DATA ANALYSIS

Week 14: Dependent Data

logistics

- cumulative final on May 3rd (next Friday)
- ALL practice materials are up on Canvas
- same format (total worth 20% of final grade)
  - conceptual (40 points)
  - computational (60 points)
  - 25% if you opt-out of PS7
- all videos up on course website

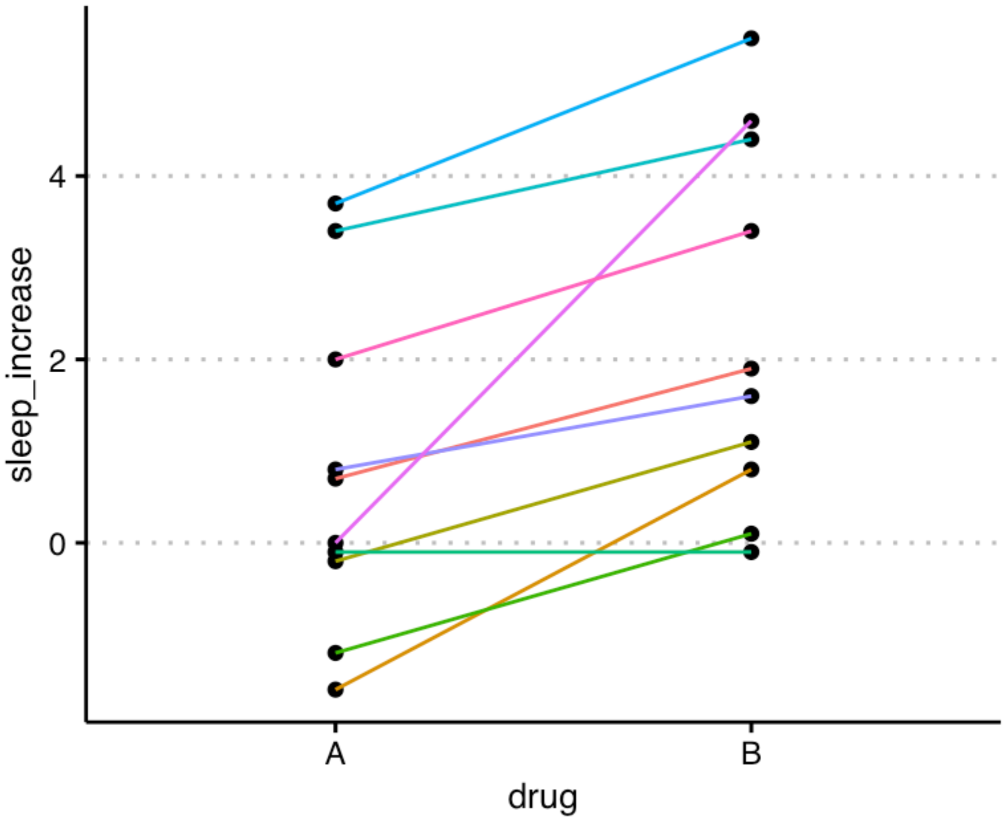
▾ Cumulative Final PRACTICE

- Weeks 1-5 Practice
- Weeks 6-12 Practice
- Weeks 13-15 Practice
- Practice Final (Conceptual)
  40 pts
- Practice Final (Computational)

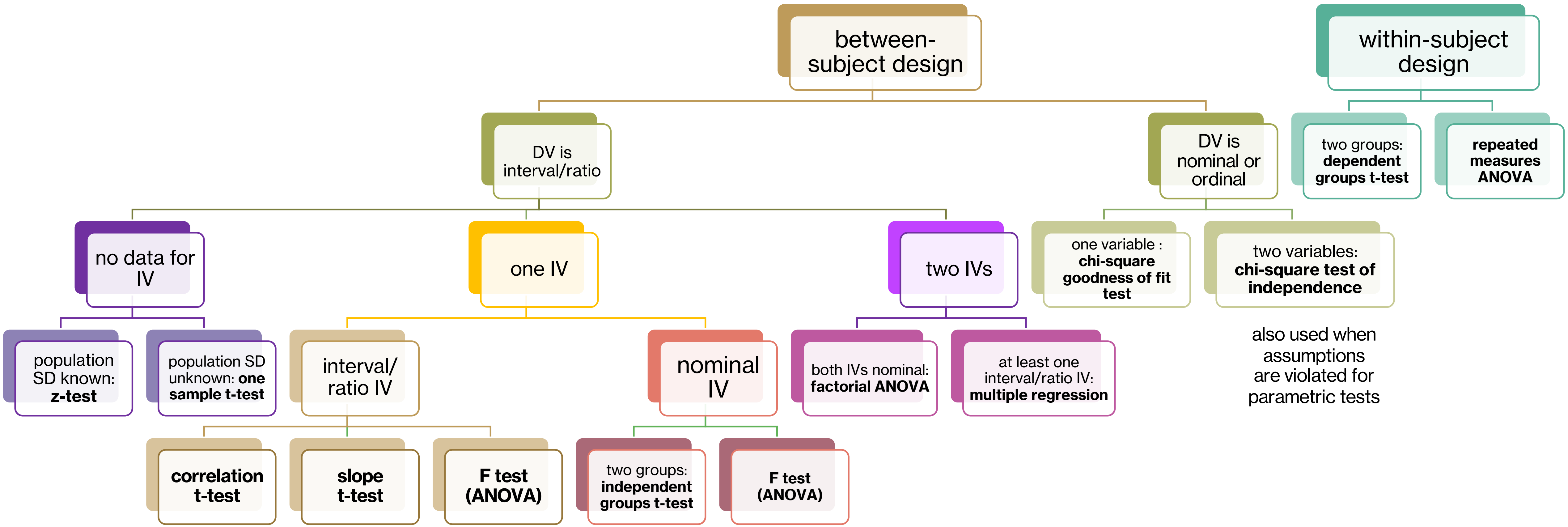
14	W: April 24, 2024	W14: Miscellaneous Data
14	F: April 26, 2024	W14 continued…
15	T: April 30, 2024	**Problem Set 7 due / Opt-out Deadline**
15	W: May 1, 2024	W15: Odds and Ends
15	T: May 2, 2024	**Data Around Us / Practice Questions due**
15	F: May 3, 2024	**Conceptual Final (In Class)**
16	T: May 7, 2024	**Computational Final Computational due**
16	T: May 7, 2024	**Last Class Survey due**
16	W: May 8, 2024	**Wrapping Up!** (Last Class)
17	T: May 14, 2024	**PS7 Revisions due**
17	M: May 14, 2024	**ALL late work due**

