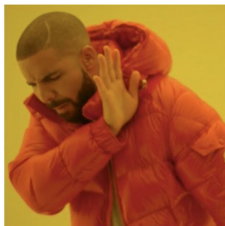


STA220H1: The Practice of Statistics I

Elena Tuzhilina

March 28, 2023

Please turn on your videos :)



prove
the null
hypothesis



fail
to reject
the null

Announcements

1. We have two Quizzes left before the final exam.
2. The Midterm 2 grades will be released in a couple of days, you will have one week to submit your regrade request (we will use email again).
3. Final exam is on April 20 at ²~~3~~-5 pm in MY.
4. We will hold additional office hours before the final exam.

Agenda for today

- ▶ Recap: testing and connection to confidence intervals, power, type I and II error
- ▶ Statistical testing for two samples: matching and non-matching pairs

Statistical testing and confidence intervals

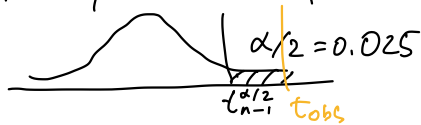
$$\left[\bar{x} - \frac{s}{\sqrt{n}}, \bar{x} + \frac{s}{\sqrt{n}} \right]$$

There is a connection between statistical testing and CI.

$$CI = \left[\bar{x} - t_{n-1}^{\alpha/2} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1}^{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right] \quad 1-\alpha$$

- ▶ If CI does not cover μ_0 , then we can reject $H_0 : \mu = \mu_0$ in favor of $H_a : \mu \neq \mu_0$ at significance level α
- ▶ If CI covers μ_0 , we do not have enough evidence to reject H_0 at significance level α

$t_{n-1}^{\alpha/2}$ - quantile of t_{n-1}



95%

$\alpha = 0.05$

90%

$\alpha = 0.1$

Example

If 90% confidence interval is $[2, 10]$ what can we say about the hypotheses $H_0 : \mu = 0$ vs. $H_a : \mu \neq 0$?

$0 \notin [2, 10] \Rightarrow$ reject $H_0 \Rightarrow \mu \neq 0$ with
90% Confidence
($\alpha = 0.1$)

If 90% confidence interval is $[2, 10]$ what can we say about the hypotheses $H_0 : \mu = 5$ vs. $H_a : \mu \neq 5$?

$5 \in [2, 10] \Rightarrow$ fail to reject H_0
 ~~$\mu = 5$~~ $\mu = 5$

Two-sided confidence interval

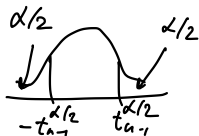
Two-sided confidence interval is $[a, b]$ that covers μ .

Standardization:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \text{ approximately } \sim t_{n-1}$$

Distribution table: $\alpha = 0.05 \Rightarrow 2.5\%$

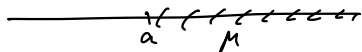
$$P\left(-t_{n-1}^{\alpha/2} \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{n-1}^{\alpha/2}\right) = 1 - \alpha$$



With probability $1 - \alpha$, the population parameter μ belongs to

$$\left[\bar{X} - t_{n-1}^{\alpha/2} \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1}^{\alpha/2} \frac{S}{\sqrt{n}} \right]$$

One-sided confidence interval



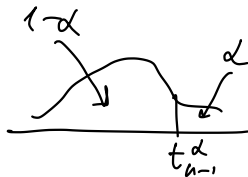
Upper one-sided confidence interval is $[a, +\infty)$ that covers μ .

Standardization:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \text{ approximately } \sim t_{n-1}$$

Distribution table:

$$P\left(\frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{n-1}^{\alpha}\right) = 1 - \alpha$$



With probability $1 - \alpha$, the population parameter μ belongs to


$$\left[\bar{X} - t_{n-1}^{\alpha} \frac{S}{\sqrt{n}}, +\infty\right)$$

$$\begin{aligned} \bar{X} - \mu &\leq t_{n-1}^{\alpha} \cdot \frac{S}{\sqrt{n}} \\ \bar{X} - t_{n-1}^{\alpha} \cdot \frac{S}{\sqrt{n}} &\leq \mu \end{aligned}$$

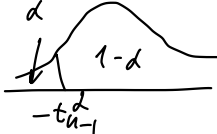
One-sided confidence interval

Lower one-sided confidence interval is $(-\infty, a]$ that covers μ .

