

EPSE 592: Design & Analysis of Experiments

Ed Kroc

University of British Columbia

ed.kroc@ubc.ca

February 6, 2020

Last Time

- Type I and type II errors
- Multiple testing problems
- Orders of magnitude rule for interpretation of p-values
- One-way (one factor) fixed effects ANOVA model

- ANOVA as a way to *partition* variance in the response into two pieces:
 - (1) variance explained by average treatment effects, and
 - (2) leftover variation (i.e. variation within each treatment group)
- Fundamental equation of ANOVA:

$$\begin{aligned}SS_{total} &= SS_{between} + SS_{within} \\ &= SS_{treatment} + SS_{error}\end{aligned}$$

- Constructing an F -test for an ANOVA from the *mean squares*; i.e. sample variances
- Post-hoc vs. pre-planned pairwise comparisons

The One-way, Fixed Effects ANOVA Model

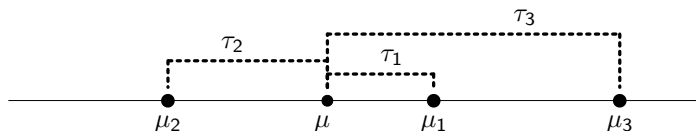
The one-way (one-factor), fixed effects ANOVA model:

$$Y = \mu + \tau_X + \varepsilon$$

- Y is the continuous response of interest
- X is the categorical variable, with observations in all categories, used to explain variation in Y
- μ is the *grand mean*; i.e. the average of all Y values
- τ_X is the *average treatment effect* of X on Y ; i.e. the average of all $Y - \mu$ values for each fixed value of X
- ε is the leftover error; i.e. the variation in Y *unexplained* by μ and τ_X .
- Used to test the hypothesis $H_0 : \tau_X = 0$ for *all* X values simultaneously.

The One-way, Fixed Effects ANOVA Model

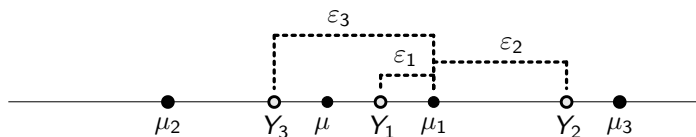
Understanding the treatment effect encoded by τ_X :



- In general, $\tau_X = \text{average of all } Y - \mu \text{ values for each fixed } X \text{ value}$
- Expressed another way, $\tau_X = \mu_X - \mu$
- So, if all treatments have the same effect, then they all equal the grand mean μ and $\tau_X = 0$ for all fixed values of X .

The One-way, Fixed Effects ANOVA Model

Understanding the individual error encoded by ε : suppose we have data points on Y (continuous response) and X , a categorical variable with 3 levels. Suppose observations Y_1 , Y_2 , and Y_3 belong to group X_1 .



- In general, ε can be different for every observation/individual; it is the difference between the observed response Y and the group mean μ_X
- Explicitly, $\varepsilon = Y - \mu_X$

The One-way, Fixed Effects ANOVA Model

- The one-way (one-factor), fixed effects ANOVA model:

$$Y = \mu + \tau_X + \varepsilon$$

- Using the previous two slides, this model can be rewritten as:

$$Y - \mu = (\mu_X - \mu) + (Y - \mu_X)$$

- In practice, we do not observe μ or μ_X , but we do observe the *sample* grand mean and *sample* group means.
- Can use these sample statistics to estimate the above equation and then test the hypothesis that $H_0 : \mu_X = \mu$ for all fixed values of X .

Example: education levels vs. anxiety

Recall our data on self-reported anxiety levels:

Bachelor's ($j = 1$)	Master's ($j = 2$)	PhD ($j = 3$)
$Y_{1,1} = 6.2$	$Y_{1,2} = 6.2$	$Y_{1,3} = 6.9$
$Y_{2,1} = 5.8$	$Y_{2,2} = 6.9$	$Y_{2,3} = 9.0$
$Y_{3,1} = 6.0$	$Y_{3,2} = 6.2$	$Y_{3,3} = 7.7$
$Y_{4,1} = 5.9$	$Y_{4,2} = 7.7$	$Y_{4,3} = 9.1$
$Y_{5,1} = 6.6$	$Y_{5,2} = 6.8$	$Y_{5,3} = 8.3$
$Y_{6,1} = 6.2$	$Y_{6,2} = 7.9$	$Y_{6,3} = 8.0$
$\bar{Y}_{.1} = 6.12$	$\bar{Y}_{.2} = 6.95$	$\bar{Y}_{.3} = 8.12$

$$\bar{Y}_{..} = 7.08$$

Partitioning the ANOVA model into variance components

Our goal is to partition the observed variation in our response Y into two distinct pieces:

