EPSE 592: Design & Analysis of Experiments

Ed Kroc

University of British Columbia

ed.kroc@ubc.ca

February 13, 2020

Last Time

- Assumptions of simple fixed effects ANOVA models
- ANOVA model diagnostics
- Two-way (one factor) fixed effects ANOVA model

Today

- More on two-way (one factor) fixed effects ANOVA model
- Interpreting two-way interactions
- Generic *n*-way ANOVA models

Two-way (two factor) ANOVA, with interaction

	Female	Male
Bachelor's	6.2, 6.6, 6.2	5.8, 6.0, 5.9
Master's	6.9, 7.7, 7.9	6.2, 6.2, 6.8
PhD	6.9, 7.7, 8.3	9.0, 9.1, 8.3

Table: Anxiety data (10 point scale) vs. highest education attained cross sex.

- Two-way, fixed effect ANOVA model, with interaction:

$$Y_{anx} = \mu + \tau_{edu} + \tau_{sex} + \tau_{edu \times sex} + \varepsilon$$

- This model will allow us to separate the *marginal* effects of Education and Sex from any potential *interaction* effect of the two.

Two-way (two factor) ANOVA, with interaction

Two-way, fixed effect ANOVA model, with interaction:

$$Y_{anx} = \mu + \tau_{edu} + \tau_{sex} + \tau_{edu \times sex} + \varepsilon$$

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	28.842	<.001
Sex	0.109	1	0.109	0.492	0.496
Education * Sex	3.694	2	1.847	8.354	0.005
Residuals	2.653	12	0.221		

- Can uncover marginal and interaction effects *simultaneously*.
- Notice: same main effect SSs as in one-way ANOVAs, and as in two-way ANOVA without interaction (consult notes from previous class).

Two-way (two factor) ANOVA, with interaction

Two-way, fixed effect ANOVA model, with interaction:

$$Y_{anx} = \mu + \tau_{edu} + \tau_{sex} + \tau_{edu \times sex} + \varepsilon$$

ANOVA

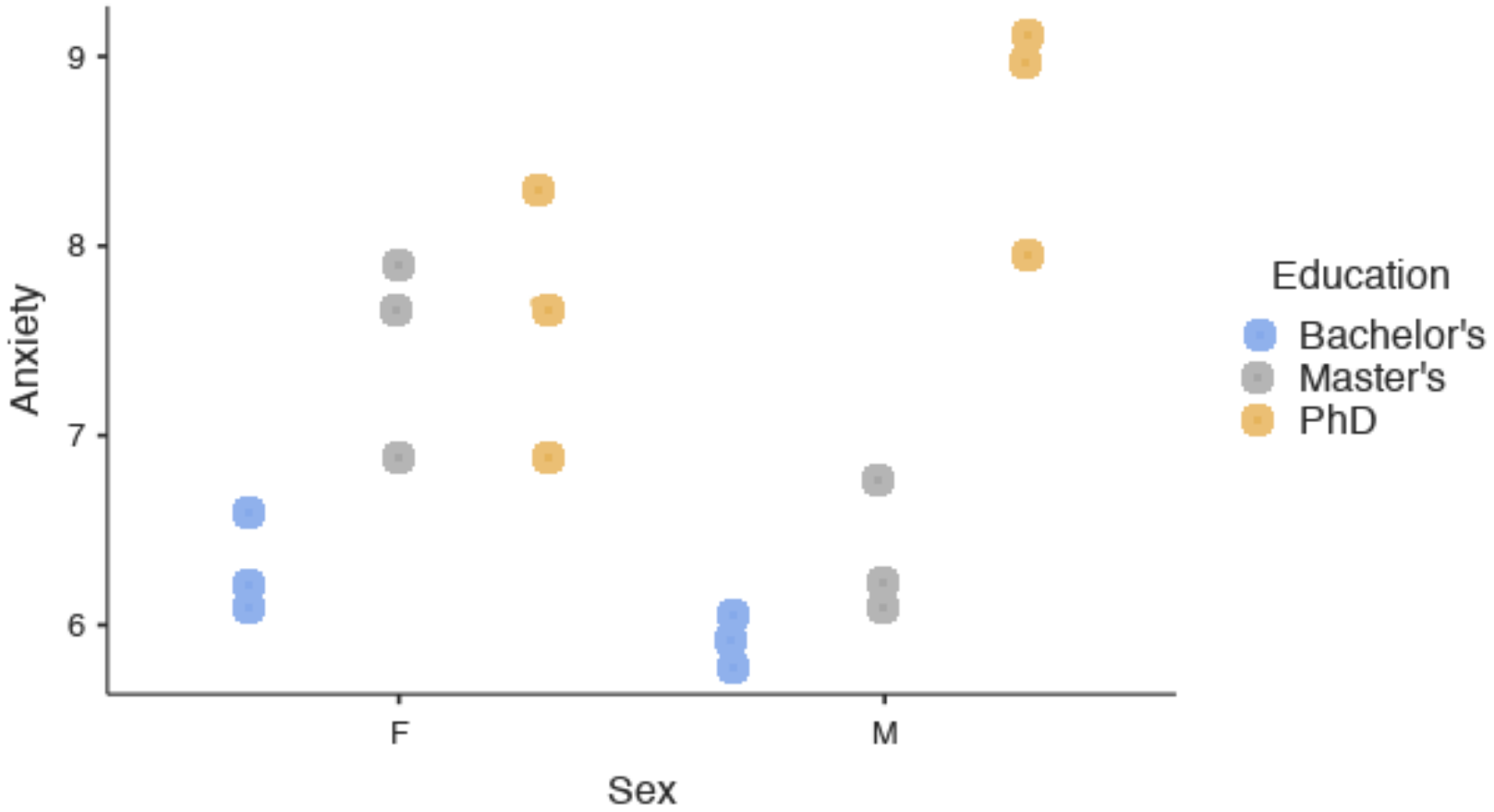
	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	28.842	<.001
Sex	0.109	1	0.109	0.492	0.496
Education * Sex	3.694	2	1.847	8.354	0.005
Residuals	2.653	12	0.221		

