Section 6.5:

Paired Difference in Means

Outline

- Paired data
- Confidence interval for difference in means based on paired data
- Hypothesis test for difference in means based on paired data

Paired Data

- In Section 6.4, we consider inference for a difference in means when the data consist of two separate samples.
- What about situations such as matched pairs experiments (discussed in Section 1.3) where the data being compared consist of pairs of data values?
- Paired data examples:
  - Two measurements on each case (compare each case to themselves under different treatments)
  - Twin studies
  - Each case is matched with a similar case, and one case in each pair is given each treatment for comparison

Pheromones in Tears

- Do pheromones (subconscious chemical signals) in female tears affect testosterone levels in men?
- Cotton pads had either real female tears or a salt solution that had been dripped down the same female's face
- 50 men had a pad attached to their upper lip twice, once with tears and once without, order randomized.

Paired Data!

- Response variable: testosterone level

Gelstein, et. al. (2011) "Human Tears Contain a Chemosignal," *Science,* 1/6/11.

Paired Data

- **Separate samples**:
  - Some men would get real tears, and a *separate* group of men get fake tears
  - Can list the entire response variable in one column.

	Tears	Testosterone
Case 1	Real	141
Case 2	Fake	162
Case 3	Fake	155
Case 4	Real	157
Case 5	Fake	139

- **Paired Data**:
  - Each man gets *both* real and fake tears
  - 2 measurements for each man
  - Real tear response data in one column, fake tear response data in another column

	Testosterone (Real)	Testosterone (Fake)
Case 1	127	141
Case 2	140	162
Case 3	142	155
Case 4	138	157
Case 5	111	139
Case 6	162	161

Should data from the following situation be analyzed as paired data or separate samples? To study the effect of sitting with a laptop computer on one’s lap on scrotal temperature, 29 men have their scrotal temperature tested before and then after sitting with a laptop for one hour.

A. Paired Data
B. Separate Samples

Should data from the following situation be analyzed as paired data or separate samples? A study investigating the effect of exercise on brain activity recruits sets of identical twins in middle age, in which one twin is randomly assigned to engage in regular exercise and the other doesn't exercise.

A. Paired Data
B. Separate Samples

Should data from the following situation be analyzed as paired data or separate samples? **In a study to determine whether the color red increases how attractive men find women, one group of men rate the attractiveness of a woman after seeing her picture on a red background and another group of men rate the same woman after seeing her picture on a white background.**

A. Paired Data
B. Separate Samples

Should data from the following situation be analyzed as paired data or separate samples? To measure the effectiveness of a new teaching method for math in elementary school, each student in a class getting the new instructional method is matched with a student in a separate class on IQ, family income, math ability level the previous year, reading level, and all demographic characteristics. At the end of the year, math ability levels are measured.

A. Paired Data
B. Separate Samples

Analyzing Paired Data

- For a matched pairs experiment, we look at the ***difference between responses*** for each unit (pair), rather than just the average difference between two treatment groups
  - Get a new variable of the differences, and do inference for the difference as you would for a single mean
- Rather than doing inference for **difference** in **means** $(\mu_1 - \mu_2)$, do inference for the **mean difference** (μ_d)

Analyzing Paired Data

Case	Treatment 1	Treatment 2	Difference, *d*
1	77	85	-8
2	82	84	-2
3	94	91	3
4	62	73	-11
5	70	77	-7
Average for Group	77	82	-5
SD for Group	12.124	7.071	5.523

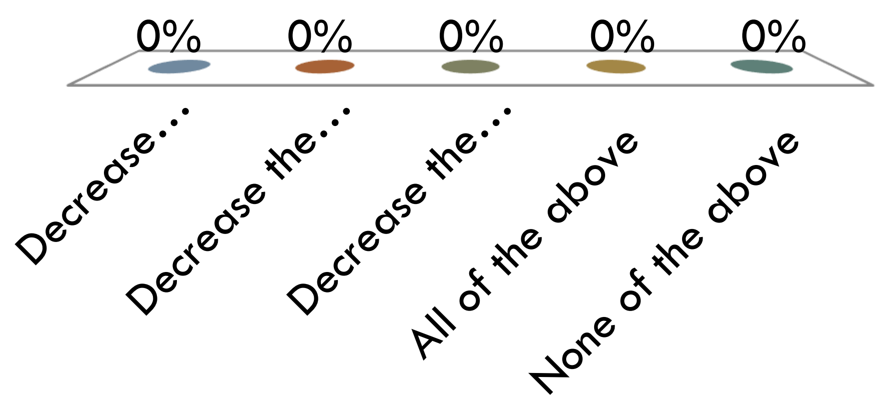
The mean of the differences, $\bar{x}_d$, is also −5.

