R Lab - Day 4
Inference, t-test

MF9130E V24
2024.04.11

Chi Zhang
Oslo Center for Biostatistics and Epidemiology
chi.zhang@medisin.uio.no

Outline

8:30-9:15	Review: statistical inference, t-test
9:15 -	Demonstration in R
Practice	**Practice (exercise 1, 2)**
	Summary and wrap up

Lab notes for today:
(under *R Lab and Code* tab)

t-test

Link to *R Lab and Code*

https://ocbe-uio.github.io/teaching_mf9130e/lab/lab_ttest.html

Statistical inference

Research question:
Compare measurements from 2 groups

Height between men and women;

Outcome between treatment and control groups;

…

It can also be more than 2 groups; but we focus on 2 groups in this course.

In the abstract of papers, you often see the following expressions:

"Significantly different with $p < 0.001$"

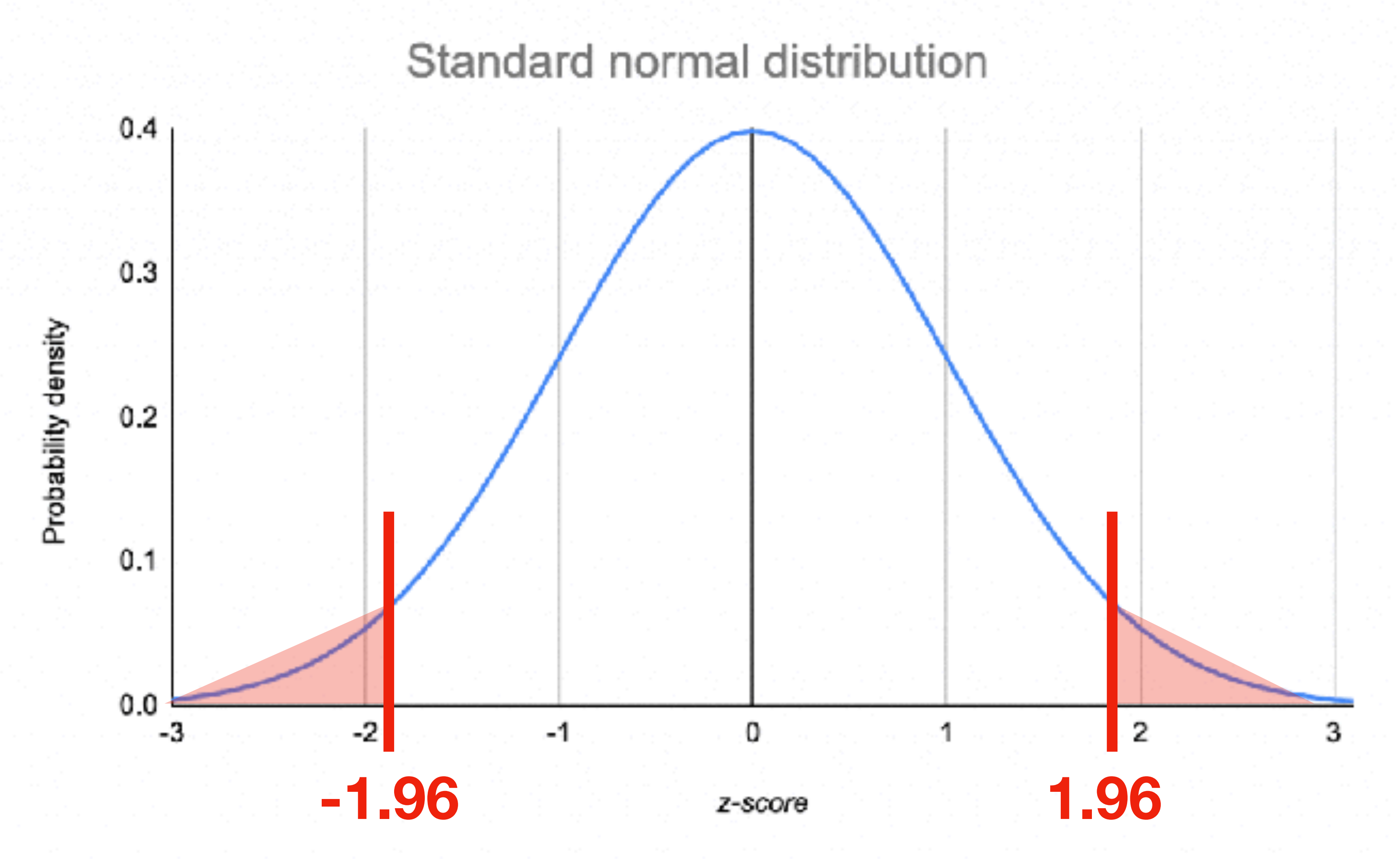
"Confidence interval is (1.2, 2.5)"

OR (odds ratio) of smoking is 2.5 (1.1, 6.7)

