

T-Test: One-Sample

Data Analysis for Psychology in R 1
Semester 2, Week 6

Dr Emma Waterston

Department of Psychology
The University of Edinburgh

Course Overview

Exploratory Data Analysis	Research design and data
	Describing categorical data
	Describing continuous data
	Describing relationships
	Functions
Probability	Probability theory
	Probability rules
	Random variables (discrete)
	Random variables (continuous)
	Sampling

Foundations of inference	Confidence intervals
	Hypothesis testing (p-values)
	Hypothesis testing (critical values)
	Hypothesis testing and confidence intervals
	Errors, power, effect size, assumptions
Common hypothesis tests	One sample t-test
	Independent samples t-test
	Paired samples t-test
	Chi-square tests
	Correlation

Learning Objectives

- Understand when to use a one sample t -test
- Understand the null hypothesis for a one sample t -test
- Understand how to calculate the test statistic
- Know how to conduct the test in [R](#)

Topics for Today

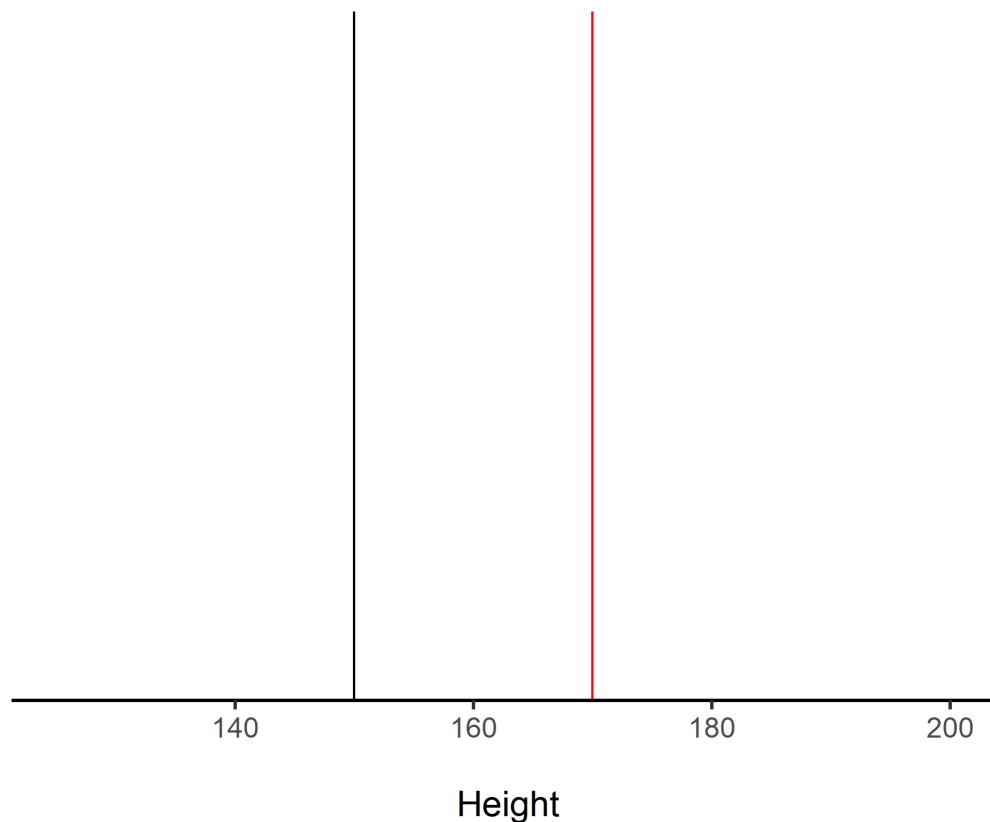
- Introduce the three types of t -test
- One-sample t -test example
- Inferential tests for the one-sample t -test
- Assumptions and effect size

