T-tests:
One sample practice

Research Methods for Human Inquiry

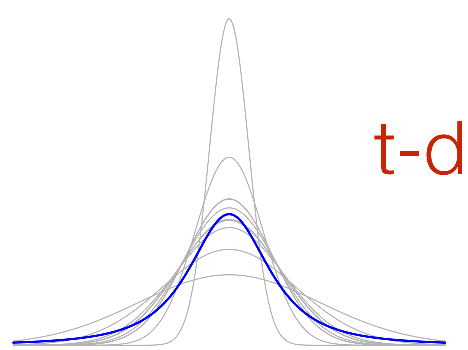
Andrew Perfors

So far we've constructed our test…

1) A diagnostic test statistic, *T*

$$t = \frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{N}}$$

2) Sampling distribution of *T* if the null is true

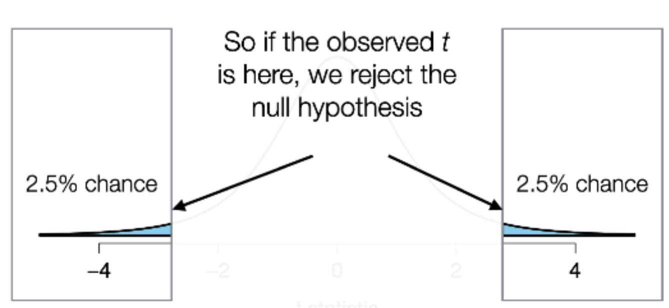


t-distribution with N-1 degrees of freedom

3) The observed *T* in your data

6.86

4) A rule that maps every value of *T* onto a decision (accept or reject H0)



Here's the data

```
> dcr <- read_csv(file=here("crscores.csv"))
> dcr
# A tibble: 34 x 2
   name            cr
   <chr>        <dbl>
 1 bunny           70
 2 gladly          69
 3 flopsy          63
 4 doggie          87
 5 nosey           67
 6 cuddly paws     56
 7 shadow          56
 8 pink bunny      59
 9 purple bunny    68
10 blue bunny      55
# … with 24 more rows
```

One-sample t-test

```
> t.test( x = dcr$cr, mu = 50 )
```

The variable containing the raw data

True population mean, according to the null hypothesis

One-sample t-test

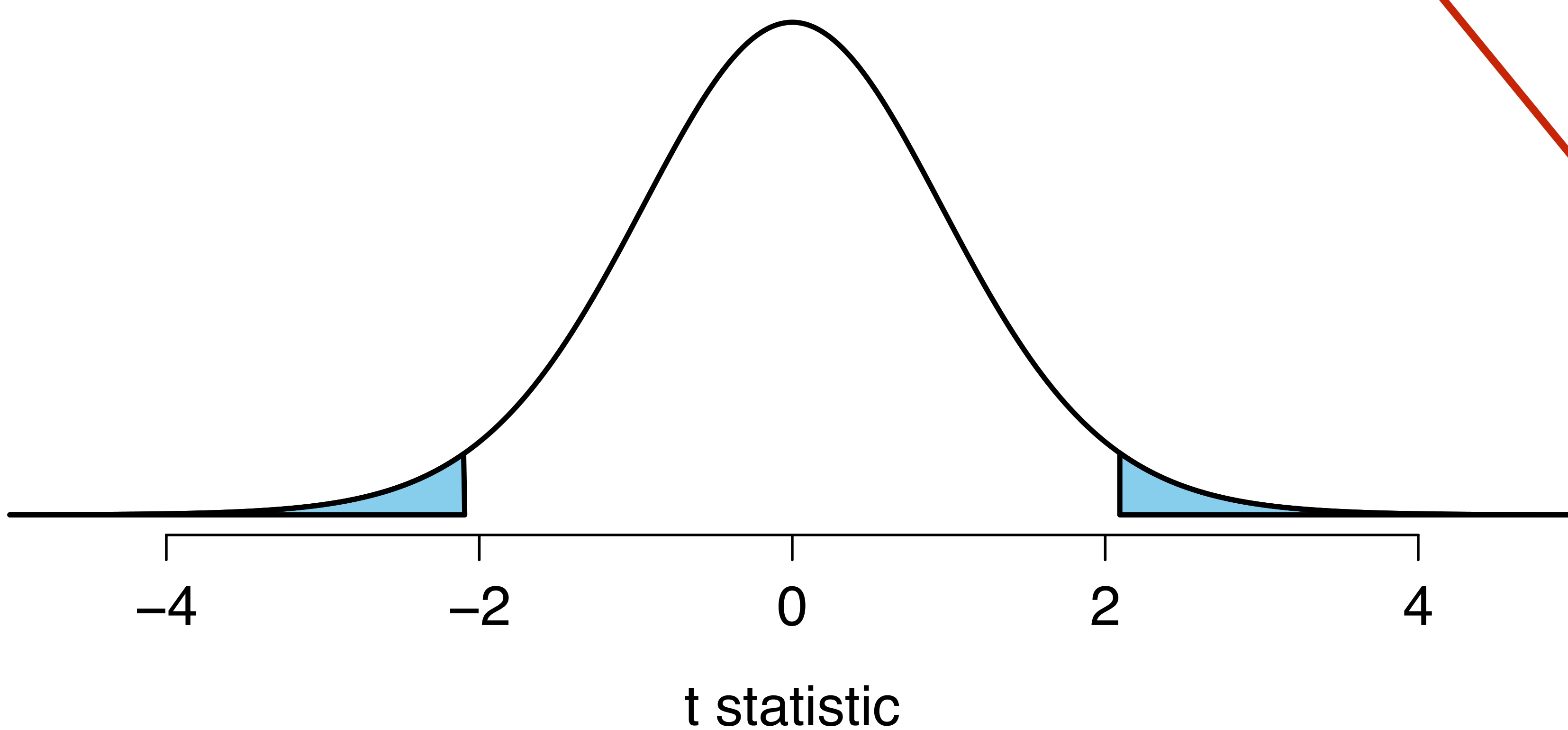
```
> t.test( x = dcr$cr, mu = 50 )

	One Sample t-test

data:  dcr$cr
t = 6.862, df = 33, p-value = 7.817e-08
```