Standardization:


$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \text{ approximately } \sim t_{n-1}$$

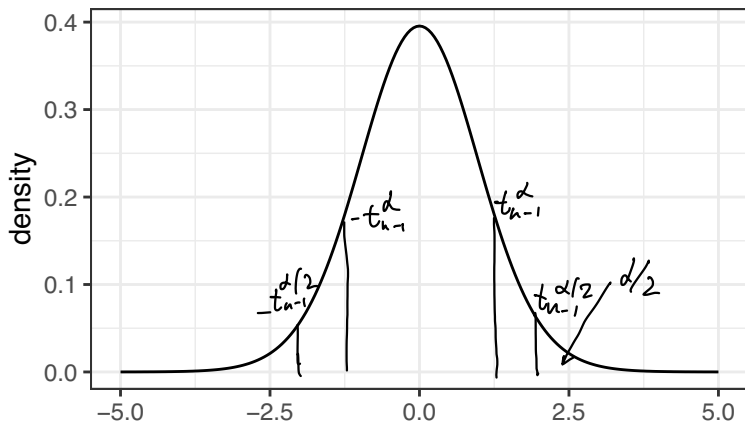
Distribution table:


$$P\left(\frac{\bar{X} - \mu}{S/\sqrt{n}} \geq -t_{n-1}^{\alpha}\right) = 1 - \alpha$$

With probability $1 - \alpha$, the population parameter μ belongs to

$$\left(-\infty, \bar{X} + t_{n-1}^{\alpha} \frac{S}{\sqrt{n}}\right]$$

One-sided vs two-sided quantiles



$$t_{n-1}^{\alpha} < t_{n-1}^{\alpha/2}$$

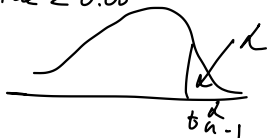
One-sided confidence intervals

$$t_{\text{obs}} = \frac{\bar{X} - \mu}{S/\sqrt{n}} > t_{n-1}^{\alpha/2} \Leftrightarrow \text{pvalue} < 0.05$$



To find **one-sided confidence interval**

- ▶ Use sample x_1, \dots, x_n to find \bar{x}, s
- ▶ Find quantile t_{n-1}^{α} that corresponds to the upper α -tail
- ▶ Compute $\left[\bar{x} - t_{n-1}^{\alpha} \frac{s}{\sqrt{n}}, +\infty \right)$ or $\left(-\infty, \bar{x} + t_{n-1}^{\alpha} \frac{s}{\sqrt{n}} \right]$



What if σ^2 was known?

$$\left[\bar{x} - z^{\alpha} \frac{\sigma}{\sqrt{n}}, +\infty \right)$$

Statistical testing and confidence intervals

There is a connection between one-sided statistical testing and one-sided CI.

$$CI = \left[\bar{x} - t_{n-1}^{\alpha} \frac{s}{\sqrt{n}}, +\infty \right)$$

- ▶ If CI does not cover μ_0 , then we can reject $H_0 : \mu = \mu_0$ in favor of $H_a : \mu > \mu_0$ at significance level α