Introduction

T-Test: Purpose

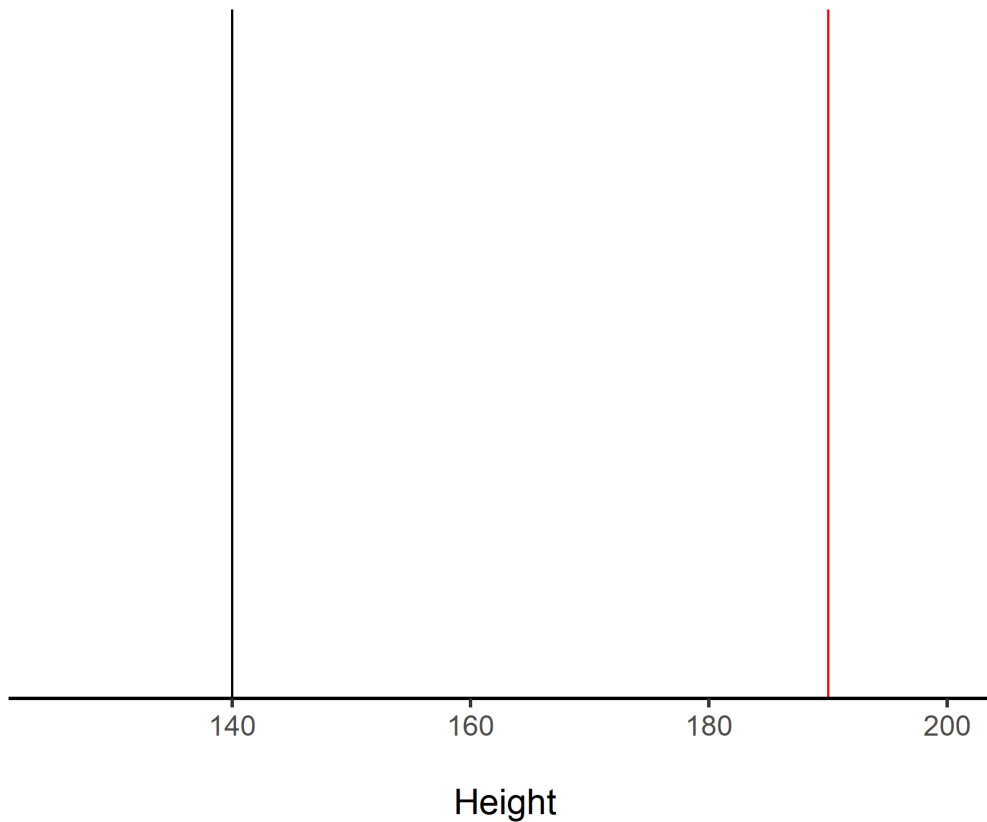
- t -tests (generally) concern testing the difference between two means
 - Another way to state this is that the scores of two groups being tested are from the sample underlying population distribution
- One-sample t -tests compare the mean in a sample to a known mean
- Independent t -tests compare the means of two independent samples
- Paired sample t -tests compare the mean from a single sample at two points in time (repeated measurements)
- We will look in more detail at these tests over the next three weeks
 - But let's start by thinking a little bit about the logic t -tests
 - For the next few slides, have a bit of paper and a pen handy

Are These Means Different?



- Write down whether you think these means (two lines) are different. Write either:
 - Yes
 - No
 - It depends

What About These?

