Module 2:

Inference for a Normal Population

Different flavours of t tests

Hypothesis testing for means using t tests Agenda

1. **Why do we use Student t-tests instead of Z scores?**
2. **What are the three types of t-tests**

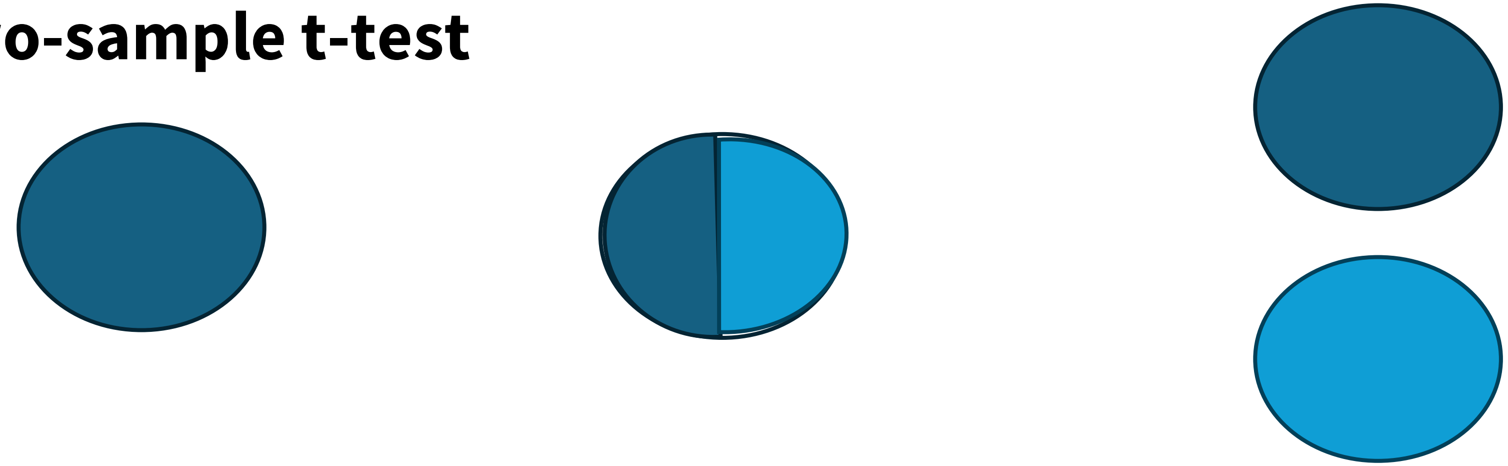
- **One sample t tests**
  - ❑ Assumptions
  - ❑ When assumptions not met, use median and rank → **Signed test**
- **Paired t test**
  - ❑ Assumptions
- **Two sample t test**
  - ❑ Assumptions
  - ❑ When variances aren't equal → **Welch's approximate t test**
  - ❑ Other assumptions not met: median and rank → **Mann Whitney U test**

We won't have time to cover everything in detail – nor every example I give - so here is another reference that outlines the different t tests:

https://statisticsbyjim.com/hypothesis-testing/t-tests-1-sample-2-sample-paired-t-tests/

Part 2: **What** t tests? We will look at the following t-tests:

1. Comparing one mean:
   a. **One-sample t-test**
2. Comparing two means:
   a. Paired t-test
   b. Two-sample t-test



one sample paired two sample

---

*Each of the above tests have **slightly different assumptions** which allow our conclusions to be supported. We will investigate what happens when these assumptions are violated and how robust our various t-tests are to violations.*

Comparing Two Means

- Tests with one categorical and one numerical variable
  - Goal: compare the mean of a numerical variable among different groups