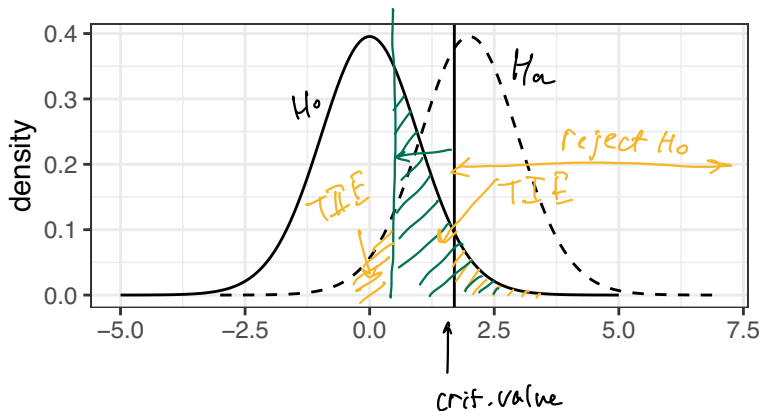
$$CI = \left(-\infty, \bar{x} + t_{n-1}^{\alpha} \frac{s}{\sqrt{n}} \right]$$

- ▶ If CI does not cover μ_0 , then we can reject $H_0 : \mu = \mu_0$ in favor of $H_a : \mu < \mu_0$ at significance level α

Statistical testing: Type I type II errors and power

Type I error: we rejected H_0 when H_0 was true.

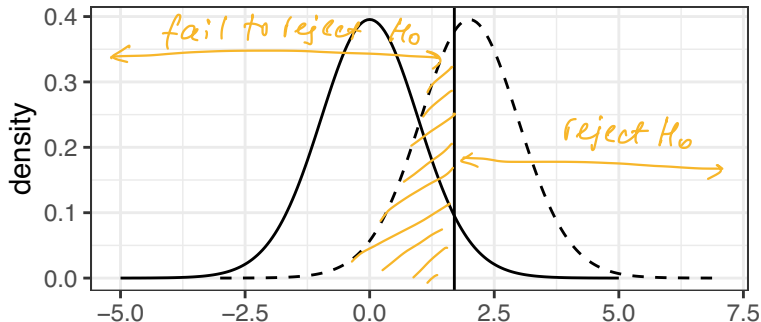
$$\alpha = P(\text{reject } H_0 | H_0 \text{ is true})$$



Statistical testing: Type I type II errors and power

Type II error: we failed to reject H_0 when H_a was true.

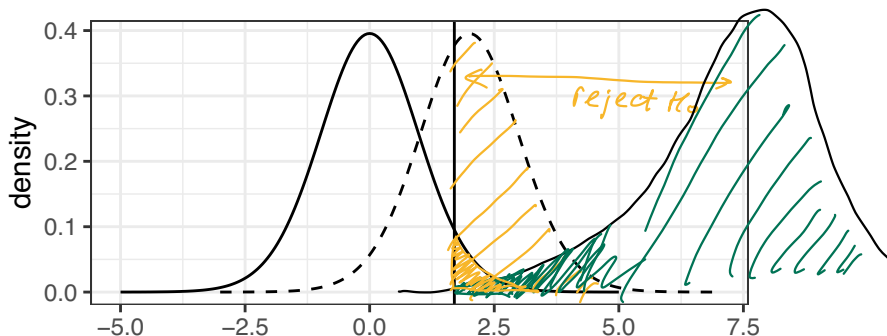
$$\beta = P(\text{fail to reject } H_0 | H_a \text{ is true})$$



Statistical testing: Type I type II errors and power

Power: chances to correctly reject H_0 when H_a is true.

$$1 - \beta = P(\text{reject } H_0 | H_a \text{ is true})$$



Statistical testing for two groups: matching samples

Paired t-test: compare two samples x_1, \dots, x_n and y_1, \dots, y_n with **matching observations**.

- ▶ Create a sample that shows the difference in measurements:

$$d_1, \dots, d_n \text{ where } d_i = x_i - y_i$$

- ▶ Perform statistical test on differences testing $H_0 : \mu_d = 0$ vs $H_a : \mu_d \neq 0$.

Example

30 chickens were fed with sunflower seeds for 1 month. Their weight gain (in gramms) was recorded.

The same chickens were fed with corn for 1 month. The new weight gain was recorded.