- (1) variation explained by the different factor levels (treatments)
- (2) leftover (residual) variation
- Recall: our ANOVA model can be written as:

$$Y - \mu = (\mu_X - \mu) + (Y - \mu_X) \quad (\textit{theoretical model})$$

- Since we do *not* observe μ_X or μ , we replace them by their sample estimates $\bar{Y}_{\cdot j}$ and $\bar{Y}_{\cdot\cdot}$.
- Also, replace the generic Y by our observed Y_{ij} values:

$$Y_{ij} - \bar{Y}_{\cdot\cdot} = (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}) + (Y_{ij} - \bar{Y}_{\cdot j}) \quad (\textit{sample estimate of model})$$

Partitioning the ANOVA model into variance components

- Now we square both sides of the equation:

$$\begin{aligned}(Y_{ij} - \bar{Y}_{..})^2 &= [(\bar{Y}_{.j} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{.j})]^2 \\ &= (\bar{Y}_{.j} - \bar{Y}_{..})^2 + (Y_{ij} - \bar{Y}_{.j})^2 + 2(\bar{Y}_{.j} - \bar{Y}_{..})(Y_{ij} - \bar{Y}_{.j})\end{aligned}$$

- Now sum over all observations:

$$\begin{aligned}\sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2 &= \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2 \\ &\quad + 2 \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})(Y_{ij} - \bar{Y}_{.j})\end{aligned}$$

- Last time: showed the cross term always equals 0.

Partitioning the ANOVA model into variance components

- Left with the *fundamental equation of analysis of variance*.

$$\sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2$$

- This equation says that the sample variance in the response variable is equal to the sample variance in the *average response for each treatment* plus the sample variance of the responses *within each treatment*.
- This is typically written as a *sum of squares (SS)* equation:

$$SS_{total} = SS_{treatment} + SS_{error}$$

Or:

$$SS_{total} = SS_{between} + SS_{within}$$

Partitioning the ANOVA model into variance components

- Left with the *fundamental equation of analysis of variance*.

$$\sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2$$

- Notice how each term is a sum of squared differences from the grand (terms 1 and 2) or treatment (term 3) means. *This is exactly how we always measure variability*, up to a constant multiple.
- Notice: the variance in the response is partitioned into variability *explained by the average treatment effect* (term 2) plus variability *leftover* (term 3).

Examples to clarify the math: Ex. 1

- Suppose we have these sample data on Y over a categorical variable X with 3 factor levels:

$X = 1$	$X = 2$	$X = 3$
$Y_{1,1} = 1$	$Y_{1,2} = -1$	$Y_{1,3} = 5$
$Y_{2,1} = 1$	$Y_{2,2} = -1$	$Y_{2,3} = 5$
$Y_{3,1} = 1$	$Y_{3,2} = -1$	$Y_{3,3} = 5$

Then:

$$\bar{Y}_{.1} = 1, \quad \bar{Y}_{.2} = -1, \quad \bar{Y}_{.3} = 5$$

And $\bar{Y}_{..} = 1.67$.

- Now plug into the fundamental equation of ANOVA:

Examples to clarify the math: Ex. 1

Fundamental equation of ANOVA:

$$\sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2$$

- Notice that the last term equals zero!

$$\begin{aligned} \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2 &= (1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2 \\ &\quad + (-1 + 1)^2 + (-1 + 1)^2 + (-1 + 1)^2 \\ &\quad + (5 - 5)^2 + (5 - 5)^2 + (5 - 5)^2 = 0 \end{aligned}$$

- So, as expected, all the variability in the response is explained by the different treatment/factor levels of X .

