Testing the hypothesis that something important happened

Gemma Wallace & Neil Yetz PSY 653 Module 11 Lab Apr 22, 2020

Goals for this activity:

- 1) Practice identifying minimum effects for rejecting specific hypotheses
- 2) Using the sleep hygiene data, conduct regular and Bayesian ANOVAs to determine whether gender and occupation are related to mental health
 - a) Interpret and compare results from the two analytic approaches

Though not shown in the demo, remember that it's best practice to examine descriptives and visualize your data before conducting analyses:)

Part 1: Minimum effects practice

Minimum Effects Testing (MET)

 It tests the hypothesis that the effect of treatments falls somewhere in an interval between zero and some number

 Rather than testing if an effect is precisely zero, we can test if it falls above a range of values (The minimum effect you are testing) The authors used Right Wing Authoritarianism (RWA) to predict differences in response time to in-group and out-group faces, and found a squared correlation of .07, which was significant, with F(1,161) = 4.81

Identifying minimum effects: Kevin's minimum effects code

```
# 1a Applied to Bret et al. (Right Wing Authoritarianism)
```{r}
dfhyp=1
dferr=161
alpha=.05
effect=.01 #enter the minimum effect you are testing
sse=100
mse=(((1/effect)-1)*sse)/dferr
noncen=sse/mse
qf((1-alpha),dfhyp,dferr,noncen)
[1] 8.683388
```

```
F(1,161) = 4.81,
\alpha = .05,
Minimum effect = 1%

Don't change
```

Detecting minimum effects: Kevin's minimum effects

code

```
1a Applied to Bret et al. (Right Wing Authoritarianism)
```{r}
dfhyp=1
dferr=161
alpha=.05
effect=.01 #enter the minimum effect you are testing
                                                                           F(1,161) = 4.81
sse=100
mse=(((1/effect)-1)*sse)/dferr
noncen=sse/mse
qf((1-alpha),dfhyp,dferr,noncen)
 [1] 8.683388
```

Compare the resulting minimum F-value needed to test an effect of 1% or more to the obtained F-value.

In this case we did NOT reach the threshold F-value to have a significant effect at 1%. The authors need an F-value of 8.68 to obtain a significant effect at a minimum effect of 1%.

Identifying minimum effects via Murphy, Myors & Woloch (2014) "One Stop" F-table (Appendix B)

Minimum F required (Appendix B)

Comparison to a Nil effect

With a $DF_{hyp} = 1$ and $DF_{err} = 161$, the Authors need an F-value of 3.89 or more to obtain a significant effect.