- Again, each F-statistic corresponds to a different test of hypothesis:
  - $F_{edu} = MS_{edu}/MS_{res}$ tests $H_0 : \tau_{edu} = 0$
  - $F_{sex} = MS_{sex}/MS_{res}$ tests $H_0 : \tau_{sex} = 0$
  - $F_{edu \times sex} = MS_{edu \times sex}/MS_{res}$ tests $H_0 : \tau_{edu \times sex} = 0$

Types of sums of squares

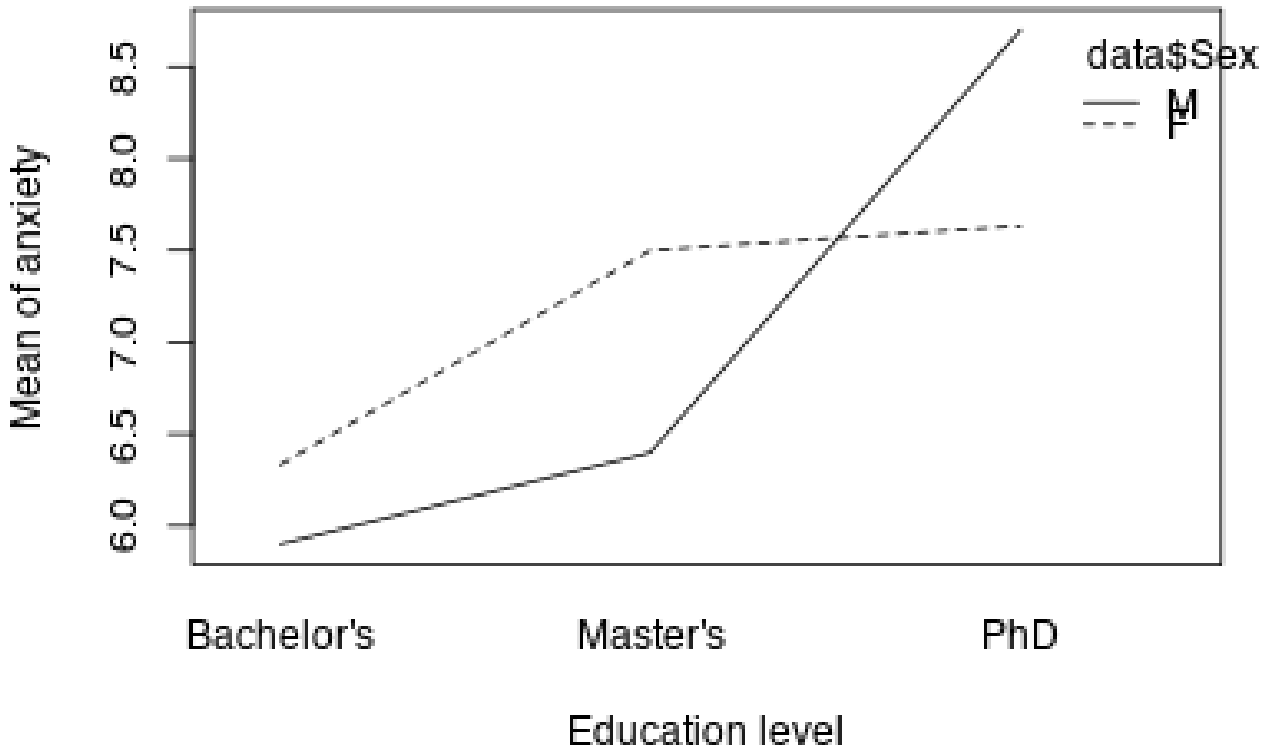
- Actually though, for ANOVAs with at least two factors, there is more than one way to partition a total sum of squares and to define a reasonable F-test on marginal and interaction effects.
- We will *not* get into the math behind this.
- **Generally, always default to the Type 3 sum of squares** (Jamovi defaults to this).
  - Type 1: some nice mathematical properties, but order of variables matters (bad)
  - Type 2: more powerful when interactions not present (unlikely in practice)
  - Type 3: good mathematical properties
  - Type 4: preferable in certain types of experimental designs

Two-way (two factor) ANOVA, with interaction



Notice the apparent interaction effect: both sexes report higher anxiety levels with higher education, but the *rate* of increase seems to be higher for males.

Two-way (two factor) ANOVA, with interaction



- Plotting the *group means* and connecting them with a line for each level of one category produces an "interaction plot."
- Very useful for visualizing interaction effects.
- But note: whether or not the lines cross is irrelevant! We are only assessing if the apparent *trends* are dissimilar.

Two-way (two factor) ANOVA, with interaction

- To make the previous interaction plot in Jamovi:
  - Use the "Estimated marginal means" tab in the "ANOVA" window
  - Drag variables you want to plot together over to the "Marginal means" window
  - Can also click option to produce tables of fitted values

Two-Way ANOVA, post-hoc pairwise comparisons

Two-Way fixed effects ANOVA model:

$$Y = \mu + \tau_A + \tau_B + \tau_{A\times B} + \varepsilon.$$

- Can perform post-hoc tests, *taking care to adjust for the multiple comparisons problem*: recall Scheffé and Tukey's HSD.