Examples to clarify the math: Ex. 2

- Now, suppose we have these sample data instead on Y over a categorical variable X with 3 factor levels:

$X = 1$	$X = 2$	$X = 3$
$Y_{1,1} = -1.1$	$Y_{1,2} = -4.2$	$Y_{1,3} = 0.5$
$Y_{2,1} = 0.5$	$Y_{2,2} = -0.1$	$Y_{2,3} = 0.6$
$Y_{3,1} = 2.4$	$Y_{3,2} = 6.1$	$Y_{3,3} = 0.7$

Then:

$$\bar{Y}_{.1} = 0.6, \quad \bar{Y}_{.2} = 0.6, \quad \bar{Y}_{.3} = 0.6$$

And $\bar{Y}_{..} = 0.6$.

- Now plug into the fundamental equation of ANOVA:

Examples to clarify the math: Ex. 2

Fundamental equation of ANOVA:

$$\sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2$$

- Notice that the second term equals zero now!

$$\begin{aligned} \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 &= (0.6 - 0.6)^2 + (0.6 - 0.6)^2 + (0.6 - 0.6)^2 \\ &\quad + (0.6 - 0.6)^2 + (0.6 - 0.6)^2 + (0.6 - 0.6)^2 \\ &\quad + (0.6 - 0.6)^2 + (0.6 - 0.6)^2 + (0.6 - 0.6)^2 = 0 \end{aligned}$$

- So, as expected, all the variability in the response is explained by the variation *within* each treatment/factor level of X .

Mean sum of squares

- The *fundamental equation of analysis of variance*:

$$\sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2$$

- Notice how each term is a sum of squared differences from the grand (terms 1 and 2) or treatment (term 3) means. *This is exactly how we always measure variability*, up to a constant multiple.
- Recall: to define the sample variance, we had to *divide* by a constant:

$$S^2 = \frac{1}{N-1} \sum_{\ell=1}^N (Y_{\ell} - \bar{Y})^2$$

- The same applies for the ANOVA equation:

Mean sum of squares

- An unbiased estimator of the *total variance* is the *total mean square*:

$$MS_{total} = \frac{1}{N-1} \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{..})^2$$

- An unbiased estimator of the *between treatment variance* is the *treatment mean square*:

$$MS_{treatment} = \frac{1}{K-1} \sum_{j=1}^K \sum_{i=1}^{n_j} (\bar{Y}_{.j} - \bar{Y}_{..})^2$$

- An unbiased estimator of the *within treatment variance* is the *error mean square*:

$$MS_{error} = \frac{1}{N-K} \sum_{j=1}^K \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2$$

Testing the ANOVA null hypothesis

- Recall that the null hypothesis that a one-way (fixed factor) ANOVA model is designed to test is:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_K,$$

where μ_j is the mean response over the j th category of the explanatory factor X , $1 \leq j \leq K$.

- Under this null hypothesis*, we have that:

$$\frac{MS_{treatment}}{MS_{error}} \sim F(K - 1, N - K)$$

- That is, the ratio of the between and within treatment sample variance estimators derived from the ANOVA model give an F -statistic under the null hypothesis.
- Thus, we can use this ratio as a test statistic and calculate p-values!

Example: education vs. anxiety

Bachelor's	Master's	PhD
6.2	6.2	6.9
5.8	6.9	9.0
6.0	6.2	7.7
5.9	7.7	9.1
6.6	6.8	8.3
6.2	7.9	8.0

- The “treatments” (i.e. factor levels) here are different levels of education attained.
- An ANOVA will decompose the total variance in the data into one piece that captures variability *within* each factor level, and one piece that captures variability of the *average response between all factor levels*.

Example: education vs. anxiety

In Jamovi:

- Enter data as two columns (anxiety, education) just as you would for an *independent samples t-test* in Jamovi.
- Click “Analyses” tab, then “ANOVA”, then “ANOVA” again.
- Assign “Anxiety” to dependent variable.
- Assign “Education” to fixed factor(s).
- Sums of squares, degrees of freedom, mean squares, F -statistic, and p-value will appear in output.
- Click on appropriate boxes to produce assumption checks, pre-planned contrasts, or post-hoc comparisons.

Example: education vs. anxiety

ANOVA

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.75	2	6.377	14.8	<.001
Residuals	6.46	15	0.430		

- Very small p-value (equivalent to very large F -statistic) means data are inconsistent with the null hypothesis.
- Thus, seem to have evidence that *at least one* factor level of education produces different average anxiety responses than the other levels of education.

