# 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**

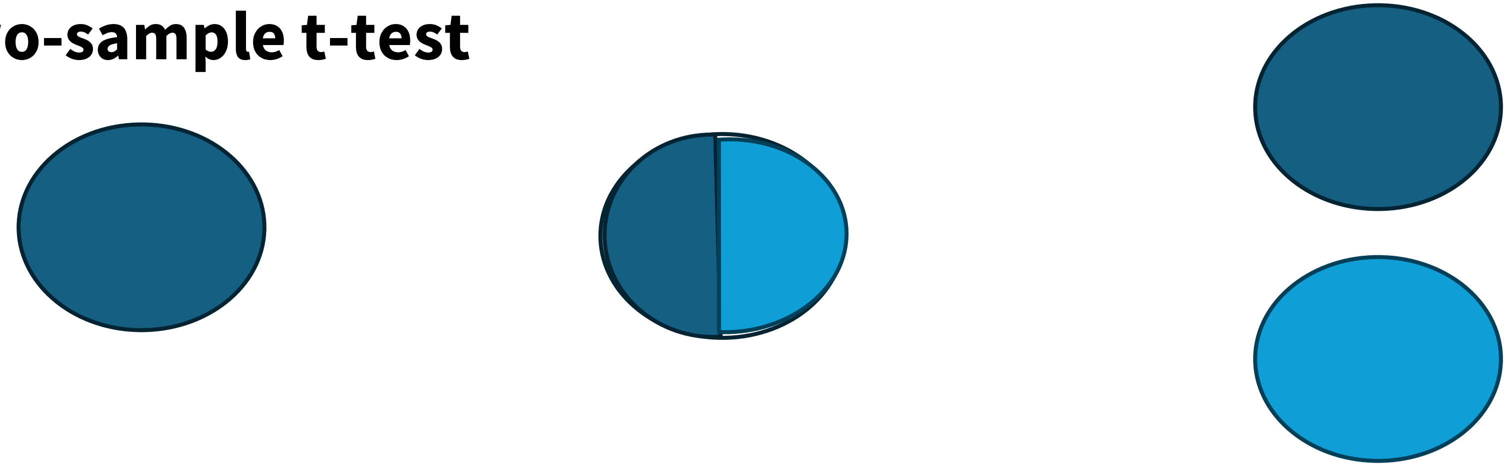
- **One sample t tests**
  - ❑ Assumptions
  - ❑ When assumptions not met, use median and rank → **Signed test**
- **Paired t test**
  - ❑ Assumptions
- **Two sample t test**
  - ❑ Assumptions
  - ❑ When variances aren't equal → **Welch's approximate t test**
  - ❑ Other assumptions not met: median and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