Test statistic, degrees of freedom, and the p-value

The t-statistic falls here so we reject the null



One-sample t-test

```
> t.test( x = dcr$cr, mu = 50 )

	One Sample t-test

data:  dcr$cr
t = 6.862, df = 33, p-value = 7.817e-08
alternative hypothesis: true mean is not equal to 50
95 percent confidence interval:
 59.60087 67.69325
sample estimates:
mean of x
 63.64706
```

Test statistic, degrees of freedom, and the p-value

Alternative hypothesis

Confidence interval and estimate for the population mean

Measuring effect size with Cohen's *d*

Cohen's *d*

- "Simple" measure of effect size:

$$d = \frac{\text{“mean 1”} - \text{“mean 2”}}{\text{“std dev”}}$$

  - Difference in means divided by an estimate of the standard deviation.

Cohen's d	rough interpretation
0.2	small
0.5	medium
0.8	large

Great visualisation: http://rpsychologist.com/d3/cohend/

Calculating Cohen's d in R

```
> library(lsr)
> cohensD( dcr$cr ,mu=50)
[1] 1.17683
```

Note that it takes the raw data and the comparison mu

```
> (mean(dcr$cr)-50)/sd(dcr$cr)
[1] 1.17683
```

Could also do it manually if you want to be masochistic

Writing up the results?

statement of null hypothesis

descriptive statistics

People in Bunnyland had a mean CR score of 63.6 (SD = 11.6). To test the null hypothesis that the data came from a population with true mean 50, we ran a one-sample t-test and obtained a significant result, $t(33) = 6.086$, $p < .0001$. This suggests that Bunnyland residents' CR scores were significantly higher than the typical mean on this test.

whether significant

interpretation

what test was run

Writing up the results?

People in Bunnyland had a mean CR score of 63.6 (SD = 11.6). To test the null hypothesis that the data came from a population with true mean 50, we ran a one-sample t-test and obtained a significant result $t(33) = 6.086$, $p < .0001$. This suggests that Bunnyland residents' CR scores were significantly higher than the typical mean on this test.

