

Stat 411/511

# ASSUMPTIONS OF THE T-TOOLS

Feb 1 2012

Charlotte Wickham

[stat511.cwick.co.nz](http://stat511.cwick.co.nz)

# HW #3 Q.2.

I forgot R reports one-sided confidence intervals. (you'll see Inf or -Inf)

We won't talk about one-sided intervals.

Use a two-sided test to get the 95% confidence interval.

Report the p-value from the one-sided test.

# Back to the creativity example

*t*-tests are motivated by **sampling from a population**, but also work well when the mechanism of chance is **random group allocation**.

**Null:** The mean treatment effect on the creativity score of being assigned the intrinsic motivation questionnaire compared to the extrinsic motivation questionnaire is zero.

$$t\text{-statistic} = (\bar{Y}_2 - \bar{Y}_1 - 0) / SE_{\bar{Y}_2 - \bar{Y}_1} = 4.14 / 1.42 = 2.92$$

$$(1 - pt(2.92, 23 + 24 - 2)) = 0.0027 \text{ (one-sided p-value)}$$

# t-tools

paired  $t$ -test

two sample  $t$ -test

# Assumptions of the $t$ -tools

1. Normality of population
2. Equality of population standard deviations
3. Independence of subjects

never see the  
population  
examine  
samples for  
evidence

A test is **robust** to an assumption if the test is valid even if the assumption is not met.

This is going to quite a high level summary.  
Read Section 3.2 for all the details.

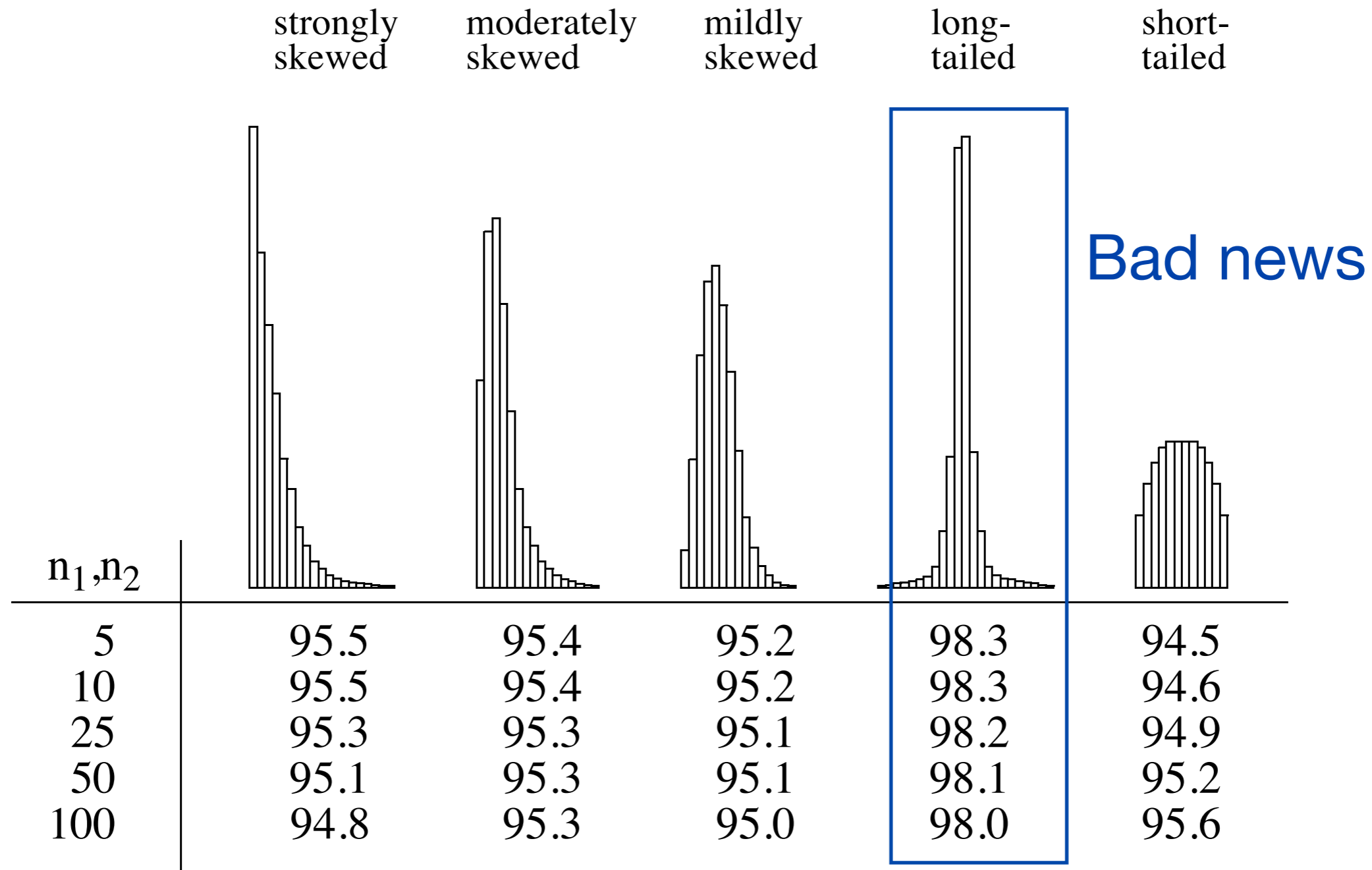
# Normality of population

# Normality of the population

The **larger** the sample size, the **less** you have to worry, thanks to the Central Limit Theorem.

What is a **large** sample? It depends...

Percentage of 95% confidence intervals that are successful when the two populations are non-normal (but same shape and SD, and equal sample sizes); each percentage is based on 1,000 computer simulations



Bad news

# Generalizations

If everything else is **the same** for the two populations:

- Skew is **ok**
- Long tailed is **not ok**

If everything else is **not the same** for the two populations:

- Skew is **not ok**
- Long tailed is **not ok**

Remember: things get better with larger samples

# Normality of population

**Check:** by looking at histograms of samples.

**Remedy:**

use a transformation (Weds) **or**

use a non-parametric test (next week)

# Equal standard deviations

# Equal population standard deviations

We assumed that  $\sigma_1 = \sigma_2$  and estimated the value with the pooled standard deviation.

If your **sample sizes** are roughly **equal**, the *t*-test will still be fairly valid.

If your **sample sizes are not equal**, the *t*-test will not be valid.

# Equal population SDs

**Check:** by looking at histograms of samples.

**Remedy:**

use a transformation (Weds) **or**

use a Welch *t*-test (next week)

# Independence

# Independence

Subjects must be independent of one another for  $t$ -tools to be valid.

If knowledge about one subject allows us to make a better guess about another subjects response, there is a lack of independence.

If there is dependence,  $t$ -tools give misleading results.

# Two common types of dependence

## **Cluster effects**

There is some kind of subgroup within groups, and subjects in the same subgroup are more similar.

## **Serial effects**

Measurements are made in time (or space) and observations made close in time (or space) are more similar.

# Independence

**Check:** read study design carefully, and look for sources of dependence.

**Remedy:**

use more complicated methods  
(ST512)