Obtained F = 4.81

We have a significant effect

								One	Stop F									
		1	2	3	4	5	6	7	dfH 8	YP 9	10	12	15	20	30	40	60	120
	dfErr 90 nil .05 nil .01 pow .50 pow .80 18 .05 18 .01 pow .50 pow .80 58 .05 58 .01 pow .50 pow .80	6.92 3.86 7.97 6.97 11.29 6.86 12.12 15.17 21.57 14.87	4.95 4.37 6.64 3.74 6.62 8.31 1.59 7.58	3.75 3.48 5.06 2.66 4.74 6.02 8.25	2.18 3.19 4.88 6.59 3.92	3.23 1.43 2.69 2.71 3.77 1.83 3.19 4.15	2.20 3.01 1.26 2.38 2.52 3.43 1.58 2.78 3.70 4.92 2.78 4.23	2.84 1.21 2.18 2.38 3.19 1.47 2.52 3.37	2.00 2.26 3.01 1.33 2.28 3.12 4.08 2.14	2.61 1.01 1.86 2.18 2.86 1.21 2.10 2.93 3.80 1.92	2.52 0.94	2.39 0.89 1.58 2.00 2.56 1.03 1.75 2.54	2.24 0.81 1.40 1.88 2.38 0.87 1.51 2.30 2.90	0.70 1.20 1.76 2.18	1.58 1.91 0.59 0.98 1.63 1.97 0.65 1.04 1.83 2.21 0.85 1.27		1.46 1.72 0.46 0.73 1.48 1.74 0.49 0.76 1.58 1.85 0.58	1.39 1.60 0.40 0.59 1.40 1.61 0.40 0.60 1.44 1.66 0.45
	100 nil .05 nil .01 pow .50 pow .80 18 .05 18 .01 pow .80 58 .05 58 .01 pow .80 120 nil .05 nil .01 pow .50 pow .80 18 .01 pow .50 pow .80 18 .01 pow .50 pow .80 18 .01 pow .50 pow .80 58 .05	6.89 3.85 7.95 7.24 11.60 7.11 12.45 16.18 22.59 15.62 23.49 3.91 6.85 3.84 7.93 7.76 12.20 7.58 13.10 17.88 24.59 17.37	4.82 2.52 4.94 4.49 6.76 3.84 6.76 8.81 2.08 7.93 2.03 3.07 4.79 4.74 7.02 4.74 7.05 9.64 3.05 8.79	3.98 1.92 3.73 3.55 5.13 2.71 4.827 8.57 5.51 8.22 2.68 3.95 3.71 3.66 5.28 4.98 6.89 9.20	4.30 2.22 3.83 5.05 6.82 4.19 6.30 2.45 3.48 1.56 3.13 4.40	3.21 1.43 2.67 2.74 3.80 1.85 3.22 5.76 3.40 5.14 2.29 3.17 2.65 2.81 3.89 3.29 4.64 6.12 6.74	2.17 2.96 1.25 2.34 2.59 3.50 1.63 2.85 4.09 5.35 3.15	2.49 3.81 2.09 2.79 1.11 2.12 2.43 3.24 1.52 2.56 3.70 4.80	2.29 3.43 2.01 2.66 1.09 1.97 2.31 3.04 1.36 2.32 3.41 4.38 2.41	2.59 1.01 1.84 2.19 2.87 1.22 2.11 3.00 3.88 2.06 3.10 1.96 2.56 1.50 1.82 2.21 2.84 2.12 3.17 4.06 2.25	2.50 0.93 1.72 2.11 2.75 1.12 1.95 2.84 3.65 1.87 2.83 1.91 2.47 0.92 1.70 2.13 2.76 1.19 1.97 2.13 2.76 2.13 2.76 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.13 2.70 2.70 2.70 2.70 2.70 2.70 2.70 2.70	1.56 2.00 2.56 1.04 1.75 2.60 3.29 1.58 2.43 1.83 2.34 0.87 1.54 2.01 2.56 6.05 1.75 2.71 3.41 1.72	1.36 2.06 1.75 2.19 0.74 1.34 1.89 2.36 0.87 1.50 2.43 3.02 1.47	2.07 0.70 1.18 1.76 2.17 0.74 1.26 2.09 2.58 1.11 1.67 1.66 2.03 0.64 1.14 1.75 2.15 2.15 2.15 2.15 2.15	1.62 1.96 0.61 1.01 1.84 2.21 0.86 1.28	1.80 0.50 0.83 1.55 1.84 0.55 0.88 1.71 2.03 0.70 1.06	1.45 1.69 0.46 0.71 1.72 0.49 1.57 1.84 0.86 1.65 0.43 0.68 1.45 1.69 0.46 1.57 1.83 0.59	1.37 1.57 0.39 0.59 1.38 1.58 0.39 0.58 1.43 1.64 0.63 1.35 1.35 0.36 0.53 1.35 0.37 0.37 0.37 0.42 0.64 0.64 0.64
C	150 nil .05 nil .01 pow .50 pow 80 1% .05 1% .01 pow .50 pow .80 5% .05 5% .01 pow .50 pow .80	6.80 3.83 7.90 8.61 13.04 8.26 14.11 20.52 27.47 19.73	4.75 2.50 4.R9 5.01 7.40 4.42 7.43 0.86 4.46 0.24	1.90 3.09 3.86 5.51 3.09 5.21 7.64 10.12	3.45 1.55 3.02 3.28 4.56 2.40 4.09 6.06 7.95 5.19	1.41 2.63 2.92 3.98 1.98 3.40 5.11 6.65 4.19	2.92 1.24 2.32 2.66 3.59 1.78 2.96 4.48 5.78 3.52	2.76 1.10 2.09 2.49 3.31 1.56 2.62 4.03 5.10 3.04	2.63 1.08 1.94 2.36 3.09 1.40 2.37 3.69 4.69 2.67	0.99 1.80 2.25 2.93 1.27 2.16 3.41 4.33 2.48	2.44 0.92 1.68 2.17 2.79 1.15 2.00 3.20 4.04 2.25	2.30 0.86 1.52 2.03 2.58 1.06 1.77 2.88 3.60 1.90	2.16 0.73 1.31 1.90 2.37 0.88 1.51 2.57 3.17 1.61	0.63 1.11 1.76 2.15 0.74 1.25	1.83 0.54 0.90 1.61 1.92 0.61 0.98 1.92 2.28 0.93		1.62 0.40 0.64 1.44 1.66 0.43 0.68 1.59 1.83 0.59	1.32 1.49 0.33 0.49 1.34 1.51 0.34 0.51 1.41 1.59 0.41 0.58

