Stat 411/511

Statistical Testing

Oct 5 2015

Charlotte Wickham

stat511.cwick.co.nz

Quiz #1

Study guide posted on class website.

Opens noon Friday, closes noon Monday.

Two parts:

- Multiple Choice - 30 mins
- Short Answer - 30 mins

Sample quiz posted on canvas.

(not stat related questions, just to see how canvas quizzes work)

Today

- Statistical testing (with a one sample t-test as our example)

- p-values

My fake Example

Do textbooks cost more at the bookstore than on Amazon?

What is the mean difference in price between Amazon and the bookstore, for **all textbooks used at OSU**?

The average price difference in our sample of size, n=100, is $10 with a sample standard deviation of $5.

$$\overline{X} = 10, \quad s = 5, \quad n = 100$$

$$\text{SE}_{\overline{X}} = s/\sqrt{n} = 5/\sqrt{100} = 0.5$$

$$t_{n-1}(0.975) = t_{99}(0.975) = 1.98$$ `qt(p = 0.975, df = 99)`

$$10 \pm 1.98 \times 0.5 = (9, 11)$$

With 95% confidence μ is between $9 and $11. no context

With 95% confidence we estimate that OSU textbooks on Amazon cost on average between $9 and $11 less than at the bookstore. in context

Hypothesis tests

Is this mean different from zero?

There's a correspondence between tests and confidence intervals.

If a **95%** confidence interval does not contain the hypothesized value, then the hypothesis would be rejected at the **5%** significance level.

substitute X% and 100 - X%, 95% and 5%

Hypothesis: the mean is zero, $\mu = 0$

Our 95% confidence interval: (9, 11)

Zero is not in the interval, we reject this hypothesis at the 5% significance level.

We conclude it is unlikely the mean is zero.

Quantified by a p-value, to get a p-value we need to know a little more about statistical testing.

The statistical justice system

Hypothesis testing is a lot like running a trial.

1. Set up the null hypothesis
(and alternative hypothesis)

what are the charges?
(the null defines innocence, the alternative guilt)

2. Calculate the **test statistic**

what's the evidence against our data?

3 Evaluate the evidence against the null hypothesis by comparing the test statistic to test statistics expected under the null hypothesis, the **null distribution**.

How does the evidence on our data compare to the evidence we would expect from innocent data?

The evidence is summarized by a **p-value**, the probability we would see such an extreme test-statistic if the null hypothesis is true.

4. **If the p is low, the null must go!**

Reject or fail to reject the null hypothesis

Innocent until proven guilty.

An example of the statistical process

in the random sampling model

We have a question.

Do textbooks cost more at the bookstore than on Amazon?

We translate this to a question about a **population** distribution.

What is the mean difference in price between Amazon and the bookstore, for **all textbooks required in OSU classes**? Is this mean bigger than zero?

We can't/won't/don't collect data on the whole population, but instead get a sample from the population and use properties of the sample to estimate properties of the population.

The average price difference in our sample of size, n=100, is $10 with a sample standard deviation of $5. With 95% confidence we estimate that OSU textbooks on Amazon cost between $9 and $11 less than at the bookstore.

1. Set up null and alternative hypotheses.

Translates our question of interest into a statement about **parameters** in a probability model.

our example

In the **random sampling** model, the statement is about the **population distribution**.

Null Hypothesis: the population mean is zero, $\mu = 0$.

Alternative Hypothesis: the population mean is not zero, $\mu \neq 0$.

Doesn't necessarily have to capture all of the other possibilities, more on this later...

2. Calculate the test-statistic

The test-statistic summarizes the data into a single number that gives evidence against the null.

Different types of test arise from considering different test statistics with different probability models (assumptions).

our example

one-sample *t*-test: t-ratio is our test-statistic

$$t\text{-statistic} = \frac{10 - 0}{0.5} = 20$$

3. Compare to the null distribution

The **null distribution** is the distribution of test-statistics you would expect to see if the null hypothesis was true and you could repeat your study many times.