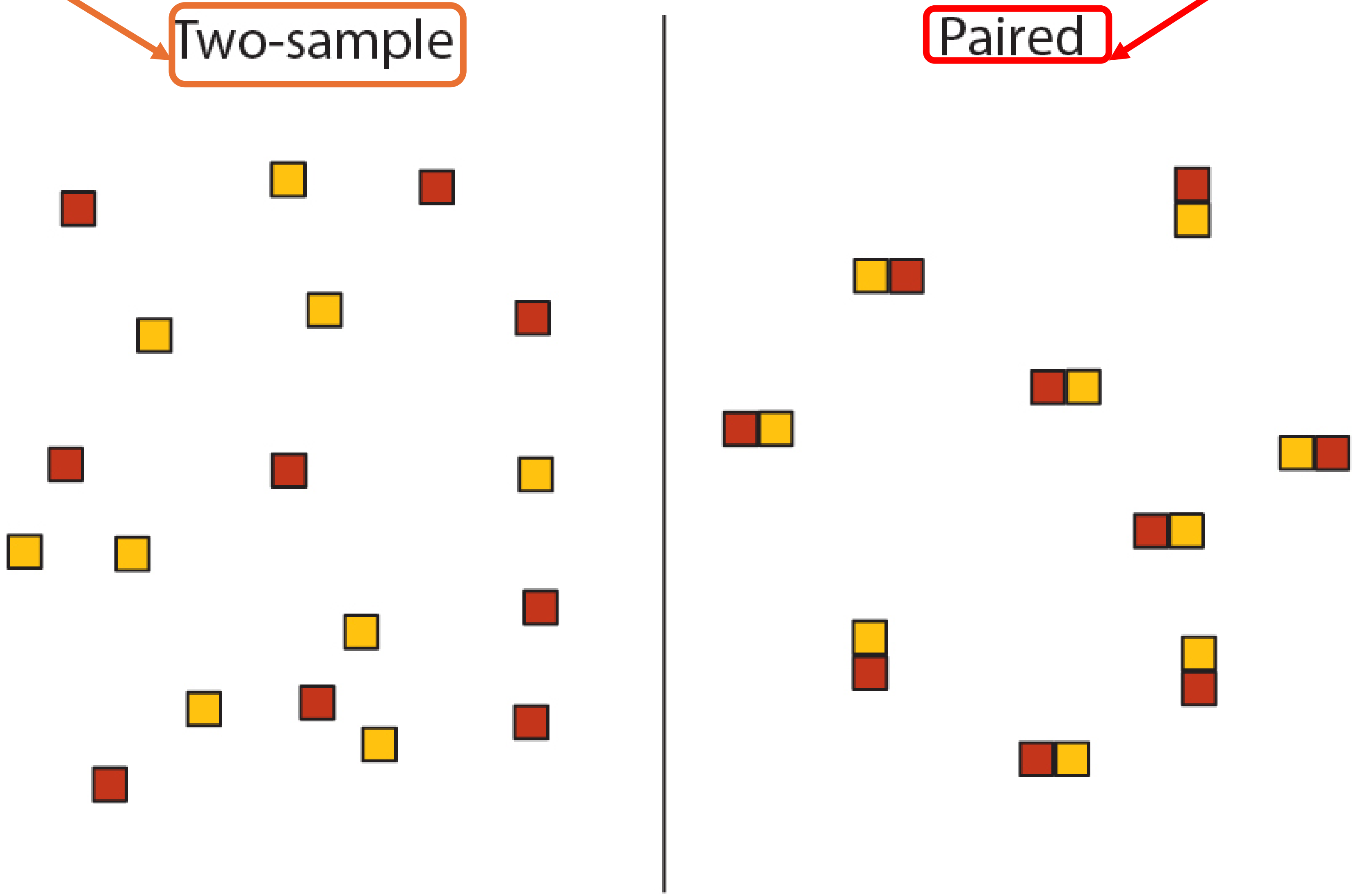
- Examine two major study design comparisons:
  - Paired Design
  - Two-Sample Design

Comparing Two Means

Easier

Better

Paired vs. Two Sample Comparisons



Paired Design:

- Data from two groups are paired
- Each member of the pair shares everything in common with the other *except* for the tested categorical variable

→ Reduces effects of (hidden) confounding variables

- One-to-one correspondence between the individuals in the two groups
- Example of a type of experimental design called "Blocking"

Examples:

- **Before** and **After** treatments
  - Mouse weight before and after high fat diet
- **Identical twins**: one with treatment and one without
- One arm given treatment (sunscreen) the other arm is not on the **same individual**
- Testing effects of treatment (smoking) in a sample of patients, each of which is compared to a non-treatment (nonsmoker) closely **matched** by age, weight, ethnic background and socioeconomic condition*

Paired Design:

- The sampling unit ***is the pair*: *one member with a treatment and a second with a different treatment***

- two measurements must be reduced to a single number which is the mean of the difference between the two measurements

  - Ex: If there are 46 individuals grouped into 23 pairs and ***n* = ?**

Paired Design:

- Strategy: Compares the mean of the differences to a value given in the null hypothesis
- Often tests the null hypothesis that the mean difference of paired measurements is equal to "0"