Minimum F required (Appendix B)

Comparison to a 1% effect

With a $DF_{hyp} = 1$ and $DF_{err} = 161$, the Authors need an F-value of 8.61 or more to obtain a significant effect.

Obtained F = 4.81

We do NOT have a significant effect

			\						dfH	VID								
	dfErr	1	2	3	4	5	6	7	8	9	10	12	15	20	30	40	60	120
	90 nil .05 nil .01	3.94 6.92	3.10 4.85	2.71 4.01											1.58 1.91			1.39
	pow .50			1.93 3.75													0.46	
	pow .80 1% .05	6.97	4.37			2.69			2.26		2.11		1.88				1.48	
	1% .01																1.74	
	08. woq	12.12	3.74 6.62		2.18			2.52	2.28	2.10	1.12						0.49	
	5% .05	15.17	8.31	6.02	4.88	4.15	3.70	3.37	3.12	2.93	2.77	2.54	2.30	2.07	1.83	1.70	1.58	1.44
	5% .01 pow .50	21.57 14.87			6.59 3.92		4.92 2.78	4.44		3.80 1.92		3.24			2.21 0.85		1.85	
	pow .80	22.43															0.87	
	100 nil .05	3.93	3 00	2.70	2 46	2 30	2 10	2 10	2 03	1 97	1 92	1 85	1 77	1 67	1 57	1 51	1.45	1 37
	nil .01			3.98	3.51	3.21	2.99	2.82	2.69			2.37	2.22	2.07			1.69	
	pow .50		2.52		1.66							0.88		0.70			0.46	
	pow .80 1% .05	7.95 7.24	4.49		3.10				1.99	2.19	1.72	1.56	1.38	1.76	1.62		1.47	
	1% .01	11.60		5.13	4.30	3.80	3.45	3.21	3.02	2.87		2.56	2.37		1.96	1.84	1.72	1.58
	08. woq 08. woq	7.11 12.45		2.71 4.82				1.49 2.53		1.22 2.11	1.12 1.95						0.49	
	5% .05	16.18	8.81		5.05	4.32	3.83	3.49	3.21	3.00	2.84	2.60	2.34	2.09	1.84	1.71	1.57	1.43
	5% .01	22.59 15.62				5.76						3.29					1.84	
	90. woq 90. woq	23.49															0.86	
	120 nil .05	3.91	2 07	2.68	2.45	2.29	2.17	2 00	2 01	1.96	1 01	1 02	1 75	1 66	1 55	1 40	1.43	1 25
	nil .05			3.95													1.43	
	pow .50		2.51			1.42	1.25	1.11	1.09	1.00	0.92	0.87	0.74	0.64	0.54	0.50	0.43	0.36
	pow .80 1% .05	7.93 7.76	4.91	3.71	3.05	2.65		2.12			1.70						0.68 1.45	
	1% .01	12.20	7.02	5.28	4.40	3.86	3.50	3.24	3.04	2.88	2.76	2.56	2.36	2.15	1.93	1.81	1.69	1.55
	08. woq 08. woq	7.58 13.10	4.04 7.05		2.29 3.93				1.36	1.24	1.13		0.87 1.50				0.46	
	5% .05									3.17							1.57	
	5% .01	24.59			7.28		5.35		4.38		3.80		3.02		2.23	2.03	1.83	
	08. woq 08. woq							2.73 4.06					1.47 2.16		0.87 1.29		0.59	
	150 .1 05	2 22	2 26	0.65	2 42	0.07	2 16	0.05	2 00	1 04	1 00		1 70	1 64	1 50		1 40	1.32
	150 nil .05 nil .01	6.80	3.06 4.75	3.92	2.43	2.27	2.16	2 07 2.76	2.00	1.94	1.89	1.81	1.73 2.16	1.64	1.53	1.47	1.40	1.49
	рож .50	3.83	2.50	1.90	1.55	1.41	1.24	1.10	1.08	0.99	0.92	0.86	0.73	0.63	0.54	0.45	0.40	
ſ	1% .05	8.61	5.01	3.86	3.28			2.49			2.17		- 200		1.61		1.44	
	18 .01 pow .50	8.26	7.40 4.42	3.09		3.98 1.98											0.43	
	pow .80	14.11	7.43	5.21	4.09	3.40	2.96	2.62	2.37	2.16	2.00	1.77	1.51	1.25	0.98	0.83	0.68	0.51
	5% .05 5% .01		0.86 4.46	7.64		5.11 6.65	4.48		3.69	3.41	3.20	2.88	2.57 3.17	2.24	1.92		1.59	
	pow .50	19.73		6.86		4.19				2.48			1.61	1.23			0.59	
	08. woq	28.49	4.57	9.81	7.46	6.05	5.10	4,43	3.92	3.56	3.24	2.76	2.31	1.81	1.35	1.09	0.85	0.58

