All handouts for this class: **https://tinyurl.com/IST772crowston**

IST772 Associations (Week 7)

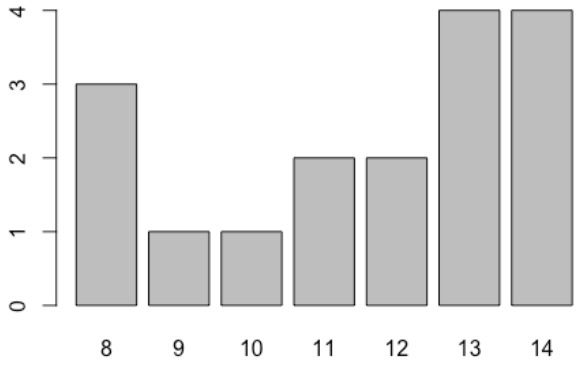
Pre-class activity:

- Open the file Week7correlationCalcsInExcel.xlsx from the handouts area.
- Why are r and b the same? When might they be different? Report your answer on the chat.

1

General Exam Feedback

- Grading: No perfect scores; modal scores were 14 and 13 with a long left tail



2

Observations

- Overall: Considerable headway concerning the various modes of inferential testing: confidence intervals, frequentist NHST, Bayesian Estimation
- Most common shortcomings
  - Not tackling the contradiction between the frequentist and Bayesian results with some kind of explanation to steer the client towards understanding the findings and its implications
  - Not explaining the tests in plain English (which is admittedly hard to do)
- A few persistent terminology/language problems:
  - Describing the meaning of an observed p-value
  - Describing the "long run" nature of a confidence interval (if we had 100 CIs they'd all be different)

3

Week7correlationCalcsInExcel

Try three of the following (or your own experiment):

1. Keeping the same set of values, modify the order of observations in the Heat vector to create a (non-zero) positive correlation.
2. Keeping the same set of values, modify the order of observations in the Heat vector to create a (non-zero) negative correlation.
3. Keeping the same set of values, modify the order of observations in the Heat vector to create a perfect correlation, either plus one or minus one.
4. Add a cell that calculates R squared from the displayed r value.

4

2. Week7corrExercise.R

Question: Is there a significant correlation between CO_2 concentration and CO_2 uptake by plants?

Evidence:

1. NHST: r is transformed into a t distributed variable and used to test the null hypothesis that $\rho = 0$
2. CI for ρ from the t test
3. MCMC posterior distribution of ρ
4. Bayes Factor for model with $\rho \neq 0$

5

Breakout 1 – Freq and Bayesian Chi-Square

- Work with the built-in HairEyeColor data set, a 3D contingency table that we will squash down to a 4x4
- Conduct an initial chi-square
- Look at residuals to focus on a 2x2 block
- Conduct Bayesian estimation on the proportions
- Plot difference in proportion between rows
- Share your code on https://codeshare.io/aJDyRX

All handouts for this class: **https://tinyurl.com/IST772crowston**

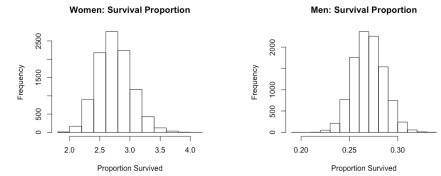
6

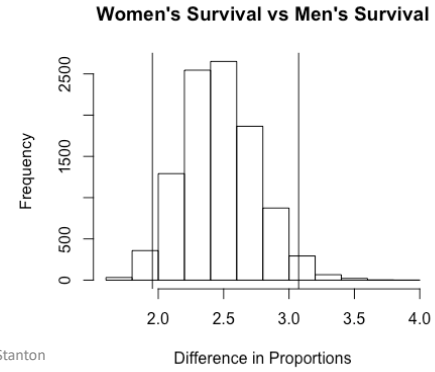
Breakout: Bayesian Modeling of Proportions

```
     Survived     No  Yes
Sex
Male             1364  367
Female            126  344

Bayes factor analysis
--------------
[1] Non-indep. (a=1) : 4.62125 ±0%

Against denominator:
  Null, independence, a = 1
---
Bayes factor type:
BFcontingencyTable, poisson
```





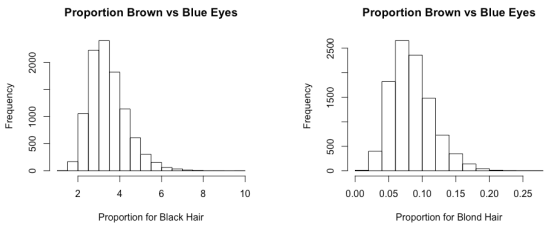
7

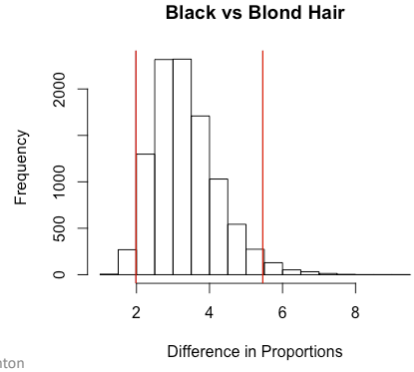
Breakout 1 Results

```
            Eye
Hair    Brown Blue
  Black    68   20
  Blond     7   94

Bayes factor analysis
--------------
[1] Non-indep. (a=1) :
6.296454e+22 ±0%

Against denominator:
  Null, independence, a = 1
---
Bayes factor type:
BFcontingencyTable, poisson
```





8

Interpreting a Correlation Matrix

- See the handout: Week7interpretCorrMatrix.pdf
- This is the final example in the handout.
- What happened to the upper triangle?
- What are the dashes?
- What are the asterisks?

Table 3
Correlations Between the Abridged Job in General (AJIG) Scale, Withdrawal Behaviors, and Organizational Commitment

	M	*SD*	AJIG	Prep1	Act1	Prep2	Act2	AC	CC
AJIG	17.52	6.83	—						
Preparatory job search behavior, Time 1 (Prep1)	6.99	4.45	–0.21*	—					
Active job search behavior, Time 1 (Act1)	4.25	5.01	–0.19*	0.65**	—				
Preparatory job search behavior, Time 2 (Prep2)	7.34	4.79	–0.16	0.72**	0.55**	—			
Active job search behavior, Time 2 (Act2)	4.17	4.95	–0.18	0.60**	0.82**	0.71**	—		
Affective commitment (AC)	22.71	7.25	0.59*	–0.24**	–0.15	–0.24*	–0.26**	—	
Continuance commitment (CC)	19.52	6.44	–0.24**	0.12	0.16	0.09	0.33**	–0.04	—

Cc

*$p < .05$, two-tailed. **$p < .01$, two-tailed.

9

Paper of the Week – Wetzels & Wagenmakers, 2012

- Explains the use of the Bayes Factor for evaluation correlation coefficients
- Provides a table with the Jeffreys (1961) guidance about the different "strengths" of Bayes Factors
- Discusses choice of priors and contains R code

A default Bayesian hypothesis test for correlations and partial correlations

Ruud Wetzels · Eric-Jan Wagenmakers

10

Homework

- The homework for week 7 is based on exercises 3, 4, 8, 9, 10 on pages 155-156 but with changes as noted in the notebook.
- Note that exercises 8, 9, and 10 should be done on the same subset of the admissions data, i.e., UCBAdmissions[,,1]
- This is the only remaining week without a practice exam! Enjoy!