sleep dataset

- data about “the effect of two soporific drugs (increase in hours of sleep compared to control) on 10 patients”
- this dataset contains repeated observations from the same patient and therefore, the data are not independent
  - also called a within-subject or within-participant design
- we have not covered any statistical tests that we can use to analyze such data! ☹



final hypothesis chart



assuming independence

- how would we have proceeded if the data were independent?
- we have 2 groups of scores, so we could have conducted an independent groups t-test

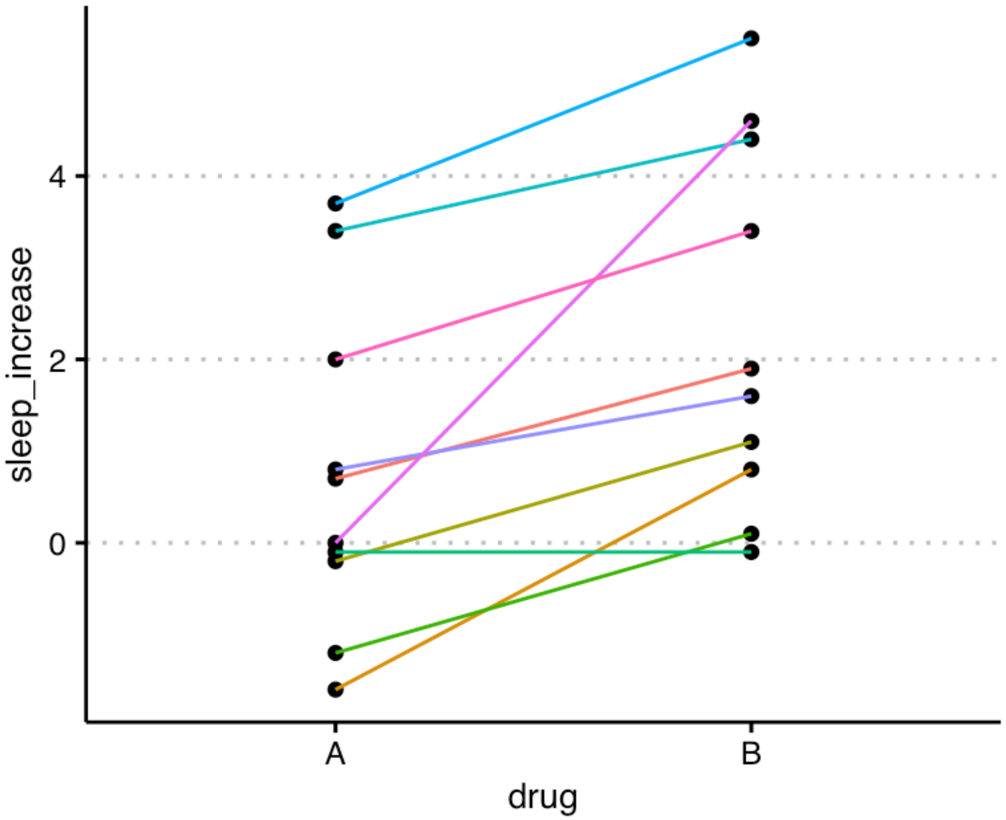
$$t_{df} = \frac{sample\ statistic\ (b)\ - population\ parameter\ (\beta)}{standard\ error} = \frac{(M_2 - M_1)\ - 0}{s_{M_2 - M_1}}$$

$$s_{M_2 - M_1} = \sqrt{\frac{{s_p}^2}{n_1} + \frac{{s_p}^2}{n_2}}$$

- we could also conduct the F-test where we substitute the group means and compare to the grand mean model

paired / dependent groups t-test

- similar idea but now we compute difference scores for each participant
  - $D = X_A - X_B$
  - if the drug had no effect on a participant, what should the value of D be?
- compute the mean of these differences
  - $M_D = \frac{\sum D}{n}$
  - if the drug had no effect overall, what should M_D be?



hypothesis testing (paired t-test)

- **step 1: state the hypotheses**
  - H_0: $\mu_D = 0$
  - H_1: $\mu_D \neq 0$
- **step 2: set criteria for decision**

$$t_{df} = t_{critical}$$

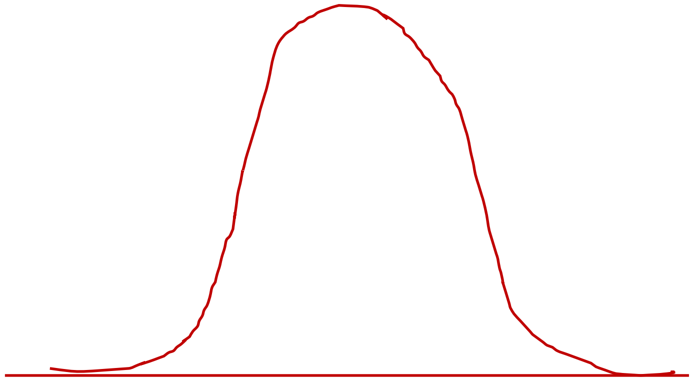
- **step 3: collect data**

$$t_{observed} = \frac{sample\ statistic\ (M_D)\ - population\ parameter\ (\mu_D)}{standard\ error}$$

$$t_{observed} = \frac{M_D - \mu_D}{SE} = \frac{M_D - \mu_D}{s_{M_D}}$$

- **step 4: make a decision!**

null hypothesis
sampling distribution
of ALL mean differences



NHST for two dependent groups (paired t-test)