One Stop F Table

Identifying minimum sample size via Murphy, Myors & Woloch (2014)

"One Stop" PV-table (Appendix C)

Minimum sample size needed (Appendix C)

Comparison to a Nil effect

The authors need a minimum sample size of 122 to have power of .80 with the effect they obtained.

Obtained $R^2 = .07$

						One		PV Tabl	Le						
	1	2	3	4	5	6	dfHy 7	MD 8	9	10	12	15	20	30	40
dfErr	-	-	3	7	5	0	,		3	10	12	13	20	30	40
80 nil .05	0.047	0.072	0.093	0.111	0.127	0.142	0.157	0.170	0.183	0.196	0.219	0.251	0.298	0.375	0.435
nil .01	0.080	0.109	0.131	0.151	0.169	0.185	0.201	0.215	0.229	0.242	0.266	0.299	0.346	0.421	0.480
pow .50			0.068												
pow .80			0.124												
1% .05			0.113												
1% .01			0.158												
08. wog	0.130		0.089												
5% .05	0.130														
5% .03	0.204														
pow .50	0.147														
pow .80	0.211	0.216	0.221	0.225	0.229	0.233	0.239	0.243	0.247	0.250	0.260	0.270	0.287	0.319	0.347
90 nil .05	0.042														
nil .01			0.118												
pow .50	0.041														
pow .80 1% .05			0.104												
18 .01	0.072														
pow .50	0.071														
08. wog	0.119	0.128	0.136	0.144	0.151	0.156	0.164	0.169	0.174	0.178	0.189	0.201	0.223	0.258	0.284
5% .05	0.144	0.156	0.167	0.178	0.188	0.198	0.208	0.217	0.227	0.235	0.253	0.277	0.315	0.378	0.431
5% .01	0.193														
pow .50	0.142														
pow .80	0.200	0.204	0.208	0.212	0.216	0.220	0.223	0.227	0.230	0.235	0.241	0.253	0.269	0.298	0.320
100 nil .05	0.038	0.58	0.075	0 000	0.103	0 116	0 128	0 140	0 151	0 161	0 182	0 200	0 251	0 320	0 377
nil .01	0.064														
pow .50	0.037														
pow .80	0.074														
1° .05	0.067	0.082	0.096	0.108	0.121	0.132	0.143	0.154	0.164	0.174	0.193	0.220	0.260	0.327	0.383
1% .01			0.133												
pow .50	0.066														
pow .80			0.126												
5% .05 5% .01			0.158												
50 .01 00w .50	0.135														
08. wog	0.190														
-															
120 nil .05	0.032														
nil .01	0.054	0.074	0.090	0.104	0.117	0.129	0.140	0.151	0.161	0.171	0.189	0.215	0.253	0.317	0.370
2011 80	0.062	076	0.005	0.002	0.000	0.105	0.110	0 116	0.120	0.124	0.122	0.142	0.150	0 100	0.212
pow .80	0.062														
18 .01			0.117												
pow .50			0.068												
08. woq			0.111												
5% .05			0.147												
5% .01			0.187												
pow .50			0.129												
08. woq	0.175	D.I/8	0.181	0.184	0.187	0.189	0.192	0.194	0.198	0.200	0.204	0.213	0.222	0.244	0.265
		7													

