FINAL EXAMINATION

- When and where:
  - Convocation final: 4:00-6:00 PM, Wednesday June 4th, in Eckhart 133.
  - Everybody else: 6:30–8:30 PM Monday, June 9th, in Eckhart 133.
- Covers the entire course, with emphasis on material since the midterm:
  - FPPA Chapters 16–23, 26 (omitting the t-test), 27–29, and the corresponding class lectures.
- Format:
  - Similar to the Midterm, but not as long relative to the given amount of time.
  - No major questions on Correlation, Regression, or Chance.
- Bring:
  - 1 or 2 double-sided pages of notes (i.e., your sheet for the Midterm, plus a new sheet for the Final).
  - Working calculator.

SUGGESTIONS FOR PREPARING FOR THE FINAL

- Compare your homework and quiz solutions to those posted in the Stat 200 display case.
- Reread parts of FPPA covered since the Midterm.
  - Especially the chapter summaries.
- Review your lecture notes.
  - Annotated copies of slides used in class are on reserve in Eckhart Library (directly above E133).
- Work problems from the Exercise Sets in FPPA.
  - Save time by skipping the arithmetic; concentrate on getting the problem set up.
  - Check your answers against those in the book.
- Make sure your sheet of notes for the final covers the list of review topics on pages 5–8 of this handout.
- See the instructors and/or TAs for help.
  - I will be available most of this week. Email me any time, or call ahead to arrange a specific time.

CHAPTER 29: MORE ON TESTS OF SIGNIFICANCE

• What is the role of the box model in a statistical analysis?

  • Setting up the box model makes you define the population (the tickets) and how the sample was taken (the draws).

  • The box model determines the chances of how the sample might have turned out:

    • SE's, P-values, Confidence intervals.

  • If the model is inappropriate, so are the conclusions based on it.

• If a z-test concludes that some difference is statistically significant, is that difference necessarily big and important?

  • No. A statistically significant difference is one that is hard to explain away as chance variation.

    • It's too many SEs away from what was expected under the null.

    • But with a large sample size, that can happen even when the difference itself is small and unimportant, just because the large sample size makes the SE small.

  • How should you assess how big a difference is?

    • With a confidence interval.

  • How should you assess how important a difference is?

    • By thinking about what it means in the population.

• Does rejecting the null hypothesis (of chance variation) necessarily imply the stated alternative?

  • No, something else may have caused the difference.

  • The study needs to be well-designed and well-conducted.

• Is it a good idea to reject or accept the null hypothesis?

  • No, not unless you have to. In general it makes more sense to assess the strength of the evidence against the null hypothesis by reporting the P-value, and let people draw their own conclusions.

• Is a statistically significant result (P-value $< 5\%$) necessarily real?

  • No. When the null hypothesis is true, there is a 5% chance of getting a result that is statistically significant, and a 1% chance of getting a result that is highly statistically significant.

  • If you make 100 tests, you can expect to get about 5 statistically significant results, and 1 highly statistically significant result, even when the null hypotheses are true in every case.

    • Those results are just flukes.

• Is it a good idea to decide what tests to make after looking at the data?

  • No. You're likely to discover the flukes.

• Is it a good idea to decide to formulate the alternative hypothesis after looking at the data?

  • No. That's like letting the data suggest what tests to make.

  • When should the alternative be formulated?

    • At the design stage.

• There are a number of other important issues covered in Chapter 29: be sure to read it.

REVIEW — SOME BIG IDEAS

- Box models are a way of modeling chance processes.
- Arguing from the box to the draws — Probability theory
  - You know what's in the box.
  - The draws are yet to be made.
  - Goal is to say something about how the draws are likely to turn out.
  - Sample result: When drawing at random with replacement, the average of the draws will be about the average of the box, give or take

    $$\text{SE for the average of the draws} = \frac{\text{SD of the box}}{\sqrt{\text{number of draws}}}$$

    or so. Moreover, the normal approximation can be used to figure chances about the average of the draws, provided sufficiently many draws will be made from the box.
- Arguing from the draws to the box — Statistical inference
  - The contents of the box are partially or completely unknown.
  - The draws have been made.
  - Goal is to say something about the tickets in the box.
  - Sample result: If the draws were made at random with replacement, then you can assert, with 95% confidence, that the (unknown) average of the box is within 2 SE's of the (known) average of the draws.
    - Here "SE" refers to the standard error for the average of the draws.
    - There need to be enough draws.

STANDARD ERRORS

- For numbers that are generated by a chance process, one can write

  $$\text{number} = \text{expected value} + \text{chance error}.$$

- The likely size of the chance error in the number is called the standard error (SE).
  - The SE (as well as the expectation) depends on the chance process involved.
- We have formulas for several chance processes which involve drawing tickets at random with replacement from a box:

SE for the	*Formula*
sum of draws	$\sqrt{\text{number of draws}} \times (\text{SD of box})$
average of draws	$\dfrac{\text{SE for sum}}{\text{number of draws}}$
count	SE for sum, from a 0–1 box
percent	$\dfrac{\text{SE for count}}{\text{number of draws}} \times 100\%$

- When reasoning forward from the box to the draws, the SD of the box can be computed from the known composition of the box.
- When reasoning backwards from the draws to the box, you usually have to estimate the SD of the box from the sample.
  - For 0–1 boxes, use the bootstrap method.
  - For other boxes, estimate the SD of the box by the SD of the sample.

DIFFERENCES

- The difference

$$\text{1st quantity} - \text{2nd quantity}$$

between two independent chance quantities:

- has expected value equal to

$$(\text{exp'd value of 1st quantity}) - (\text{exp'd value of 2nd quantity});$$

- has SE equal to

$$\sqrt{(\text{SE of 1st quantity})^2 + (\text{SE of 2nd quantity})^2};$$

- follows the normal curve if each of the two quantities does.

SOME TOOLS OF STATISTICAL INFERENCE

- Confidence interval
  - Range of plausible values for some population parameter, given how the sample turned out.
  - Under appropriate conditions,

$$\text{Estimate} \pm (\text{SE of estimate})$$

is an approximate 68% confidence interval, and

$$\text{Estimate} \pm 2 \times (\text{SE of estimate})$$

is an approximate 95% confidence interval.

- Test of significance
  - Used to decide if an observed difference is real, or just chance variation.
  - The P-value is the chance, computed under the null hypothesis, of getting a difference as extreme, or more so, than the observed difference.
    - The P-value is measure of how consistent the data are with the null hypothesis.

SOME SCENARIOS TO RECOGNIZE AND KNOW HOW TO DEAL WITH

- Make an inference (confidence interval and/or significance test) about the average of a box, based on a SRS from the box.
  - One sample z.
- Make an inference about the difference between the averages of two boxes, based on independent SRSs from each box.
  - Two sample z.
- As above, but for proportions instead of averages.
- Test whether the contents of the box follow some fully specified model, based on a SRS from the box.
  - χ^2-goodness of fit test.
- Test whether two attributes are independent, based on a SRS.
  - χ^2-test for independence.

CAVEAT EMPTOR

- Know when the methods apply, and when they don't.