…

Normal distribution, t-distribution

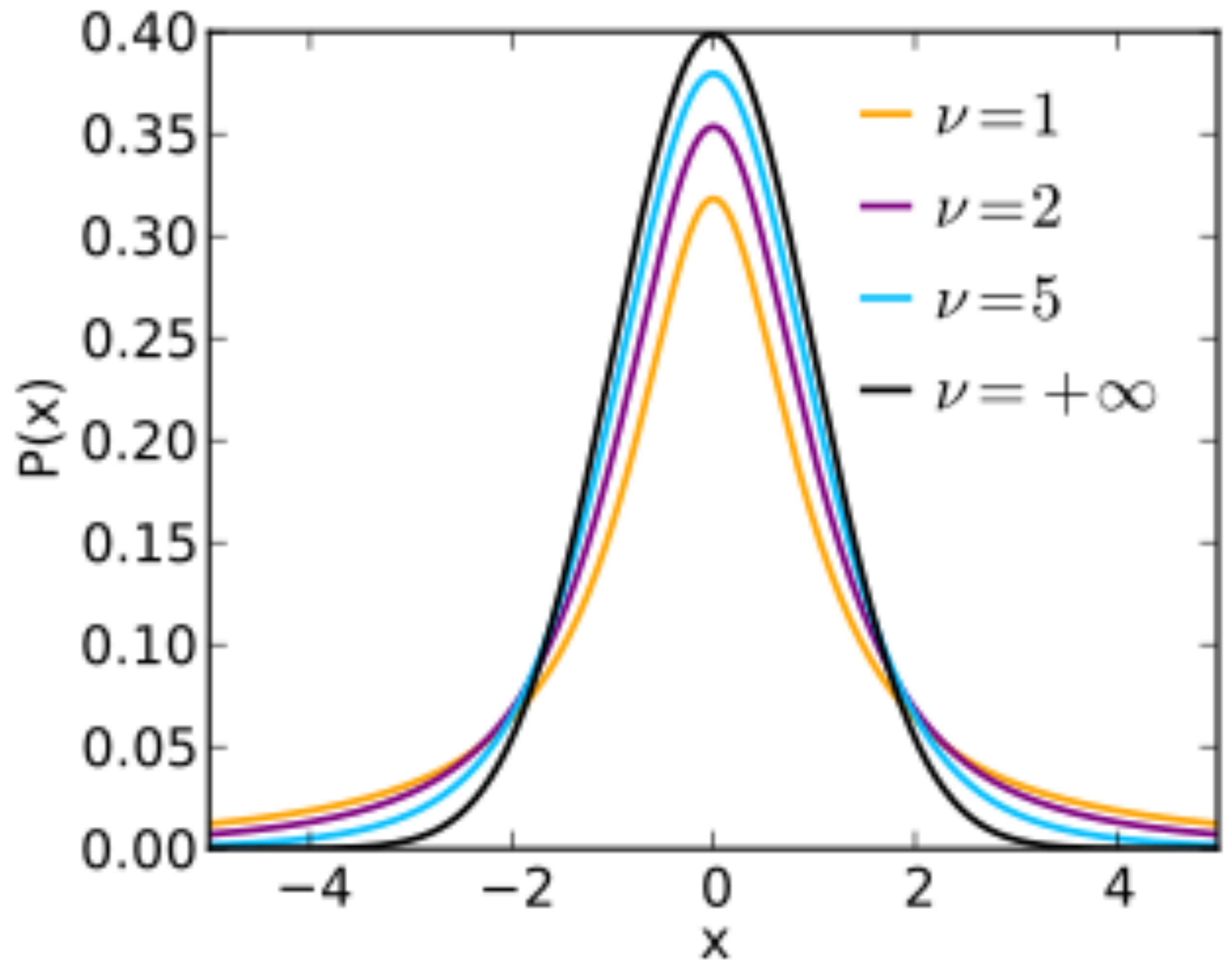
$$X \sim N(\mu, \sigma^2)$$



-1.96, 1.96 are 2.5% and 97.5% **quantile** for N(0,1)

P(X > 1.96) = 0.025; P(X < 1.96) = 0.975
P(X < -1.96) = 0.025

Probability distribution of t-distribution for different **degrees of freedom** (v)



When v is big (around 30), t-distribution is close to normal distribution

However the smaller v, the more different t-dist is from a normal distribution.

Statistical inference

We focus on **continuous** measurements today

Sample mean

Confidence interval

Test for mean (t-test):
one sample
paired samples
two samples

(Example 2a, 2b in t-test lab notes)
Lung data (PEFH98-english)

Height for women:
We want to compare the **average height** of women with a fixed value.
How **confident** are we about the conclusion?

	age	gender	height	weight	pefsit1	pefsit2	pefsit3	pefsta1	pefsta2	pefsta3	pefsitm	pefstam	pefmean
1	20	female	165	50	400	400	410	410	410	400	403.3333	406.6667	405.0000
2	20	male	185	75	480	460	510	520	500	480	483.3333	500.0000	491.6667
3	21	male	178	70	490	540	560	470	500	470	530.0000	480.0000	505.0000
4	21	male	179	74	520	530	540	480	510	500	530.0000	496.6667	513.3333
5	20	male	196	95	740	750	750	700	710	700	746.6667	703.3333	725.0000
6	20	male	189	83	600	575	600	600	600	640	591.6667	613.3333	602.5000
7	32	male	173	65	740	760	720	705	690	680	740.0000	691.6667	715.8333
8	22	male	196	94	720	720	700	700	730	800	713.3333	743.3333	728.3333
9	21	female	173	66	480	530	540	520	520	530	516.6667	523.3333	520.0000
10	23	female	173	65	400	430	420	430	430	430	416.6667	430.0000	423.3333
11	22	female	169	65	500	510	540	520	580	530	516.6667	543.3333	530.0000
12	23	male	185	75	730	630	700	700	700	710	686.6667	703.3333	695.0000
13	21	male	194	84	630	690	670	680	700	690	663.3333	690.0000	676.6667
14	21	female	170	55	360	360	370	370	360	360	363.3333	363.3333	363.3333