step 1: state the hypotheses

$H_0: \mu_D = 0$
$H_1: \mu_D \neq 0$

step 2: set criteria for decision

$\alpha = .05$
find $t_{critical}$ based on one vs. two tailed test and degrees of freedom
$df = n - 1$

step 3: collect data

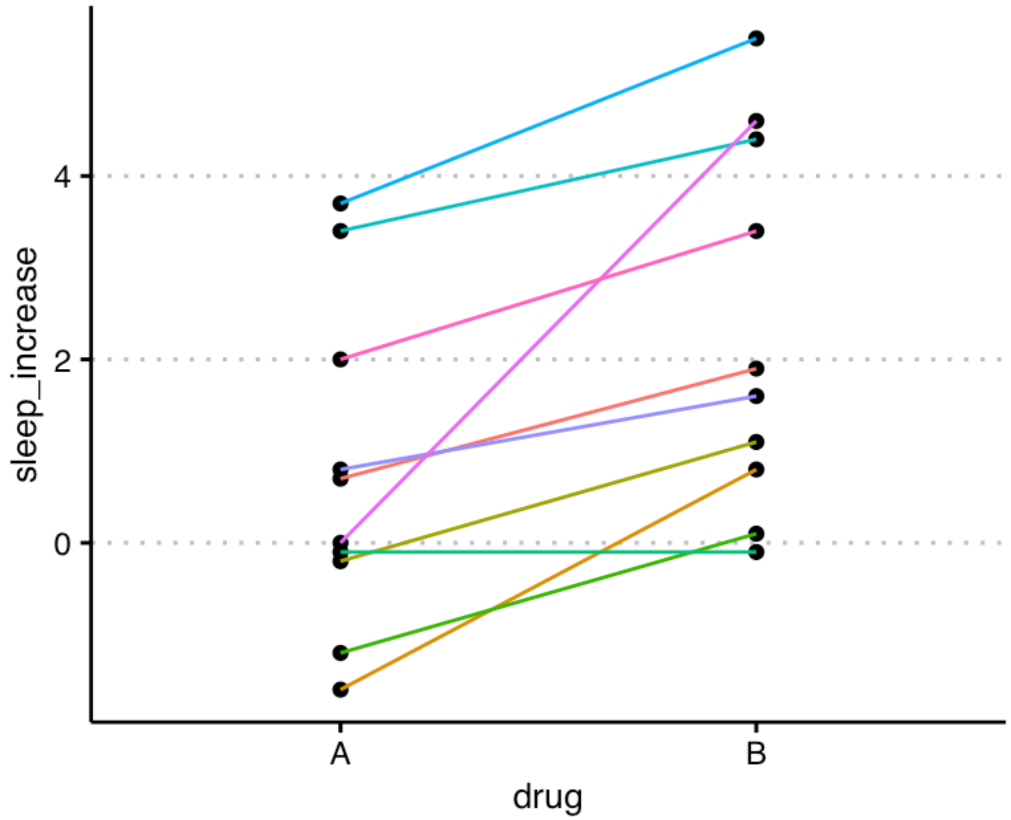
(1) compute $s_{M_D} = \frac{s_D}{\sqrt{n}}$
(2) compute $t_{observed} = \frac{M_D - \mu_D}{s_{M_D}}$
(3) find p-value for t-score

step 4: make a decision!

check whether $t_{observed}$ is beyond $t_{critical}$ and p-value < .05. if so, reject null hypothesis!

activity: conduct the t-test

- sleep data
- find $t_{critical}\ (n-1)$ = ± 2.2621
- compute the differences for each participant
- compute M_D = -1.58
- compute $s_{M_D} = \frac{s_D}{\sqrt{n}}$ = 0.388
- compute $t_{observed} = \frac{M_D - \mu_D}{s_{M_D}} = -4.062$
- compute p-value = 0.0028
- decide!



Sheets solution
Video tutorial

effect size for paired t-test

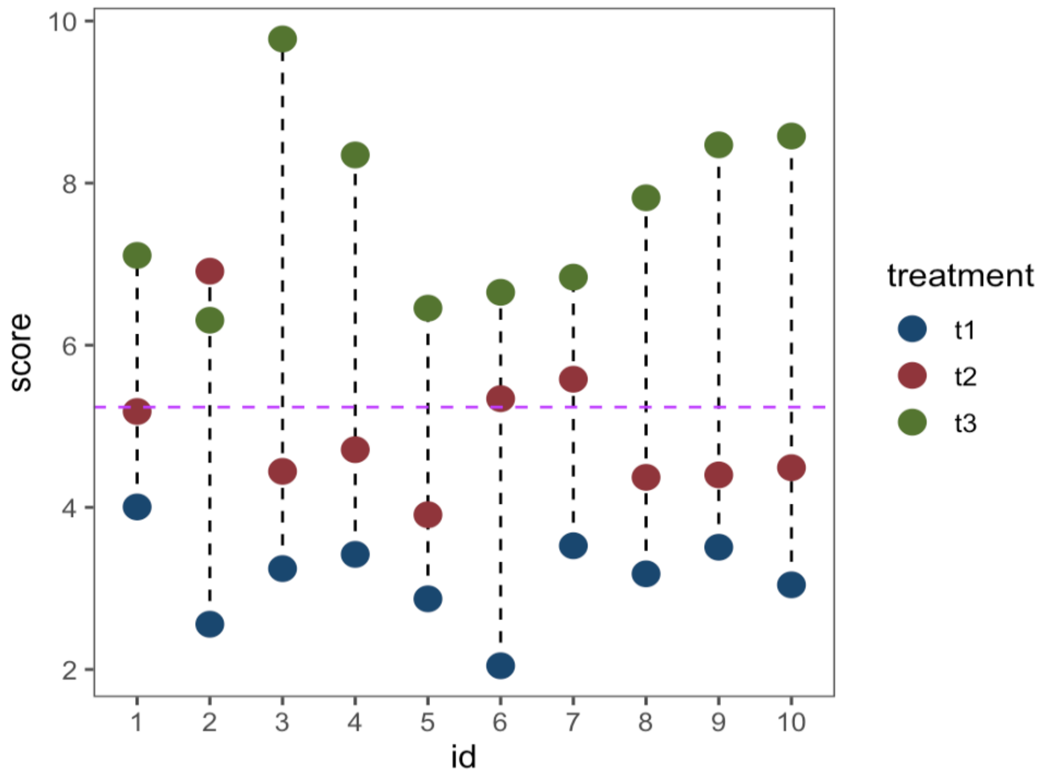
- effect sizes represent how extreme is the mean difference relative to the "standard" difference that is to be expected under the null hypothesis
- $d = \frac{\mu_D}{\sigma_D}$
- when the original population standard deviation is unknown, we estimate it using the sample standard deviation of mean differences
- estimated $d = \frac{M_D}{s_D}$