In the **random sampling** model, the **null distribution,** is the sampling distribution of the test-statistic assuming the null hypothesis was true.

our example

If the null hypothesis was true, our t-ratio would be $\frac{\overline{X} - 0}{\mathrm{SE}_{\overline{X}}}$

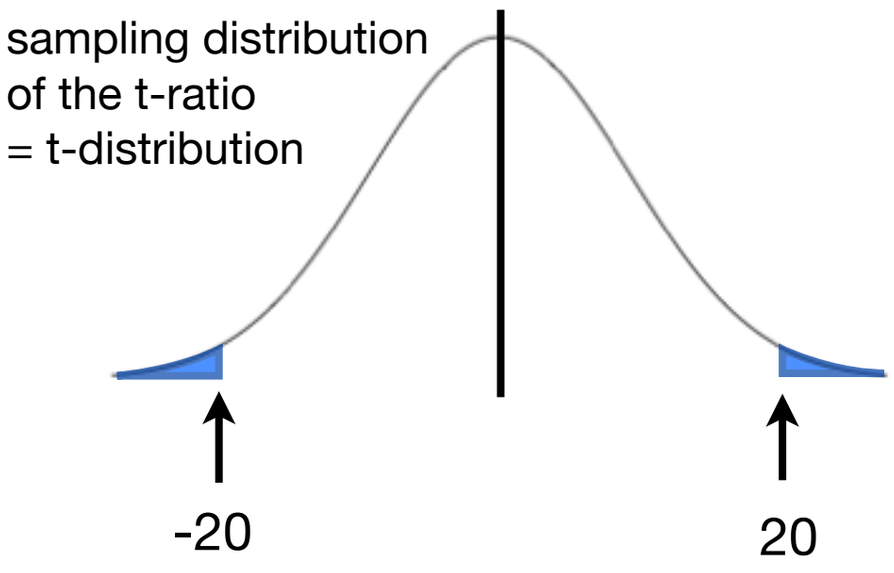
and we know this has a *t*-distribution with n-1 degrees of freedom

4. The p-value

The **p-value** is the probability we would see a test-statistic as extreme or more extreme than the one we observed, if the null hypothesis was true.

It is **not** the probability the null is true

our example

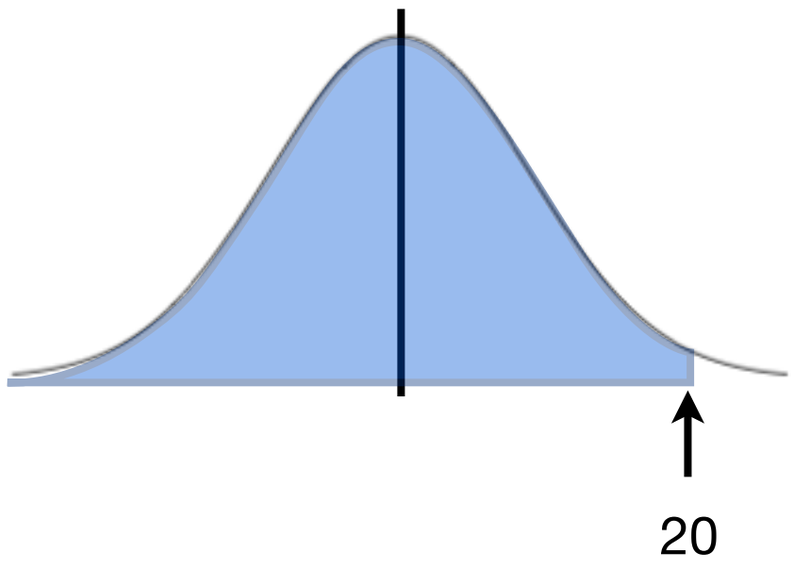


If the null was true,

$$t\text{-statistic} = \frac{10 - 0}{0.5} = 20$$

would be like a sample of size one from the sampling distribution of the t-ratio

What's the probability of seeing a more extreme number than 20 from a t-distribution with 99 degrees of freedom?



$$t_{99}(?) = 20$$

What quantile is 20?

our t-stat

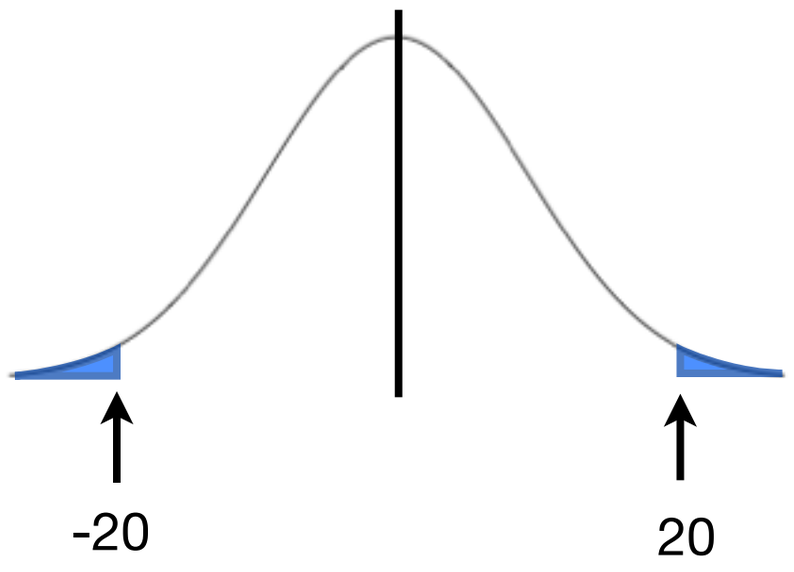
```
pt(q = 20, df = 99)
```

The function `pt` gives the area to the left of the first argument.