Example: education vs. anxiety

ANOVA

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.75	2	6.377	14.8	<.001
Residuals	6.46	15	0.430		

- Notice: mean squares equal sum of squares divided by appropriate degrees of freedom.
- Notice: F -statistic is quotient of $MS(treatment)$ and $MS(error)$.
- TERMINOLOGY: we usually refer to *observed errors* as *residuals*.

Post hoc, pairwise treatment comparisons

- In the previous example, we found strong evidence that the data are inconsistent with

$$H_0 : \tau_1 = \tau_2 = \tau_3 = 0$$

- But *which* treatments actually exhibit significant differences?
- Post hoc comparisons:
 - Testing for pairwise (or other composite) differences *after* you have observed a significant F -statistic in the ANOVA.
 - This is different than *pre-planning* a specific pairwise (or other composite) comparison *before* you run your ANOVA.
 - Statistical procedure: run a bunch of t -tests on the groups you want to compare.
 - But what about the multiple testing issue (i.e. inflated Type I error)?...

Post hoc, pairwise treatment comparisons

- There are *many* different techniques that have been developed to conduct post hoc comparisons while adjusting for Type I error inflation:
 - Bonferroni (bad: never use for post hoc comparisons of an ANOVA)
 - Tukey's honest significant differences
 - Scheffe's correction
 - ...and many others
- All these methods examine different types of t -tests and/or impose different p -value adjustments.
- The output of all these methods is interpreted *the same*: *exactly how you would interpret a standard t -test*.

Post hoc, pairwise treatment comparisons

- In Jamovi, click on “Post Hoc Tests” tab;
- Identify the variable/factor you want to make comparisons on;
- Select which kind of adjustment you want to make.

Post Hoc Comparisons - Education

Comparison		Mean Difference	SE	df	t	P _{Tukey}	P _{Scheffe}
Education	Education						
Bachelor's	- Master's	-0.833	0.379	15.0	-2.20	0.104	0.123
	- PhD	-2.050	0.379	15.0	-5.41	<.001	<.001
Master's	- PhD	-1.217	0.379	15.0	-3.21	0.015	0.020

Post hoc, pairwise treatment comparisons

- Scheffe's is most common: widely applicable, but conservative.
- Tukey's is more powerful than Scheffe's when *groups are close to balanced*; i.e. when the number of responses in each treatment group are about the same.

Post Hoc Comparisons - Education

Comparison		Mean Difference	SE	df	t	P _{Tukey}	P _{Scheffe}
Education	Education						
Bachelor's	- Master's	-0.833	0.379	15.0	-2.20	0.104	0.123
	- PhD	-2.050	0.379	15.0	-5.41	<.001	<.001
Master's	- PhD	-1.217	0.379	15.0	-3.21	0.015	0.020

- weak evidence of a Master's vs. PhD effect
- no evidence of a Bachelor's vs. Master's effect
- strong evidence of a Bachelor's vs. PhD effect

Post hoc vs. pre-planned comparisons

- Post hoc comparisons are experimentally (and so statistically) *different* from pre-planned comparisons. Why?
 - A post hoc comparison is made *because you first observe a significant ANOVA statistic*; i.e. your decision to make a follow-up comparison *depends* on the outcome of your first statistical test (the ANOVA)
 - A pre-planned comparison is decided upon in the experimental design phase, *before you collect any data or observe any statistics*; thus, it is not dependent on the outcome of any other statistics (e.g. an ANOVA)
 - Post hoc comparisons are *more* susceptible to inflated Type I errors
 - Pre-planned comparisons will provide stronger evidence, since you are not just “looking for any significant effects”
- In most applied research, all pairwise comparisons are post hoc
- In deeper, more controlled experiments (e.g. Phase II and III clinical trials), pre-planned comparisons are required

