STAT 113
Analytic Intervals and Tests for Differences Between Two Groups

Colin Reimer Dawson

Oberlin College

November 9, 2017

Cases to Address

We will need standard errors to do CIs and tests for the following parameters:

1. Single Proportion (Last Time)
2. Single Mean (Wrap Up Today)
3. Difference of Proportions (Today)
4. Difference of Means (Today)
5. Mean of Differences (Next Week)

Example: Penguins Again!

Penguin Breeding

The scientists who studied whether metal bands were harmful to penguin survival also examined whether they affected the penguins' breeding patterns. For the metal-band group, 39 of 122 penguin-seasons resulted in offspring (32%). In the control group, 70 out of 160 penguin-seasons 32% of 122 breeding seasons (combined across penguins), whereas the controls had offpsring in 70 of 160 breeding seasons (44%).

- If we want to construct a confidence interval or do a test about the difference in the proportion "breeding opportunities" that were successful, what is the relevant *population parameter*?
- What is the relevant sample statistic?

Outline

Variability of a Difference

CI and Test for Difference of Proportions

CI and Test for Difference of Means

Variance and Standard Error of Differences

- With two *independent* samples, A and B, then quantities such as $\hat{p}_A - \hat{p}_B$ that depend on both random samples have two *independent* sources of variability.
- So the difference is *more variable* than either sample statistic alone.
- Specifically, the *variance* of the difference is the *sum* of the separate variances:

$$s^2_{\hat{p}_A - \hat{p}_B} = s^2_{\hat{p}_A} + s^2_{\hat{p}_B}$$

Variance and Standard Error of Proportions

- Recall: across all random samples, the *standard deviation* of the sample proportions (i.e., the standard error) is

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

  where p is the population proportion and n is the sample size.
- The *variance* of $\hat{p}$ is the square of this; i.e., the same thing without the square root.

Standard Error of Difference of Proportions

So the *variance* of the difference between two independent sample proportions is

$$s^2_{\hat{p}_A - \hat{p}_B} = \frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}$$

and the *standard deviation* (i.e., standard error) of the difference is

$$s_{\hat{p}_A - \hat{p}_B} = \sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}$$

Standard Error of Difference of Means

The exact same reasoning applies to the standard error of a difference between means of two independent samples:

$$s_{\bar{x}_A} = \frac{\sigma_A}{\sqrt{n_A}} \qquad s_{\bar{x}_B} = \frac{\sigma_B}{\sqrt{n_B}}$$

$$s^2_{\bar{x}_A} = \frac{\sigma_A^2}{n_A} \qquad s^2_{\bar{x}_B} = \frac{\sigma_B^2}{n_B}$$

$$s^2_{\bar{x}_A - \bar{x}_B} = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}$$

$$s_{\bar{x}_A - \bar{x}_B} = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$$

Analytic Approximations of Sampling Distributions

Param.	Stat.	Randomization	Theory SE	Test Dist.
p	$\hat{p}$	Simulate from p_0	$\sqrt{\frac{p_0(1-p_0)}{n}}$	Normal
μ	$\bar{x}$	Bootstrap + shift	$\frac{s}{\sqrt{n}}$	t_{n-1}
$p_A - p_B$	$\hat{p}_A - \hat{p}_B$	Scramble groups	$\sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}$	Normal
$\mu_A - \mu_B$	$\bar{x}_A - \bar{x}_B$	Scramble groups	$\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$	$t_{\min(n_A,n_B)-1}$
μ_D	$\bar{x}_D$	Flip pairs*	$\frac{s_D}{\sqrt{n_D}}$	t_{n_D-1}
ρ	r	Scramble pairings	$\sqrt{\frac{1-r^2}{n-2}}$	t_{n-2}

$$CI: \quad \text{Statistic} \pm \text{Critical Value} \times \widehat{SE}$$

$$\text{Sandardized Test Statistic}: \quad \frac{\text{Statistic} - \text{Null Param.}}{\widehat{SE}}$$

Outline

Variability of a Difference

CI and Test for Difference of Proportions

CI and Test for Difference of Means

Distribution of $\hat{p}_A - \hat{p}_B$

- Condition: The sampling distribution of $\hat{p}_A - \hat{p}_B$ is approximately Normal with at least 10 cases from all four combinations:

$$n_A p_A \geq 10 \qquad n_A(1-p_A) \geq 10$$
$$n_B p_B \geq 10 \qquad n_B(1-p_B) \geq 10$$