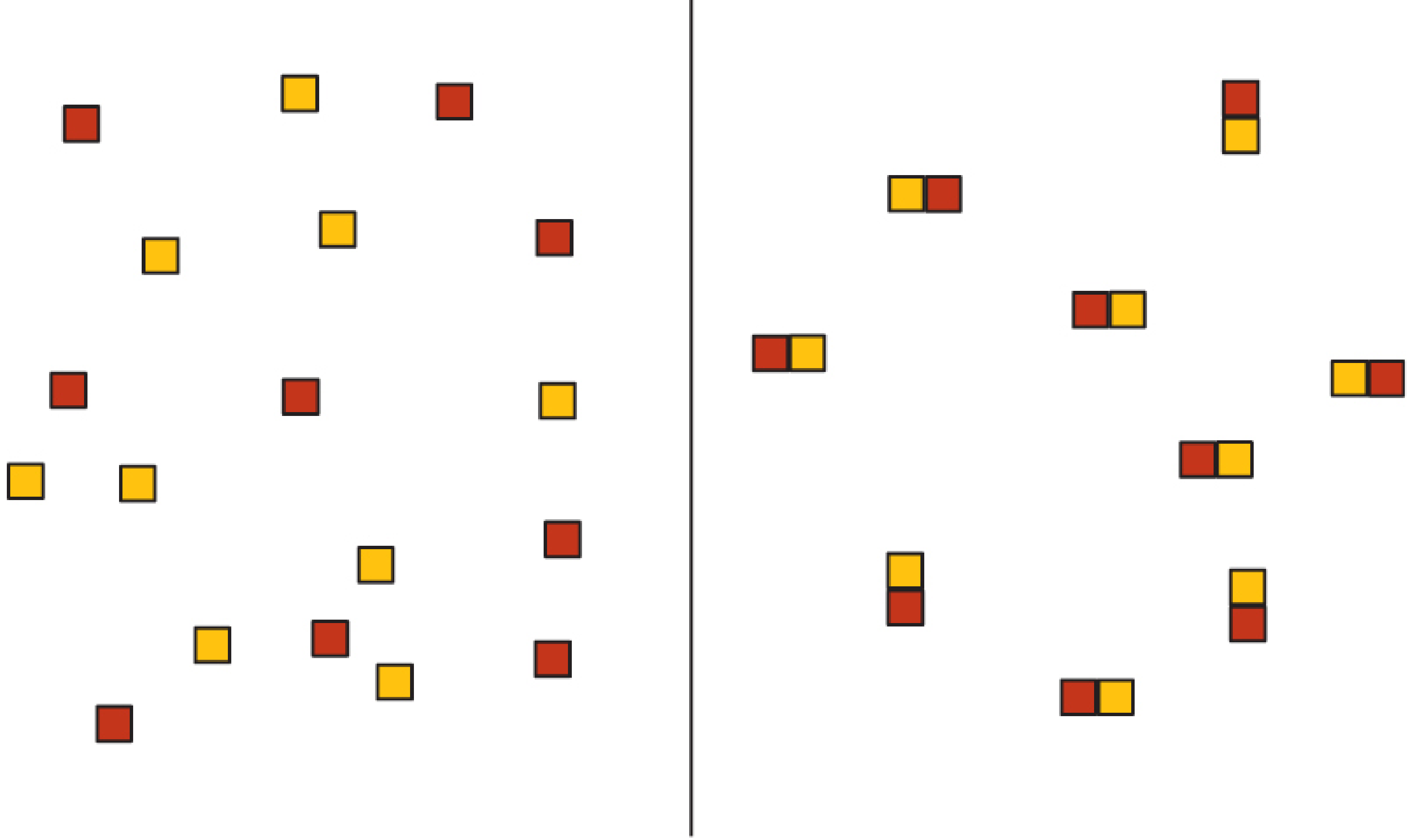
Easier

Better

Paired vs. Two Sample Comparisons

Two-sample

Paired



## Two-sample Design:

- **Assumptions:**
  - **Random Sample** is **normally distributed** in both populations --> sampling distribution for difference between sample means is also normal
  - **Standard deviation** is the same in both populations --> if this is not true, use Welch's approximate t-test instead*

- Strategy:
  - Unlike in a paired t-test, there are two variables from two entirely different populations. Instead of one variable describing the difference, $\bar{d}$, you have two: $\bar{Y}_1 - \bar{Y}_2$
  - Standard Error of $\bar{Y}_1 - \bar{Y}_2$ is **pooled!**

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

- *two-sample t-test is robust to violations of assumptions if n is similar between the two groups.*

# Two-sample Design:

## Pooled sample variance:

- Weighted average; the average of the variances of the samples weighted by their degrees of freedom

$$s_p^2 = \frac{df_1 s_1^2 + df_2 s_2^2}{df_1 + df_2}$$

- tangent: what is the "pooled" variance doing?
  - allowing us to access and use the additional information that is in our sample
  - We will something similar in ANOVA
  - This is why the variances must be approximately equal
    - Behrens-Fisher problem (illustrated on the next slide).