Assumptions of One-way, Fixed Effects ANOVA Model

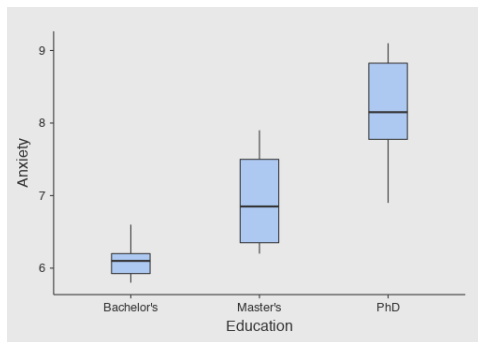
- Independence of observations \Leftrightarrow independence of errors
 - By far, the *most* important assumption.
 - If this fails, you cannot trust your analysis for anything.
 - Good study design will ensure this holds (e.g. random sampling).
 - Maybe check residuals vs. time of measurement plot
- Equal variances across factor levels (homoskedasticity)
 - Moderately important assumption.
 - If this fails, you can lose a lot of power to detect effects.
 - Can formally check with Levene's test in Jamovi; inspect groupwise boxplots; inspect residuals vs. fitted values plot
- Errors should be normally distributed, $\varepsilon \sim N(0, \sigma^2)$
 - Least important (but still important!) assumption.
 - ANOVA usually quite robust to moderate departures from normality (CLT helps).
 - Check qq-plot; check residuals vs. fitted values plot

Assumption 1: independence of observations

- If study required completely random sampling, then Assumption 1 should hold.
- Things to consider in general:
 - Are the data a completely random sample?
 - Could there be an *order of measurement* effect? For example, was the same interviewer measuring anxiety in all respondents and was their “mood” always the same? Or maybe they were tired and snippy at the end of the day, so could have made the respondents more stressed out later in the day? This could produce *autocorrelation in time*.
 - Could there be a *proximity* effect? For example, was half of the random sample generated at the library (high stress environment) and the other half at a coffee shop (low stress environment)? This could produce *autocorrelation in space*.
- If Assumption 1 is violated, need to *add random effects* (will discuss later) and/or *discuss with a statistician*

Assumption 2: equal variances across factor levels

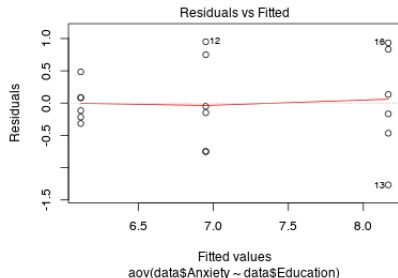
- In Jamovi, check with Levene's test (F -test)
- Also a good idea to check groupwise boxplots (under "Exploration" \Rightarrow "Descriptives" option in Jamovi):



- But can be difficult to tell if variances are similar for skewed data...

Assumption 2: equal variances across factor levels

- Check plot of residuals (observed errors) vs. fitted values (sample treatment means):



- Literally, compares the errors made by the model with the predictions (or “fits”) made by the model for your set of observations.
- These plots should *always* be checked when fitting *any* kind of statistical model (e.g. ANOVA, regression)

Residuals vs. Fitted Values Plots

- Ideally, for ANOVA:
 - average residual should = 0 for *each* fixed fitted value (look at the red line in the plot: gives the mean residuals),
 - residuals should look approx. normally distributed for *each* fixed fitted value,
 - residual variation should be about the same for all fixed fitted value levels,
 - if all this holds, then $\varepsilon \sim N(0, \sigma^2)$ is a reasonable assumption.

Residuals vs. Fitted Values Plots

- To plot residuals (observed errors) vs. fitted values (sample treatment means) in Jamovi:
 - Click on “Modules” in upper right and then install “Rj - Editor to run R code inside Jamovi” [only have to do this step the first time]
 - Click the “R” tab, and then “Rj Editor”
 - Copy the following code into the editor window (in general, you could replace “Anxiety” or “Education” with the names of other variables):

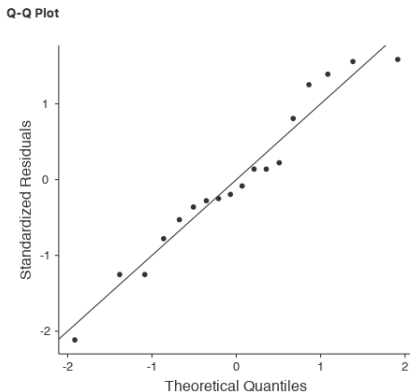
```
mod1 <- aov(formula=data$Anxiety~data$Education)
plot(mod1)
```
 - Press the green triangle button to run the code

