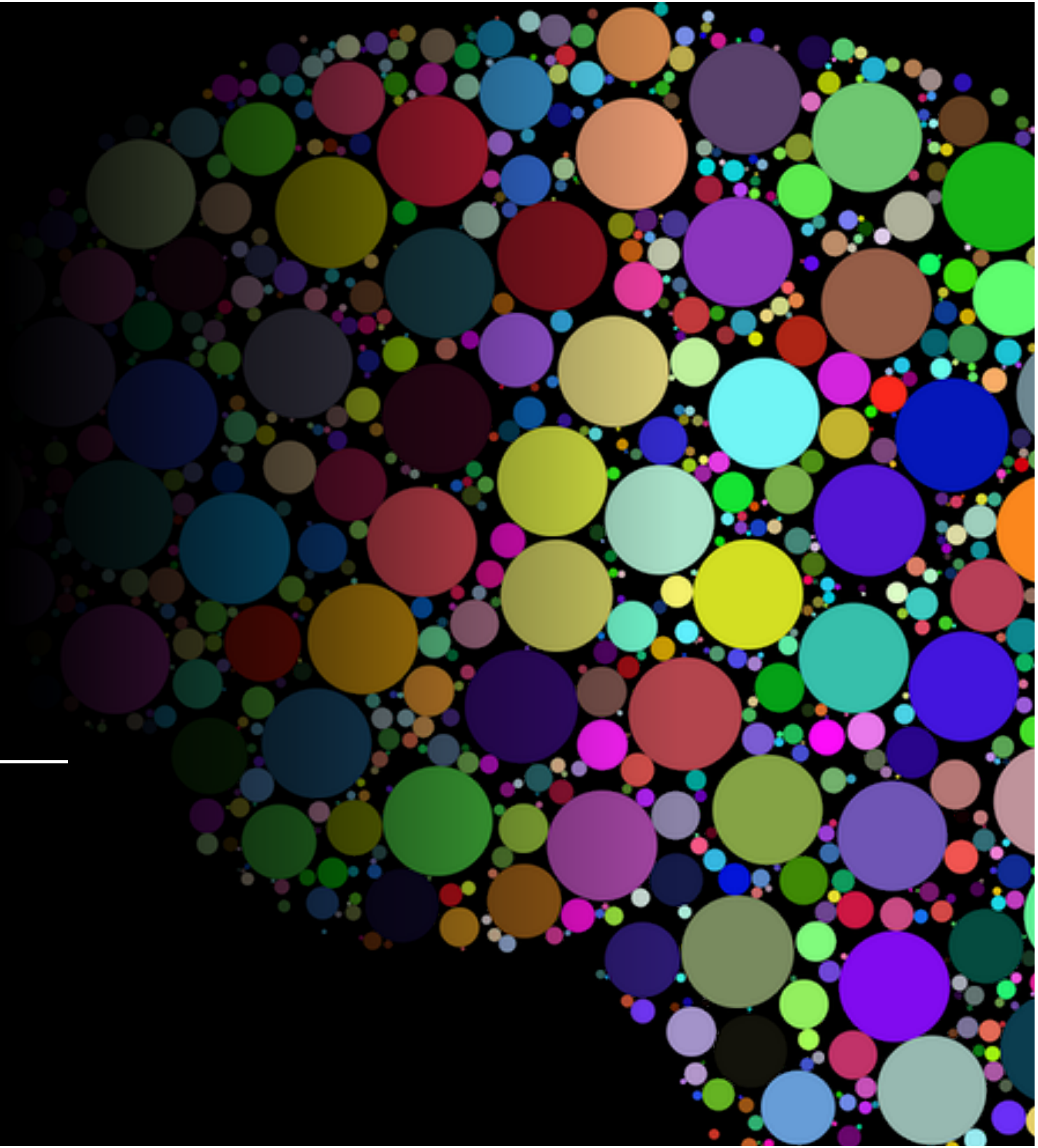
The Single-Sample *t* Test



Chapter 9

t distributions

- Sometimes, we do not have the population standard deviation, σ.
  - Very common!
- So what can we do?

t distributions

> The *t* distribution is used when we do not know the population information.