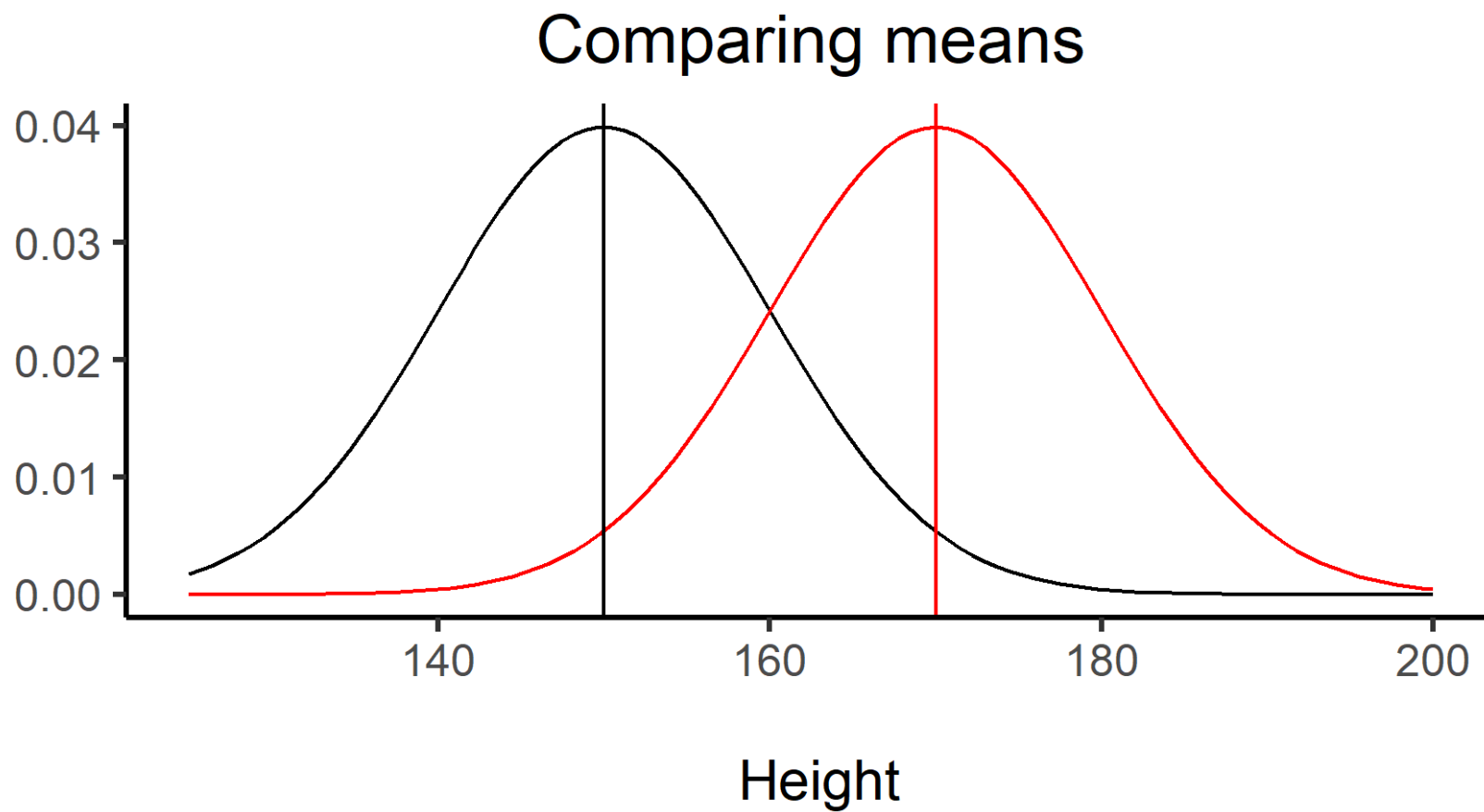


- Write down whether you think these means (two lines) are different. Write either:
 - Yes
 - No
 - It depends