The difference in the two means, $\bar{x}_1 - \bar{x}_2$, is −5.

Using matched pairs decreases the standard deviation of the response, which decreases the standard error

Why use paired data?

A. Decrease standard deviation of the response
B. Decrease the chance of a Type II error for tests
C. Decrease the margin of error for intervals
D. All of the above
E. None of the above



Matched Pairs

- Matched pairs experiments are particularly useful when responses vary a lot from unit to unit
- We can decrease standard deviation of the response (and so decrease standard error of the statistic) by comparing each unit to a matched unit

Inference for Paired Data

- To analyze the differences, we use the same formulas we already learned for a single mean:

$$\mathrm{SE} \approx \frac{s_d}{\sqrt{n_d}}$$

$$\bar{x}_d \pm t* \cdot \frac{s_d}{\sqrt{n_d}}$$

$$t = \frac{\bar{x}_d}{\frac{s_d}{\sqrt{n_d}}}$$

- $\bar{x}_d$: sample mean of the differences
- s_d : sample standard deviation of the differences
- n_d : number of differences (number of pairs)

- If the distribution of the differences is approximately normal or n_d is large ($n_d \geq 30$), we can use a *t*-distribution with $n_d - 1$ degrees of freedom

For the 50 men, the average difference in testosterone levels between tears and no tears was −21.7 pg/ml. ("pg" = picogram = 0.001 nanogram = 10^{-12} gram). The standard deviation of these differences was 46.5. Average level before sniffing was 155 pg/ml. Do female tears lower male testosterone levels? Calculate standard error.

A. 4.79
B. 5.42
C. 6.23
D. 6.58

$$\boxed{\mathrm{SE} \approx \frac{s_d}{\sqrt{n_d}}}$$

For the 50 men, the average difference in testosterone levels between tears and no tears was −21.7 pg/ml. ("pg" = picogram = 0.001 nanogram = 10^{-12} gram). The standard deviation of these differences was 46.5. Average level before sniffing was 155 pg/ml. Do female tears lower male testosterone levels? Calculate test statistic.

A. 0.001

B. -0.47

C. -3.3

D. -3.9

$$\boxed{t = \frac{\bar{x}_d}{\frac{s_d}{\sqrt{n_d}}}}$$

For the 50 men, the average difference in testosterone levels between tears and no tears was −21.7 pg/ml. ("pg" = picogram = 0.001 nanogram = 10^{-12} gram). The standard deviation of these differences was 46.5. Average level before sniffing was 155 pg/ml. Do female tears lower male testosterone levels? Find the p-value.

A. 0.0009

B. 0.009

C. 0.09

D. 0.9

Pheromones in Tears: Hypothesis Test

$\bar{x}_D = -21.7$
$s_D = 46.5$
$n_D = 50$

1. State hypotheses: $H_0: \mu_D = 0$
$H_a: \mu_D < 0$

2. Check conditions: $n_D = 50 \geq 30$

3. Calculate standard error:

4. Calculate test statistic:

5. Compute p-value: Distribution: t with $50 - 1 = 49$ df

6. Interpret in context:

Pheromones in Tears: 95% CI

$\bar{x}_D = -21.7$
$s_D = 46.5$
$n_D = 50$

1. Check conditions: $n_D = 50 \geq 30$

2. Find *t:** t with $50 - 1 = 49$ df, 95% CI:
$\Rightarrow t^* = 2.01$

3. Calculate standard error:

4. Calculate CI: $statistic \pm t^* \cdot SE = \bar{x}_d \pm t^* \cdot \frac{s_d}{\sqrt{n_d}}$

$=$

5. Interpret in context:

Summary

- **Standard error** for paired difference in means: $\frac{s_d}{\sqrt{n_d}}$
- **Inference for paired difference in means:** If the sample size is large ($n_d \geq 30$), then $\bar{x}_d \approx N\left(\mu_d, \frac{\sigma_d}{\sqrt{n_d}}\right)$. However, using s_d in place of σ_d, ***changes*** the distribution of the sample means ***to a t-distribution***.
  - The t-distribution is characterized by its ***degrees of freedom*** $= n_d - 1$
  - Conditions for the t-distribution: $n_d \geq 30$ or the distribution of the differences are approximately normal.