sampling distribution is a t-distributon

degrees of freedom

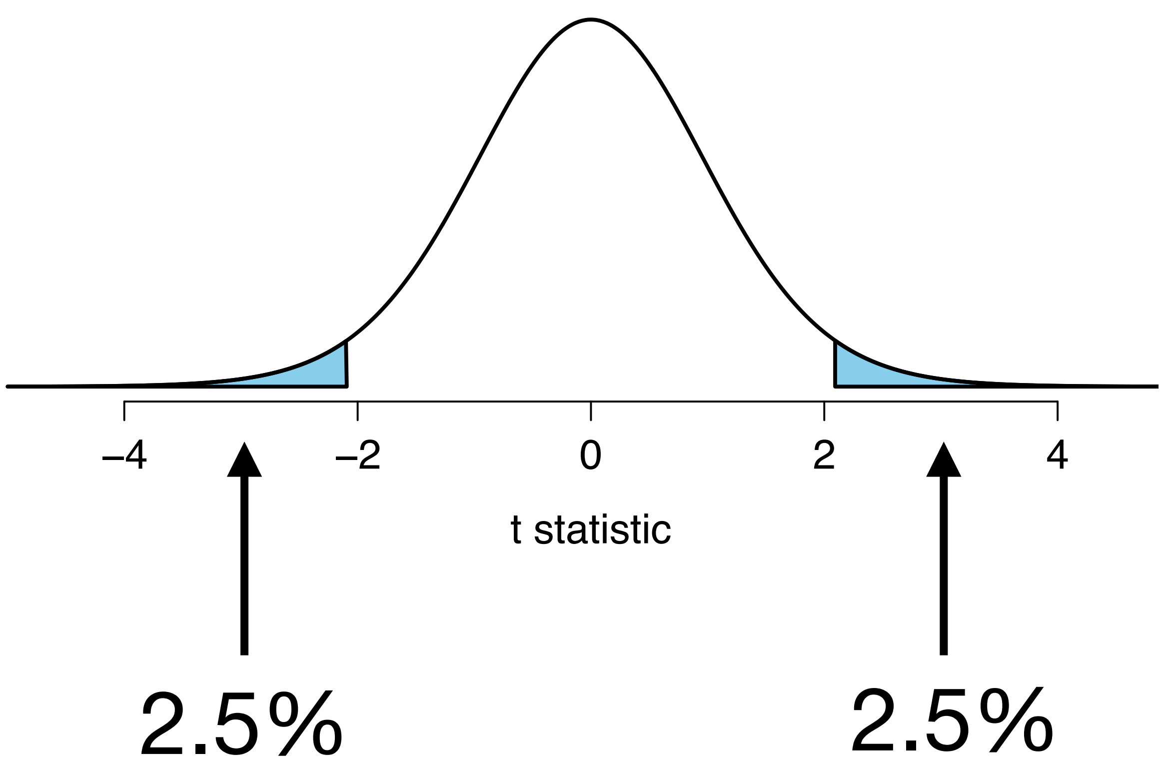
t-statistic

P-value

Two sided test

H0: true mean is 50

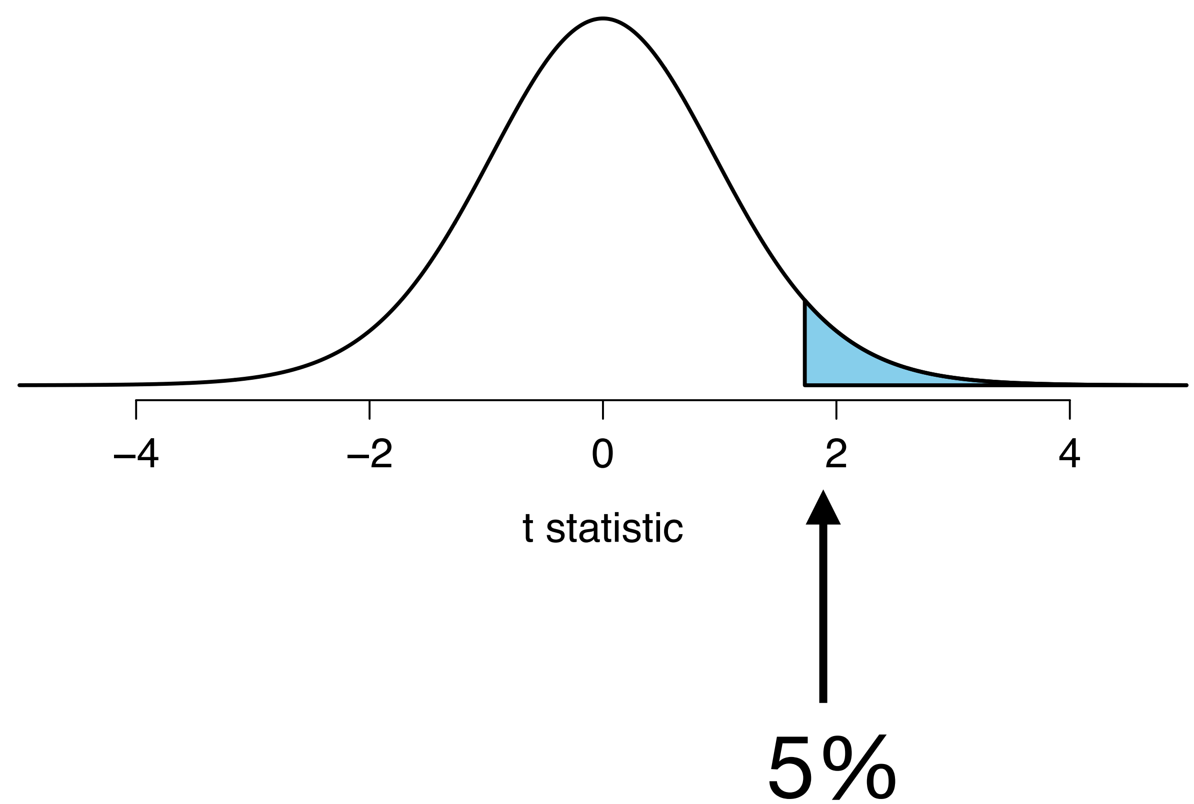
H1: true mean isn’t 50



One sided test

H1: true mean is **above*** 50

So we only care about one side of the sampling distribution in making our accept/reject decision



(*or, equivalently, H1 is below 50)

I mention these because people do them and you'll see them, but I think they are a bad idea.

The null hypothesis (and nature of the t distribution) haven't changed and NHST by definition should depend only on H0

The logic given is usually that you should do it if you genuinely believe you should not see results on one side — but if that's the case, H0 would be different and so would the sampling distribution

In practice they are a way of sneakily getting "significant" results and IMHO often result in inflated Type I error

One sided test

There is a way of calculating one-sided tests in R using the `t.test()` function with the `alternative` argument but I will not assess this or ask you to ever do it because I don't think we should use them

Exercises are in w7day1exercises.Rmd