more than two groups

- when more than two levels of the independent variable have to be compared, we cannot use a paired/dependent groups t-test
- what have we done before in this situation?
- we conduct the F-test!
- we have SS_{total} = $SS_{model} + SS_{error}$
- we will account for additional variance explained due to the same participant being measured to ultimately reduce SS_{error} further

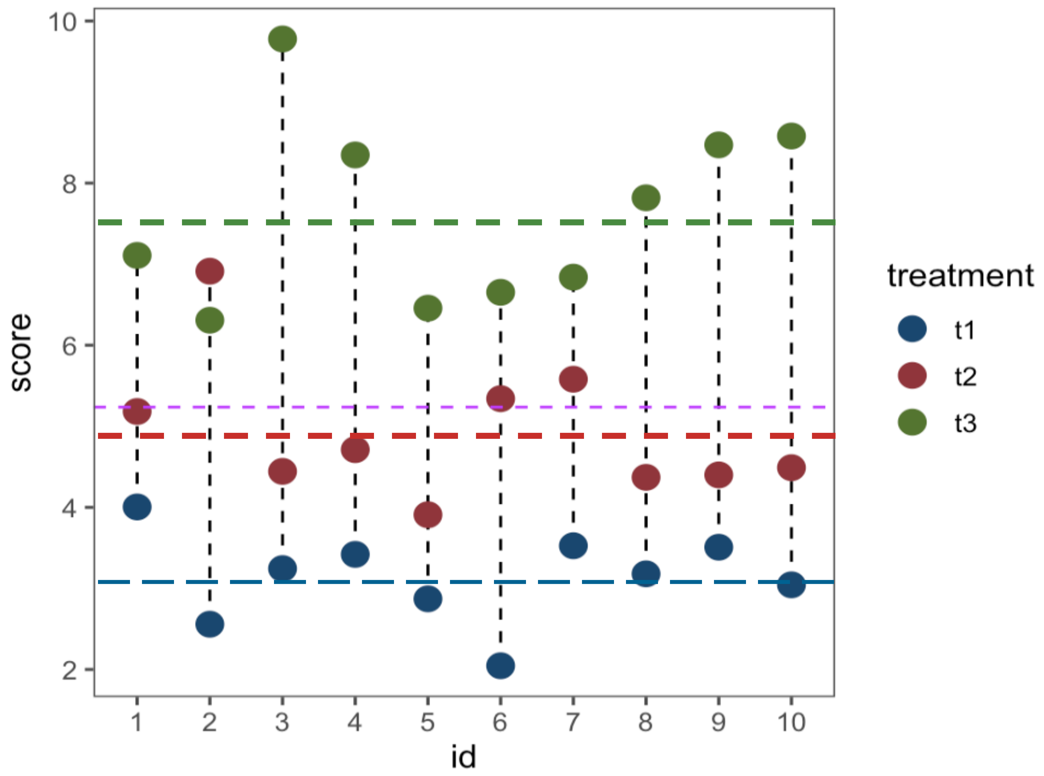
self esteem data

- the self-esteem dataset in R contains results from an experiment comparing self-esteem scores from a group of participants who were all exposed to three different treatment conditions
- **research question**: are there differences in self-esteem scores across treatments?
- how do we start building a model?



repeated-measures F-test

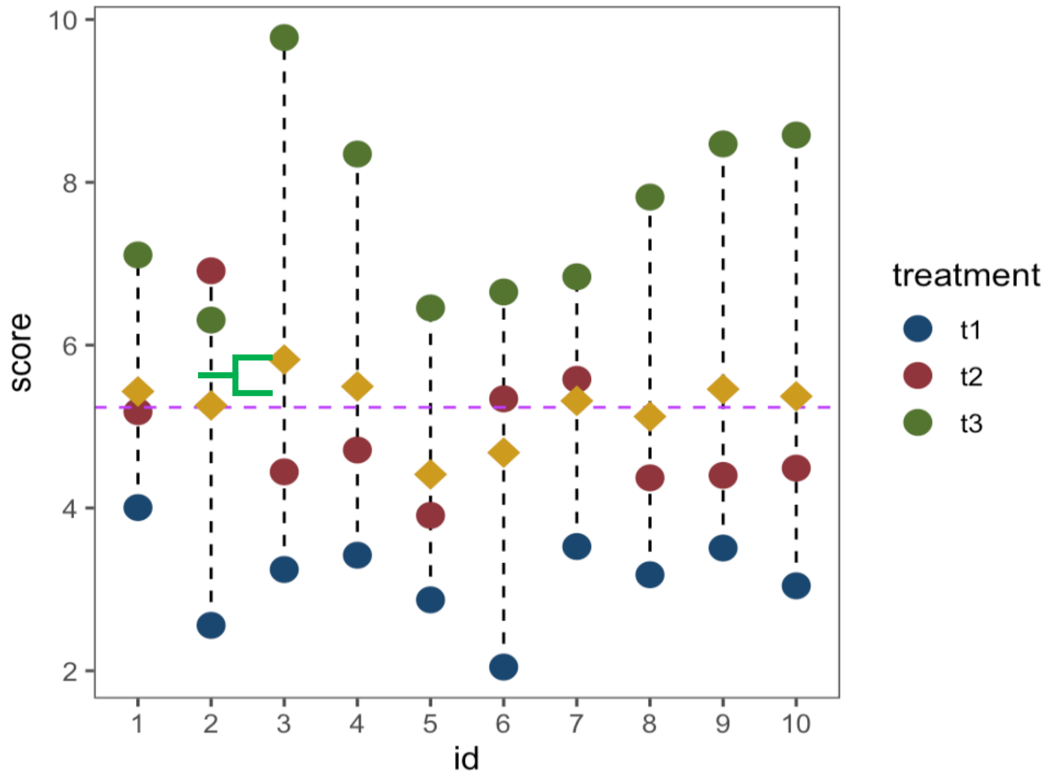
- step 1: grand mean model
  - compute grand mean = M_y = 5.2368
  - obtain $SS_{total} = \sum(Y - M_y)^2$ = 123.65
- step 2: treatment mean model
  - get $\hat{Y}$ by substituting treatment means
    - $M_{t1} = 3.14, M_{t2} = 4.93, M_{t3} = 7.64$
  - obtain $SS_{treatment\_error} = \sum(Y - \hat{Y})^2$ = 21.19
  - obtain
  - $SS_{treatment\_model} = SS_{total} - SS_{treatment\_error}$ = 102.46