Post Hoc Comparisons in Two-way, Fixed Effects ANOVA Model

Post Hoc Comparisons - Education * Sex

Comparison									
Education	Sex		Education	Sex	Mean Difference	SE	df	t	p_{tukey}
Bachelor's	F	-	Bachelor's	M	0.433	0.384	12.000	1.129	0.860
		-	Master's	F	−1.167	0.384	12.000	−3.039	0.085
		-	Master's	M	−0.067	0.384	12.000	−0.174	1.000
		-	PhD	F	−1.300	0.384	12.000	−3.386	0.048
		-	PhD	M	−2.367	0.384	12.000	−6.164	<.001
	M	-	Master's	F	−1.600	0.384	12.000	−4.167	0.013
		-	Master's	M	−0.500	0.384	12.000	−1.302	0.779
		-	PhD	F	−1.733	0.384	12.000	−4.515	0.007
		-	PhD	M	−2.800	0.384	12.000	−7.293	<.001
Master's	F	-	Master's	M	1.100	0.384	12.000	2.865	0.113
		-	PhD	F	−0.133	0.384	12.000	−0.347	0.999
		-	PhD	M	−1.200	0.384	12.000	−3.126	0.074
	M	-	PhD	F	−1.233	0.384	12.000	−3.212	0.064
		-	PhD	M	−2.300	0.384	12.000	−5.991	<.001
PhD	F	-	PhD	M	−1.067	0.384	12.000	−2.778	0.129

Assumptions of Two-way, Fixed Effects ANOVA Model

Same assumptions as one-way model!

- Independence of observations $\Leftrightarrow$ independence of errors
- Equal variances across factor levels (homoskedasticity)
- Errors should be normally distributed, $\varepsilon \sim N(0, \sigma^2)$

Check assumptions same way as one-way model.

Assumptions 1: independence

Recall that this is the most important assumption and the most difficult to check.

- In general, good study design should *ensure* independence of observations (errors).
  - Did any of our sample individuals know each other?
  - Were the sampled individuals assessed for the anxiety measure jointly (as in a focus group setting), or independently of each other?

Assumptions 2: homoskedasticity (equal variances across groups)

Many ways to check this assumption (use more than one!)

- Levene's test (F-test)
- Compare boxplots
- Compare plots of raw data

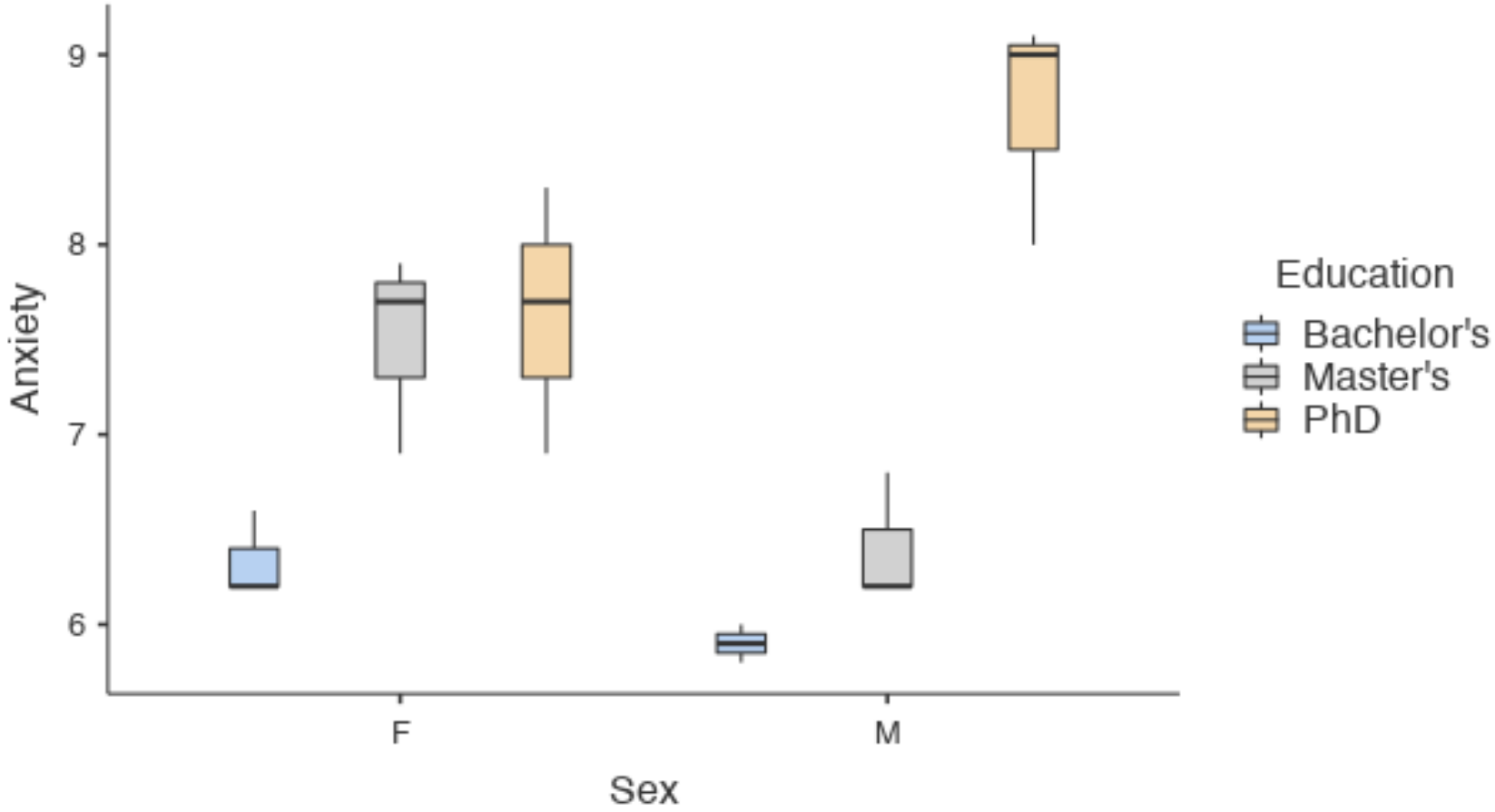
Assumptions 2: homoskedasticity (equal variances across groups)

Test for Homogeneity of Variances (Levene's)