Is any diet better for the weight gain?

chicken	diet1	diet2	difference
1	248.4738	229.5219	18.95187
2	213.6821	272.2564	-58.57428
3	285.9187	227.6527	58.26604
4	160.2935	310.4249	-150.13137
5	140.8080	219.8894	-79.08139
6	249.5029	201.6018	47.90115

Statistical testing for two groups: matching samples

To test $H_0 : \mu_d = 0$ vs $H_a : \mu_d \neq 0$ you need to

- ▶ Compute sample mean \bar{d} and sample standard deviations s_d for the differences
- ▶ Compute $t_{obs} = \frac{\bar{d}}{s_d/\sqrt{n}}$
- ▶ Get p-value from the table and interpret the results

$$t_{obs} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

This requires normal approximation!

Statistical testing for two groups: matching samples

- ▶ X_1, \dots, X_n is the outcome for diet 1
- ▶ Y_1, \dots, Y_n is the outcome for diet 2

Then for differences $D_i = X_i - Y_i$, t-test requires that

$$\bar{D} \sim \text{Normal}(\mu_d, \sigma_d^2) \rightarrow t_{\text{obs}} = \frac{\bar{d} - \mu_0}{s_d / \sqrt{n}}$$

which is true when

- ▶ n is large (CLT)
- ▶ D_i are normal

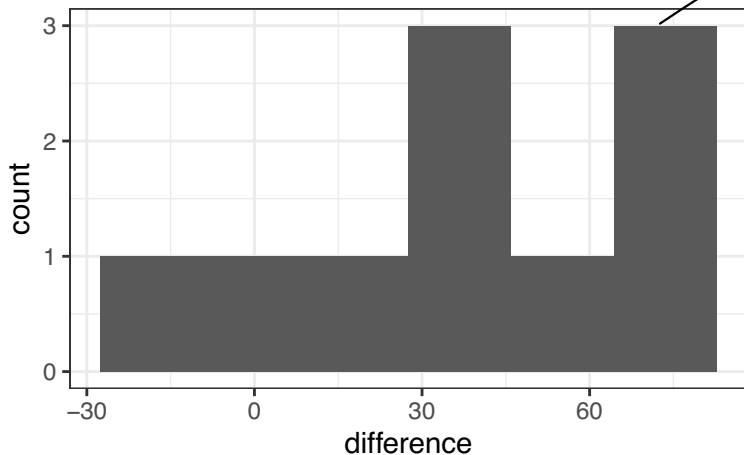
$$D_i \Rightarrow \bar{D} \sim \text{Normal}$$

Can we compare matching samples if n is small and differences are not normal?

Statistical testing for two groups: matching samples

Assume we collected the data only for 10 chickens.

~~$D_i \sim N(\dots)$~~



Statistical testing for two groups: matching samples

Let's look at the **sign of the differences**.



chicken	diet1	diet2	difference	sign
1	183.8958	158.2390	25.656797	1
2	193.4912	157.0623	36.428881	1
3	211.5568	177.4809	34.075889	1
4	241.7387	165.3641	76.374552	1
5	178.1514	180.7937	-2.642283	-1
6	240.8551	169.9080	70.947102	1
7	245.0208	178.7047	66.316034	1
8	219.4718	189.6762	29.795558	1
9	216.6203	165.2014	51.418857	1
10	165.5608	181.0978	-15.537045	-1

Statistical testing for two groups: matching samples

Signed test compares how often the difference is positive (or negative).

- ▶ New random variable $S_i = 1$ if $D_i \geq 0$ and $S_i = 0$ if $D_i < 0$
- ▶ Then $S_1, \dots, S_n \sim \text{Bernoulli}(p)$ S_i :
- ▶ If $\mu_d = 0$ then $p = 0.5$

$p =$ probability to obs positive difference

Signed test tests the hypotheses $H_0 : p = 0.5$ vs $H_a : p \neq 0.5$

- ▶ The observed statistic is $N = \sum_{i=1}^n S_i$, i.e. the **number of positive signs**
- ▶ The null distribution is $N \sim \text{Bernoulli}(n, 0.5)$
Binomial