Sample mean

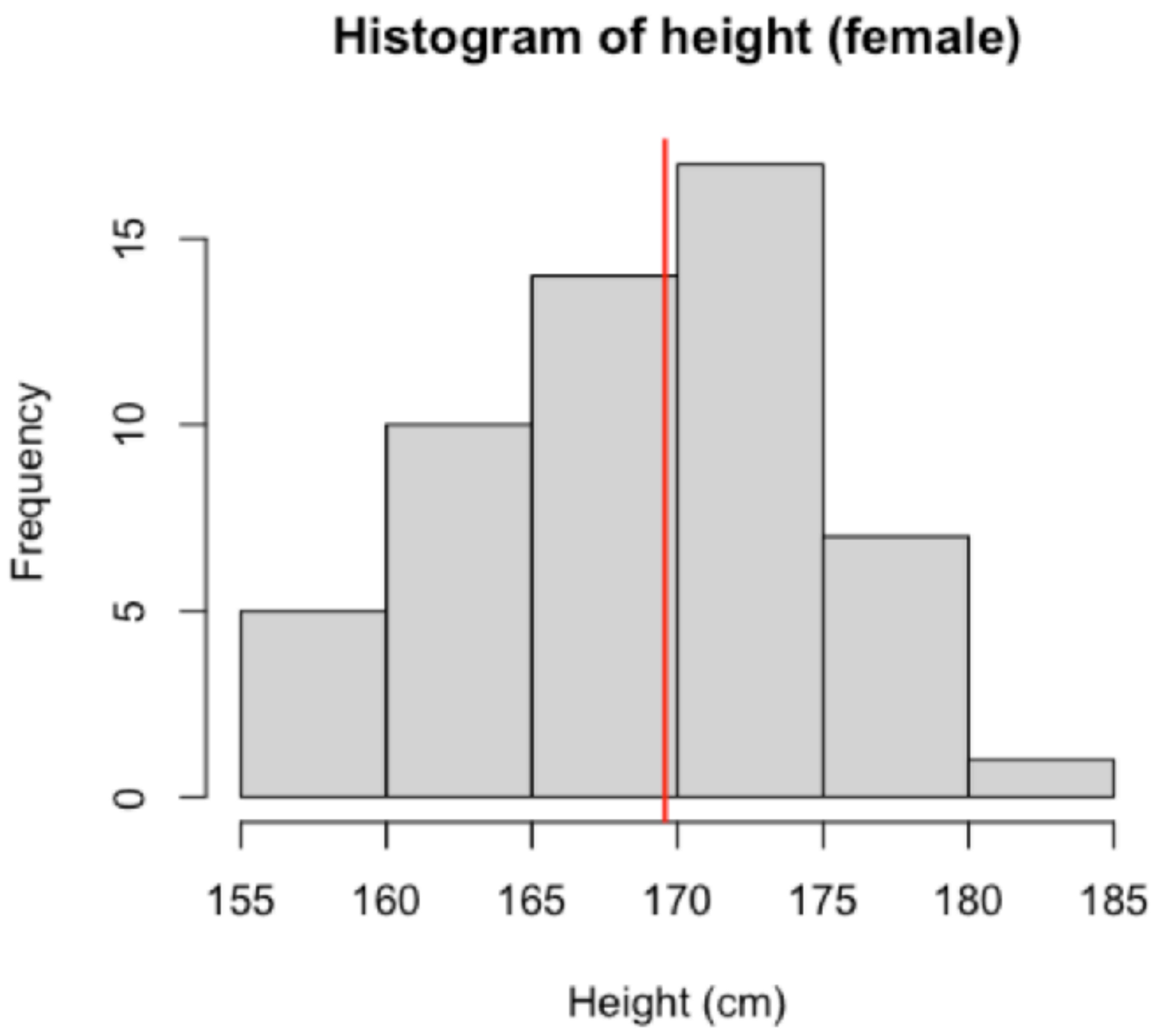
(Example 2a, 2b in t-test lab notes)
Lung data (PEFH98-english)

Height for women:
We want to compare the **average height** of women with a fixed value. (Say, 167cm; or 178 cm)
How **confident** are we about the conclusion?

Step 1: get to know your data

Descriptive statistics:
produce **mean, variance, min, max** (among others)

Visualise your data:
Histogram
Useful to plot **mean** on top of the histogram



mean (average): 169.57
variance: 32.40
min, max: 158, 183

Sample mean

(Example 2a, 2b in t-test lab notes)
Lung data (PEFH98-english)

