# Lecture 20
# Paired *t*-test

CHOONG-WAN WOO | COCOAN lab | http://cocoanlab.github.io

# Review: Key Points

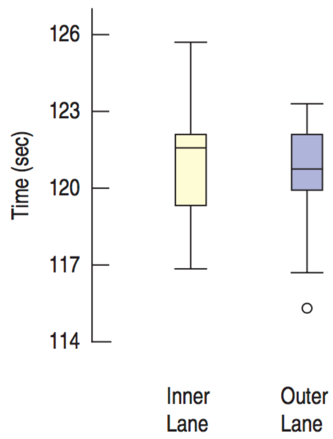
## Chapter 24: Comparing Groups

- $Var(X-Y) = Var(X) + Var(Y)$
- $SD(X-Y) = \sqrt{SD^2(X) + SD^2(Y)} = \sqrt{Var(X) + Var(Y)}$
- Confidence interval for the difference between two proportions: $(\hat{p}_1 - \hat{p}_2) \pm z^* \times SE(\hat{p}_1 - \hat{p}_2)$
  - $SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$
- Z-test for the difference between two proportions: $z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{SE_{\text{pooled}}(\hat{p}_1 - \hat{p}_2)}$
- Confidence interval for the difference between two means: $(\bar{y}_1 - \bar{y}_2) \pm ME$, where $ME = t^*_{df} \times SE(\bar{y}_1 - \bar{y}_2)$
  - $SE(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
- Two-sample t-test, $t = \frac{(\bar{y}_1 - \bar{y}_2) - \Delta_0}{SE(\bar{y}_1 - \bar{y}_2)}$

# Paired data

- The example from the textbook: Speed-skating
- Inner lane vs. outer lane
- Question: Is there a difference between two lanes?

| Inner Lane | | Outer Lane | |
|---|---|---|---|
| Name | Time | Name | Time |
| OLTEAN Daniela | 129.24 | (no competitor) | |
| ZHANG Xiaolei | 125.75 | NEMOTO Nami | 122.34 |
| ABRAMOVA Yekaterina | 121.63 | LAMB Maria | 122.12 |
| REMPEL Shannon | 122.24 | NOH Seon Yeong | 123.35 |
| LEE Ju-Youn | 120.85 | TIMMER Marianne | 120.45 |
| ROKITA Anna Natalia | 122.19 | MARRA Adelia | 123.07 |
| YAKSHINA Valentina | 122.15 | OPITZ Lucille | 122.75 |
| BJELKEVIK Hedvig | 122.16 | HAUGLI Maren | 121.22 |
| ISHINO Eriko | 121.85 | WOJCICKA Katarzyna | 119.96 |
| RANEY Catherine | 121.17 | BJELKEVIK Annette | 121.03 |
| OTSU Hiromi | 124.77 | LOBYSHEVA Yekaterina | 118.87 |
| SIMIONATO Chiara | 118.76 | JI Jia | 121.85 |
| ANSCHUETZ THOMS Daniela | 119.74 | WANG Fei | 120.13 |
| BARYSHEVA Varvara | 121.60 | van DEUTEKOM Paulien | 120.15 |
| GROENEWOLD Renate | 119.33 | GROVES Kristina | 116.74 |
| RODRIGUEZ Jennifer | 119.30 | NESBITT Christine | 119.15 |
| FRIESINGER Anni | 117.31 | KLASSEN Cindy | 115.27 |
| WUST Ireen | 116.90 | TABATA Maki | 120.77 |



# Paired *t*-test

- Two-sample *t*-test?
  - NO. The races are run in pairs, so the columns are not independent.

- Instead, we should focus on the difference between each pair.
  - It's not a problem, paired data provides an opportunity!
  - We need to take advantage of the paired data structure.

- "Paired" *t*-test:
  - Use *pairwise* differences!
  - Ignore original two columns
  - One-sample *t*-test on the pairwise differences

| Skating Pair | Inner Time | Outer Time | Inner – Outer |
|---|---|---|---|
| **1** | 129.24 | | . |
| **2** | 125.75 | 122.34 | 3.41 |
| **3** | 121.63 | 122.12 | −0.49 |
| **4** | 122.24 | 123.35 | −1.11 |
| **5** | 120.85 | 120.45 | 0.40 |
| **6** | 122.19 | 123.07 | −0.88 |
| **7** | 122.15 | 122.75 | −0.60 |
| **8** | 122.16 | 121.22 | 0.94 |
| **9** | 121.85 | 119.96 | 1.89 |
| **10** | 121.17 | 121.03 | 0.14 |
| **11** | 124.77 | 118.87 | 5.90 |
| **12** | 118.76 | 121.85 | −3.09 |
| **13** | 119.74 | 120.13 | −0.39 |
| **14** | 121.60 | 120.15 | 1.45 |
| **15** | 119.33 | 116.74 | 2.59 |
| **16** | 119.30 | 119.15 | 0.15 |
| **17** | 117.31 | 115.27 | 2.04 |
| **18** | 116.90 | 120.77 | −3.87 |

# Assumptions and conditions

- Paired data assumption
  - You should not use methods for paired data on independent data.
  - Or methods for independent data (e.g., two-sample *t*-test) should not be used on paired data.
- Independence assumption
  - The *differences* for pairs should be independent of each other.
  - Conditions
    - Randomization condition
    - 10% condition
- Normal population assumption
  - The population of *differences* should follows a Normal model.
  - Each group doesn't need to follow a Normal model.