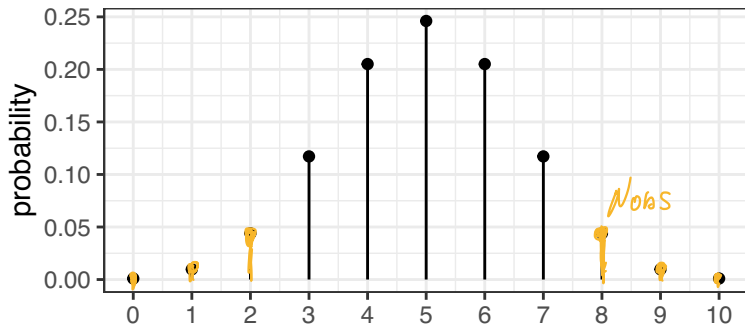
Statistical testing for two groups: matching samples

$$N_{\text{obs}} = 8$$

We observe 8/10 positive differences.

$$p\text{-value} = P(N \geq 8) + P(N \leq 2)$$

$$N \sim \text{Binomial}(10, 0.5)$$



$p\text{-value} > 0.05$
fail to reject $H_0: p = 0.5$ (null for N)
 $H_1: \mu_d = 0$

Non-matching samples: proportions

50 patients received Moderna vaccine. The hospital recorded if they got COVID during the following 6 months.

```
## [1] 0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0  
       $x_1$  . . . —  $x_{50}$ 
```

Other 30 patients received Pfizer vaccine and were monitored during the following 6 months

```
## [1] 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0  
       $y_1$  — — — —  $y_{30}$ 
```

Can we say if any vaccine is more efficient against COVID?

Non-matching samples: proportions

Let's compare the proportions of patients that got COVID.

```
table(moderna)
```

```
## moderna  
## 0 1  
## 42 8
```

$$8/50 = \bar{x}$$

```
table(pfizer)
```

```
## pfizer  
## 0 1  
## 26 4
```

$$4/30 = \bar{y}$$

Non-matching samples: proportions

We are given two samples x_1, \dots, x_n and y_1, \dots, y_m with **non-matching observations**.

- ▶ $X_1, \dots, X_n \sim \text{Bernoulli}(p)$ is the outcome for each Moderna recipient
 $p = P(\text{get Covid after Moderna})$
- ▶ $Y_1, \dots, Y_m \sim \text{Bernoulli}(q)$ is the outcome for each Pfizer recipient
 $q = P(\text{get Covid after Pfizer})$

From CLT and properties of normal distribution

$$\text{Var}(x-y) = \text{Var}(x) + \text{Var}(y)$$

$$\bar{X} - \bar{Y} \text{ approximately } \sim N\left(p - q, \frac{p(1-p)}{n} + \frac{q(1-q)}{m}\right)$$

$$\begin{aligned} \bar{X} &\sim N\left(p, \frac{p(1-p)}{n}\right) & E(x-y) &= E(x) - E(y) \\ \bar{Y} &\sim N\left(q, \frac{q(1-q)}{m}\right) & \text{Var}(-y) &= (-1)^2 \text{Var}(y) \\ & & &= \text{Var}(y) \end{aligned}$$

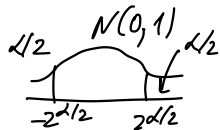
Non-matching samples: confidence intervals for proportions

Standardization:

$$Z = \frac{(\bar{X} - \bar{Y}) - (p - q)}{\sqrt{\frac{p(1-p)}{n} + \frac{q(1-q)}{m}}} \text{ approximately } \sim N(0, 1)$$

With probability $1 - \alpha$ the difference is $p - q$ in

$$(\bar{X} - \bar{Y}) \pm z^{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n} + \frac{q(1-q)}{m}}$$