For each pair, calculate the difference. *The paired t-test is then simply the one-sample t-test on the differences*

Paired Design:

- Strategy: Compares the mean of the differences to a value given in the null hypothesis
- Often tests the null hypothesis that the mean difference of paired measurements is equal to "0"

For each pair, calculate the difference. *The paired t-test is then simply the one-sample t-test on the differences*

Assumptions:

- **Difference** between members of each pair have a normal distribution
  - Distribution of the single measurements on each sampling unit do not need to be normally distributed – only the differences must be normally distributed
- Pairs are chosen at random

Example: No Smoking Day in Great Britain on the second Wednesday of March. Compared to the previous Wednesday, does (voluntarily) not smoking for a day affect injury rate?

Year	Injuries Previous Wednesday	Injuries "No Smoking" Wednesday
1987	516	540
1988	610	620
1989	581	599
1990	586	639
1991	554	607
1992	632	603
1993	479	519
1994	583	560
1995	445	515
1996	522	556

Example: No Smoking Day in Great Britain on the second Wednesday of March. Compared to the previous Wednesday, does (voluntarily) not smoking for a day affect injury rate?

Step 1: formulate null hypothesis

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_A: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \neq 0$

Example: No Smoking Day in Great Britain on the second Wednesday of March. Compared to the previous Wednesday, does (voluntarily) not smoking for a day affect injury rate?

Step 1: formulate null hypothesis

H_0: There is no difference in number of injuries experienced on “No smoking” Wednesday and a regular Wednesday, $\mu_d = 0$

H_A: There is a difference in number of injuries experienced on “No smoking” Wednesday and a regular Wednesday, $\mu_d \neq 0$

Step 2: Identify test statistic and determine if assumptions are met

- paired t-test
- Assumptions: d_i are normally distributed
- Calculate $\bar{d}$, $SE_{\bar{d}}$

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

- Remember: *n* is the number of pairs, which are the independent sampling units

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \; \square \; 0$

Test statistic & assumptions: paired t-test $\boxed{t = \dfrac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$

Year	Injuries Previous Wed.	Injuries "No Smoking" Wed.	d_i
1987	516	540	24
1988	610	620	10
1989	581	599	18
1990	586	639	53
1991	554	607	53
1992	632	603	-29
1993	479	519	40
1994	583	560	-23
1995	445	515	70
1996	522	556	34

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday,

$$\mu_d = 0$$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday,

$$\mu_d \square 0$$

Test statistic & assumptions:

- paired t-test

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

Year	Injuries Previous Wed.	Injuries "No Smoking" Wed.	d_i
1987	516	540	24
1988	610	620	10
1989	581	599	18
1990	586	639	53
1991	554	607	53
1992	632	603	-29
1993	479	519	40
1994	583	560	-23
1995	445	515	70
1996	522	556	34

$\bar{d} = 25$

$SE_{\bar{d}} = 10.22$

$n = 10$

$dof = 9$

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on “No smoking” Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on “No smoking” Wednesday and a regular Wednesday, $\mu_d \square 0$

Test statistic & assumptions:

- paired t-test

$$t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}$$

$\bar{d}$ = 25
$SE_{\bar{d}}$ =10.22
dof = 9

$$t = \frac{25 - 0}{10.22} = 2.45$$

Critical Values and Significance levels:

$\alpha = 0.05$

$t_{0.05(2),\,9} > \mathbf{2.26}$

*We can reject the H_0: there **is** a difference in injury rate between days on which people smoke and those on which people don’t*

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \square 0$

Test statistic:

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

- paired t-test

$t = \frac{25 - 0}{10.22} = 2.45$

$$\boxed{\begin{aligned} \bar{d} &= 25 \\ SE_{\bar{d}} &= 10.22 \\ dof &= 9 \end{aligned}}$$

Critical Values and Significance levels:

$\alpha = 0.05$

$t_{0.05(2),\,9} > 2.26$

We can reject the H_0: there is a difference in injury rate between the days on which people smoke and those on which people don't. We can calculate the Confidence Interval of this difference to support our conclusion. It does not include 0 so there is a difference.

$$\boxed{\bar{d} - t_{\alpha(2),df} SE_{\bar{d}} < \mu_d < \bar{d} + t_{\alpha(2),df} SE_{\bar{d}}}$$

$$1.9 < \mu_d < 48.1$$