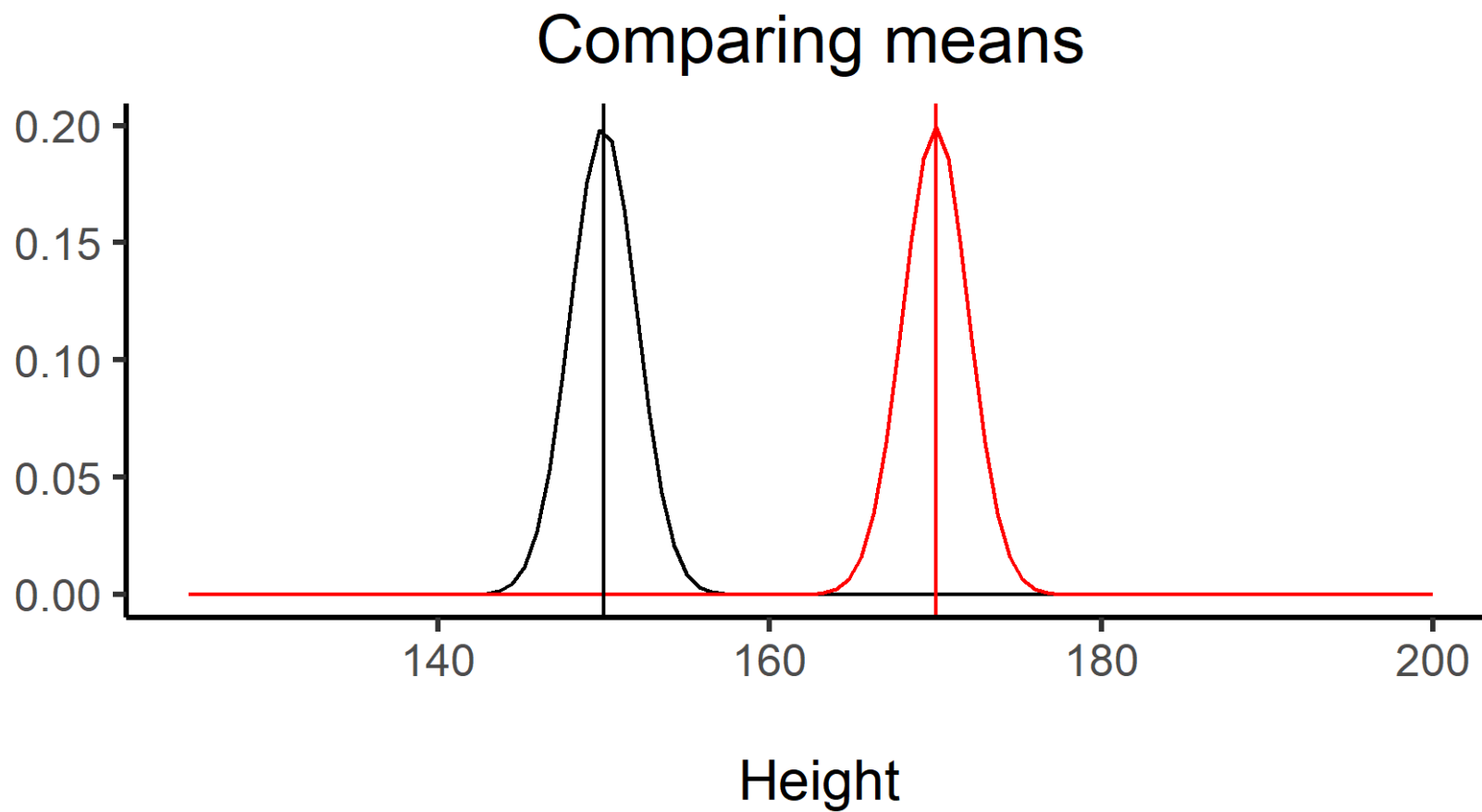
Differences in Means

- OK, now please write down:
 1. Why you wrote the answers you did?
 2. If you wrote, "It depends", why can we not tell whether they are different or not?
 3. What else might we want to know in order to know whether not the group means could be thought of as coming from the same distribution?

All the Information



All the Information



Questions?

T-Test

t-statistic

- Recall when talking about hypothesis testing:
 - We calculate a test statistic that represents our question
 - We compare our sample value to the sampling distribution under the null
- Here the test statistic is a t -statistic

t-statistic

$$t = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}} \quad \text{where} \quad SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

- The numerator = a difference in means
 - where
 - \bar{x} = sample mean
 - μ_0 = hypothesized value
 - $SE_{\bar{x}}$ = standard error of the mean
- The denominator = a estimate of variability
 - where
 - s = sample estimated standard deviation of x
 - n = sample size
- t = a standardized difference in means

Example

- Suppose I want to know whether the retirement age of Professors at my University is the same as the national average
- The national average age of retirement for Prof's is 65
- So I look at the age of the last 40 Prof's that have retired at Edinburgh and compare against this value

Data

```
## # A tibble: 40 × 2
##   ID      Age
##   <chr> <dbl>
## 1 Prof1    76
## 2 Prof2    66
## 3 Prof3    58
## 4 Prof4    68
## 5 Prof5    79
## 6 Prof6    74
## 7 Prof7    75
## 8 Prof8    50
## 9 Prof9    69
## 10 Prof10   70
## # i 30 more rows
```

Hypotheses

- When we are testing whether the population mean (μ) is equal to a hypothesized value (μ_0)

$$H_0 : \mu = \mu_0$$

- Note this is identical to saying:

$$H_0 : \mu - \mu_0 = 0$$

Alternative Hypotheses

- Two-tailed:

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$$

- One-tailed:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

$$H_1 : \mu > \mu_0$$

Hypotheses

- Let's assume a priori we have no idea of the ages the Prof's retired
- So I specify a two-tailed hypothesis with $\alpha = .05$
- So I am simply asking, does my mean differ from the known mean

Calculation

$$t = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}} \quad \text{where} \quad SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

- Steps to calculate t :
 - Calculate the sample mean (\bar{x})
 - Calculate the standard error of the mean ($\frac{s}{\sqrt{n}}$)
 - Calculate the sample standard deviation (s)
 - Check I know my sample size (n)
 - Use all this to calculate t

Calculation

$$t = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}} \quad \text{where} \quad SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

```
dat |>
  summarise(
    mu0 = 65,
    xbar = mean(Age),
    s = sd(Age),
    n = n()
  ) |>
  mutate(
    se = s/sqrt(n)
  ) |>
  kable(digits = 2) |>
  kable_styling(full_width = FALSE)
```

mu0	xbar	s	n	se
65	66.3	10.01	40	1.58

Calculation

mu0	xbar	s	n	se
65	66.3	10.01	40	1.58

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{66.30 - 65.00}{\frac{10.01}{\sqrt{40.00}}} = \frac{1.30}{1.58} = 0.82$$