To get the (two-sided) **p-value,**

```
2*(1 - pt(q = 20, df = 99))
```

What if our t-stat was negative?



```
> 2*(1 - pt(q = 20, df = 99))
[1] 0
```

This probability isn't really zero, but it is so small, it is beyond the precision of R.
I.e. it is smaller than 2.220446×10^{-16}

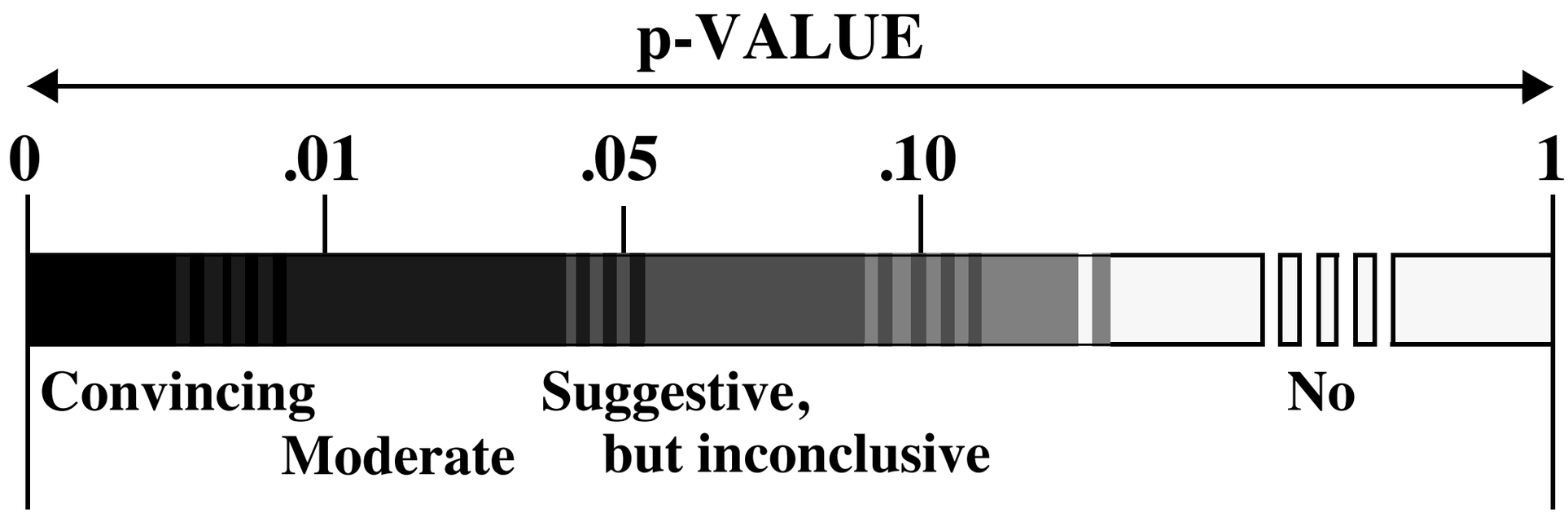
Probabilities are always between 0 and 1.

Translate the number into a level of evidence and report the p-value (read 2.5.1 & 2.5.3)

To make a concrete decision, we would decide on a threshold of the p-value: the **significance level, α**

Display 2.12

Interpreting the size of a p-value



0 - 0.01	Convincing or strong evidence
0.01 - 0.05	Moderate or some evidence
0.05 - 0.10	Suggestive or weak evidence
0.10 - 1.00	No evidence

against the null

We have convincing evidence that the population mean difference in price for textbooks at OSU between the bookstore and Amazon is not equal to zero (one sample t-test, p-value $< 2.2 \times 10^{-16}$).

We estimate that OSU textbooks cost on average $10 less at Amazon.

With 95% confidence we estimate that OSU textbooks on Amazon cost on average between $9 and $11 less than at the bookstore.

A Statistical Summary