# Steps of paired $t$-test

- $H_0$: $\mu_d = \Delta_0$ (usually, $\Delta_0 = 0$)
- $t_{n-1} = \frac{\bar{d}-\Delta_0}{SE(\bar{d})}$
- $df = n - 1$
- $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$

# Confidence intervals for matched pairs

- $\bar{d} \pm t^*_{n-1} \times SE(\bar{d})$
- $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$

# Example of paired $t$-test: Speed skating

$H_0$: Neither lane offered an advantage:

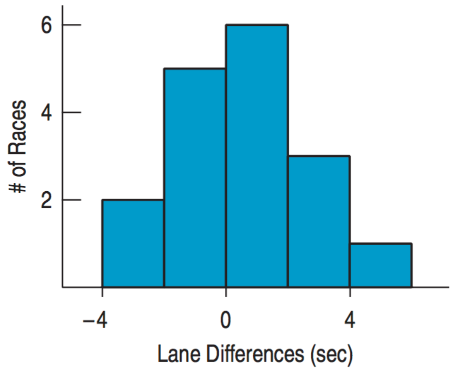
$$\mu_d = 0.$$

$H_A$: The mean difference is different from zero:

$$\mu_d \neq 0.$$

| Skating Pair | Inner Time | Outer Time | Inner – Outer |
|---|---|---|---|
| **1** | 129.24 | | . |
| **2** | 125.75 | 122.34 | 3.41 |
| **3** | 121.63 | 122.12 | −0.49 |
| **4** | 122.24 | 123.35 | −1.11 |
| **5** | 120.85 | 120.45 | 0.40 |
| **6** | 122.19 | 123.07 | −0.88 |
| **7** | 122.15 | 122.75 | −0.60 |
| **8** | 122.16 | 121.22 | 0.94 |
| **9** | 121.85 | 119.96 | 1.89 |
| **10** | 121.17 | 121.03 | 0.14 |
| **11** | 124.77 | 118.87 | 5.90 |
| **12** | 118.76 | 121.85 | −3.09 |
| **13** | 119.74 | 120.13 | −0.39 |
| **14** | 121.60 | 120.15 | 1.45 |
| **15** | 119.33 | 116.74 | 2.59 |
| **16** | 119.30 | 119.15 | 0.15 |
| **17** | 117.31 | 115.27 | 2.04 |
| **18** | 116.90 | 120.77 | −3.87 |

- ✔ **Independence Assumption:** Each race is independent of the others, so the differences are mutually independent.
- ✔ **Paired Data Assumption:** The data are paired because racers compete in pairs.
- ✔ **Randomization Condition:** Skaters are assigned to lanes at random. Repeating the experiment with different pairings and lane assignments would give randomly different values.
- ✔ **Nearly Normal Condition:** The histogram of the differences is unimodal and symmetric:




The conditions are met, so I'll use a Student's $t$-model with $(n - 1) = 16$ degrees of freedom, and perform a **paired $t$-test.**

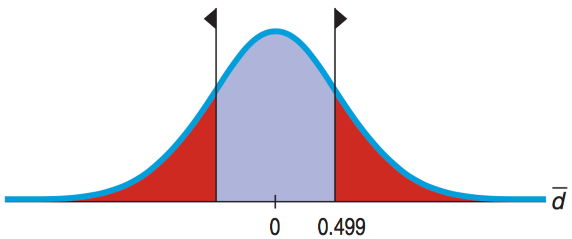
The data give

$$n = 17 \text{ pairs}$$
$$\bar{d} = 0.499 \text{ seconds}$$
$$s_d = 2.333 \text{ seconds.}$$

I estimate the standard deviation of $\bar{d}$ using

$$SE(\bar{d}) = \frac{s_d}{\sqrt{n}} = \frac{2.333}{\sqrt{17}} = 0.5658$$

$$\text{So } t_{16} = \frac{\bar{d} - 0}{SE(\bar{d})} = \frac{0.499}{0.5658} = 0.882$$




$$P\text{-value} = 2P(t_{16} > 0.882) = 0.39$$

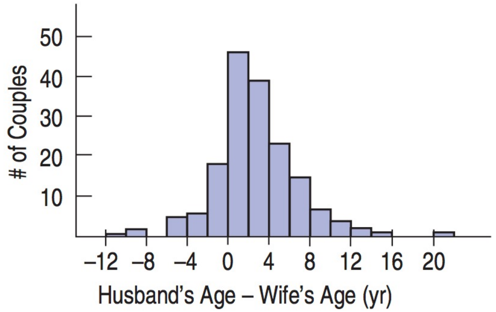
# Example of confidence intervals

| Wife's Age | Husband's Age | Difference (husband – wife) |
|---|---|---|
| 43 | 49 | 6 |
| 28 | 25 | −3 |
| 30 | 40 | 10 |
| 57 | 52 | −5 |
| 52 | 58 | 6 |
| 27 | 32 | 5 |
| 52 | 43 | −9 |
| ⋮ | ⋮ | ⋮ |