- Mean: $p_A - p_B$
- Standard deviation (standard error):

$$SE_{\hat{p}} = \sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}$$

Confidence Interval for a Difference of Proportions

CI Summary: Difference of Proportions

To compute a confidence interval for a difference of two proportions when the sampling distribution for $\hat{p}_A - \hat{p}_B$ is approximately Normal (see the last slide for conditions)

1. Find the standardized endpoints, Z^*, for the confidence level, using a standard Normal
2. "Destandardize" to get the endpoints

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

Why do we use $\hat{p}_A$ and $\hat{p}_B$ in the standard error, again?

P-values for a difference of two sample proportions from a Standard Normal

Computing P-values when the null sampling distribution is approximately Normal (see previously stated conditions) is the reverse process:

1. Convert $\hat{p}_A - \hat{p}_B$ to a z-score within the theoretical null sampling distribution (i.e., using its mean and standard deviation).

$$Z_{observed} = \frac{\hat{p}_A - \hat{p}_B - 0}{?}$$

2. Find the relevant area beyond $Z_{observed}$ using a Standard Normal

Null Standard Error

- Problem: H_0 states what the *difference* is, but the standard error depends on *each* population proportion:

$$s_{\hat{p}_A - \hat{p}_B} = \sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}$$

- This is not a function of the difference.
- But, H_0 says that p_A and p_B are the *same thing*, so we can estimate this single number using $\hat{p}_{combined}$, the proportion of the relevant category across *both* groups.
- Note: hold this proportion constant already when doing a randomization test.

P-values for a difference of two proportions

Computing P-values when the null sampling distribution is approximately Normal (see previously stated conditions) is the reverse process:

1. Convert $\hat{p}_A - \hat{p}_B$ to a z-score within the theoretical null sampling distribution (i.e., using its mean and standard deviation).

$$Z_{observed} = \frac{\hat{p}_A - \hat{p}_B - 0}{\sqrt{\frac{\hat{p}_{combined}(1-\hat{p}_{combined})}{n_A} + \frac{\hat{p}_{combined}(1-\hat{p}_{combined})}{n_B}}}$$
$$= \frac{\hat{p}_A - \hat{p}_B - 0}{\sqrt{\hat{p}_{combined}(1-\hat{p}_{combined})(\frac{1}{n_A} + \frac{1}{n_B})}}$$

2. Find the relevant area beyond $Z_{observed}$ using a Standard Normal

Example: Penguins Again!

Penguin Breeding

The scientists who studied whether metal bands were harmful to penguin survival also examined whether they affected the penguins' breeding patterns. For the metal-band group, 39 of 122 penguin-seasons resulted in offspring (32%). In the control group, 70 out of 160 penguin-seasons 32% of 122 breeding seasons (combined across penguins), whereas the controls had offpsring in 70 of 160 breeding seasons (44%).

$n_{metal} =$ ______ $n_{control} =$ ______

$\hat{p}_{metal} =$ ______ $\hat{p}_{control} =$ ______

$\hat{p}_{metal} - \hat{p}_{control} =$ ______

Penguins Breed Confidence (Interval)

Is the Normal approximation reasonable?

$$CI: \text{point estimate} \pm Z^* \cdot SE$$

$$SE = \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$

$$= \underline{\hspace{4cm}}$$

$$Z^* = \underline{\hspace{4cm}}$$

Find a 90% CI for $p_A - p_B$

Do metal bands reduce breeding chances?

$$Z_{observed} = \frac{\text{observed difference} - \text{null difference}}{\text{standard error}}$$

$$SE = \sqrt{\hat{p}_{combined}(1 - \hat{p}_{combined})(\frac{1}{n_A} + \frac{1}{n_B})}$$

$\hat{p}_{combined} =$ ______

$SE =$ ______

$Z_{observed} =$ ______

P-value = ______

Decision? ____________

Outline

Variability of a Difference

CI and Test for Difference of Proportions

CI and Test for Difference of Means

Distribution of Difference of Sample Means

Distribution of $\bar{x}$ (σ known)