Confidence interval for $p - q$: $(1 - \alpha)$ $\bar{x} \approx p$ $\bar{y} \approx q$

$$(\bar{x} - \bar{y}) \pm z^{\alpha/2} \cdot \sqrt{\frac{\bar{x}(1-\bar{x})}{n} + \frac{\bar{y}(1-\bar{y})}{m}}$$

$$\bar{x} \pm z^{\alpha/2} \sqrt{\frac{\bar{x}(1-\bar{x})}{n}}$$

Exercise

↙ quantile 2.5%

$$(\bar{x} - \bar{y}) \pm z^{\alpha/2} \cdot \sqrt{\frac{\bar{x}(1-\bar{x})}{n} + \frac{\bar{y}(1-\bar{y})}{m}}$$

Find 95% confidence interval for $p - q$ in the vaccine example.

$$\bar{x} = 8/50 = 0.16$$

$$\bar{y} = 4/30 = 0.133$$

$$\begin{aligned} CI &= 0.16 - 0.133 \pm 1.96 \cdot \sqrt{\frac{0.16 \cdot (1-0.16)}{50} + \frac{0.133(1-0.133)}{30}} \\ &= [-0.132, 0.185] \end{aligned}$$

Statistical testing for two groups: non-matching samples

Do we observe significant difference in x_1, \dots, x_n and y_1, \dots, y_m ?

Ultimate goal: test hypotheses $H_0 : p = q$ vs. $H_a : p \neq q$

$$z_{obs} = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}} \quad H_0 : p = p_0$$

Statistical testing for two groups: non-matching samples

$$Z = \frac{(\bar{X} - \bar{Y}) - \cancel{(p - q)}}{\sqrt{\frac{p(1-p)}{n} + \frac{q(1-q)}{m}}} \text{ approximately } \sim N(0, 1)$$

$$H_0: \frac{\bar{X} - \bar{Y}}{\sqrt{p(1-p) \left(\frac{1}{n} + \frac{1}{m} \right)}}$$

Test hypotheses $H_0 : p = q$ vs. $H_a : p \neq q$

- ▶ Test statistic is constructed under the null hypothesis
- ▶ Need to use both x and y samples to find a common estimate for p and q

Idea: "pool" two samples into $x_1, \dots, x_n, y_1, \dots, y_m$ and approximate both by $p, q \approx \frac{n\bar{x} + m\bar{y}}{n+m}$

$$x_1 + \dots + x_n + y_1 + \dots + y_m = n \cdot \bar{x} + m \cdot \bar{y}$$

$$n \cdot \bar{x} = \frac{x_1 + \dots + x_n}{n} \quad \cancel{n} = x_1 + \dots + x_n$$

$$m \cdot \bar{y} = y_1 + \dots + y_m$$

$p, q \approx$ average (pool sample)

$$\frac{x_1 + \dots + x_n + y_1 + \dots + y_m}{n + m}$$

Statistical testing for two groups: non-matching samples

To test hypotheses $H_0 : p = q$ vs. $H_a : p \neq q$ use test statistic

$$Z_{obs} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{n\bar{x} + m\bar{y}}{n+m} \left(1 - \frac{n\bar{x} + m\bar{y}}{n+m}\right) \left(\frac{1}{n} + \frac{1}{m}\right)}}$$

pooled \hat{p}, \hat{q} pooled \hat{p}, \hat{q}

Then, use normal table to find p-value and draw the conclusion...

Exercise

$$\alpha = 0.05$$

Use statistical testing to check if there is a difference between Pfizer and Moderna vaccines. What will change if we want to check that Pfizer has higher success rate than Moderna?

$$Z_{obs} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{n\bar{x} + m\bar{y}}{n+m} \left(1 - \frac{n\bar{x} + m\bar{y}}{n+m}\right) \left(\frac{1}{n} + \frac{1}{m}\right)}} = \frac{0.16 - 0.133}{\sqrt{0.15(1-0.15) \left(\frac{1}{50} + \frac{1}{30}\right)}} = 0.328$$

pooled p, q pooled p, q

$$\bar{x} \approx 0.16 = \frac{8}{50} \quad \bar{y} = 0.133 = \frac{4}{30}$$

$$\text{pooled } p, q = \frac{8 + 4}{50 + 30} = \frac{12}{80} = 0.15$$

$$p\text{-value} = 0.37 \cdot 2 > 0.05 \Rightarrow \text{fail to reject } H_0: p = q$$

Non-matching samples: means

A normal pregnancy can range from 38 to 42 weeks.