Minimum sample size needed (Appendix C)

Comparison to a 1% effect

The authors need a minimum sample size of 302 to have power of .80 with the effect they obtained.

Obtained $R^2 = .07$

One Stop PV Table dfErr 0.052 0.057 0.061 0.064 0.068 0.071 0.074 0.076 0.081 0.088 0.098 0.114 0.130 0.072 0.078 0.084 0.089 0.095 0.100 0.111 0.127 0.151 0.195 0.051 0.052 0.053 0.054 0.055 0.059 0.060 0.062 0.067 0.070 0.084 0.092 0.148 0.153 0.158 0.163 0.168 0.172 0.177 0.182 0.191 0.204 0.225 0.263 0.298 0.123 0.123 0.124 0.126 0.126 0.127 0.127 0.128 0.129 0.130 0.133 0.136 0.142 0.149 0.112 0.112 0.113 0.113 0.113 0.114 0.114 0.116 0.116 0.118 0.119 0.124 0.128 0.015 0.019 0.021 0.023 0.025 0.026 0.028 0.029 0.030 0.031 0.033 0.036 0.039 0.047 0.052 0.030 0.031 0.031 0.032 0.033 0.033 0.033 0.036 0.037 0.040 0.043 .80 0.103 0.103 0.104 0.104 0.104 0.104 0.106 0.106 0.106 0.106 0.107 0.108 0.110 0.112 0.116

Part 2: Practice with bayesian analyses

A quick introduction to bayesian statistics

- Increasingly popular in psychology
- Allows us to make probabilistic statements about events/hypotheses (tells us the probability of specific events given the observed data included in the analysis)
- Models account for background knowledge (not discussed in detail in this lab)
- May allow us to overcome several limitations of NHSTs (e.g., the reproducibility crisis in psychology and reliance on large sample sizes)
- All common statistical analyses can be conducted in a bayesian framework (ANOVAs, regression, correlation, factor analysis, etc.)
- Interpretations are often more concrete than significance tests (clear probability statements)

Note: While bayesian analyses are on the rise and may overcome several limitations of frequentist methods, use of classical statistics has been argued for as well (e.g., this NY Times article from 2014: https://www.nytimes.com/2014/09/30/science/the-odds-continually-updated.html?_r=1)

Compare bayesian and frequentist statistics

	Frequentist statistics	Bayesian statistics
Definition of the <i>p</i> value	The probability of observing the same or more extreme data assuming that the null hypothesis is true in the population	The probability of the (null) hypothesis
Large samples needed?	Usually, when normal theory-based methods are used	Not necessarily
Inclusion of prior knowledge possible?	No	Yes
Nature of the parameters in the model	Unknown but fixed	Unknown and therefore random
Population parameter	One true value	A distribution of values reflecting uncertainty
Uncertainty is defined by	The sampling distribution based on the idea of infinite repeated sampling	Probability distribution for the population parameter
Estimated intervals	Confidence interval: Over an infinity of samples taken from the population, 95% of these contain the true population value	Credibility interval: A 95% probability that the population value is within the limits of the interval

This table is from: Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & Van Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. Child development, 85(3), 842-860.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4158865/

Table 1. Evidence Categories for p Values (adapted from Wasserman, 2004, p. 157), for Effect Sizes (as proposed by Cohen, 1988), and for Bayes Factor BF_{A0} (Jeffreys, 1961)

