

# Statistics with Recitation: TA Session

Danny Po-Hsien Kang (康柏賢)

November 25, 2025

# Today's agenda

## 1 Hypothesis Testing: t-test

- One-Sample t-test
- Paired Sample t-test
- Two Independent Sample t-test

# Today's Dataset

- Please download the three datasets from the OpenIntro website.
  - `china.csv`: Weekly time spent on child care in China.
  - `textbooks.csv`: Prices of textbooks for various courses at UCLA Bookstore and Amazon.com.
  - `ncbirths.csv`: A 2004 dataset from North Carolina on births, including mothers' habits and practices, with 1,000 cases.
- After that, import the data

```
china <- read.csv("data/china.csv")
book <- read.csv("data/textbooks.csv")
birth <- read.csv("data/ncbirths.csv")
```

# One-Sample t-test: `t.test()`

- **Syntax:**

```
t.test(data$variable, mu,  
       alternative, conf.level)
```

- **Example 1:**

```
china_male <- china %>% filter(gender == 1)  
t.test(china_male$child_care,  
       mu = 25,  
       conf.level = 0.95)
```

- Default value for parameter `alternative` is "two.sided"
  - Use "less" or "greater" for one-sided tests.

# One-Sample t-test: t.test()

- Output:

```
One Sample t-test

data: china_male$child_care
t = -13.232, df = 434, p-value < 2.2e-16

alternative hypothesis: true mean is not equal to 25

95 percent confidence interval:
-9.429757 -0.524266

sample estimates:
mean of x
-4.977011
```

# One-Sample t-test: `t.test()`

- Example 2:

```
t.test(china$child_care,  
       mu = 25,  
       alternative = "less",  
       conf.level = 0.95)
```

# One-Sample t-test: `t.test()`

- **Output:**

```
One Sample t-test

data: china$child_care
t = -14.697, df = 1238, p-value < 2.2e-16

alternative hypothesis: true mean is less than 25

95 percent confidence interval:
-Inf 5.249741

sample estimates:
mean of x
2.758676
```

# Paired Sample t-test: t.test()

- **Syntax:**

```
t.test(data$var1, data$var2,  
       paired = TRUE, conf.level)
```

- **Example 1:**

```
t.test(book$ucla_new,  
       book$amaz_new,  
       paired = TRUE,  
       coef.level = 0.95)
```

# Paired Sample t-test: t.test()

- Output:

```
Paired t-test

data: book$ucla_new and book$amaz_new
t = 7.6488, df = 72, p-value = 6.928e-11

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:
 9.435636 16.087652

sample estimates:
mean difference
12.76164
```

# Two Independent Sample t-test: `t.test()`

- **Syntax:**

```
t.test(data$ValueVar ~ data$GroupVar,  
       conf.level)
```

- **Example 1:**

```
t.test(birth$weight ~ birth$habit,  
       conf.level = 0.95)
```

# Two Independent Sample t-test: `t.test()`

- Output:

```
Welch Two Sample t-test
```

```
data: birth$weight by birth$habit  
t = 2.359, df = 171.32, p-value = 0.01945
```

alternative hypothesis: true difference in means between group nonsmoker and group smoker **is not equal** to 0

95 percent confidence interval:

0.05151165 0.57957328

**sample estimates:**

mean in group nonsmoker	mean in group smoker
7.144273	6.828730