# BEHRENS-FISHER Problem

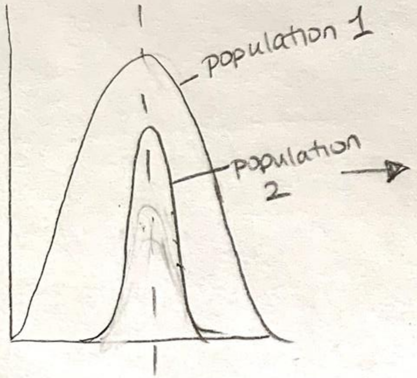
- when variances of two populations are not equal
- we can illustrate the problem with two extreme situations

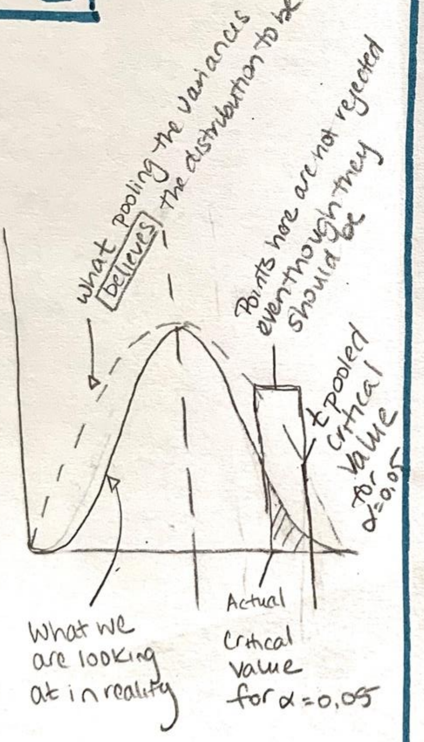
- Two sample t-test is not robust when $s_1^2 \neq s_2^2$ and $n_1 \neq n_2$

## Situation 1

larger sample has larger variance:

$\boxed{n_1 >> n_2 \ ; \ \sigma_1 > 3\sigma_2}$



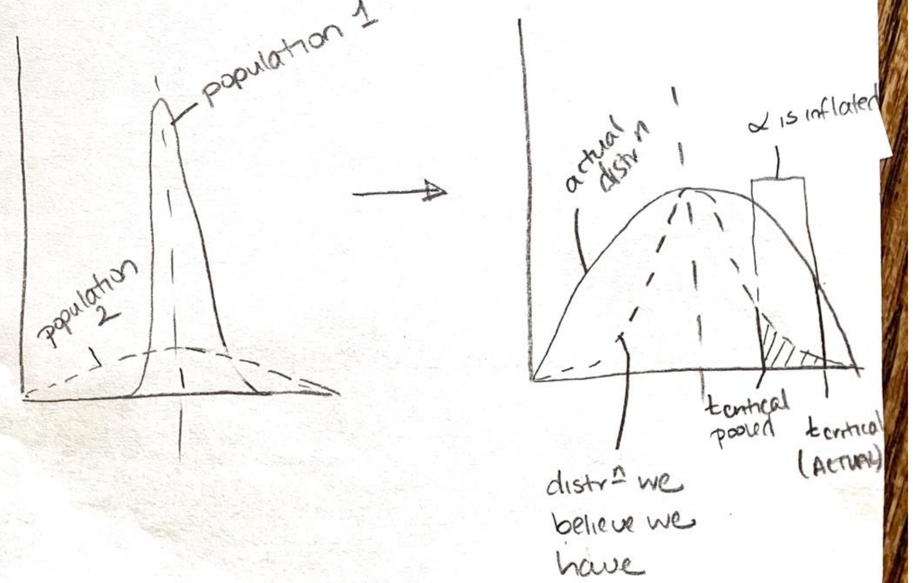




- pooled variance allows the larger sample, with its much larger variance, to contribute more

## Situation 2

larger sample has smaller variance

$\boxed{n_1 >> n_2 \ ; \ \sigma_2 > 3\sigma_1}$



# Two-sample Design:

Student's t-distribution of two-sample design:

- Compares the means of a numerical variable between two populations

$$\boxed{t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{SE_{\bar{Y}_1 - \bar{Y}_2}}}$$

Total degrees of freedom:

$$df = df_1 + df_2 = n_1 + n_2 - 2$$

- Two means are estimated, so subtract 2

# Two-sample Design:

2 genotypes of lettuce: *susceptible* and *resistant*. Do these genotypes differ in fitness in the absence of aphids.