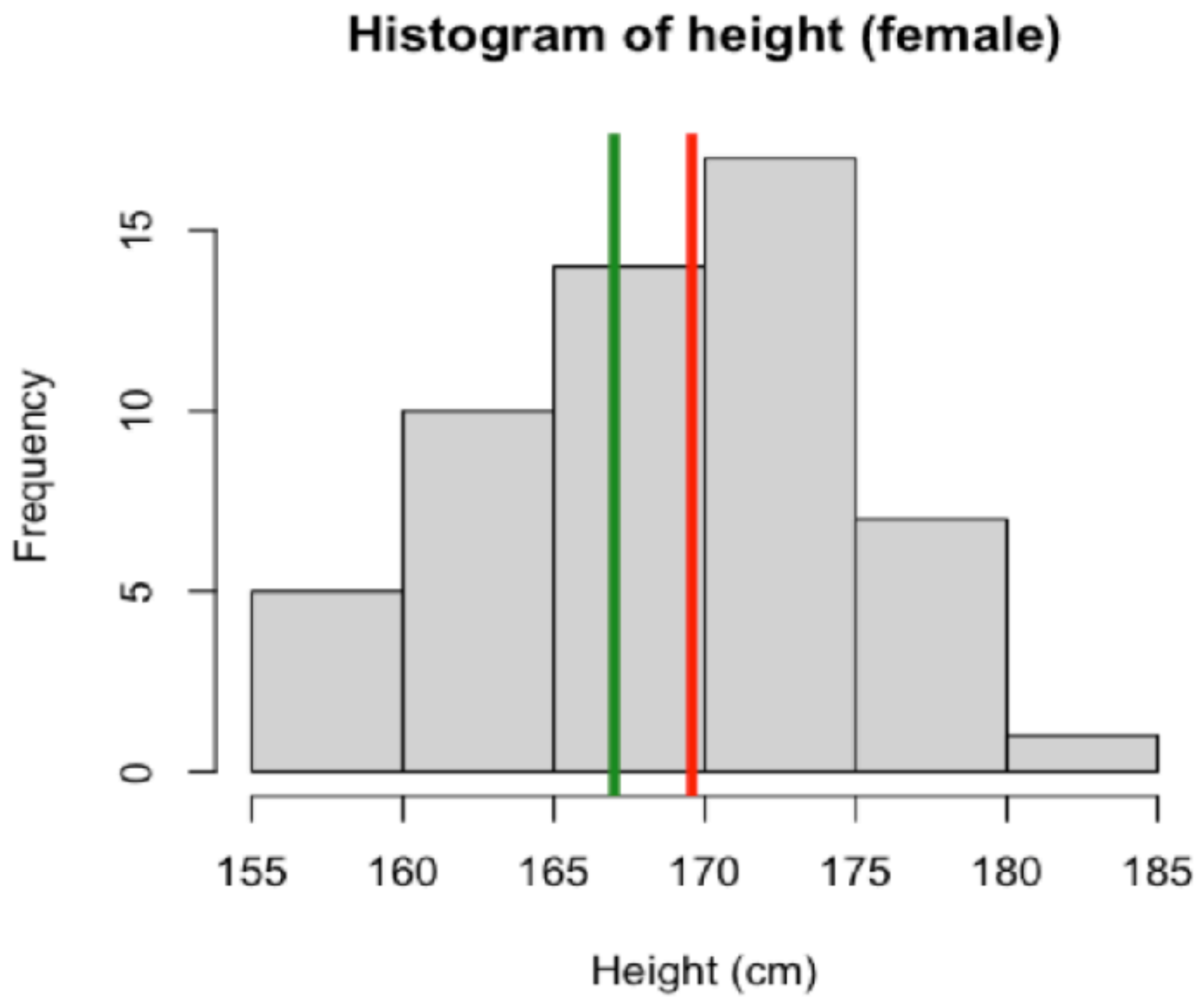
Height for women:
We want to compare the **average height** of women with a fixed value. (Say, 167cm)
How **confident** are we about the conclusion?

Step 2: understand what you are comparing

Sample mean (average height: 169.57)
A fixed value (167)

Are these two lines sufficiently different? ("significantly")

Need to know what possible values the ***sample mean*** *could take (with probability)*



Confidence interval (of mean)

(Example 2a, 2b in t-test lab notes)
Lung data (PEFH98-english)

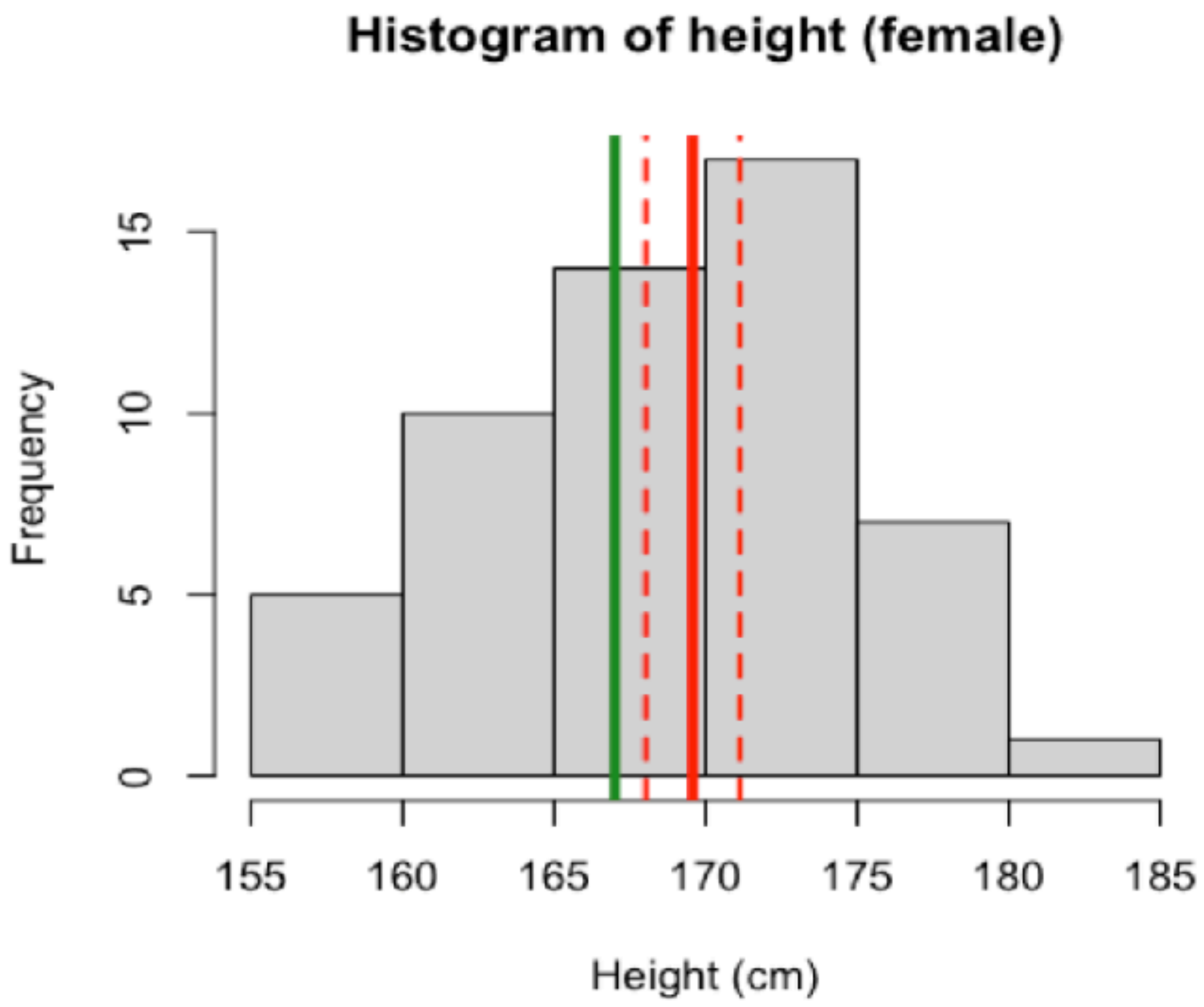
Height for women:
We want to compare the **average height** of women with a fixed value. (Say, 167cm)
How **confident** are we about the conclusion?

