I'll send you a note later today about this but...

↳ via canvas !! ↳ April 4th

↳ your late pass count is up to date through HW 6

↳ you have 9 total HWs in this class

↳ Canvas grades are up to date through:

Test 2, HW 5, ICA 16

↳ TBD - late passes + mini-project

chi-square tests, multiple comparisons

What is a scenario where we might want to do a two-tailed t-test at a significance level $\alpha = .2$?

↳ p-value to reject H_0

↳ we've looked at .1 or .05 before, → 90% + 95% confidence

↳ 80% in this case

checking a diff, rather than < or >

$H_1: \mu_1 \neq \mu_2$

t-tests - summary

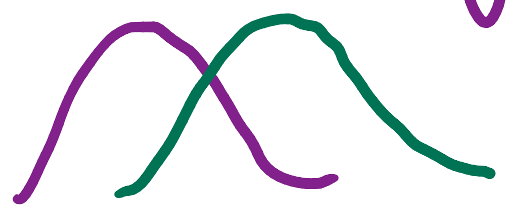
- Do "morning people" and "night people" have differences in how long they sleep?
    - two-tailed test
    - observe a sample from each population of how many hours they sleep a day
- Does tire A last longer than tire B?
    - one-tailed
    - measure tire treads for tire A vs. tire B at the same time (e.g. after one year)
    - could do multiple t-tests (see end of lecture)

t-tests

- When do we use a t-test?

↳ compare two groups*

↳ variances of the groups are similar



↳ the observations are indep.

* there are other kinds of t-tests → future stats

- These are limited circumstances—there are other kinds of tests for other scenarios

↳ chi-squared

Reading t-tables

df = 4
t value = 1.72
$\alpha = 0.05$
two-tailed

- Let's do a t-tables example (because you'll often see these referenced for other statistics as well, and it's good to know a way to tell this w/o python/excel)

$1.72 > 2.776$?

$0.2 \leq p \leq 0.1$

t-test table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**
df								
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143

degrees of freedom

t-value -> p-value

- How is the p-value really being calculated though?
  - (the answer beyond "we asked the t-test function" or "we looked it up in the t-table")

specific p-value:
1) TTEST in excel/spreadsheet
2) python + cdf function

- what % of the way through the t-val dist. am I?

Chi-squared tests

- When do we use a chi-squared test?
  - want to ask whether a certain variable follows an expected distribution

- test if heights follows a norm dist
- test if a r.v. follows a poisson dist

Chi-squared tests

t - test

→ t value

- Formula!
- $\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$
- Where
- O_i is the observed values
- E_i is the expected values

Chi-squared tests

- This is the test that we would use to answer the question:
    - "Hey! Is that a loaded die???"
    - "Hey! Is that an unfair coin???"
- Data:
    - Observations about a single random variable/attribute

Chi-squared tests

- Things that we'll have again:

  - **degrees of freedom**: number of outcomes that you could have minus 1

    ↳ t-test: d of free n - 2

  - **critical value** (this like the t-values in t-tests needed for a certain p-value): what # for χ^2 do we need to have a certain p-value (threshold)?

- ~~For the flips example:~~ If we were flipping a coin:

  - degrees of freedom: 1 (heads, tails) = 2 - 1 = 1

  - critical value: we'll pick 0.05 — we want to be 95% sure that the coin actually is unfair before calling the casino cops

    ↳ corresponding X^2 value ↳ $\alpha = 0.05$

Chi-squared tests

- Data:
  - 50 flips
  - 28 heads, 22 tails - observed values
- Null hypothesis: no significant difference between the observed and expected values

Chi-squared tests

- Data:
  - 50 flips
  - 28 heads, 22 tails - observed values
  - 25 heads, 25 tails - expected values
- $\chi^2 = \sum_i \frac{(o_i - E_i)^2}{E_i}$ -> i will be *heads* then *tails*

categorical outcomes of the var.

Chi-squared tests

- Data:
  - 50 flips
  - 28 heads, 22 tails - observed values
  - 25 heads, 25 tails - expected values → because we assume the coin is fair
- $\chi^2 = \sum_i \frac{(o_i - E_i)^2}{E_i}$ -> $\frac{(28-25)^2}{25} + \frac{(22-25)^2}{25} = \frac{18}{25} = 0.72$
- That's our chi-squared value!

Chi-squared tests

- $\chi^2 = \sum_i \frac{(o_i - E_i)^2}{E_i}$ ->

$$\frac{(28-25)^2}{25} + \frac{(22-25)^2}{25} = \frac{9}{25} + \frac{9}{25} = \frac{18}{25} = 0.72$$

$0.72 > 3.841?$
↳ no, don't reject H_0

- That's our chi-squared value! Is our number higher than our critical value?

