	L. Limit	U. Limit
p1 - p2	986	2253	923	2629	0.08655464	0.013996254	0.059122488	0.1139868

Minitab

Session

Test and CI for Two Proportions

```
Sample      X     N  Sample p
1         986  2253  0.437639
2         923  2629  0.351084

Difference = p (1) - p (2)
Estimate for difference:  0.0865546
95% CI for difference:  (0.0591225, 0.113987)
Test for difference = 0 (vs not = 0):  Z = 6.18  P-Value = 0.000
```

"Plus four" CI for two proportions

The "plus four" method again produces more accurate confidence intervals. We act as if we had four additional observations: one success and one failure in each of the two samples. The new combined sample size is $n_1 + n_2 + 4$, and the proportions of successes are:

$$\tilde{p}_1 = \frac{X_1 + 1}{n_1 + 2} \quad \text{and} \quad \tilde{p}_2 = \frac{X_2 + 1}{n_2 + 2}$$

An approximate level C confidence interval is:

$$CI: \quad (\tilde{p}_1 - \tilde{p}_2) \pm z^* \sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1 + 2} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2 + 2}}$$

Use this when C is at least 90% and both sample sizes are at least 5.

Cholesterol and heart attacks

Let's now calculate the plus four CI for the difference in percentage middle-aged men who suffer a heart attack (placebo – drug).

	H. attack	n	$\tilde{p}$
Drug	56	2051	2.78%
Placebo	84	2030	4.18%

$$\tilde{p}_1 = \frac{X_1 + 1}{n_1 + 2} = \frac{56+1}{2051+2} \approx 0.0278 \quad \text{and} \quad \tilde{p}_2 = \frac{X_2 + 1}{n_2 + 2} = \frac{84+1}{2030+2} \approx 0.0418$$

Standard error of the population difference $p_1 - p_2$:

$$SE = \sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1+2} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2+2}} = \sqrt{\frac{0.0278(0.9722)}{2053} + \frac{0.0418(0.9582)}{2032}} = 0.0057$$

The confidence interval is $(\tilde{p}_1 - \tilde{p}_2) \pm z^*SE$

➔ (0.0418 − 0.0278) ± 1.645*0.00573 = 0.014 ± 0.0094

We are 90% confident that the percentage of middle-aged men who suffer a heart attack is 0.46% to 2.34% lower when taking the cholesterol-lowering drug.

Test of significance

If the null hypothesis is true, then we can rely on the properties of the sampling distribution to estimate the probability of drawing two samples with proportions $\hat{p}_1$ and $\hat{p}_2$ at random.

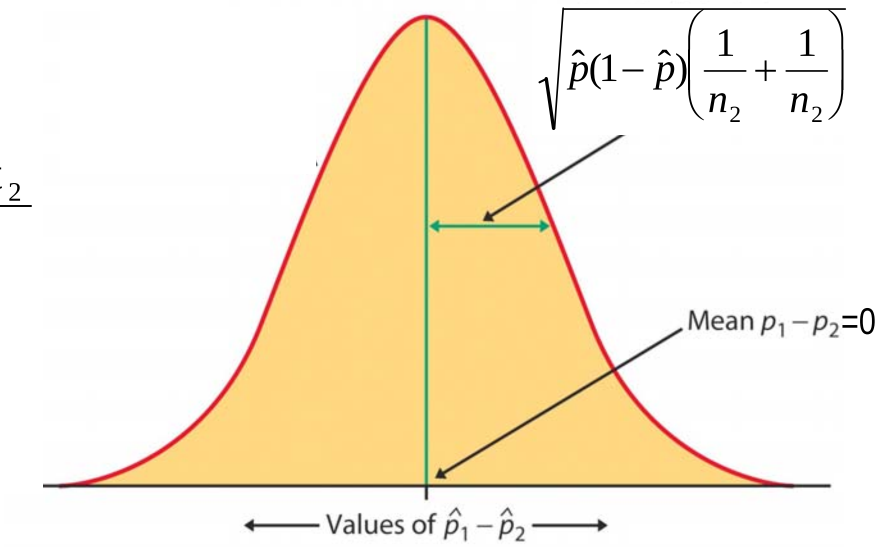
$$H_0 : p_1 = p_2 = p$$

Our best estimate of p is $\hat{p}$,

the pooled sample proportion

$$\hat{p} = \frac{\text{total successes}}{\text{total observations}} = \frac{\text{count}_1 + \text{count}_2}{n_1 + n_2}$$

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



This test is appropriate when all counts are at least 5 (number of successes and number of failures in each sample).

Gastric Freezing

Gastric freezing was once a treatment for ulcers. Patients would swallow a deflated balloon with tubes, and a cold liquid would be pumped for an hour to cool the stomach and reduce acid production, thus relieving ulcer pain. **The treatment was shown to be safe, significantly reducing ulcer pain,** and so was widely used for years.

A randomized comparative experiment later compared the outcome of gastric freezing with that of a placebo: 28 of the 82 patients subjected to gastric freezing improved, while 30 of the 78 in the control group improved.

$$H_0: p_{gf} = p_{placebo}$$
$$H_a: p_{gf} > p_{placebo}$$

$$\hat{p}_{pooled} = \frac{28+30}{82+78} = 0.3625$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.341-0.385}{\sqrt{0.363*0.637\left(\frac{1}{82}+\frac{1}{78}\right)}} = \frac{-0.044}{\sqrt{0.231*0.025}} = -0.499$$

Conclusion: The gastric freezing was no better than a placebo (p-value 0.69), and this treatment was abandoned. **ALWAYS USE A CONTROL!**