Step 3: compute the **confidence interval** (95%) for sample mean

95%: out of 100 experiments (random sampling), 95 times, the sample mean falls within this range

This range is computed based on **mean** and **standard error,** along with **quantiles** of a distribution (t or normal)

(However, in practice such as doing a t-test you do not need to compute by hand, standard statistical tests implements this for you)



mean (average): 169.57
95% CI: 168.02; 171.13

t-test (one sample)

We can formally do a **hypothesis test** and compute a **p-value** to express our confidence in the results

Student's t-test (in this case, **one-sample**) formally tests whether a **sample mean** is equal to a pre-specified value

State your hypothesis

Null hypothesis H0: mean height of female is equal to 167

Alternative hypothesis Ha: mean height of female is NOT equal to 167

Compute test statistic T0 <u>under H0</u>

$$T_0 = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{169.57 - 167}{5.69/\sqrt{54}} = 3.323$$

Compare with **critical values** at a certain level of significance (0.05)

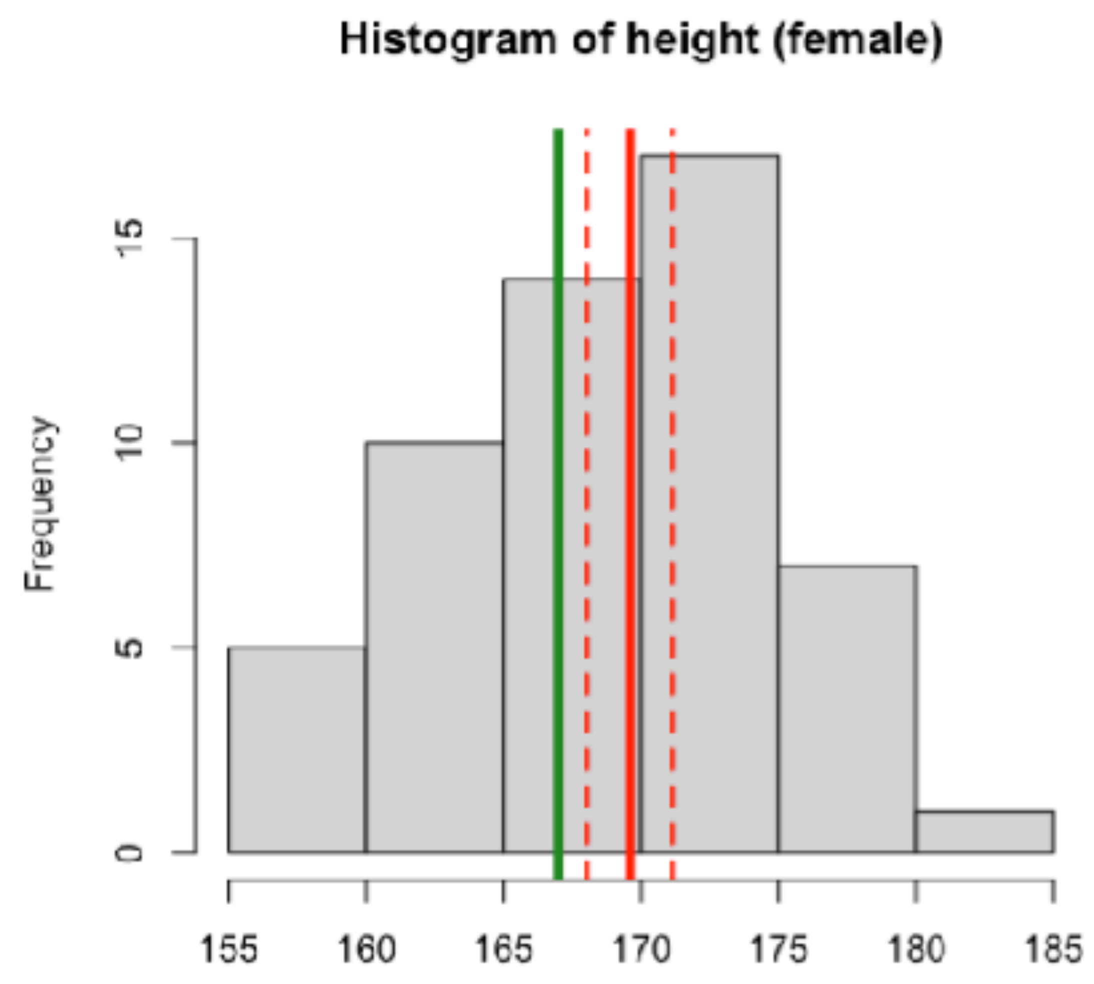
If T0 is more extreme than critical value, it means it is unlikely to be observed -> reject H0

