

Testing a Claim about a Parameter

Definition: Statistically significant. A set of data in a statistical study is called statistically significant if it is unlikely to have occurred by chance.

For each of these events, does it seem to you to be statistically significant, and why?

1. In terms of average annual rainfall in Austin, Texas (approximately 34 inches), six of the years between 2000 and 2009 exceeded the average.
2. In 20 tosses of a fair coin, you get 9 heads.
3. In 60 rolls of a fair die, you get a three only twice.
4. The major league baseball team with the worst win-loss record defeats the team with the best win-loss record.
5. In terms to the global average temperature, five of the years between 1990 and 1999 were the five hottest years of the 20th century.

Review of Ch. 11. (In the left margin here, or on another sheet of paper, draw and label appropriate pictures as part of your solution.) The Survey of Student Habits and Attitudes (SSHA) is a psychological test that measures students' study habits and attitudes toward school. Scores are normally distributed and range from 0 to 200. The mean score for college students is 115 and the standard deviation is 30.

- a. What is the probability that an individual score in this distribution is above 125.8?
- b. What is the probability that the mean of a SRS of size 25 has mean above 125.8?

Chapter 15: Hypothesis testing

In that Ch. 11 review problem, suppose you had a sample of size 30 whose mean was 125.8 from this population of college students. Would that surprise you? Would you consider that "statistically significant?"

Suppose it turned out that it wasn't a simple random sample from all college students, but it was a simple random sample from all older college students. What would that suggest to you?

How low would this probability you computed in part b. have to be for you to say that this event is “statistically significant?” (Think about your answer to our activity at the beginning of class. If you are having trouble thinking of one, you may choose among these: 1.00, 0.90, 0.50, 0.20, 0.10, 0.05, 0.01 or you may choose a different probability. This is YOUR significance level – at least for now.) Less than or equal to _____

When we think about a “statistically significant” event, we have to be comparing to some “baseline” or some view of the relevant world. When we clarify what the view is, we call it the “null hypothesis.” For this problem, we would consider that **the population of older college students has some unknown population mean, called μ** and our original idea is that would be **Ho: $\mu = 115$** . (Notice that we don’t need a symbol for the population mean of ALL college students because we have been told that is 115.)

When we’re thinking of whether this data is far enough away from that “baseline” which is described by the null hypothesis, our typical idea is that it is far away in one direction (because, of course, that’s what we’re seeing.) So, in the statistical situation of hypothesis testing, we write that as an alternative hypothesis: **Ha: $\mu > 115$** .

Now let’s think back to our Ch. 11 review calculation, where we found the probability of getting a sample mean of 125 or greater from a sample of size 30 if the population mean is really 115. That probability is called the **p-value of the data**.

Write a statement in words about whether this p-value is small enough for you to consider the result to be statistically significant.

So to clarify your thoughts here, fill in YOUR significance level in both blanks and then circle the appropriate conclusion for your significance level and the given data from the sample of older students. (This is the appropriate wording for a conclusion to a hypothesis test.)

- These data provide significant evidence, at the ____ level, that the mean SSHA score for older college students is greater than 115. (p = ____)
- These data DO NOT provide significant evidence, at the ____ level, that the mean SSHA score for older college students is greater than 115. (p = ____)

Read Exercise 15.1. Do you see that we did one part of it? ____ For the other part, sketch something on your picture at the beginning of this handout. Then answer 15.1 b.

This discussion above summarizes exercises 15.1, 15.3, and 15.11 and most of the discussion in Chapter 15.

Now, you try exercise 15.14.

- write hypotheses (and define the unknown parameter)
- choose your significance level (before looking at the data in detail)
- sketch the distribution of the sampling distribution of \bar{X} if the Ho is true
- put the data value on it
- shade in the p-value of the data

- compare the p-value to your significance level
- write a conclusion, in context, to your hypothesis test

Comments on reading the text:

- The picture of the sampling distribution of \bar{X} if the H_0 is true, **must be included** as part of your solution – right there in the “Solve” part. This is, of course, the second picture in the CLT pictures you learned in Ch. 11.
- DO NOT use the applet or other software to compute the p-value. Use the normal table. (It’s not that those other ways are wrong – it is just that there is a lot to think about in these chapters without finding p-values in multiple ways.)

General comments:

We will be doing hypothesis tests (and confidence intervals) every day until the end of the course. You will have quite a bit of time to absorb all the details. For now, try to understand the big picture – the meaning.

In most of my classes, in this part of the course, most of the students end each class day thinking that they understand how to do hypothesis testing, and then they get home and find themselves stuck. So we do another problem the next class. And the same thing happens again. After about four times through that, students begin to feel more comfortable with it. So don’t feel frustrated if that’s what you feel too!

A quick summary:

Estimating a parameter: Find an interval that is likely to contain the true value of the parameter, and say how likely it is. (95% confident that μ lies in ___ to ___) This is about a “typical value” in the sampling distribution of the statistic.

Testing a claim: Decide whether the data is an “unusual value” if the null hypothesis is true.

Many students of statistics over the years have learned to do all these methods and not really understood how all the conclusions come from the sampling distribution that we learned about in Ch. 11. (The “nickels” distribution.) It may turn out that you will be one of those students. I hope not, because I think that it is actually much easier to do the exercises and solve statistics problems if you really “get it” about how the whole analysis depends on the variability of the sample statistic in repeated samples. That, of course, is why I emphasize drawing the pictures.

“Next week” (after spring break) we will spend about 1 ½ of the two days of that week going over various additional exercises to help you understand all of this better.

Quiz 8: Due Wed. March 20 at the beginning of class.

15.38 (just the p-values and conclusions), 15.40 (all four steps), 16.6, 16.9 (the main point is the three pictures), 16.12