- So in our example $t = 0.82$

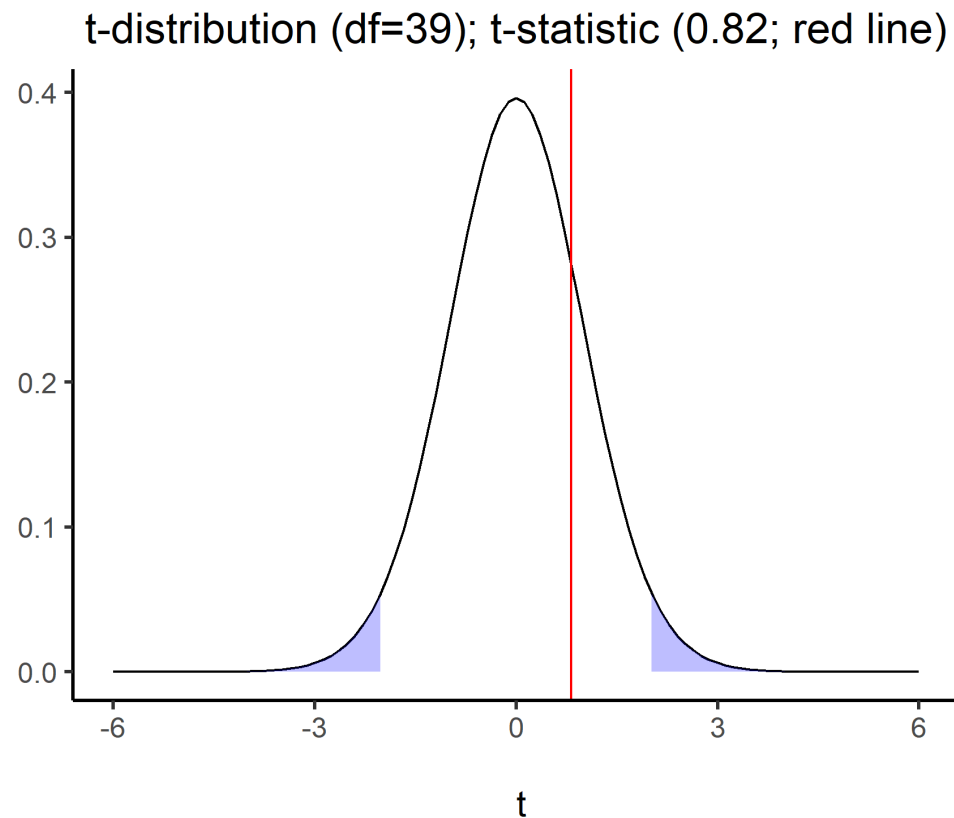
Is our Test Significant?

- The sampling distribution for t -statistics is a t -distribution
- The t -distribution is a continuous probability distribution very similar to the normal distribution
 - Key parameter = degrees of freedom (df)
 - df are a function of n
 - As n increases (and thus as df increases), the t -distribution approaches a normal distribution
- For a one sample t -test, we compare our test statistic to a t -distribution with $n - 1$ df

Is our Test Significant?

- So we have all the pieces we need:
 - Degrees of freedom = $n-1 = 40-1 = 39$
 - We have our t -statistic (0.82)
 - Hypothesis to test (two-tailed)
 - α level (.05)
- So now all we need is the critical value from the associated t -distribution in order to make our decision

Is our Test Significant?



