

## Class 34: Review of Key Concepts

**Held:** Wednesday, 23 April 2008

### Notes:

- I got enough questions on Topic 25 that I will not hold a quiz today.
- If you plan to take the makeup exam tomorrow, please let me know when asap. Makeups will be in 3813.
- I will be around that room to answer questions on the project all morning, too.
- Projects remain due Friday. I will look at them over the weekend and, if I have suggested changes, may email those suggestions to you.
- A peer-review form is now available, and gives hints as to the kinds of things I will expect.
- On Friday, we return to the book, and will work on topic 20.

### Overview:

- Notes on chi-square tests.
- Questions and answers on the exam.
- More fun with the project.

## Chi-Square Tests, Revisited

- We have seen two kinds of chi-square tests.
  - One is used with a single categorical variable (goodness of fit)
  - One is used with a pair of categorical variables (more general)
- They have two different kinds of hypothesis:
  - For goodness-of-fit, the hypothesis is "The distribution of the different responses is as follows."
  - For two-variable, the hypothesis is "The two variables are independent."
- They have the same technique for finding the test statistic
  - Compute an expected value for each entry.
  - Compute the scaled squared difference between observed and expected.
  - Sum those values.
- They use the same table for computing probabilities.
- Given the different hypotheses, they have different techniques for computing expected values:
  - For goodness-of-fit, it's "expected proportion time sample size"
  - For two-sample, it's "sum of row \* sum of column / sample size"
- Why is the two-sample one given that way?
  - Well, if the two variables are independent, the proportions in any row should be the same as the proportions within the population.
  - We compute the proportion in the population by summing the column and dividing by the sample size.

- This works the other way, too: The proportions in any column should be the same as the proportion in the population.
- We compute the proportion in the population by summing the row and dividing by the sample size.

## Questions and Answers on Exam 2

- I'll spend as much or as little time as necessary going over any questions you may have on exam 2.
- Note that I grade without names, so I generally don't know what any particular person got wrong. (I do know scores, but that doesn't tell me much.)

## Some Project Notes

- The draft poster peer review form should give you a sense of things to work on.
- I have not found a good way to join rows or columns, other than "by hand". However, there are still good reasons to merge.
- I'll do my best to answer any other questions you may have.

---

Copyright © 2008 Samuel A. Rebelsky. This work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/2.5/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.