F	df1	df2	p
2.142	5	12	0.130

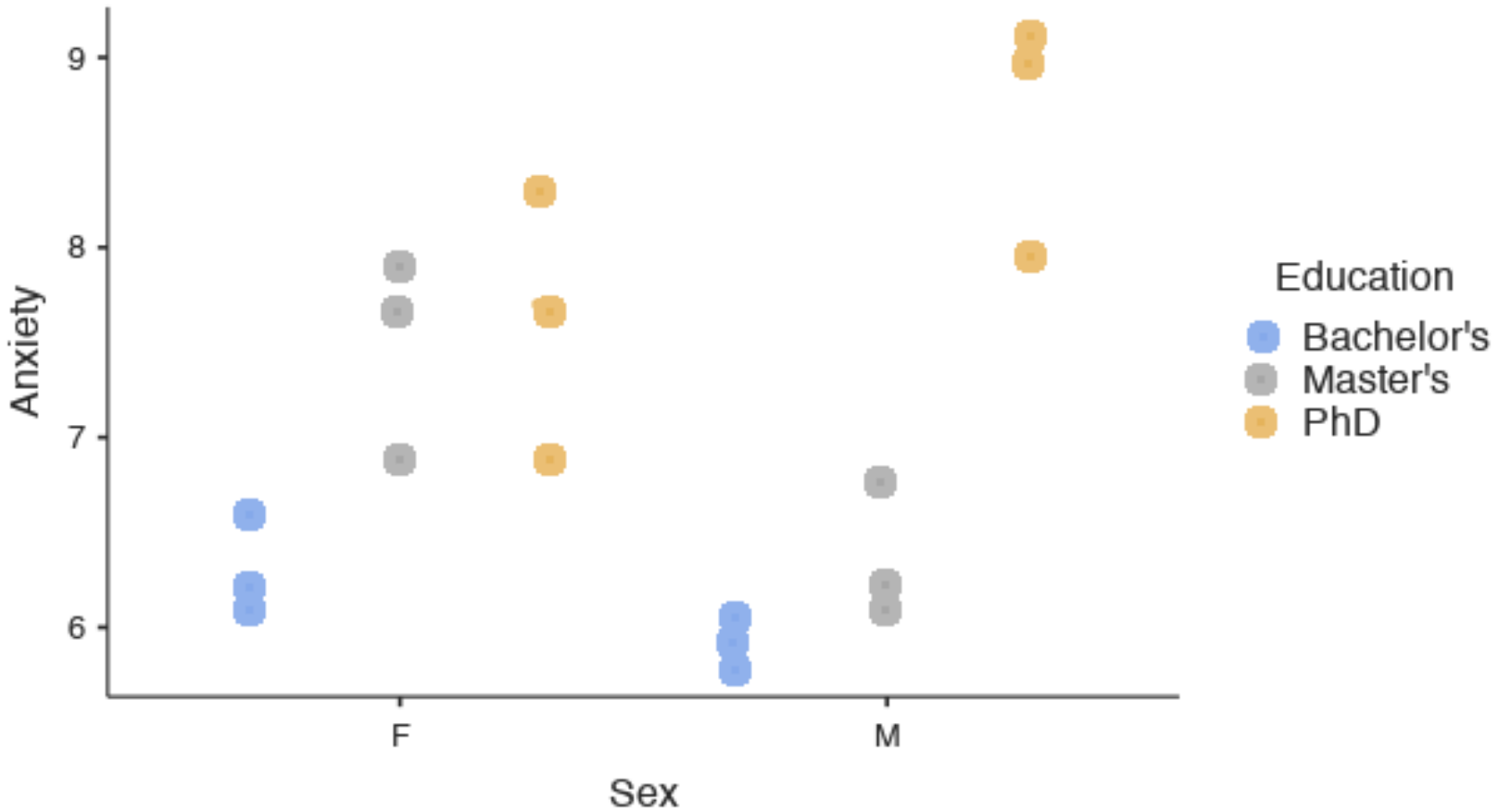
- Levene's test not significant; thus, gives no evidence of heteroskedasticity.
- Lack of power? See plots...

Assumptions 2: homoskedasticity (equal variances across groups)