*The proxy for fitness that is measured are number of buds*.

# Two-sample Design:

2 genotypes of lettuce: *susceptible* and *resistant*. Do these genotypes differ in fitness in the absence of aphids.

| | ***Susceptible*** | **Resistant** |
|---|---|---|
| **Mean number of buds** | **720** | **582** |
| **SD of number of buds** | **223.6** | **277.3** |
| **Sample size** | **15** | **16** |

Both distributions are normally distributed

## Two-sample Design:

2 genotypes of lettuce: *susceptible* and *resistant*.

Do these genotypes differ in fitness in the absence of aphids.

| | ***Susceptible*** | **Resistant** |
|---|---|---|
| **Mean number of buds** | **720** | **582** |
| **SD of number of buds** | **223.6** | **277.3** |
| **Sample size** | **15** | **16** |

$H_0$: There is no difference between the number of buds in susceptible and resistant plants ($\mu_1 = \mu_2$)

$H_A$: There is a difference between the number of buds in susceptible and resistant plants ($\mu_1 \neq \mu_2$)

## Two-sample Design:

Example: 2 genotypes of lettuce: *susceptible* and *resistant*. Do these genotypes differ in fitness in the absence of aphids.

| | *Susceptible* | Resistant |
|---|---|---|
| **Mean number of buds** | **720** | **582** |
| **SD of number of buds** | **223.6** | **277.3** |
| **Sample size** | **15** | **16** |

$H_0$: There is no difference between the number of buds in susceptible and resistant plants ($\mu_1 = \mu_2$)

$H_A$: There is a difference between the number of buds in susceptible and resistant plants ($\mu_1 \neq \mu_2$)

## t-test:

df = 15+16 - 2 = 29

$\alpha = 0.05$

$$s_p^2 = \frac{df_1 s_1^2 + df_2 s_2^2}{df_1 + df_2} = \frac{14(223.6)^2 + 15(277.3)^2}{14+15} = 63909.9$$

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{63909.9\left(\frac{1}{15} + \frac{1}{16}\right)} = \sqrt{8255.02} = 90.86$$

**Two-sample Design**:

Example: 2 genotypes of lettuce: *susceptible* and *resistant*. Do these genotypes differ in fitness in the absence of aphids.

| | ***Susceptible*** | **Resistant** |
|---|---|---|
| **Mean number of buds** | **720** | **582** |
| **SD of number of buds** | **223.6** | **277.3** |
| **Sample size** | **15** | **16** |

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{SE_{\bar{Y}_1 - \bar{Y}_2}} = \frac{(720 - 582)}{90.86} = 1.52$$

**$H_0$:** There is no difference between the number of buds in susceptible and resistant plants ($\mu_1 = \mu_2$)

**$H_A$:** There is a difference between the number of buds in susceptible and resistant plants ($\mu_1 \neq \mu_2$)

**Two sample t-test:**

**Assumptions have been met**

## Confidence Interval: Two-sample Design:

$$\boxed{(\bar{Y}_1 - \bar{Y}_2) - t_{\alpha(2),df} SE_{\bar{Y}_1 - \bar{Y}_2} < \mu_1 - \mu_2 < (\bar{Y}_1 - \bar{Y}_2) + t_{\alpha(2),df} SE_{\bar{Y}_1 - \bar{Y}_2}}$$

$$138 - 2.05(90.86) < \mu_1 - \mu_2 < 138 + 2.05(90.86)$$

$$-48.21 < \mu_1 - \mu_2 < 324.26$$

Note: this interval includes 0 which supports our conclusion (FTR)

# Assumptions of parametric tests:

- Random Samples
- Populations are normally distributed
- for two sample t-test: Populations have equal(ish) variances
  - if not → Welch's approximate t-test
  - How do we tell when populations don't have equal variances?