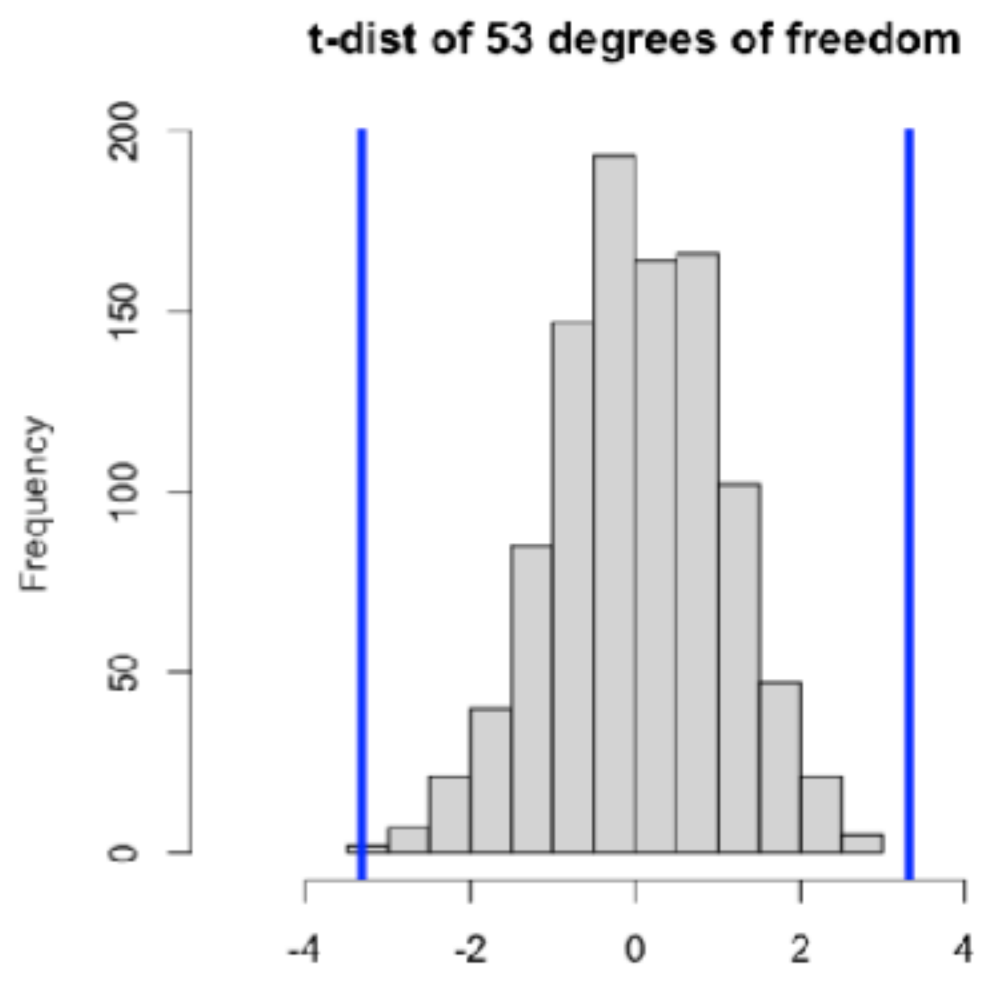
$$t_{53,0.975} = 2.005$$

You can also compute a **p-value** (of observing T0 = 3.323) under the null hypothesis (t-distribution of 53 degrees of freedom)

If p-value < 0.05, it means T0 is unlikely to be observed -> reject H0
(p = 0.0016 -> reject H0)