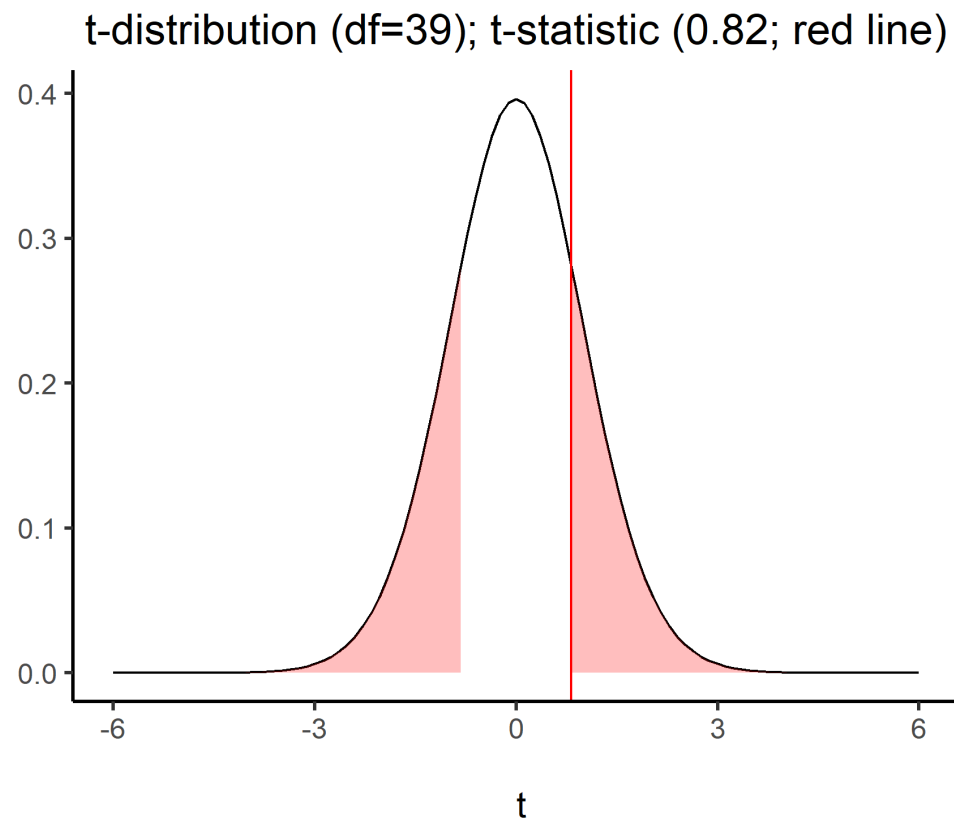
```
tibble(  
  LowerCrit = round(qt(0.025, 39), 2),  
  UpperCrit = round(qt(0.975, 39), 2),  
)
```

```
## # A tibble: 1 × 2  
##   LowerCrit UpperCrit  
##   <dbl>      <dbl>  
## 1     -2.02       2.02
```

Is our Test Significant?

- So our critical value is 2.02
 - Our t -statistic (0.82) is closer to 0 than this
 - So we **fail to reject the null hypothesis**

Exact p-value



```
tibble(  
  Exactp = round(2*(1-pt(0.821, 39)),2)  
)
```

```
## # A tibble: 1 × 1  
##   Exactp  
##   <dbl>  
## 1    0.42
```

Specifying Hypotheses in R

- `alternative` = refers to the direction of our alternative hypothesis (H_1)
 - $\mu < \mu_0$: `alternative="less"`
 - Our Edinburgh Prof's have a younger retirement age than the national average
 - $\mu > \mu_0$: `alternative="greater"`
 - Our Edinburgh Prof's have a older retirement age than the national average
 - $\mu \neq \mu_0$: `alternative="two.sided"`
 - Our Edinburgh Prof's have a different retirement age than the national average

```
t.test(dat$Age, mu=65, alternative="_____")
```

One-Sample T-Test in R

```
t.test(dat$Age, mu=65, alternative="two.sided")
```

```
##  
##      One Sample t-test  
##  
## data:  dat$Age  
## t = 0.82152, df = 39, p-value = 0.4163  
## alternative hypothesis: true mean is not equal to 65  
## 95 percent confidence interval:  
##  63.09922 69.50078  
## sample estimates:  
## mean of x  
##      66.3
```

Write Up

A one-sample t -test was conducted to determine if there was a statistically significant ($\alpha = .05$) mean difference between the average retirement age of Professors and the age at retirement of a sample of 40 Edinburgh Professors. Although the sample had a higher average age of retirement ($M = 66.3, SD = 10.01$) than the population ($M = 65$), this difference was not statistically significant ($t(39) = 0.82, p = .416, two - tailed$).

Questions?

Data Requirements & Assumptions

Data Requirements

- A numeric variable
- A known mean that we wish to compare our sample to
- A sample of data from which we calculate the sample mean

Assumption Checks Summary

	Description	One-Sample t-test	Independent Samples t-test	Paired Samples t-test
Normality	Numeric variable (or difference) is normally distributed OR sample size is sufficiently large.	Yes (variable). Sample size guideline: $n \geq 30$	Yes (variable in each group). Sample size guideline: $n_1 \geq 30$ and $n_2 \geq 30$	Yes (difference). Sample size guideline: number of pairs ≥ 30
Tests:	Descriptive Statistics and Plots; QQ-Plot; Shapiro-Wilks Test			
Independence	Observations are sampled independently.	Yes	Yes (within and across groups)	Yes (across pairs)
Tests:	None. Design issue.			
Homogeneity of variance	Population standard deviation is the same in both groups.	NA	Yes*	NA
Tests:	F-test			
Matched Pairs in data	For paired sample, each observation must have matched pair.	NA	NA	Yes
Tests:	None. Data structure issue.			