$df = 2-1 = 1$ $\alpha = 0.05$

Chi-square (χ^2) Distribution Table						
α / df	0.1	0.05	0.025	0.01	0.005	0.001
1	2.706	$\boxed{3.841}$	5.024	6.635	7.879	10.828
2	4.605	5.991	7.378	9.21	10.597	13.816
3	6.251	7.815	9.348	11.345	12.838	16.266
4	7.779	9.488	11.143	13.277	14.86	18.467
5	9.236	11.07	12.833	15.086	16.75	20.515

Chi-squared tests

- For the flips example:
  - degrees of freedom: 1
    - Why 1 degree of freedom? → 2 options $- 1 = 1$
- critical value: we'll pick α = 0.05 — we want to be 95% sure that the coin actually is unfair before calling the casino cops, so this is the corresponding chi-squared value (3.841)
  - Threshold to reject the null hypothesis

Chi-squared tests

- Data:
  - 100 flips
  - 64 heads, 3~~7~~6 tails - observed values
  - 50 heads, 50 tails - expected values
- $\chi^2 = \sum_i \frac{(o_i - E_i)^2}{E_i}$ -> $\frac{(64-50)^2}{50} + \frac{(36-50)^2}{50} = \frac{392}{50} = 7.84$
- That's our chi-squared value!

Chi-squared tests

reject H_0 if $p < \alpha$, corresponding to $\chi^2 >$ the value associated w/ α

- Data:
  - 100 flips
  - 64 heads, 3~~7~~6 tails - observed values
  - 50 heads, 50 tails - expected values

reject H_0, this is an unfair coin

p between 0.01 + 0.005

- $\chi^2 = \sum_i \frac{(o_i - E_i)^2}{E_i}$ -> $\frac{(64-50)^2}{50} + \frac{(3\cancel{7}6-50)^2}{50} = \frac{196}{50} + \frac{196}{50} = \frac{392}{50} = 7.84$

Chi-square (χ^2) Distribution Table

α / df	0.1	0.05	0.025	0.01	0.005	0.001
1	2.706	$\boxed{3.841}$	5.024	6.635	7.879	10.828
2	4.605	5.991	7.378	9.21	10.597	13.816
3	6.251	7.815	9.348	11.345	12.838	16.266
4	7.779	9.488	11.143	13.277	14.86	18.467

ICA Question 2: chi-squared

Say you have 36 4-sided dice.

```
import random
rolls = [random.randint(1, 4) for i in range(36)]
```

$$\frac{0}{9} + \frac{1}{9} + \frac{0}{9} + \frac{1}{9} = \frac{2}{9} =$$

0.222

Fill in the following table, then calculate your chi-squared value:

	1	2	3	4
Expected	9	9	9	9
Observed	9	8	9	10

ICA Question 2: chi-squared

Say you have 36 4-sided dice.

d.o.f = 3

`If our chi-squared value is:` 12.32 → reject.

`If our chi-squared value is:` 0.2245 → don't reject!

`If our chi-squared value is:` 11.333 → don't reject!

Do we reject the null hypothesis? (if α = 0.01)

Chi-square (χ^2) Distribution Table

α / df	0.1	0.05	0.025	0.01	0.005	0.001
1	2.706	3.841	5.024	6.635	7.879	10.828
2	4.605	5.991	7.378	9.21	10.597	13.816
3	6.251	7.815	9.348	11.345	12.838	16.266
4	7.779	9.488	11.143	13.277	14.86	18.467
5	9.236	11.07	12.833	15.086	16.75	20.515

Calculating Chi-squared in python

- What are we actually calculating here?
- The percentage of the way that we are through a chi-square distribution
  - (just like in a t-test we calculate the % of the way that you are through a t-value distribution)

Chi-squared and some "real world" (non casino) data

- Where are chi-squared tests used in the real world?
  - There is equal number of riders ride the Orange Line each weekday.
  - The relative species distribution for 3 sub-species of bees in Massachusetts is x%, y%, z%.
  - The number of honks that Felix hears on their way to work follows a poisson distribution.

↳ you do need to be able to eval. categories
  ↳ translate into buckets

Chi-squared and some "real world" (non casino) data

- Where are chi-squared tests used in the real world?
  - There is equal number of riders ride the Orange Line each weekday.
  - The relative species distribution for 3 sub-species of bees in Massachusetts is x%, y%, z%.
  - The number of honks that Felix hears on their way to work follows a poisson distribution.
- (we can also use a slightly different chi-squared test to determine if two variables are independent)

Amount of data needed

- Chi-squared
    - ~30 data points

- T-tests
    - No minimum sample size
    - When you get > 40 samples, other tests become more appropriate

ICA Question 3: Chi-squared tests

Say we want to know if the number of goals scored in a game of soccer follows a Poisson distribution where $\lambda = 1$ (number of goals/game).