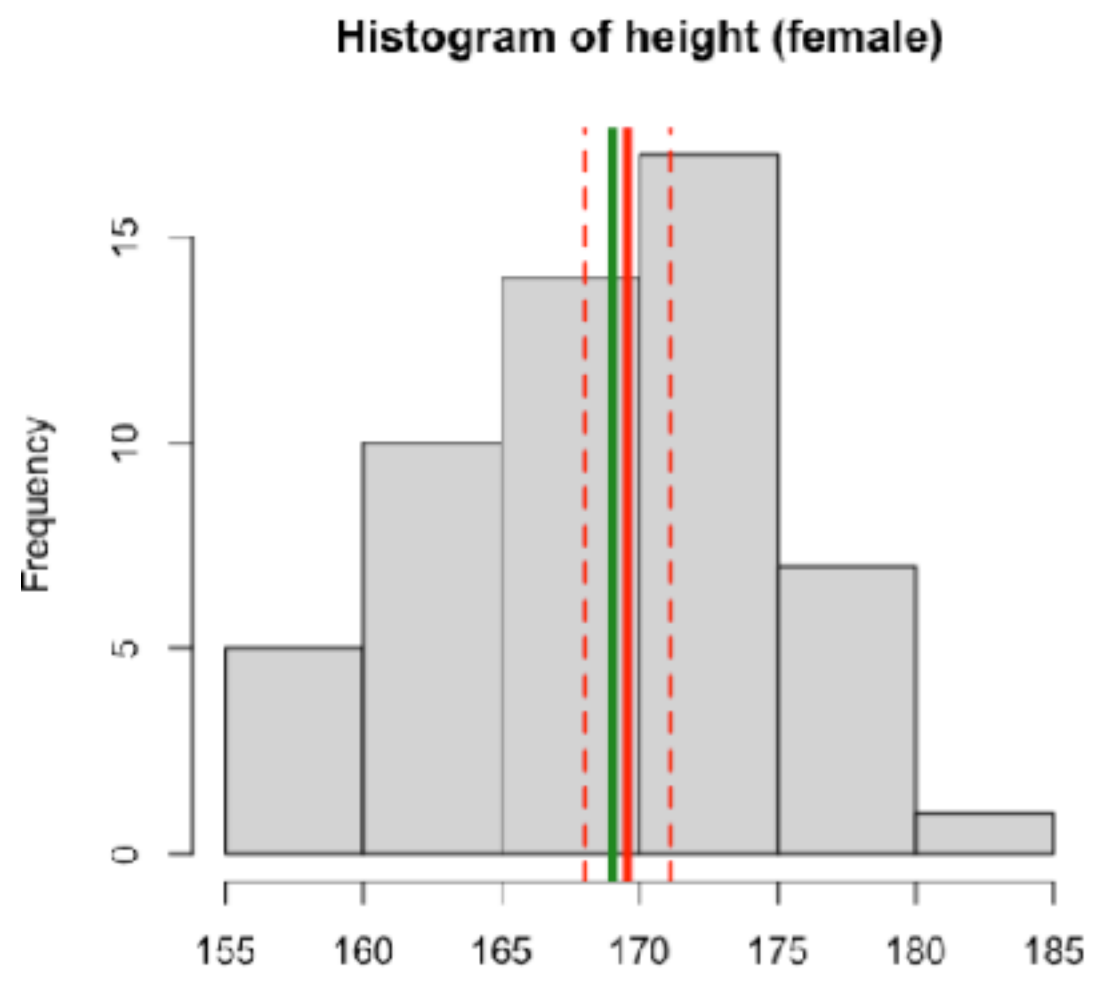
t-test (one sample)

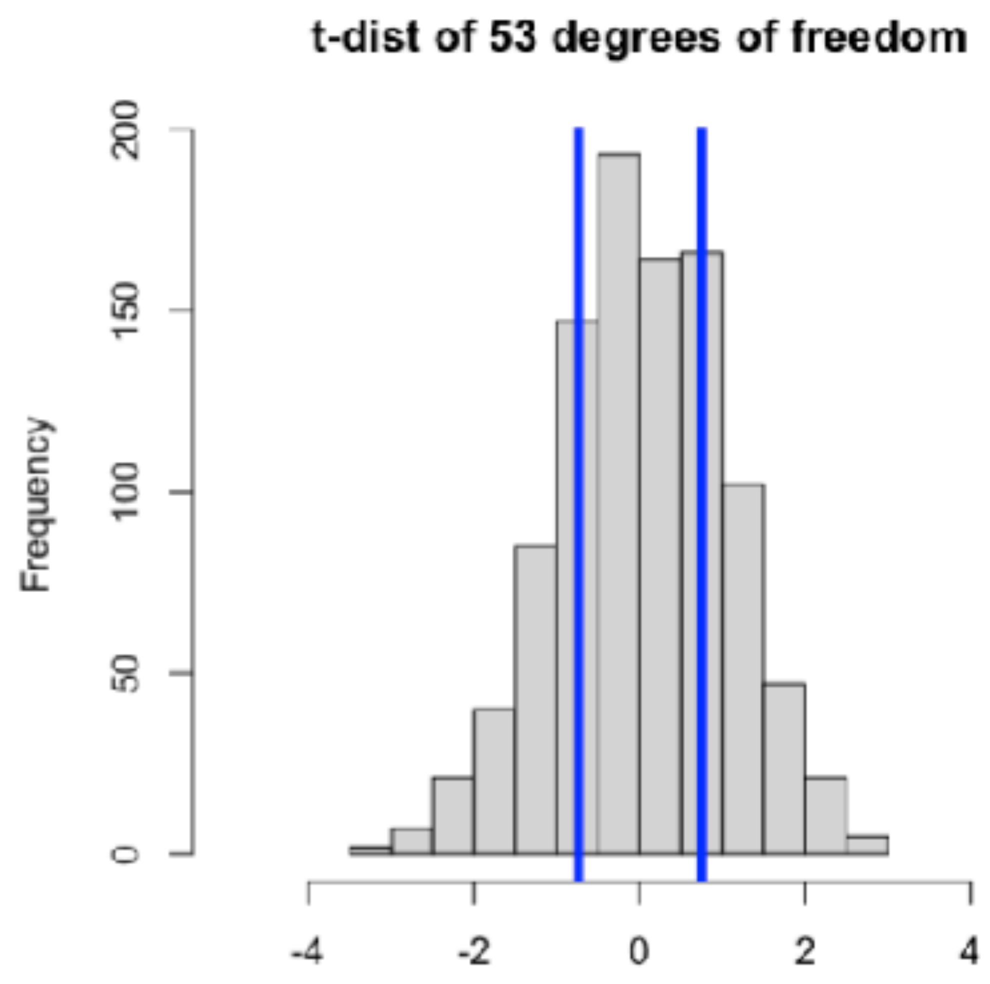




Mean height (169.57) vs 167
T = 3.23
p = 0.0016

(Prob beyond +- 3.23 under t53)