* Welch t-test is available if this is not met

Data Requirements & Assumptions: How to Check/Test

- DV is numeric
 - The dependent variable should be measured on a interval/ratio/integer scale
- Independence
 - More of a study design issue, and cannot directly test
- Normality
 - Can be checked with descriptive statistics, visually with plots, and with a Shapiro-Wilks Test

Normality: Skew

- Skew is a descriptive statistic informing us of both the direction and magnitude of asymmetry
 - Below are some rough guidelines on how to interpret skew
 - No strict cuts for skew - these are loose guidelines

Verbal label	Magnitude of skew in absolute value
Generally not problematic	$ \text{Skew} < 1$
Slight concern	$1 > \text{Skew} < 2$
Investigate impact	$ \text{Skew} > 2$

Skew in R

```
library(psych)
dat |>
  summarise(
    skew = round(skew(Age),2)
  )
```

```
## # A tibble: 1 × 1
##   skew
##   <dbl>
## 1 -0.63
```

- Skew is low (< 1), so we would conclude that it is not problematic

Normality: Visual Assessment

- We can visually assess normality by plotting the distribution of our variable:
 - Histograms
 - The count (or frequency) of data points that fall within specified intervals/bins
 - Density Plots
 - The probability density (or proportion of values) of data points at each value of the observed variable
 - QQ-Plots (Quantile-Quantile plot):
 - Plots the sorted quantiles of one data set (distribution) against sorted quantiles of data set (distribution)
 - Quantile = the percent of points falling below a given value
 - For a normality check, we can compare our own data to data drawn from a normal distribution

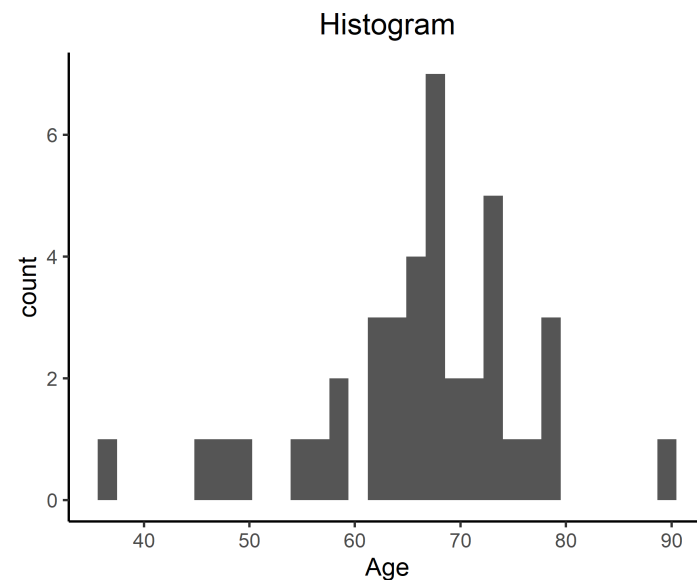
Histogram & Density Plots in R

```
ggplot(data = dat, aes(x=Age)) +  
  geom_density() +  
  labs(title = "Density Plot")
```



- Our density plot looks relatively normal

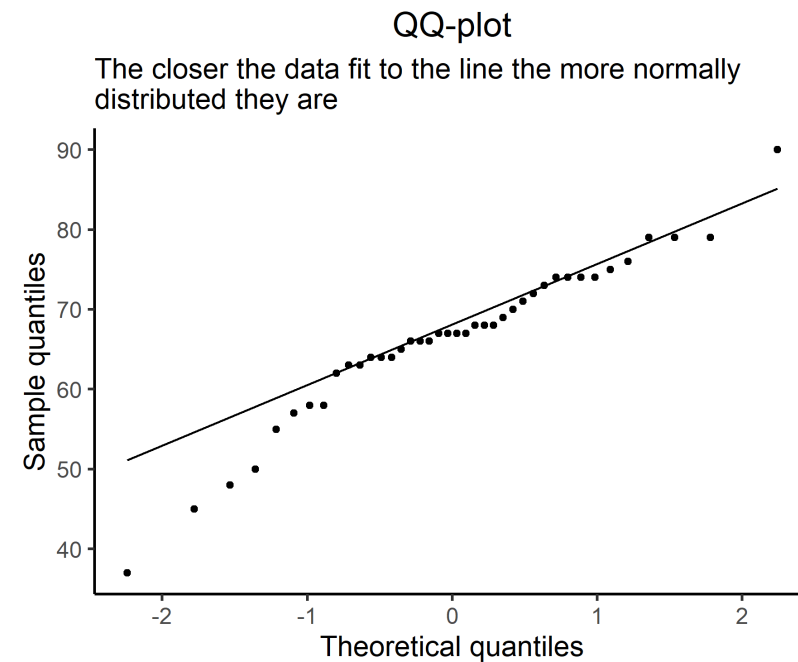
```
ggplot(data = dat, aes(x=Age)) +  
  geom_histogram() +  
  labs(title = "Histogram")
```