50 pregnant patients got COVID in the first trimester. The hospital recorded the gestational age of each born baby (in weeks).

```
## [1] 36.3 36.9 37.9 39.5 36.0 39.5
```

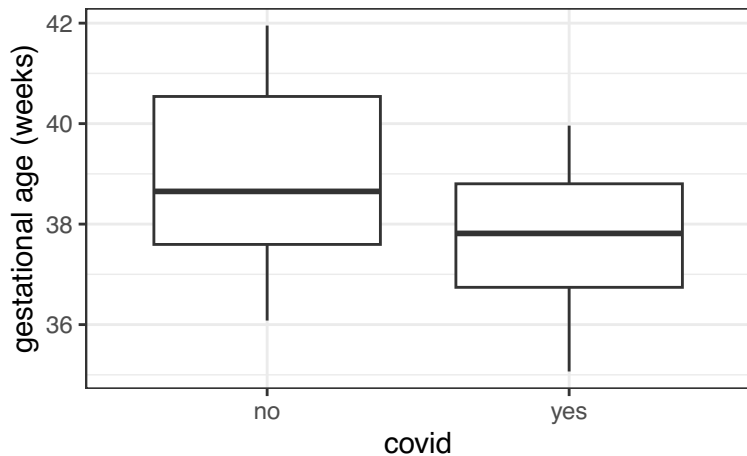
Other 30 pregnant patients did not get COVID in the first trimester.

```
## [1] 37.6 38.2 39.4 41.4 37.2 41.4
```

Does COVID lead to a preterm birth?

Non-matching samples: means

Let's also compare the boxplots.



Does COVID lead to a preterm birth?

Non-matching samples: means

- ▶ $X_1, \dots, X_n \sim (\mu_x, \sigma_x^2)$ is the outcome patients with COVID
- ▶ $Y_1, \dots, Y_m \sim (\mu_y, \sigma_y^2)$ is the outcome patients without COVID

From CLT and properties of normal distribution

$$\bar{X} - \bar{Y} \text{ approximately } \sim N\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}\right)$$
$$\bar{X} \sim N\left(\mu_x, \frac{\sigma_x^2}{n}\right) \quad \bar{Y} \sim N\left(\mu_y, \frac{\sigma_y^2}{m}\right)$$

Non-matching samples: means

Standardization:

$$\frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \text{ approximately } \sim N(0, 1)$$

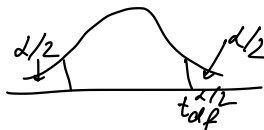
But, we do not know σ_x and σ_y

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \text{ approximately } \sim t_{df}$$

t_{n-1}

With probability $1 - \alpha$ the difference is $\mu_x - \mu_y$ in

$$(\bar{X} - \bar{Y}) \pm t_{df}^{\alpha/2} \cdot \sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}$$



Non-matching samples: confidence intervals for means

Confidence interval for $\mu_x - \mu_y$:

$$(\bar{x} - \bar{y}) \pm t_{df}^{\alpha/2} \cdot \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

Well, this is not the end of the story! You need to estimate $df \dots$:(

Non-matching samples: confidence intervals for means

Confidence interval for $\mu_x - \mu_y$:

$$(\bar{x} - \bar{y}) \pm t_{df}^{\alpha/2} \cdot \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

“Pooled” degrees of freedom are estimated as

$$df = \frac{\left(\frac{s_x^2}{n} + \frac{s_y^2}{m}\right)^2}{\frac{(s_x^2/n)^2}{n-1} + \frac{(s_y^2/m)^2}{m-1}}$$

Exercise

Find df for the pregnancy data example.

```
sd(nocovid)
```

```
## [1] 1.77168
```

```
sd(covid)
```

```
## [1] 1.361195
```

Find 95 confidence interval for $\mu_x - \mu_y$.

```
mean(nocovid)
```

```
## [1] 39.05731
```

```
mean(covid)
```

```
## [1] 37.66296
```

Statistical testing for two groups: non-matching samples

Do we observe significant difference in x_1, \dots, x_n and y_1, \dots, y_m ?