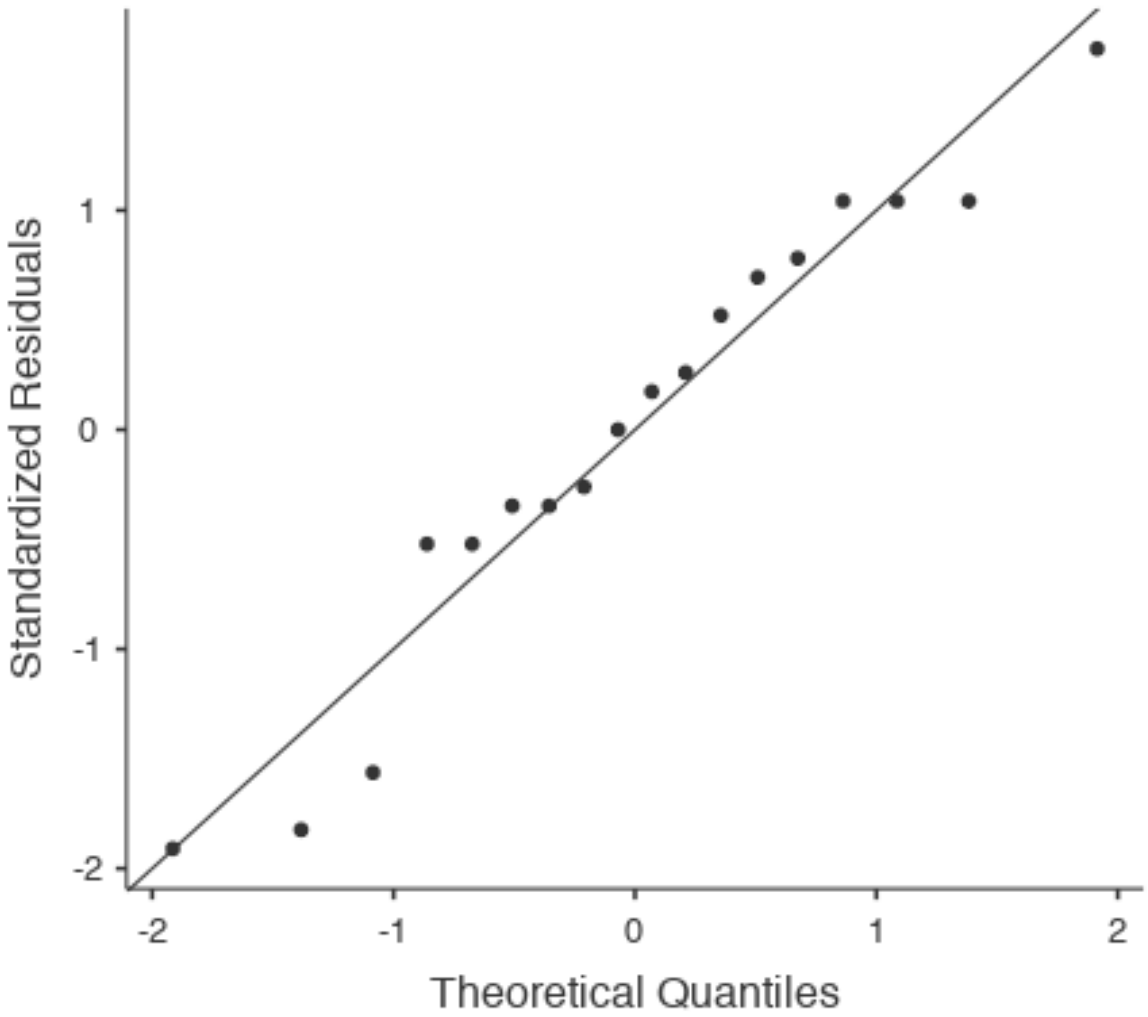
- Some visual evidence of heteroskedasticity, but...

Assumptions 2: homoskedasticity (equal variances across groups)



- ...only based on three data points per group. Not really enough data per group to adequately assess the homoskedasticity assumption.

Assumptions 3: normality of errors



- Looks pretty good! But again, only 18 data points total....

Two-Way ANOVA, in practice

Two-way fixed effects ANOVA (full) model:

$$Y = \mu + \tau_A + \tau_B + \tau_{A\times B} + \varepsilon$$

- # of obs. in each category can be different. If all the same, then the design is said to be "balanced".
- Balanced analyses have higher power and are more robust to unequal variances across categories (i.e. violations of Assumption 2). They are *very* robust to moderate departures from normality; i.e. skewness not a big problem, but multiple modes or outliers can be.
- The interaction term, $\tau_{A\times B}$, is often of the greatest interest.
- However, need lots of data to detect meaningful interaction effects.

Generic *n*-way ANOVAs

Nothing special about two factors; can write models with as many explanatory factors as we like.

- For example, three-way fixed effects ANOVA (full) model:

$$Y = \mu + \tau_A + \tau_B + \tau_{A\times B} + \tau_C + \tau_{A\times C} + \tau_{B\times C} + \tau_{A\times B\times C} + \varepsilon$$

- Or, for example, a four-way ANOVA with two pairwise interactions:

$$Y = \mu + \tau_A + \tau_B + \tau_C + \tau_D + \tau_{A\times C} + \tau_{B\times D} + \varepsilon$$

- Theoretically, the possibilities are endless.

Generic *n*-way ANOVAs

- However, in practice, the more complicated your model:
  (1) the more data you need to detect effects
  (2) the better experimental control you need to make sure you are isolating the effects of interest
  (3) the harder it is to diagnose your model and check your assumptions (need lots more data!)
  (4) the easier it is to fool yourself into thinking "complicated answer" means the same thing as "right answer"

Two-Way ANOVA: partitioning the variance

Return to the two-way fixed effects ANOVA (full) model:

$$Y = \mu + \tau_A + \tau_B + \tau_{A\times B} + \varepsilon$$

- Recall that we worked out mathematically how a one-way ANOVA model *partitions* the observed variance in our *response variable* into two pieces:
  - (1) variance explained by the (average) differences between the explanatory (categorical) variable,
  - (2) variance leftover (attributable to within-group/individual differences).
- An analagous kind of partitioning happens when we work with a more complicated ANOVA model....

Extreme examples to clarify partitions of variance: Ex. 1

- Suppose we have these sample data on Y over two categorical variables X and Z with 2 factor levels each:

	$X = M$	$X = F$
$Z = No$	2.0, 2.5, 2.3	1.9, 2.3, 2.6
$Z = Yes$	1.5, 1.6, 1.1	1.6, 1.7, 0.9

- Then:

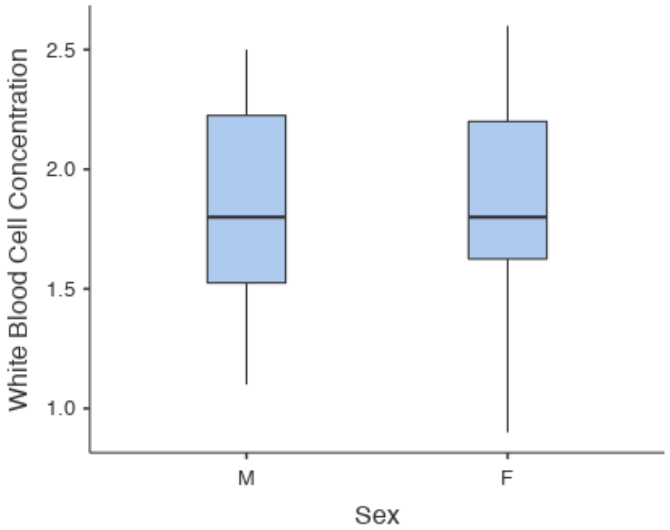
$$\overline{X_M} = 1.83,\ \overline{X_F} = 1.83,\ \overline{Z_N} = 2.27,\ \overline{Z_Y} = 1.40$$