data

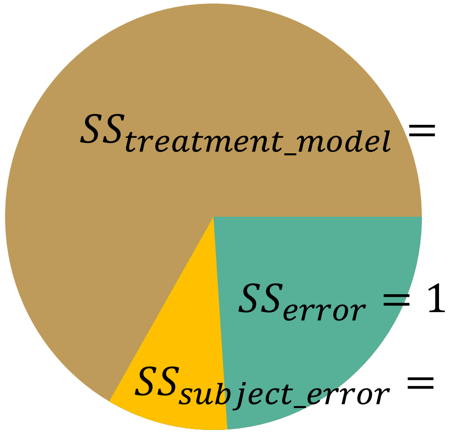
building a subject-level model

- our goal is to further reduce $SS_{treatment\_error}$ by utilizing information about the subject
- we start by calculating a mean for each subject $M_{subject_i}$
- next, we look at how much we gain by using a subject-level mean relative to the grand mean
    - $SS_{subject} = \sum k \ (M_{subject_i} - M_y)^2$
    - k: number of levels of IV

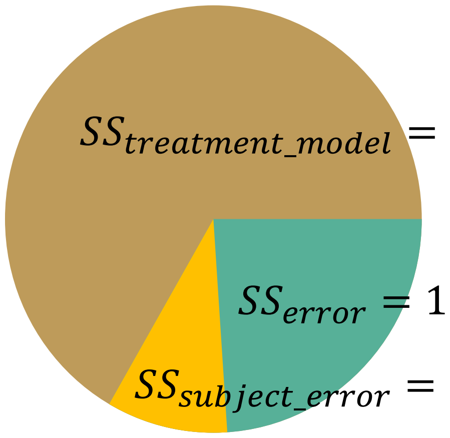


factoring out $SS_{subject}$

- step 1: from grand mean model
  - $SS_{total} = \sum(Y - M_y)^2$ = 123.65
- step 2: from drug mean model
  - $SS_{treatment\_error} = \sum(Y - \hat{Y})^2$ = 21.19
  - $SS_{treatment\_model} = SS_{total} - SS_{treatment\_error}$ = 102.46
- step 3: subject-level model
  - $SS_{subject} = \sum k\ (M_{subject_i} - M_y)^2$ = 4.57
- step 4: remove this estimate from remaining error
  - final SS_{error} = $SS_{treatment\_error} - SS_{subject}$ = 16.62373649



F table



- n: number of observations (data points)
- $n_{subjects}$: number of subjects or participants

	SS	df	MS	F	p-value
between-subjects (treatment)	102.46	$k-1=3\ -1=2$	51.23	55.47	<.001
within-subjects					
• subject error	4.57	$n_{subjects}-1=10-1=9$			
• residual error	16.62	$(k-1)(n_{subjects}-1)=(2)(9)=18$	0.92		
total	123.65				

Sheets solution
Video tutorial

NHST for repeated measures ANOVA

step 1: build grand mean model

(1) "summarizing" data using a single grand mean (ignoring all group labels)

(2) compute

$SS_{total} = \sum (Y - M_y)^2$

step 2: build group means model

(1) find group means M_{group}

(2) compute $SS_{model\_error} = \sum (Y - M_{group})^2$

(3) compute $SS_{model} = SS_{total} - SS_{model\_error}$

step 3: build subject means model

(1) find subject-level means ($M_{subject_i}$)

(2) compute $SS_{subject} = \sum k \, (M_{subject_i} - M_y)^2$

k: number of levels of IV

(3) compute final $SS_{error} = SS_{model\_error} - SS_{subject}$

step 4: conduct F test

(1) create F table

(2) find $F_{critical}$

(3) compute $F_{observed} = \frac{MS_{model}}{MS_{error}}$

(4) find p-value for F-score

(4) decide!

RM-ANOVA assumptions

- interval/ratio dependent variable
- normality
- sphericity: the variances of the differences between all possible pairs of within-subject conditions are equal
  - Mauchly's test is typically performed to test for sphericity