Statistic Interpretation

Decisive evidence against H_0 Substantive evidence against H_0

Positive evidence against H_0 No evidence against H_0

Decisive evidence for H_0

Effect size < 0.2 Small effect size

b value <.001

.001 - .01

.01 - .05

<1/100

>.05

0.2 - 0.5Small to medium effect size 0.5 - 0.8Medium to large effect size

8.0 Large to very large effect size Bayes factor

>100 Decisive evidence for H_A 30-100 Very strong evidence for H_A

10-30 Strong evidence for H_A 3-10 Substantial evidence for H_{Δ} 1-3Anecdotal evidence for H_A No evidence

1/3-1Anecdotal evidence for H_0 1/10-1/3 Substantial evidence for H_0 1/30-1/10 Strong evidence for H_0 1/100-1/30 Very strong evidence for Ho

Note: For the Bayes factor categories, we replaced the label "worth no more than a bare mention" with "anecdotal." Also, in contrast to p values, the Bayes factor can quantify evidence in favor of the null hypothesis.

Rules of Thumb for Bayes Factor **Interpretations**

Bayes factors are indices of *relative* evidence of one model (or hypothesis) over another

Evidence for alternative hypothesis (compared to null hypothesis)

Evidence for null hypothesis (compared to alternative hypothesis)

Wetzels et al., 2011

https://journals.sagepub.com/doi/abs/10.1177/1745691611406923

Load libraries and read in data

```
# Load libraries
 ``{r}
library(tidyverse)
library(BayesFactor)
# read in retirement.csv
 ``{r}
retirement <- read_csv("retirement.csv")</pre>
Parsed with column specification:
 cols(
   occupation = col_double(),
   sex = col_double().
   mental = col_double()
```

The retirement dataset has N=1910 for 3 variables:

Occupation: 1 = professor, 2 = manager, 3 = non manual worker, 4 = skilled worker, 5 = semi-skilled worker, 6 = unskilled worker

Sex: 1 = female, 2 = male

Mental: Continuous outcome variable indicating mental health (1-5)

Factor the categorical variables

Tell R to read sex identity and occupation as categorical when conducting analyses

1) Conduct a regular ANOVA in which gender, occupation, and the interaction between the two predict mental health

```
66 - # Normal ANOVA
67 - 11 {r}
    summary(aov(lm(mental~ sex.f*occupation.f, data = retirement)))
                         Df Sum Sq Mean Sq F value
                                                    Pr(>F)
                                            0.256
     sex.f
                                    0.206
                                                     0.613
                                     9.187 11.396 7.44e-11 ***
     occupation.f
     sex.f:occupation.f
                               6.0
                                     1.202
                                           1.491
                                                     0.189
     Residuals
                    1898 1530.0
                                     0.806
     Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results indicate that only the main effect of occupation significantly predicts mental health. This effect has the largest eta squared value, suggesting that occupation explains ~3% of the variance in mental health.

2) Conduct a bayesian ANOVA, starting with all models

```
m1 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "all") m1
```

```
100%
Bayes factor analysis
[1] sex.f
                                                 : 0.05966566
                                                                ±0%
                                                                \pm 0.01\%
[2] occupation.f
[3] sex.f:occupation.f
                                                                \pm 0.1\%
[4] sex.f + occupation.f
                                                                \pm 2.36\%
                                                 : 4852553
[5] sex.f + sex.f:occupation.f
                                                : 0.0006680688 ±3.05%
[6] occupation.f + sex.f:occupation.f
                                                : 1648748
                                                                ±1.45%
[7] sex.f + occupation.f + sex.f:occupation.f : 122412.5
                                                                ±4.21%
Against denominator:
  Intercept only
Bayes factor type: BFlinearModel, JZS
```

Gives us the bayes factor for each possible model (i.e., all combinations of effects that we can test). In this notation, a:b denotes an interaction term.

2) Conduct a bayesian ANOVA, starting with all models