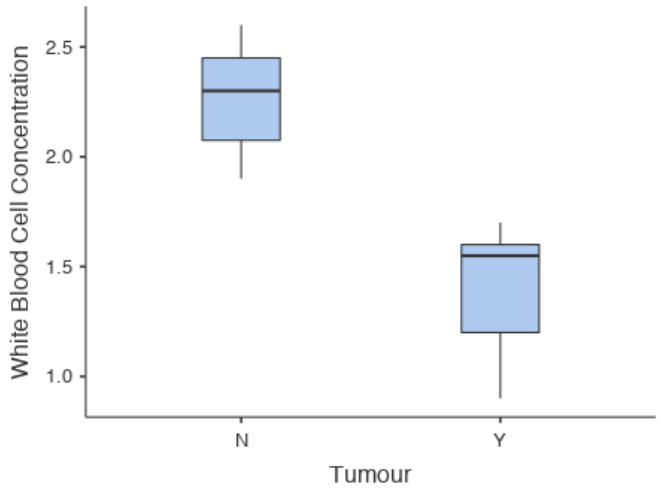
- And

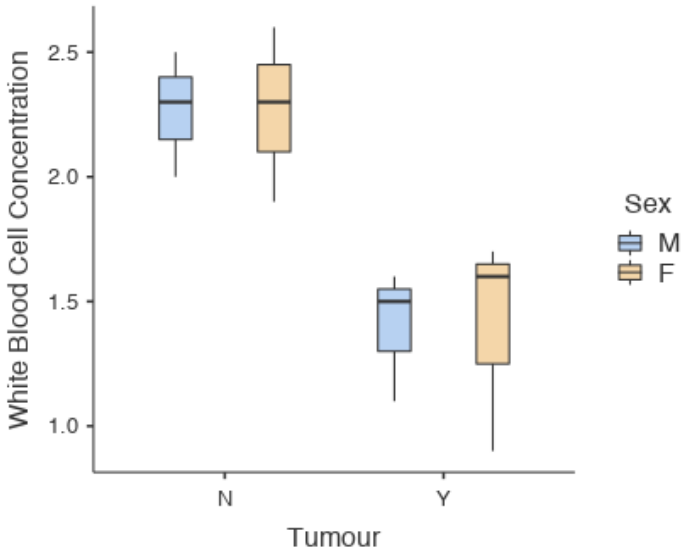
$$\overline{X_M Z_N} = 2.27,\ \overline{X_M Z_Y} = 1.40,\ \overline{X_F Z_N} = 2.27,\ \overline{X_F Z_Y} = 1.40$$

- X denotes sex and Z denotes presence of a tumour; response is white blood cell count.

Extreme examples to clarify partitions of variance: Ex. 1







Extreme examples to clarify partitions of variance: Ex. 1

ANOVA

	Sum of Squares	df	Mean Square	F	p
Sex	0.000	1	0.000	4.152e−30	1.000
Tumour	2.253	1	2.253	20.179	0.002
Sex ✻ Tumour	0.000	1	0.000	4.312e−30	1.000
Residuals	0.893	8	0.112		

- No variation explained by averaging over sex
- Clear variation explained by averaging over tumour presence
- No additional variation explained by averaging over sex$\times$tumour factor levels
- Leftover (residual) variation from individual observations within each fixed factor level

Extreme examples to clarify partitions of variance: Ex. 2

- Now response is measure of blood pressure; experimental design assigns people to no drug, Drug A, Drug B, or both.

	Drug A No	Drug A Yes
Drug B No	2.0, 2.5, 2.3	1.6, 1.7, 0.9
Drug B Yes	1.5, 1.6, 1.1	1.9, 2.3, 2.6

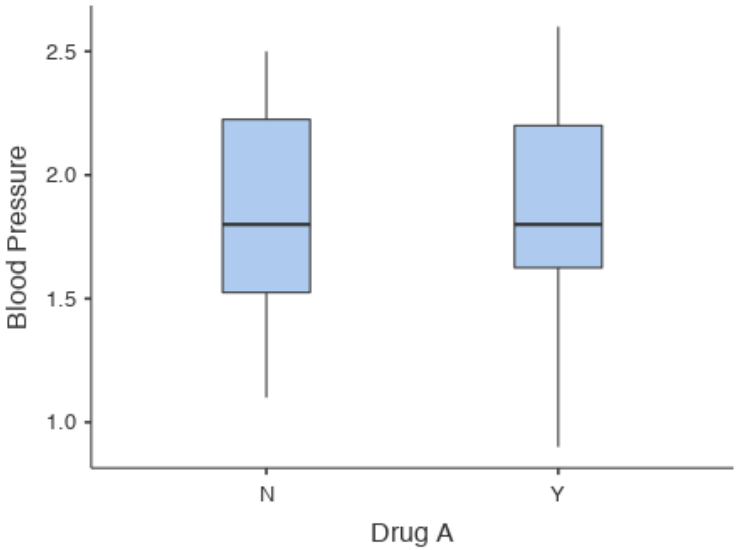
- Then:

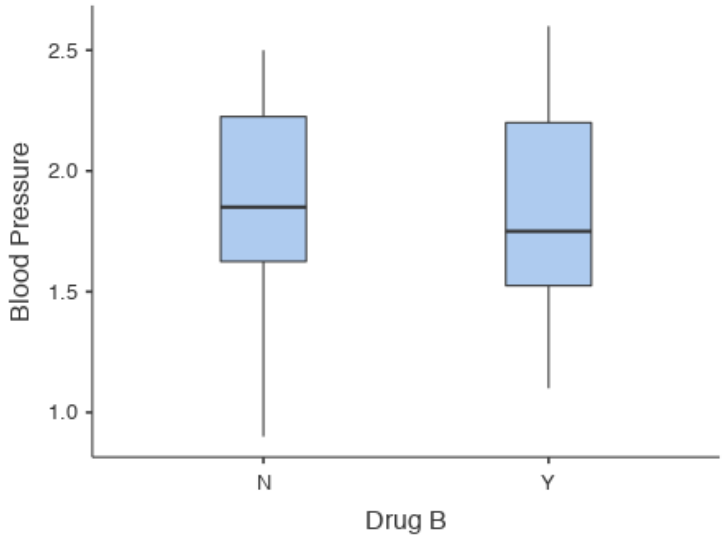
$$\overline{A_N} = 1.83,\ \overline{A_Y} = 1.83,\ \overline{B_N} = 1.83,\ \overline{B_Y} = 1.83$$

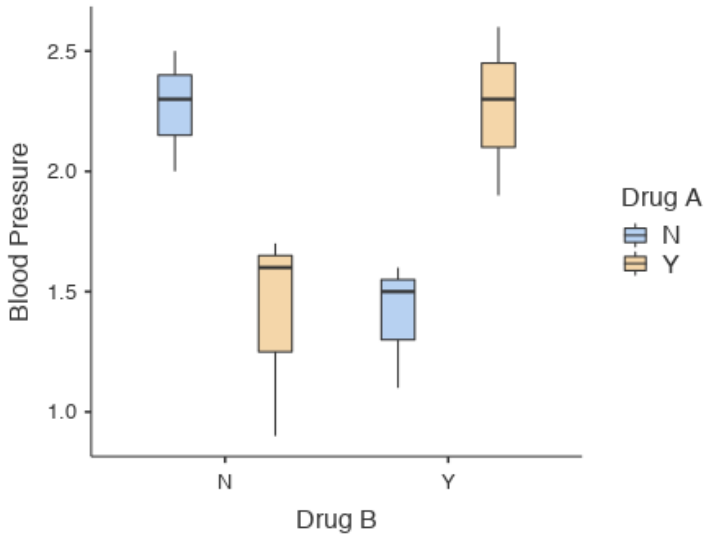
- And

$$\overline{A_N B_N} = 2.27,\ \overline{A_N B_Y} = 1.40,\ \overline{A_Y B_N} = 1.40,\ \overline{A_Y B_Y} = 2.27$$

Extreme examples to clarify partitions of variance: Ex. 2







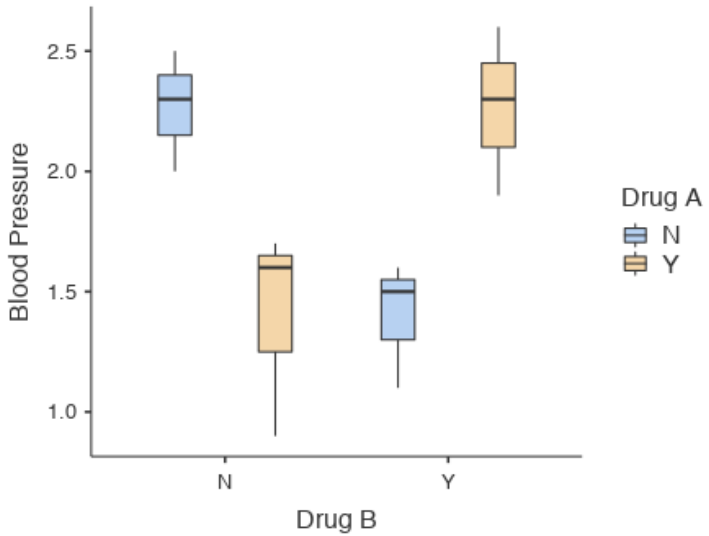
Extreme examples to clarify partitions of variance: Ex. 2

ANOVA

	Sum of Squares	df	Mean Square	F	p
Drug A	0.000	1	0.000	4.760e−30	1.000
Drug B	0.000	1	0.000	3.140e−30	1.000
Drug A ✻ Drug B	2.253	1	2.253	20.179	0.002
Residuals	0.893	8	0.112		

- No variation explained by taking Drug A (marginally)
- No variation explained by taking Drug B (marginally)
- Clear variation explained by taking *both* Drug A and Drug B
- Leftover (residual) variation from individual observations within each fixed factor level

Extreme examples to clarify partitions of variance: Ex. 2



- Those who took Drug A *only* saw blood pressure go down.
- Those who took Drug B *only* saw blood pressure go down.
- But those who took *both* drugs (or neither) have high blood pressure; drugs seem to be interacting to negate effects of treatment.

Extreme examples to clarify partitions of variance: Ex. 3

- Now suppose we have these sample data on Y over two categorical variables X and Z with 2 factor levels each:

	$X = A$	$X = B$
$Z = 1$	2.0, 2.5, 2.3	-1.2, 4.9, 3.1
$Z = 2$	0.7, 3.0, 3.1	1.9, 2.3, 2.6

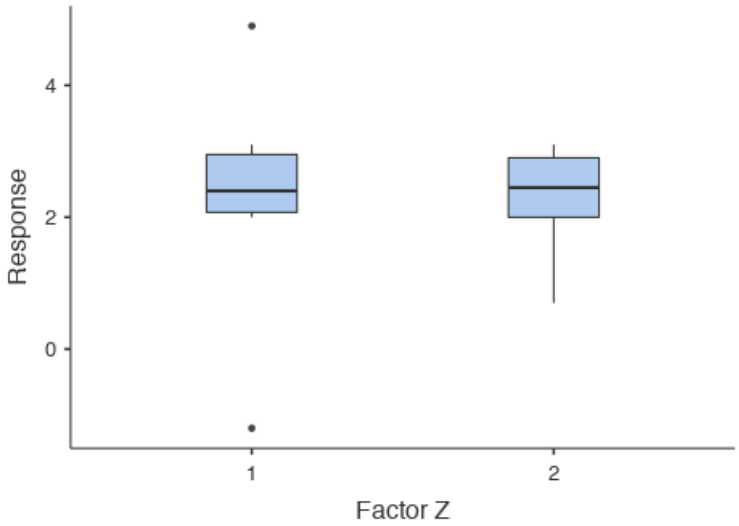
- Then:

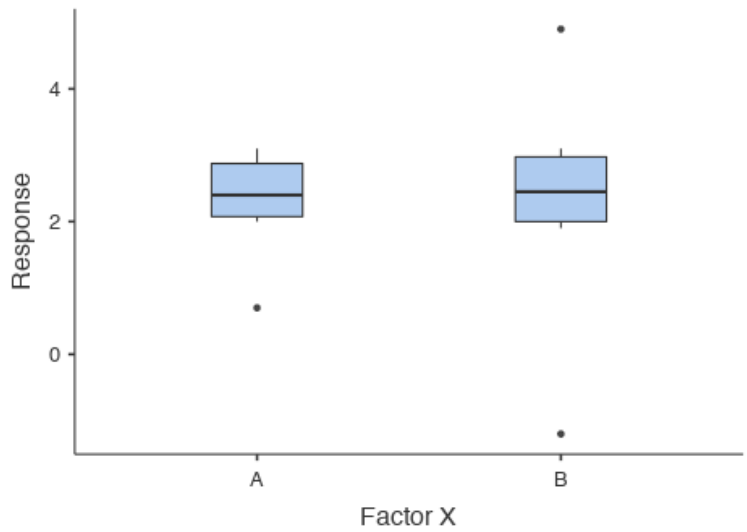
$$\overline{X_A} = 2.27,\ \overline{X_B} = 2.27,\ \overline{Z_1} = 2.27,\ \overline{Z_2} = 2.27$$

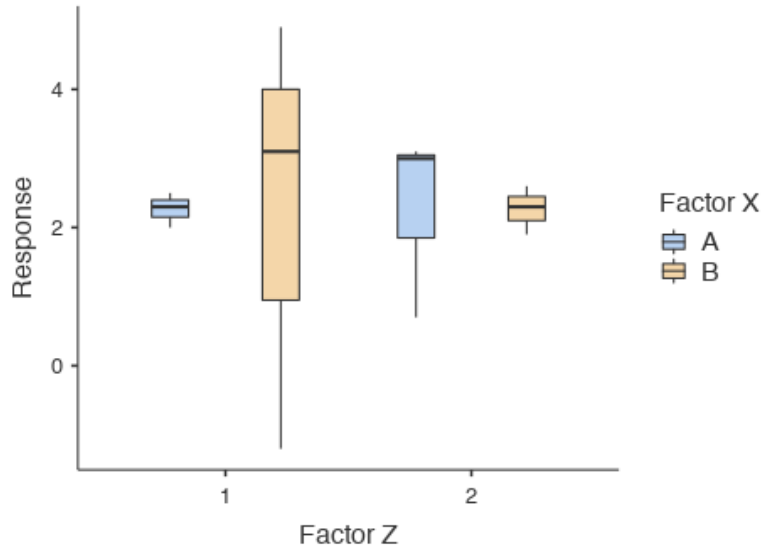
- And

$$\overline{X_A Z_1} = 2.27,\ \overline{X_A Z_2} = 2.27,\ \overline{X_B Z_1} = 2.27,\ \overline{X_B Z_2} = 2.27$$

Extreme examples to clarify partitions of variance: Ex. 3







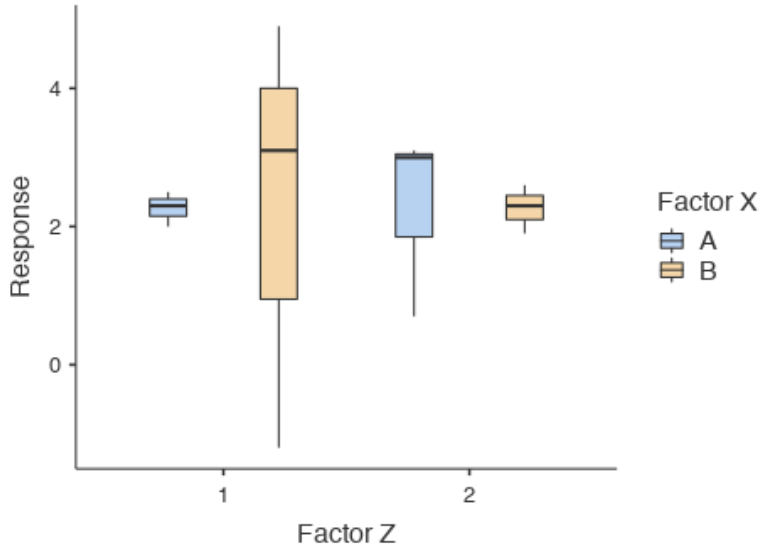
Extreme examples to clarify partitions of variance: Ex. 3

ANOVA

	Sum of Squares	df	Mean Square	F	p
Factor X	0.000	1	0.000	1.664e−32	1.000
Factor Z	0.000	1	0.000	9.359e−33	1.000
Factor X ✻ Factor Z	0.000	1	0.000	1.664e−32	1.000
Residuals	23.707	8	2.963		

- No variation explained by averaging over X factor levels
- No variation explained by averaging over Z factor levels
- No variation explained by averaging over $X \times Z$ factor levels
- All variation is residual variation from individual observations within each fixed factor level

Extreme examples to clarify partitions of variance: Ex. 3



Notice: obviously there are differences between the $X \times Z$ groups, but not *average* differences.

- ANOVAs are only able to detect *average differences* between groups.
- But there are many ways groups can be different, e.g. different variance, skewness, kurtosis, etc.
- This is why it is always important to **look at your data**; don't just rely on statistical tests of hypotheses.