Ages of husbands and wives

from 170 couples

- ✔ **Paired Data Assumption:** The data are paired because they are on members of married couples.
- ✔ **Independence Assumption:** The data are from a randomized survey, so couples should be independent of each other.
- ✔ **Randomization Condition:** These couples were randomly sampled.
- ✔ **10% Condition:** The sample is less than 10% of the population of married couples in Britain.
- ✔ **Nearly Normal Condition:** The histogram of the husband – wife differences is unimodal and symmetric:




The conditions are met, so I can use a Student's $t$-model with $(n - 1) = 169$ degrees of freedom and find a **paired $t$-interval.**

$n = 170$ couples
$\bar{d} = 2.2$ years
$s_d = 4.1$ years

I estimate the standard error of $\bar{d}$ as

$$SE(\bar{d}) = \frac{s_d}{\sqrt{n}} = \frac{4.1}{\sqrt{170}} = 0.31 \text{ years.}$$

The df for the $t$-model is $n - 1 = 169$.

The 95% critical value for $t_{169}$ (from the table) is 1.97.

The margin of error is

$$ME = t^*_{169} \times SE(\bar{d}) = 1.97(0.31) = 0.61$$

So the 95% confidence interval is

$$2.2 \pm 0.6 \text{ years,}$$

or an interval of (1.6, 2.8) years.

I am 95% confident that British husbands are, on average, 1.6 to 2.8 years older than their wives.

# Hands-on

| $X_1$ | $X_2$ | $X_1 - X_2$ |
|---|---|---|
| 12 | 12 | 0 |
| 12 | 9 | 3 |
| 14 | 13 | 1 |
| 12 | 10 | 2 |
| 15 | 11 | 4 |
| 14 | 13 | 1 |
| 20 | 10 | 10 |
| 20 | 10 | 10 |
| 11 | 14 | -3 |
| 18 | 10 | 8 |
| 13 | 16 | -3 |
| 15 | 17 | -2 |
| 16 | 15 | 1 |
| 20 | 11 | 9 |
| 15 | 13 | 2 |
| 20 | 9 | 11 |
| 14 | 17 | -3 |
| 18 | 16 | 2 |
| 17 | 16 | 1 |
| 16 | 10 | 6 |

## Steps of paired *t*-test

- $H_0$: $\mu_d = \Delta_0$ (usually, $\Delta_0 = 0$)
- $t_{n-1} = \frac{\bar{d}-\Delta_0}{SE(\bar{d})}$
- $df = n - 1$
- $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$

## Confidence intervals for matched pairs

- $\bar{d} \pm t^*_{n-1} \times SE(\bar{d})$
- $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$

$$n = 20$$
$$(\sqrt{20} = 4.47)$$

$$\bar{d} = 3$$
$$s_d = 4.5538$$

$$SE(\bar{d}) =$$
$$t_{19} =$$

$$P(t_{19} > \quad) =$$

$$t^*_{19} \times SE(\bar{d}) =$$

$$t^*_{19} = 2.093$$

```
>> std(d)/sqrt(20)

ans =

    1.0183
```

```
>> ste(d)

ans =

    1.0183

>> mean(d)/ste(d)

ans =

    2.9462

>> (1-tcdf(mean(d)/ste(d), 19))*2

ans =

    0.0083

>> tinv(0.975,19)*ste(d)

ans =

    2.1312
```

# Effect size

- Speed skating example again.. Though the test failed to reject the null hypothesis, is it really okay?
- If we calculate the 95% confidence interval of the difference, it's $-0.70 < \mu_d < 1.70$ seconds.
  - Examining the confidence interval is a good way to get a sense for the size of the effect.
  - Is 1.7 seconds okay?
  - What if the gap between Gold and Silver medals is smaller than this?

# Non-parametric test: Sign test

- Distribution-free methods again!
  - For the pairs with negative differences, record 0's
  - For the pairs with positive differences, record 1's
  - Ignore the pairs with difference = 0

  - Then, test the associated proportion $p = 0.5$ using a $z$-test

# Key Points

## Chapter 25: Paired Samples and Blocks

- **Paired $t$-test:** use pairwise differences, and then one-sample $t$-test on the pairwise differences
  - $H_0$: $\mu_d = \Delta_0$ (usually, $\Delta_0$ = 0), $t_{n-1} = \frac{\bar{d}-\Delta_0}{SE(\bar{d})}, df = n-1, SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$
- **Confidence interval:** $\bar{d} \pm t^*_{n-1} \times SE(\bar{d})$
- **Nonparametric sign test:**
  - Record 0 for the pairs with negative differences, record 1 for the pairs with positive differences
  - and ignore the pairs with difference = 0
  - Then, test the associated proportion $p$ = 0.5 using a $z$-test