You observe the following total goal counts for 5 games: 1, 1, 2, 0, 1.

How would you do a chi-squared test to calculate the p-value for games with 0, 1, and 2 goals scored?

any game

$$p(x=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

ICA Question 3: Chi-squared tests

You observe the following total goal counts for 5 games: 1, 1, 2, 0, 1.

cats	0	1	2	3+
Observed	1	3	1	0
Expected	$P(X=0)\times 5$			

\# of games

value from table χ^2

$\chi^2 > 4.605$

$P(X=0) \Rightarrow$ prob that one game has 0 total goals

$\alpha = .1$, deg. of freedom $= 4 - 1 - 1$ ← estimating λ

3+ column

→ 2 total d.o.f.

Multiple Comparisons (Bonferonni Correction)

- Family-wise error (for t-tests): probability of making one or more false positives (type 1 errors) when performing multiple t-tests
- We want to know whether or not using a certain fertilizer increases our crop yield on our spinach farm.
- Each week, we measure the crops in two fields and perform a t-test to determine whether no fertilizer or fertilizer is better.
- We'd like to control the Family-Wise Error rate to be under 0.1

Multiple Comparisons (Bonferonni Correction)

- We'd like to control the Family-Wise Error rate to be under 0.1 **and we have 13 weeks of data**

- If each week's t-test has a p-value that is under our chosen threshold of 0.1, what is the probability that we've made at least one type 1 error?

- $1-(1-\text{prob of type 1 error})^{\#\text{ of tests}} = 1-(1-.1)^{13} = 0.746$

- Now, we'll adjust our weekly p-value cutoff so that we can guarantee that the family-wise error rate is not above .1

- $\dfrac{\text{desired fw error rate}}{\#\text{ of tests}} = \dfrac{0.1}{13} = 0.0077$

  Bonferonni correction

Multiple Comparisons (Bonferonni Correction)

- In summary:
- When doing multiple significance tests, to guarantee a Family Wise Error rate at a certain level, we need to increase the threshold of confidence on each individual test
- $\alpha_{bonferonni} = \dfrac{\alpha_{orginal}}{n}$
- Where α are the significance levels needed and n is the number of tests that will be happening

Admin

- Four weeks in the semester left (schedule on next page)

- **Test 4:** will be **May 4th, 1 - 3pm**, in this room (this is during your final exam period scheduled by the university registrar)

  - **Change**: this test will not be cumulative, it will only* cover HW 7 -9

  - *don't be surprised if previous topics are referenced or built-upon, but there won't be questions that focus specifically on the HW 1 - 6 material

- This week would be a *great week* to get your mini-project done

  ↳ gap in HWs this week

Mini project questions

- This week would be a *great week* to get your mini-project done
  - Mini-project question:
    - Talk to a group by you—what are they thinking of doing for their mini project?

ICA passcode: "chi"

Schedule

HW8 will be released on Thursday

Turn in **ICA 19** on Canvas (make sure that this is submitted by 2pm!) - passcode is "chi"

Test 4 is May 4th @ 1-3 pm in Snell Eng. 108

Mon	Tue	Wed	Thu	Fri	Sat	Sun
April 4th Lecture 19 - chi-square test, multiple comparison correction	**Felix OH Calendly**	**Felix OH Calendly**	**Felix OH Calendly** Lecture 20 - covariance, correlation			
April 11th Lecture 21 - conditional probabilities, bayes	**Felix OH Calendly**	**Felix OH Calendly**	**Felix OH Calendly** Lecture 22 - conditional ind., bayes nets			**HW 8 due @ 11:59pm**
April 18th **No lecture - Patriot's Day**	**Felix OH Calendly**	**Felix OH Calendly**	**Felix OH Calendly** Lecture 23 - Regression: R^2 & F			
April 25th Lecture 24 - presentations, wrap-up **Mini-project due @ 11:45am**		**HW 9 due @ 11:59pm**				

Test 4

More recommended resources on these topics

- Chi-squared Test: YouTube, Bozeman Science | Chi-squared Test
- Family-Wise Error Rate & Bonferonni Correction:
  - https://riffyn.com/blog/family-wise-error-rate
  - https://www.statology.org/family-wise-error-rate/
- Amount of data for a t-test: https://stats.stackexchange.com/questions/37993/is-there-a-minimum-sample-size-required-for-the-t-test-to-be-valid
- t-table: https://cdn1.byjus.com/wp-content/uploads/2020/04/T-table.png