Assumption 2: equal variances across factor levels

- Back to our example: some evidence of heteroskedasticity in residuals vs. fitted values plot (same as in boxplots)
- But Levene's test said no evidence of heteroskedasticity!
- Common to have different diagnostics suggest different conclusions
- In practice, only severe violations of ANOVA assumptions are usually worrisome; for our example, heteroskedasticity is mild, so likely not a major worry.
- But if heteroskedasticity is extreme... you will likely have to change your model/method

Assumption 3: normality of errors

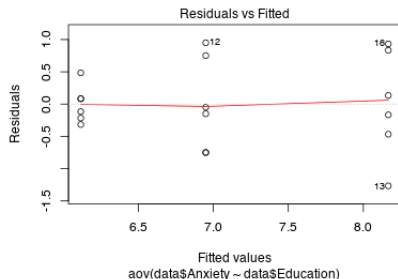
- In Jamovi, check with qq-plot of residuals:



- Compares observed quantiles (percentiles) of standardized errors (residuals) vs. theoretical quantiles (percentiles) of a standard normal distribution: all points close to diagonal = good!

Assumption 3: normality of errors

- Or: check plot of residuals (observed errors) vs. fitted values (sample treatment means) as before:



- Data should look somewhat normally distributed, with mean zero, *in each factor level*.
- In our example, this looks like a reasonable assumption.

What to do if your assumptions are violated?

- If violations are mild, then can likely still use a traditional analysis, but:
 - interpretations/conclusions should be made carefully,
 - violations of assumptions should *always* be reported.
- Otherwise, if normality and/or homoskedasticity are violated:
 - can perform *robust* tests of hypotheses (e.g. Welch tests)
 - can try to transform variables to obviate violations (not advised)
 - can use a different methodology altogether (e.g. nonparametric ANOVA)
- As always, if your observations are *not* independent (Assumption 1), then the analysis cannot be trusted. A different methodology is the only way out of this.

Building a more realistic model

- In practice, we usually have *multiple* explanatory factors/variables we would like to account for.
- This will require defining a more complex model.
- For our anxiety data, suppose we also have information on the sex of each respondent.
- We would like to then see if education level and/or sex explain variation in the anxiety response.
- Note: I'll assume sex is binary here (usually necessary in statistical practice since usually the vast majority of your sample respondents will be M/F; i.e. won't have any *power* to detect sex effects for other categories).

One-way ANOVA (Education)

Bachelor's	Master's	PhD
6.2	6.2	6.9
5.8	6.9	9.0
6.0	6.2	7.7
5.9	7.7	9.1
6.6	6.8	8.3
6.2	7.9	8.0

Table: Self-reported anxiety levels, 10 point scale. 18 respondents.

- One-way, fixed effect ANOVA model:

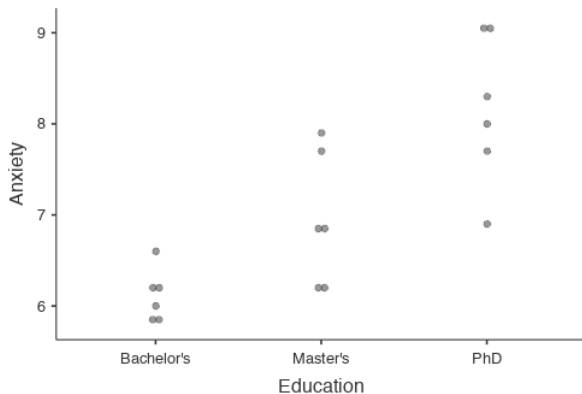
$$Y_{anx} = \mu + \tau_{edu} + \varepsilon, \quad [\text{or}]$$

$$Y_{anx} = \mu + (\mu_{edu} - \mu) + (Y_{anx} - \mu_{edu})$$

One-way ANOVA (Education)

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.815	<.001
Residuals	6.457	15	0.430		



One-way ANOVA plotting

In Jamovi:

- Install the 'scatr' module via "Modules" button in top right (first time only)
- Click on "Exploration" and then "Descriptives"
- Select the variable(s) you want to plot and which factor(s) you want to split your response by.
- Under "Plots", select "Data" and "Stacked"

Note: there is no statistical difference between the "Stacked" and the "Jittered" options when plotting. They are simply different visual choices ("jittering" adds some arbitrary horizontal space between nearby points to aid visualization).