```
m1 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "all")
m1</pre>
```

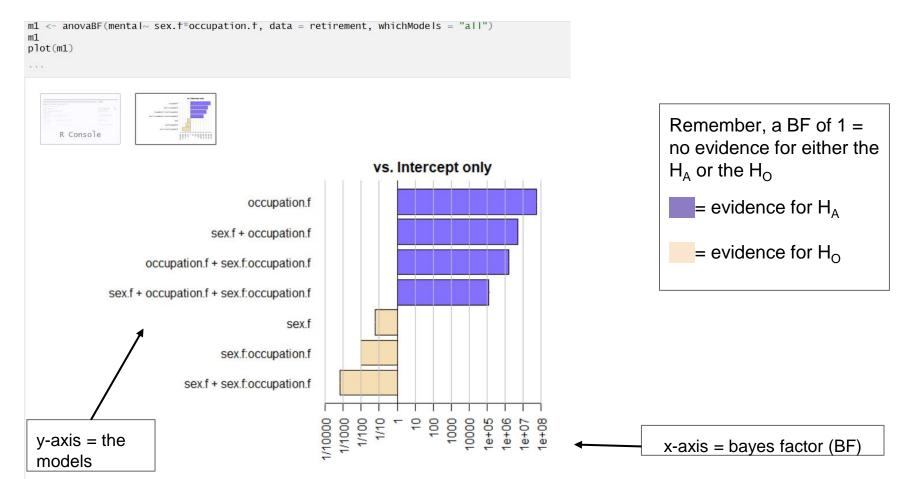
```
100%
Bayes factor analysis
[1] sex.f
                                                : 0.05966566
                                                               ±0%
                                                  54417793
                                                               \pm 0.01\%
[2] occupation.f
[3] sex.f:occupation.f
                                                  0.009737589
                                                               ±0.1%
[4] sex.f + occupation.f
                                                 4852553
                                                               ±2.36%
[5] sex.f + sex.f:occupation.f
                                                  0.0006680688 \pm 3.05\%
[6] occupation.f + sex.f:occupation.f
                                                : 1648748
                                                               ±1.45%
[7] sex.f + occupation.f + sex.f:occupation.f : 122412.5
                                                               ±4.21%
Against denominator:
 Intercept only
Bayes factor type: BFlinearModel, JZS
```

Main takeaway:

Models that include occupation as a main effect have large bayes factors

BF > 100 = "decisive evidence for H_A "

2a) Plot the bayes factor values for all models



3) Conduct a bayesian ANOVA, using a "top-down" approach

```
m2 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "top")
m2</pre>
```

Gives us the change in the model's bayes factor when each effect is eliminated one at a time

3) Conduct a bayesian ANOVA, using a "top-down" approach

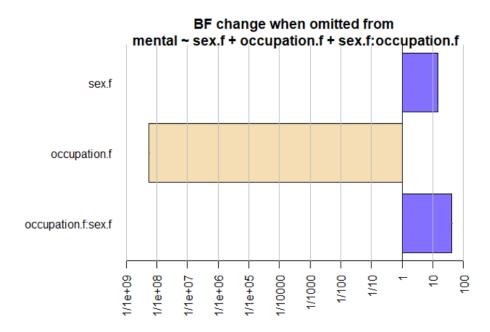
```
m2 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "top")
m2</pre>
```

Main takeaway:

Removing occupation leaves a model with a very small bayes factor

3a) Plot the change in bayes factor values for the "top-down" approach

```
m2 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "top") m2 plot(m2)
```



4) Conduct a bayesian ANOVA, using a "bottom-up" approach

```
m3 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "bottom") m3 \,
```

```
100%

Bayes factor analysis

------------

[1] sex.f : 0.05966566 ±0%

[2] occupation.f : 54417793 ±0.01%

[3] sex.f:occupation.f : 0.009737589 ±0.1%

Against denominator:
   Intercept only
---

Bayes factor type: BFlinearModel, JZS
```

Gives us the change in the model's bayes factor when each effect is added one at a time

4) Conduct a bayesian ANOVA, using a "bottom-up" approach

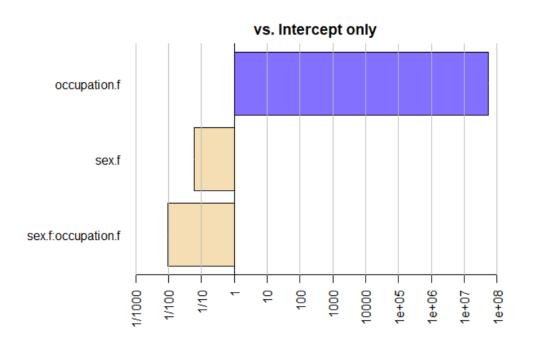
```
m3 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "bottom")
m3</pre>
```

Main takeaway:

Adding occupation increases the model's bayes factor by a large amount

4a) Plot the change in bayes factor values for the "bottom-up" approach

