

Final Practice Exam
Statistics 160 – Tom LaGatta – Fall 2006

This practice exam is **not guaranteed** to prepare you for the test! The only way to accurately prepare is to reread each section of the textbook and identify the major concepts and types of questions I could ask. This practice exam is meant to get this process started for you, *not* to exhaust all the possibilities of material I may test you on.

You may bring a normal-sized notecard to the final exam, with any pertinent information on it. If you abuse this privilege I will take it away. Use your common sense.

You should write down your answers to the below questions as complete sentences. Each of these questions should look familiar to you, but it's not until you attempt to write down an answer that you will know whether you actually know the material or not. Remember, you cannot learn mathematics by reading; the only way to learn it is by doing.

Most of the questions I've written for the practice exam deal with *concepts*, not *calculations*. Finding calculation-based questions is easy: just open up the textbook and pick random questions to answer, or look at past homework assignments and tests.

New Material

Chapter 19

- T or F: 2 sample t -procedures must be dependent.
- T or F: In 2 sample t -procedures, the standard deviation gets smaller as the sample sizes n_1 and n_2 increase.
- Here are the IQ test scores from 31 seventh grade girls in a Midwest school district.

114 100 104 89 102 91 114 114 103 105 108 130 120 132 111 128
118 119 86 72 111 103 74 112 107 103 98 96 112 112 93

Here are the IQ test scores from 47 seventh grade boys in the same district

111 109 124 93 90 107 113 126 123 114 100 128 116 79 106 107
128 127 119 115 118 119 110 111 113 97 110 97 124 102 107 112
127 110 105 104 136 120 105 106 113 123 115 77

Is there good evidence that the boys and the girls differ in their IQ scores?

- To measure the presence of spores, air samples are pumped into an agar plate and “colony forming units” are counted after an incubation period. Here is the data from the ‘kill room’ of a plant that processes 37,000 turkeys per day, taken on four different days in the summer and the winter. The units are CFUs per meter of air.

Summer: 3175, 2526, 1765, 1090

Winter: 384, 104, 251, 97

Give a 90% confidence interval to estimate how much higher the mean count is during the summer.

Chapter 20

- A student project put this question to an SRS of 172 undergraduate students at a large college. “You witness 2 people cheating on a test. Do you tell the professor?” Only 19 answered “Yes”. Give a 95% confidence interval for the proportion of all undergraduates at this university who would report cheating.
- The IRS plans to examine an SRS of individual income tax returns from each state. One variable of interest is the proportion of returns claiming itemized deductions. The total number of tax returns in a state varies from more than 15 million in California to about 240,000 in Wyoming.
 - a. Will the margin of error for estimating the population proportion change from state to state if an SRS of 2000 tax returns is selected in each state?
 - b. Will the margin of error change from state to state if an SRS of 1% of all tax returns is selected in each state?

Chapter 21

- A study by the National Athletic Trainers Association surveyed 1679 high school freshman and 1366 high school seniors in Illinois. Results showed that 34 of the freshman and 24 of the seniors had used anabolic steroids.
 - a. In order to draw conclusions about all Illinois freshman and seniors, how should the samples be chosen?
 - b. Give a 95% confidence interval for the proportion of all high school freshman in Illinois who have used steroids.
 - c. Is there a significant difference between the proportions of freshman and seniors who have used steroids?

Chapter 23

- 23.41

- 23.43

Old Material

Part I

- What are some qualitative ways to describe data? Quantitative? What are the pros and cons of these different descriptions?

Part II

- What is the difference between a sample and an experiment?

- What is “bias” in the context of samples and experiments? Give a few examples of obtaining biased data.

- What *is* probability? In particular, if I say that an event has a probability of $1/3$ of happening, what does this mean? Give at least two different answers.

- Give two different examples of probability model. Make sure to identify the sample space and a possible event. Do all the different outcomes have equal probability?

- On a roll of two fair dice, what is the probability that you roll a 7 on one face? What about a roll which sums to 7? Ask and answer two more similar questions.

- In your own words, what is a random variable?

- In the context of chapter 11, explain why \bar{x} is a random variable. Under the usual conditions (what are they?), what is the distribution of \bar{x} ? Make sure to identify its mean and standard deviation.

- What is a parameter? Give three examples of parameters we like to estimate. What is a statistic? Give three examples of statistics we like to calculate.

- Write down the general addition and multiplication rules. Under what conditions do these become the more familiar, simple addition and multiplication rules?

- Recall the Monty Hall problem. Draw a probability tree diagram to demonstrate why always switching is an optimal strategy.

- Explain how our study of binomial distributions was essential for estimating population parameters later on.

- When is the Normal approximation for binomially-distributed random variables appropriate? Give an example of a question that uses this idea.

Part III

- In a paragraph, explain the reasoning behind confidence intervals.

- A newspaper says, “The average grade in Tom LaGatta's classes is an 82.5, with a margin of error of 33 points at a 95% confidence level.” Does this mean that 95% of students in my classes get between a 79.5 and an 85.5? Why or why not?

- In a paragraph, explain the reasoning behind significance tests.

- What is a P -value? Hint: It is *not* the probability that the null hypothesis is true.
- Why do we use z -procedures for estimating population means? Hint: use the ideas in chapter 11.
- When doing a two-sided z -test, you calculate $z = 2$. Without doing any computation, explain why the P -value is approximately .05. Use Table A to find a more accurate P -value.
- When do we use t -procedures over z - ones?
- Draw some sample t -distributions. How do these compare to the z -distribution?

- What are degrees of freedom?
- Explain the reasoning behind two sample procedures.
- Why do we not use two-sample procedures for matched pairs data?
- We use the F -test to infer about standard deviations. Why should we rarely, if ever, do this?

- Fill in the table:

<i>Overview of basic inference methods</i>			
	<i>Type of Parameter</i>	<i>Number of Samples</i>	<i>Type of Test</i>
Inference About...	Population Means μ	One Sample	
			Two-sample <i>t</i> -test
			One-sample <i>z</i> -test
		Matched Pairs	
			Two Samples
	Categorical Variables	Two-Way Table	