One-way ANOVA (Sex)

Female	Male
6.2, 6.6, 6.2	5.8, 6.0, 5.9
6.9, 7.7, 7.9	6.2, 6.2, 6.8
6.9, 7.7, 8.3	9.0, 9.1, 8.3

Table: Self-reported anxiety levels, 10 point scale. 18 respondents.

- One-way, fixed effect ANOVA model:

$$Y_{anx} = \mu + \tau_{sex} + \varepsilon, \quad [\text{or}]$$

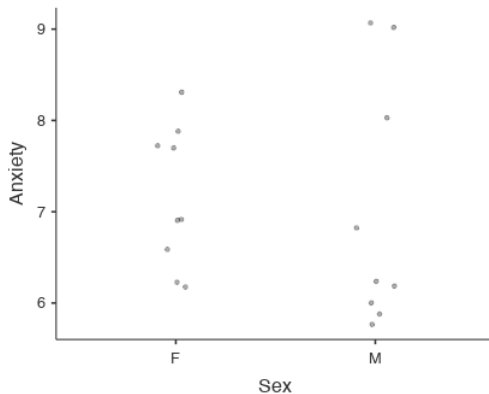
$$Y_{anx} = \mu + (\mu_{sex} - \mu) + (Y_{anx} - \mu_{sex})$$

- Note: the error, ε , will *not* be the same as in the previous model, since the explanatory factor is not the same (sex, rather than education)

One-way ANOVA (Sex)

ANOVA

	Sum of Squares	df	Mean Square	F	p
Sex	0.109	1	0.109	0.0912	0.767
Residuals	19.102	16	1.194		



Compare the two one-way ANOVAs

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.815	<.001
Residuals	6.457	15	0.430		

ANOVA

	Sum of Squares	df	Mean Square	F	p
Sex	0.109	1	0.109	0.0912	0.767
Residuals	19.102	16	1.194		

- Since we are explaining (partitioning) variance in the *same* response in both ANOVAs, the SSs add up to the *same* number.
- Education seems to explain some variation, but sex does not; what about an *interaction* effect?

One-way ANOVA on a “composite” variable

	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: Same data, now split into six distinct categories.

- One-way, fixed effect ANOVA model:

$$Y_{anx} = \mu + \tau_{edu \times sex} + \varepsilon, \quad [\text{or}]$$

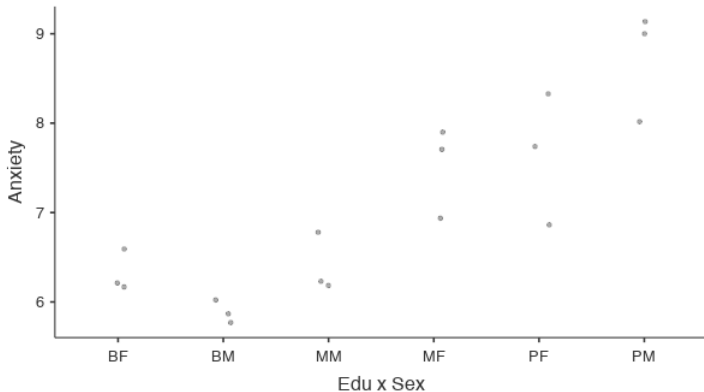
$$Y_{anx} = \mu + (\mu_{edu \times sex} - \mu) + (Y_{anx} - \mu_{edu \times sex})$$

- Note again: as always, the error, ε , is *not* the same as in the previous two models (different explanatory variable now)

One-way ANOVA on a “composite” variable

ANOVA

	Sum of Squares	df	Mean Square	F	p
Edu x Sex	16.558	5	3.312	14.977	<.001
Residuals	2.653	12	0.221		



Compare all three one-way ANOVAs

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.815	<.001
Residuals	6.457	15	0.430		

ANOVA

	Sum of Squares	df	Mean Square	F	p
Sex	0.109	1	0.109	0.0912	0.767
Residuals	19.102	16	1.194		

ANOVA

	Sum of Squares	df	Mean Square	F	p
Edu x Sex	16.558	5	3.312	14.977	<.001
Residuals	2.653	12	0.221		

- Note: again, we see SSs add up to the *same* number.

One-way ANOVA on a “composite” variable

ANOVA

	Sum of Squares	df	Mean Square	F	p
Edu x Sex	16.558	5	3.312	14.977	<.001
Residuals	2.653	12	0.221		

- This ANOVA explains more variation in our response than the previous two ANOVAs.
- But how do we interpret it?
- How to separate *marginal* effects from *interaction* effects?