To test hypotheses $H_0 : \mu_x = \mu_y$ vs. $H_a : \mu_x \neq \mu_y$

- ▶ Use test statistic

$$t_{obs} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$$

$$\mu_x < \mu_y$$
$$\mu_x - \mu_y < 0$$

- ▶ Compute df from the ugly formula
- ▶ Compute p-value from the t-distribution table

$$P(T < t_{obs})$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \text{ approximately } \sim t_{df}$$

Exercise

Use statistical testing to check if there is a difference in gestational ages of babies for COVID and non-COVID patients. What will change if we want to check that COVID leads to a premature birth?

Exercise

```
t.test(covid, nocovid, alternative = "two.sided", mu = 0,  
       paired = FALSE)
```

```
##  
## Welch Two Sample t-test  
##  
## data: covid and nocovid  
## t = -3.7043, df = 49.505, p-value = 0.0005341  
## alternative hypothesis: true difference in means is not  
## 95 percent confidence interval:  
## -2.1505816 -0.6381185  
## sample estimates:  
## mean of x mean of y  
## 37.66296 39.05731
```

*reject $H_0: \mu_x = \mu_y$
 $H_a: \mu_x \neq \mu_y$
0.05*

Statistical testing for two groups: non-matching samples

Can we simplify computations for statistical testing by "pooling" the samples?

$$H_0: \mu_x = \mu_y \quad \text{vs} \quad H_a: \mu_x \neq \mu_y$$

Yes, but only if $\sigma_x = \sigma_y$.

Idea: "pool" two samples into $x_1, \dots, x_n, y_1, \dots, y_m$ and approximate s_x^2, s_y^2 by

$$s^2 \approx \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2} \quad (n-1)s_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2$$

Then $\frac{1}{n+m-2} \sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2$

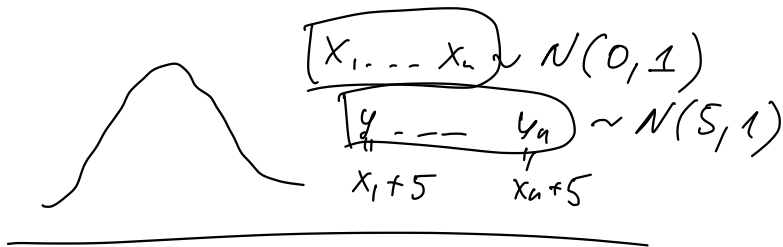
$$t_{\text{obs}} = \frac{\bar{x} - \bar{y}}{\sqrt{s^2 \left(\frac{1}{n} + \frac{1}{m} \right)}}$$

with $df = n + m - 2$.

$$t_{\text{obs}} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{\bar{x} - \bar{y}}{\sqrt{s_x^2 \left(\frac{1}{n} + \frac{1}{m} \right)}}$$

Exercise

Assuming that variances for two samples are equal, use statistical testing to check if there is a difference in gestational ages of babies for COVID and non-COVID patients.



Exercise

```
t.test(covid, nocovid, alternative = "two.sided", mu = 0,  
       paired = FALSE, var.equal = TRUE)
```

```
##  
## Two Sample t-test  
##  
## data: covid and nocovid  
## t = -3.9546, df = 78, p-value = 0.0001676  
## alternative hypothesis: true difference in means is not  
## 95 percent confidence interval:  
## -2.0963005 -0.6923995  
## sample estimates:  
## mean of x mean of y  
## 37.66296 39.05731
```

TO DO

1. Module 10. Comparing Two Groups
2. Quiz 11 due Monday (April 3) @ 11:59 PM (EST)
3. Practice Problem Set 11