- Two populations with means μ_A and μ_B, and standard deviations σ_A and σ_B
- Conditions: Sampling distribution of $\bar{x}_A - \bar{x}_B$ is Normal if
  - Populations are Normal, or
  - Sample sizes in each group are large enough (roughly can use $n_A, n_B \geq 27$)
- Mean: $\mu_A - \mu_B$
- Standard deviation (standard error):

$$SE_{\bar{x}_A - \bar{x}_B} = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$$

Estimating σ_A and σ_B

Distribution of standardized difference

- Since we don't know σ_A and σ_B, we use

$$\widehat{SE}_{\bar{x}_A - \bar{x}_B} = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

- Now, the standardized difference has a t-distribution

$$\frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} \sim t_{df}$$

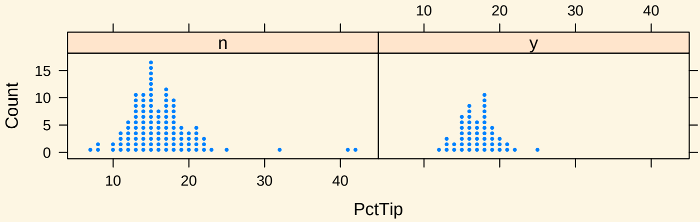
- A conservative rule is to set $df = \min(n_A, n_B) - 1$

Example: Credit Card Use and Tip Percentage

Credit Cards and Tip Percentage

We analyze the percent tip left on 157 bills from the First Crush bistro in Northern New York State. The mean percent tip left on the 106 bills paid in cash was 16.39 with a standard deviation of 5.05. The mean percent tip left on the 51 bills paid with a credit card was 17.10 with a standard deviation of 2.47.

```
library("mosaic"); library("Lock5Data"); data("RestaurantTips")
dotPlot(~PctTip | Credit, data = RestaurantTips, width = 1, cex = 1)
```



Credit Cards and Tip Percentage

We analyze the percent tip left on 157 bills from the First Crush bistro in Northern New York State. The mean percent tip left on the 106 bills paid in cash was 16.39 with a standard deviation of 5.05. The mean percent tip left on the 51 bills paid with a credit card was 17.10 with a standard deviation of 2.47.

$n_{card} =$ ______ $n_{cash} =$ ______

$\bar{x}_{card} =$ ______ $\bar{x}_{cash} =$ ______

$s^2_{card} =$ ______ $s^2_{cash} =$ ______

$\bar{x}_{card} - \bar{x}_{cash} =$ ______

Card Vs. Cash Confidence Interval

Is the Normal approximation reasonable?

$$CI : \text{point estimate} \pm T^* \cdot SE$$
$$SE = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$
$$= \_\_\_\_\_\_\_\_\_\_\_\_$$
$$T^* = \_\_\_\_\_\_\_\_\_\_\_\_$$

Find a 99% CI for $\mu_{card} - \mu_{cash}$.

Do people who pay differently tip differently?

$$T_{observed} = \frac{\text{observed difference} - \text{null difference}}{\text{standard error}}$$

$$SE = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

$$SE = \_\_\_\_\_\_$$

$$T_{observed} = \_\_\_\_\_\_$$

$$P\text{-value} = \_\_\_\_\_\_$$

Decision? ____________

Analytic Approximations of Sampling Distributions

Param.	Stat.	Randomization	Theory SE	Test Dist.
p	$\hat{p}$	Simulate from p_0	$\sqrt{\frac{p_0(1-p_0)}{n}}$	Normal
μ	$\bar{x}$	Bootstrap + shift	$\frac{s}{\sqrt{n}}$	t_{n-1}
$p_A - p_B$	$\hat{p}_A - \hat{p}_B$	Scramble groups	$\sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}$	Normal
$\mu_A - \mu_B$	$\bar{x}_A - \bar{x}_B$	Scramble groups	$\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$	$t_{\min(n_A, n_B)-1}$
μ_D	$\bar{x}_D$	Flip pairs*	$\frac{s_D}{\sqrt{n_D}}$	t_{n_D-1}
ρ	r	Scramble pairings	$\sqrt{\frac{1-r^2}{n-2}}$	t_{n-2}

$$CI: \quad \text{Statistic} \pm \text{Critical Value} \times \widehat{SE}$$

$$\text{Sandardized Test Statistic}: \quad \frac{\text{Statistic} - \text{Null Param.}}{\widehat{SE}}$$