Mean height (169.57) vs 169
T = 0.74
p = 0.462

(Prob beyond +- 0.74 under t53)

t-test (paired sample, two samples)

Hypothesis testing workflow

Step 1. State **null** and **alternative hypothesis**, H0, Ha

Step 2. Compute **test statistic** (under the null hypothesis)

Step 3. Compare with **critical values**, compute a p-value

Step 4. Decide whether to **reject** or **not reject** the null hypothesis

Paired sample:
on the same subject (before / after treatment)

Two independent samples:
different subject (case control, male female)

```
# by default, conf.level = 0.95

t.test(x) # by default, mu = 0

# whether sample mean is equal to 5
t.test(x, mu=5)

t.test(x1,x2, paired = T) # paired
t.test(x, y)    # two ind. samples
```

Check assumption

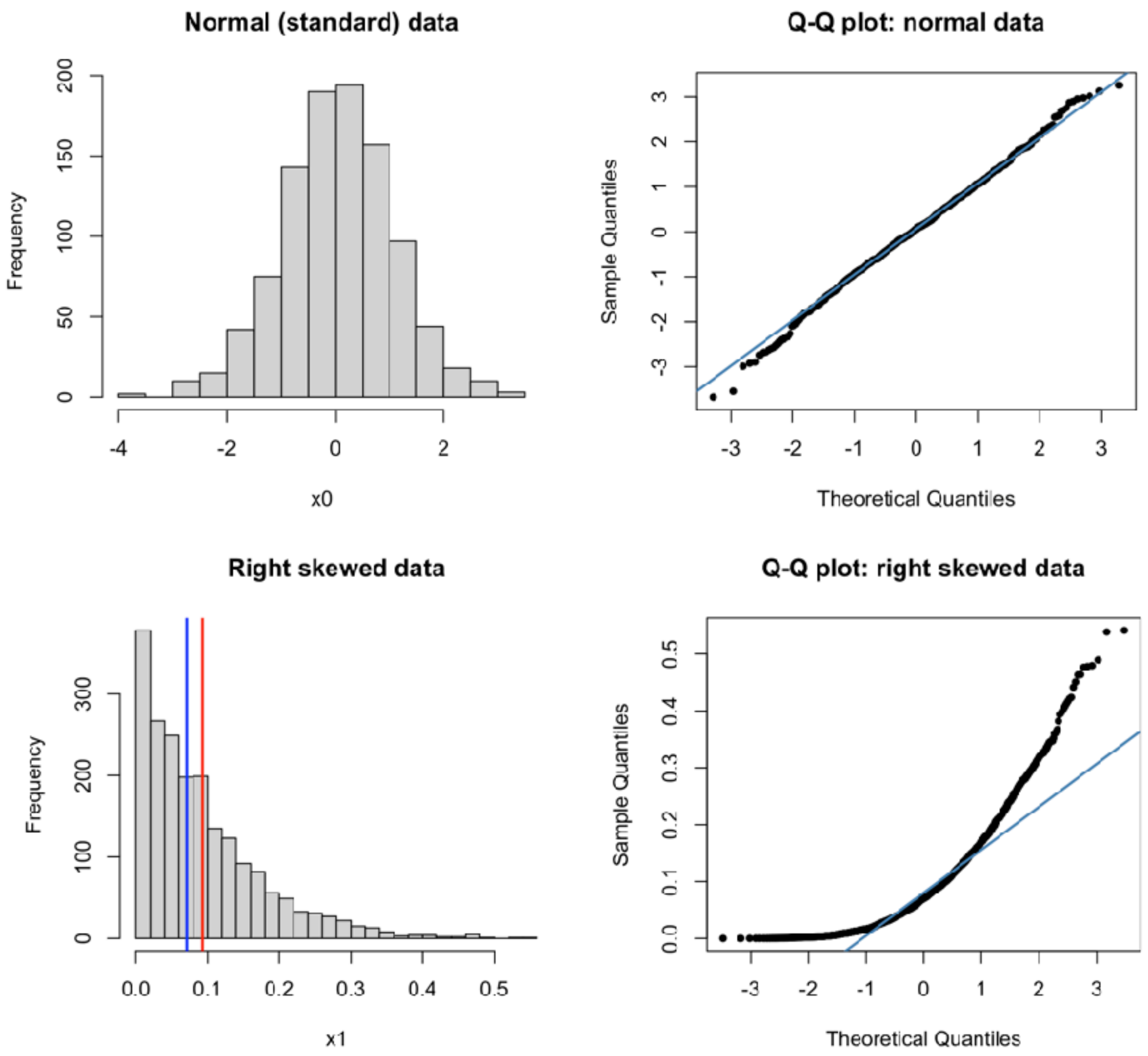
Check for normal distribution with visualization

Histogram: looks like bell shaped

Q-Q (quantile-quantile) plot: points fall approximately on a straight line

Important to do EDA! (Exploratory data analysis)

If your data histogram looks like this: t-test is NOT appropriate. (Week 2)



Demonstration