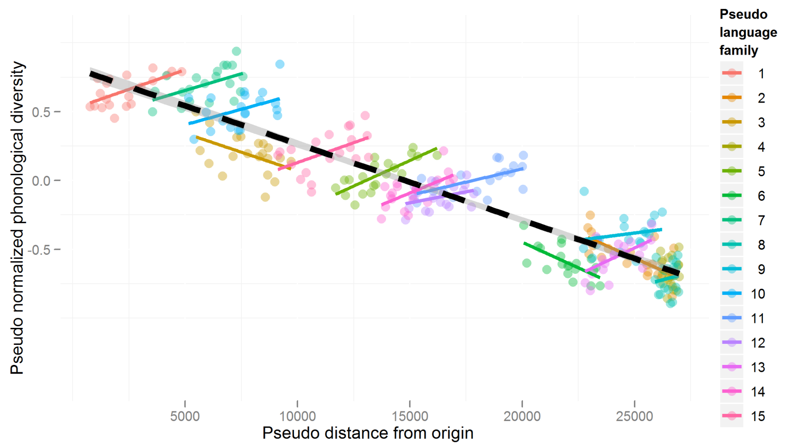
id	t1	t2	t3	t1-t2	t1-t3	t2-t3
1	4.005027	5.182286	7.107831	-1.177259	-3.102804	-1.925545
2	2.558124	6.912915	6.308434	-4.354791	-3.75031	0.604481
3	3.244241	4.443434	9.77841	-1.199193	-6.534169	-5.334976
4	3.419538	4.711696	8.347124	-1.292158	-4.927586	-3.635428
5	2.871243	3.908429	6.457287	-1.037186	-3.586044	-2.548858
6	2.045868	5.340549	6.653224	-3.294681	-4.607356	-1.312675
7	3.525992	5.580695	6.840157	-2.054703	-3.314165	-1.259462
8	3.179425	4.370234	7.818623	-1.190809	-4.639198	-3.448389
9	3.507964	4.399808	8.471229	-0.891844	-4.963265	-4.071421
10	3.043798	4.489376	8.5811	-1.445578	-5.537302	-4.091724
				var	var	var
				1.303951124	**1.155305965**	**3.081987038**

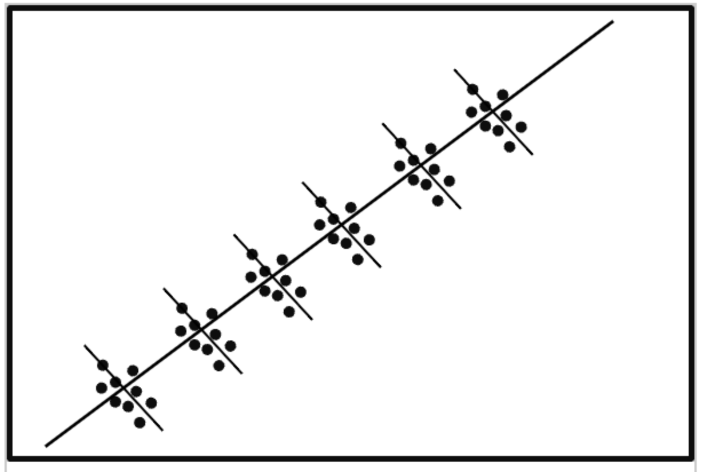
ANOVA limitations

- require simple designs, complete data, and normal residuals
- not equipped to handle missing data
- difficult to accommodate differing number of repeats/trials
- cannot capture nested/clustered designs

mixed models

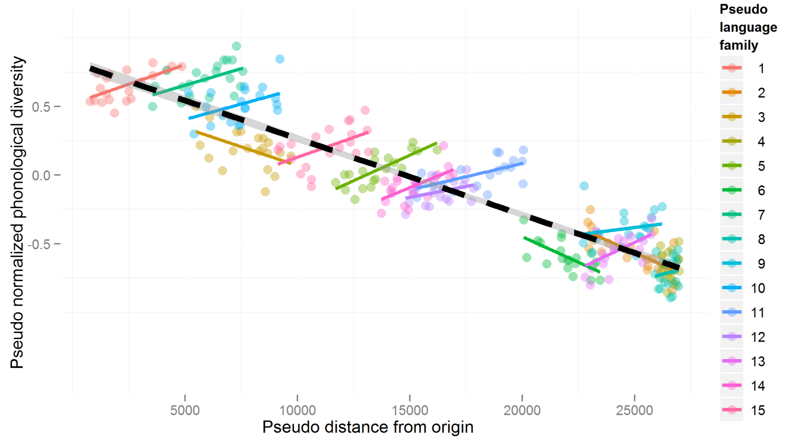
- linear/generalized mixed effects models consider the variability due to:
  - missing data
  - categorical/continuous IVs and DVs
  - unbalanced designs
  - clustered designs (no collapsing into means)
- think of them as the parent models from which special cases such as t-tests and ANOVAs are derived
- different ‘lines/curves’ are fit for each individual and for each item, with their own slope and intercept, instead of “averaging” across everyone

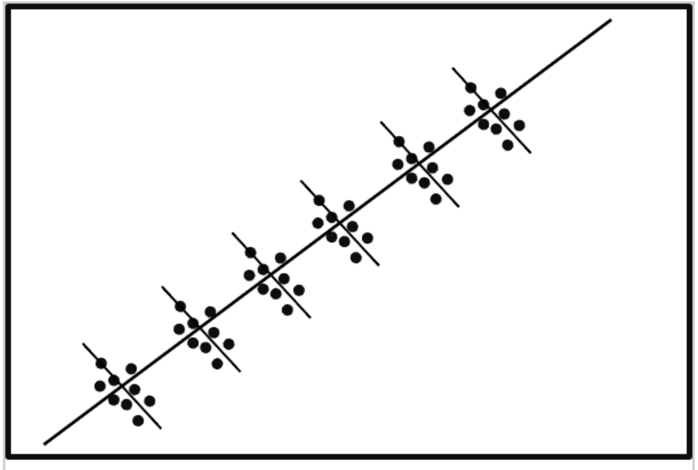




mixed effects models

- are appropriate when data are nested – when several units or levels of analysis are possible and their separate and joint influences need to be considered
- common alternative terms are multilevel models, random coefficient models, and hierarchical linear models
- data structures suitable for mixed models arise in a wide variety of common research problems
  - classes within schools within states
  - trials within subjects within age groups