- Our histogram looks "lumpy", but we have relatively low n for looking at these plots

QQ-Plots in R

```
ggplot(data = dat, aes(sample = Age)) +  
  geom_qq() +  
  geom_qq_line() +  
  labs(title="QQ-plot",  
        x = "Theoretical quantiles",  
        y = "Sample quantiles")
```



- This looks a little concerning
- We have some deviation in the lower left corner
- This is showing we have more lower values for age than would be expected

Normality: Shapiro-Wilks Test

- Shapiro-Wilks test:
 - Checks properties of the observed data against properties we would expect from normally distributed data
 - Statistical test of normality
 - H_0 : data = the sample came from a population that is normally distributed
 - $p\text{-value} < \alpha$ = reject the null, data are not normal
 - Sensitive to n as all p -values will be
 - In very large n , normality should also be checked with QQ-plots alongside statistical test

Shapiro-Wilks Test in R

```
shapiro.test(dat$Age)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  dat$Age  
## W = 0.95122, p-value = 0.08354
```

- Fail to reject the null, $p = .084$, which is $> .05$

Questions?

Effect Size

Cohen's D

- Cohen's-D is the standardized effect size for measuring the difference in means
 - Having a standardized metric is useful for comparisons across studies
 - It is also useful for thinking about power calculations
- The basic form of D is the same across the different t -tests:

$$D = \frac{\text{Difference}}{\text{Variation}}$$

Interpreting Cohen's D

- Below are some rough guidelines on how to interpret the size of the effect
- These are not exact labels, but a loose guidance based on empirical research
- Perhaps the most common "cut-offs" for D -scores:

Verbal label	Magnitude of D in absolute value
Small (or weak)	≤ 0.20
Medium (or moderate)	≈ 0.50
Large (or strong)	≥ 0.80

Cohen's D: One-Sample T-Test

- One-sample t -test:

$$D = \frac{\bar{x} - \mu_0}{s}$$

- where

- \bar{x} = sample mean
- μ_0 = hypothesised mean
- s = sample standard deviation

Cohen's D: One-Sample T-Test in R

```
library(effectsize)
cohens_d(dat$Age, mu=65, alternative="two.sided")
```

```
## Cohen's d |          95% CI
## -----
## 0.13      | [-0.18, 0.44]
##
## - Deviation from a difference of 65.
```

Write Up: Data Requirements, Assumptions, & Effect Size

The DV of our study, Age, was measured on a continuous scale, and data were independent (based on study design). The assumption of normality was visually assessed (via histograms, density plots, and a QQplot) as well as statistically via a Shapiro-Wilks test. Whilst the QQplot did show some deviation from the diagonal line, the Shapiro-Wilks test suggested that the sample came from a population that was normally distributed ($W = 0.95, p = .084$). This was inline with the histogram and density plot, which suggested that Age was normally distributed (and where skew < 1). The size of the effect was found to be small $D = 0.13[-0.18, 0.44]$.

Summary

- Today we have covered:
 - Basic structure of the one-sample t -test
 - Calculations
 - Interpretation
 - Assumption checks
 - Effect size measures (Cohen's D)

This Week



Tasks

- Attend both lectures
- Attend your lab and work on the assessed report with your group (due by 12 noon on Friday 28th of March 2025)
- Complete the weekly quiz
 - Opened Monday at 9am
 - Closes Sunday at 5pm



Support

- **Office Hours:** for one-to-one support on course materials or assessments
(see LEARN > Course information > Course contacts)
- **Piazza:** help each other on this peer-to-peer discussion forum
- **Student Adviser:** for general support while you are at university
(find your student adviser on MyEd/Euclid)