Two-way (two factor) ANOVA, no interaction (marginals only)

Two-way, fixed effect ANOVA model, no interaction:

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

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.065	<.001
Sex	0.109	1	0.109	0.240	0.632
Residuals	6.348	14	0.453		

- Each F -statistic corresponds to a different test of hypothesis:
 - F -stat on Education factor tests if all education groups have the same average response: $F = MS_{edu}/MS_{res}$
 - F -stat on Sex factor tests if both sex groups have the same average response: $F = MS_{sex}/MS_{res}$

Two-way (two factor) ANOVA, no interaction (marginals only)

Two-way, fixed effect ANOVA model, no interaction:

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

ANOVA

	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.065	<.001
Sex	0.109	1	0.109	0.240	0.632
Residuals	6.348	14	0.453		

- In practice, there is *no need* to adjust for multiple comparisons here because the ANOVA model tests each hypothesis by using the same model/error information.

Two-way (two factor) ANOVA, no interaction (marginals only)

Compare the output of the separate one-way ANOVAs:

ANOVA					
	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.815	<.001
Residuals	6.457	15	0.430		

ANOVA					
	Sum of Squares	df	Mean Square	F	p
Sex	0.109	1	0.109	0.0912	0.767
Residuals	19.102	16	1.194		

...with the output of the two-way ANOVA:

ANOVA					
	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	14.065	<.001
Sex	0.109	1	0.109	0.240	0.632
Residuals	6.348	14	0.453		

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.

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 $H_0 : \tau_{edu \times sex} = 0$

Two-way (two factor) ANOVA, with interaction

Compare the output of the two-way ANOVA with an interaction:

ANOVA

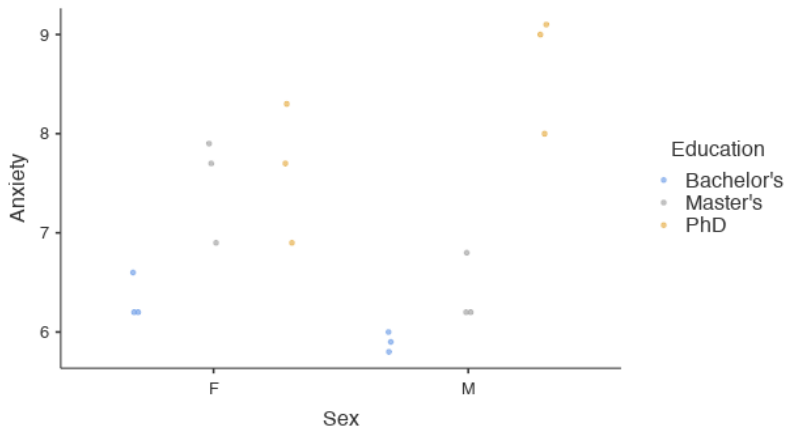
	Sum of Squares	df	Mean Square	F	p
Education	12.754	2	6.377	28.842	<.001
Sex	0.109	1	0.109	0.492	0.496
Education * Sex	3.694	2	1.847	8.354	0.005
Residuals	2.653	12	0.221		

...with the output of the one-way ANOVA on the six categories induced by the two factors:

ANOVA

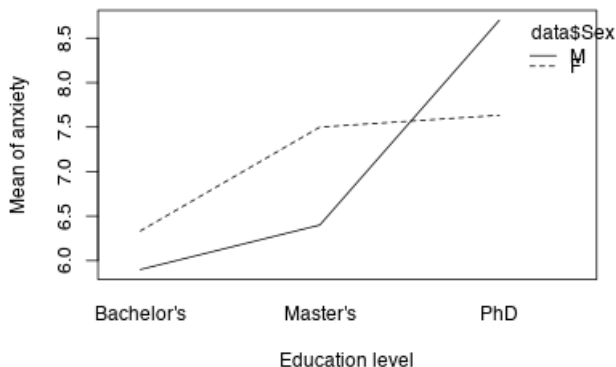
	Sum of Squares	df	Mean Square	F	p
Edu x Sex	16.558	5	3.312	14.977	<.001
Residuals	2.653	12	0.221		

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.

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