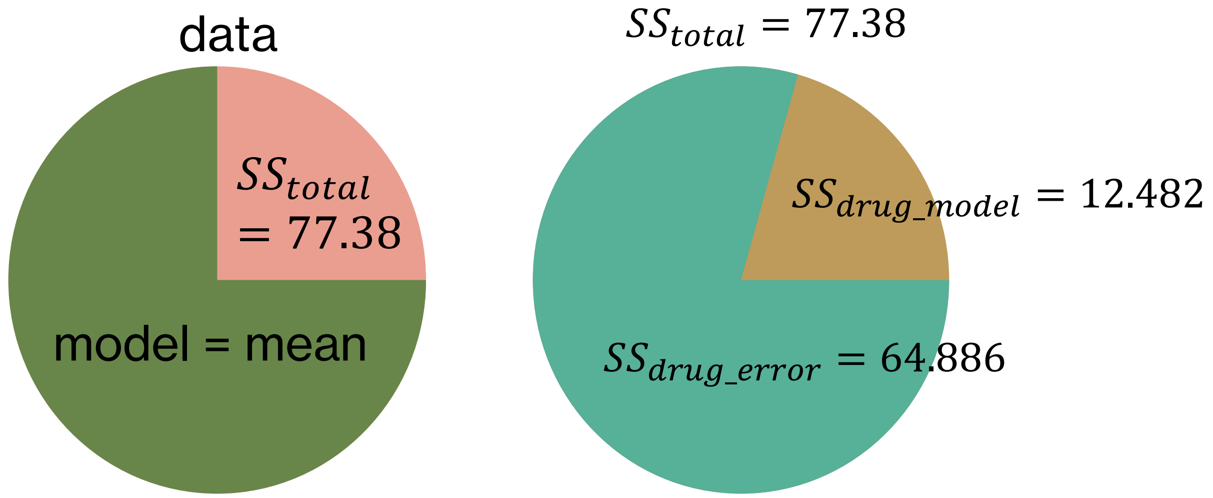


bonus content

- the following slides describe how the F-test would be conducted for the sleep data with two within-subject conditions
- NOTE: as we saw, you can do a paired t-test for these data but the same ideas of F-test will also apply here and the F-test is doable here too

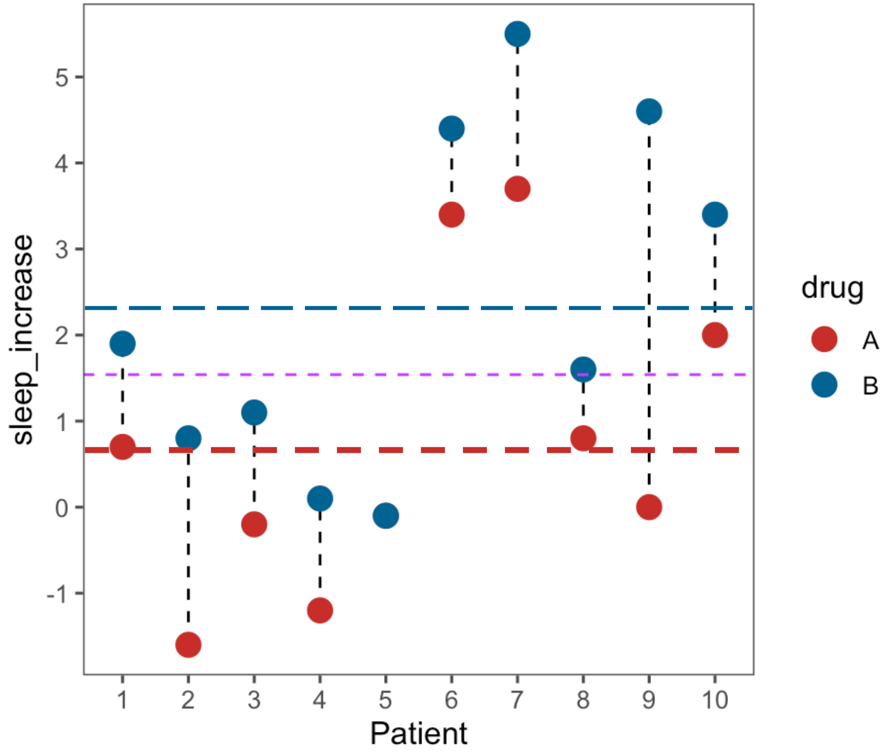
F-test for sleep data

- let's start building our model(s)
- step 1: grand mean model
  - compute grand mean = M_y = 1.54
  - obtain $SS_{total} = \sum(Y - M_y)^2$ = 77.368
- step 2: drug mean model
  - get $\hat{Y}$ by substituting drug means
    - $M_{drugA} = 0.75, M_{drugB} = 2.33$
  - obtain $SS_{drug\_error} = \sum(Y - \hat{Y})^2$ = 64.886
  - obtain $SS_{drug\_model} = SS_{total} - SS_{error}$ = 12.482