```
m3 <- anovaBF(mental~ sex.f*occupation.f, data = retirement, whichModels = "bottom")
m3
plot(m3)
```



Summary of results from the Bayesian analyses

All three bayesian analysis steps ("all models", "top-down", and "bottom-up") support the same conclusion:

- There is decisive evidence to support the hypothesis that occupation influences mental health.
- There is strong evidence to support the *null hypothesis* that sex identity influences mental health (i.e., the null hypothesis, stating that gender does *not* predict mental health, is likely to be true).
- There is decisive evidence to support the *null hypothesis* that the interaction between sex identity and occupation influences mental health (i.e., the null hypothesis is likely to be true).

Compare the regular ANOVA vs. Bayesian ANOVA

```
Bayesian "all models" ANOVA
66 - # Normal ANOVA
                                               NHST ANOVA
                                                                   Bayes factor analysis
67 - ```{r}
   summary(aov(lm(mental~ sex.f*occupation.f, data = retirement)))
                                                                                                                    0.05966566
                                                                                                                                   ±0%
                                                                       occupation.f
                                                                                                                     54417793
                                                                                                                                   \pm 0.01\%
                                                                   [3] sex.f:occupation.f
                                                                                                                     0.009737589
                                                                                                                                  ±0.1%
                                                                   [4] sex.f + occupation.f
                                                                                                                     4852553
                                                                                                                                   ±2.36%
                        Df Sum Sq Mean Sq F value
                                                   Pr(>F)
                                                                   [5] sex.f + sex.f:occupation.f
                                                                                                                    0.0006680688 ±3.05%
     sex.f
                                    0.206
                                           0.256
                                                    0.613
                                                                   [6] occupation.f + sex.f:occupation.f
                                                                                                                   : 1648748
                                                                                                                                   \pm 1.45\%
                                          11.396 7.44e-11 ***
                                    9.187
    occupation.f
                                                                      sex.f + occupation.f + sex.f:occupation.f : 122412.5
                                                                                                                                   ±4.21%
     sex.f:occupation.f
                              6.0
                                   1.202
                                           1.491
                                                    0.189
     Residuals
                                   0.806
                      1898 1530.0
                                                                   Against denominator:
                                                                    Intercept only
    Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
                                                                   Bayes factor type: BFlinearModel, JZS
```

Both analyses indicate that occupation predicts mental health, while gender and gender*occupation do not.

The bayesian approach provides information that the frequentist approach does not:

- degree of evidence for the null hypothesis
- concrete statements about how well supported each model hypothesis is compared to the alternative hypothesis (e.g., null vs. alternative)

Additional resources on bayesian approaches

An article by Etz & Vandekerckhove (2018) about basic bayesian inferences. It opens with a quote by Dumbledore, so you know you want to read it! https://link.springer.com/article/10.3758/s13423-017-1262-3

Helpful tutorials for learning bayesian analyses using the BayesFactor package:

https://richarddmorey.github.io/BayesFactor/#fixed

More great tutorials for getting started with bayesian analyses, this time from the BayestestR package: https://cran.r-project.org/web/packages/bayestestR/vignettes/bayestestR.html

The accompanying citation for the BayestestR package can be found at: https://www.theoj.org/joss-papers/joss.01541/10.21105.joss.01541.pdf

An example of using top down and bottom-up approaches with bayesian analyses:

https://datascienceplus.com/bayesian-statistics-analysis-of-health-data/

An article by Krypotos et al. (2017) that calls for increased use of Bayesian approaches (and less NHST) in experimental psychology: https://journals.sagepub.com/doi/10.5127/jep.057316