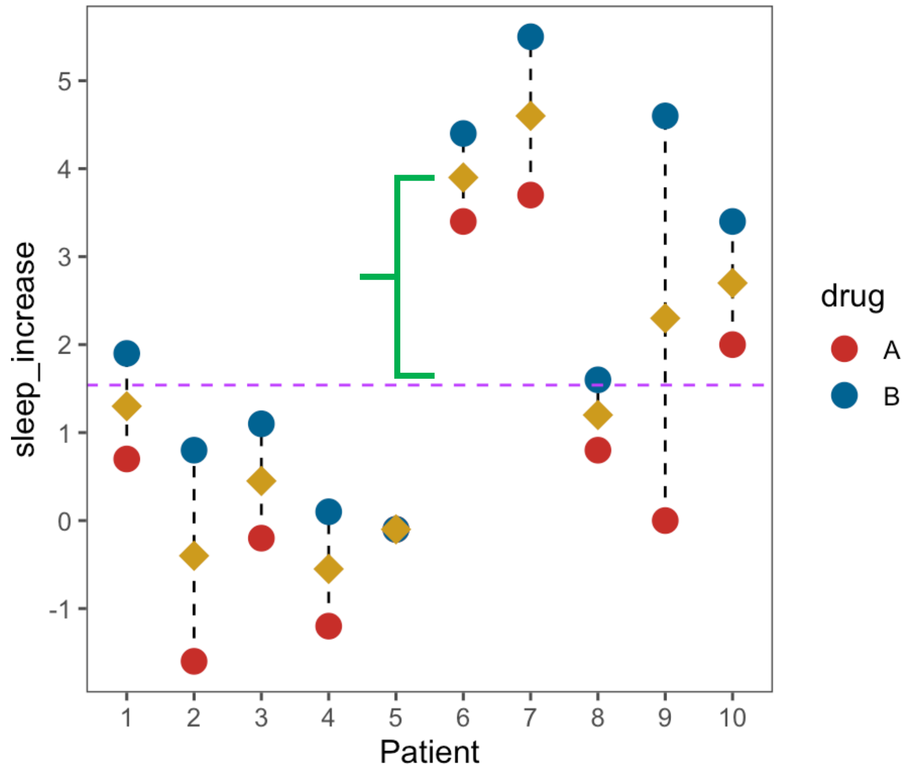
F-test for sleep data

- let's start building our model(s)
- step 1: grand mean model
  - compute grand mean = M_y = 1.54
  - obtain $SS_{total} = \sum(Y - M_y)^2$ = 77.368
- step 2: drug mean model
  - get $\hat{Y}$ by substituting drug means
    - $M_{drugA} = 0.75, M_{drugB} = 2.33$
  - obtain $SS_{drug\_error} = \sum(Y - \hat{Y})^2$ = 64.886
  - obtain $SS_{drug\_model} = SS_{total} - SS_{error}$ = 12.482



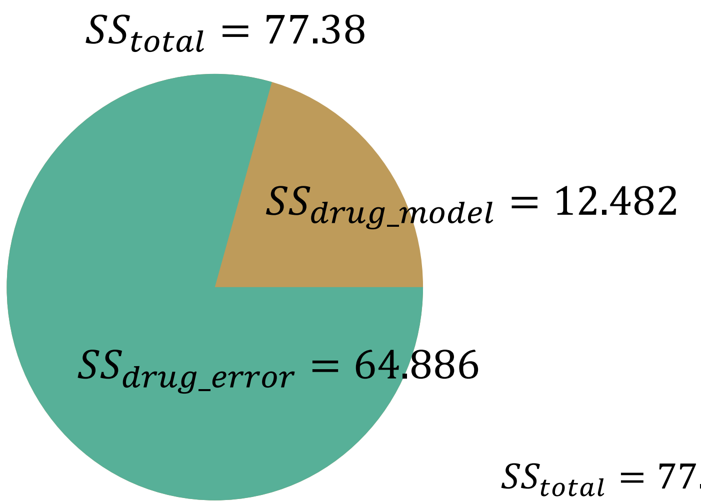
building a subject-level model

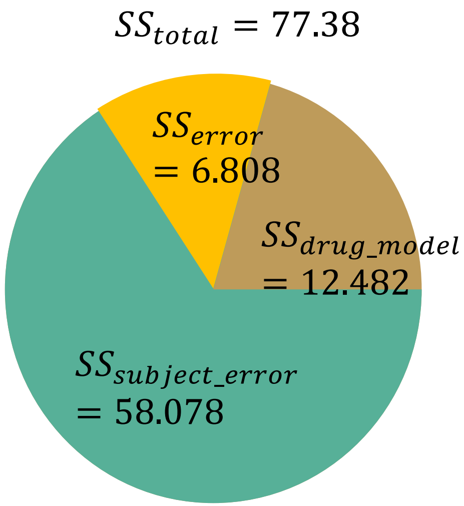
- our goal is to further reduce SS_{error} by utilizing information about the subject that is implicit in our drug model
- we start by calculating a mean for each subject $M_{subject_i}$
- next, we build a model that substitutes each value (Y_i) with the subject-level mean ($M_{subject_i}$)
  - essentially, we are looking at how much we gain by using building a model at the level of the subject
- we do this for ALL subjects across ALL groups in our data and then look at how much "error" is explained by this subject-level model relative to the grand mean
  - $SS_{subject} = \sum n \ (M_{subject_i} - M_y)^2$



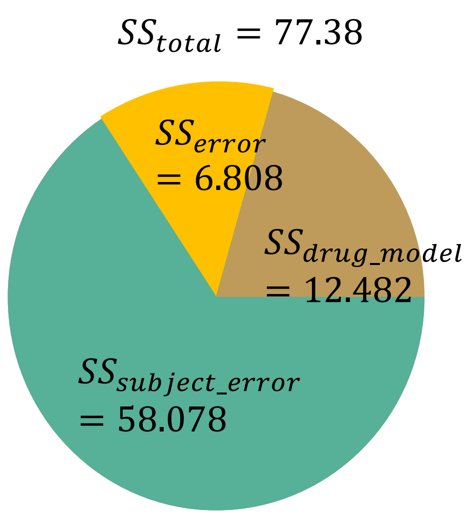
factoring out $SS_{subject}$

- step 1: from grand mean model
  - $SS_{total} = \sum(Y - M_y)^2$ = 77.368
- step 2: from drug mean model
  - $SS_{drug\_error} = \sum(Y - \hat{Y})^2$ = 64.886
  - $SS_{drug\_model} = SS_{total} - SS_{error}$ = 12.482
- step 3: subject-level model
  - $SS_{subject\_error} = \sum n\ (M_{subject_i} - M_y)^2$ = 58.078
- step 4: remove this estimate from remaining SS_{error}
  - final $SS_{error} = SS_{drug\_error} - SS_{subject\_error}$ = 6.808





F table



- n: number of observations (data points)
- $n_{subjects}$: number of subjects or participants

	SS	df	MS	F	p-value
between-subjects (drug)	12.482	$k - 1 = 2 - 1 = 1$	12.482	16.50	.0028
within-subjects					
• subject error	58.08	$n_{subjects} - 1 = 10 - 1 = 9$			
• residual error	6.808	$(k - 1)(n_{subjects} - 1) = 9$	0.756		
total	77.38				

next time

- **before** class
  - *watch:* Dependent Groups t-test [6 min]
  - *watch:* Repeated Measures ANOVA [16 min]
  - *work on:* Problem Set 7!
  - *post*: Data Around Us OR practice questions (class participation)
- **during** class
  - chi-square tests