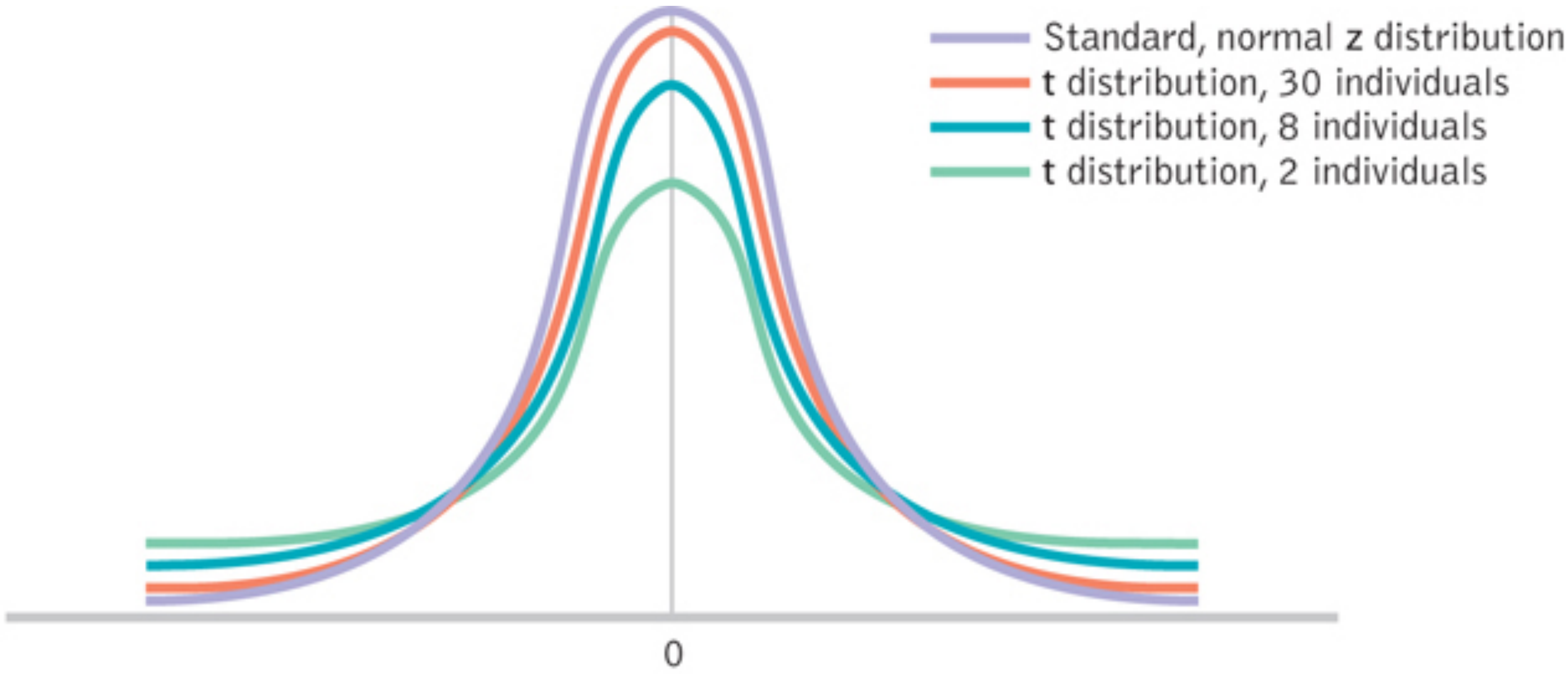
- So we use the sample to estimate the population information.
- Because we are using the sample, the *t* distribution changes based on that sample.

The t Statistic

- When sample size increases:
  - *s or sd* (the spread of *t*) → σ
  - Therefore, *t* and *z* become more equal

- The *t* distributions
  - Distributions of differences between means

Wider and Flatter *t* Distributions



Check Your Learning

- When would you use a z test?
- When would you use a t test?

Types of *t*

- Single sample *t*
  - One sample (group of people), population mean to compare against
- Dependent sample *t*
  - One sample tested <u>**twice**</u> to compare those two scores
- Independent sample *t*
  - Two <u>**samples**</u> to compare those two groups

Hypothesis Tests: The Single Sample *t* Test

> The single sample *t* test

- When we know the population mean, but not the standard deviation
- So, we will use the sample to estimate SD
- But that means we have to use the sample to estimate cut off scores too, since the distribution spread is not set.

The *t* test

- The six steps of hypothesis testing
  - 1. Identify population, distributions, assumptions
  - 2. State the hypotheses
  - 3. Characteristics of the comparison distribution
  - 4. Identify critical values

    df =N-1
  - 5. Calculate
  - 6. Decide

Hypothesis Testing: Step 1

> Assumptions:
- DV is scale
- Random selection → random assignment?
- Normal → N > 30?

Hypothesis Testing: Step 2

- Label the sample and the population
- Pick a null and research
  - Remember the combinations are:
    - Greater
    - Lesser
    - Two Tailed = Different

Hypothesis Testing: Step 3

> Label:

- Sample M
- Sample SD
- Sample SE
- Sample N
- Population u

Hypothesis Testing: Step 3

- Something new:
  - Degrees of freedom

    $df = N - 1$ where N is sample size
- This number is used in two ways:
  - Estimation of SD
  - Calculating the cut off score

t distributions

Sample Standard Deviation

$$SD = \sqrt{\frac{\sum (X - M)^2}{N}}$$

What we did before…
Biased estimate

Population Standard Deviation

$$s = \sqrt{\frac{\sum (X - M)^2}{(N - 1)}}$$

New formula…
Unbiased estimate

Based on some error

Calculating in *R*

- summary(*dataset*)
  - mean(*dataset*, na.rm = T)
- sd(*dataset$column, na.rm = T*)
- OR you can enter the data:
  - data = c(#,#,#,#,#)
  - sd(*data, na.rm = T*)

Calculating Standard Error for the t Statistic

> Using the standard error

$$S_M = \frac{s}{\sqrt{N}}$$

Calculating in *R*

> se = sd / sqrt(*N*)

- If you have all the data in R you can try:
- sd(*data*) / sqrt(length(*data*))
  > Remember the sd function gives you SD
  >
  > The sqrt() function is square root
  >
  > length() calculates the number of items or N
  >
  > Remember that *data* can be one column or *dataset$column*

Hypothesis Testing: Step 4

- Use *df* to find the cut off score
- qt(*alpha p, df*, lower.tail = T or F)
  - Alpha p is set at .05 or .01
  - Remember, for a two tailed test, do alpha / 2

TABLE 9-1. Excerpt from the *t* Table

When conducting hypothesis testing, we use the *t* table to determine critical values for a given *p* level, based on the degrees of freedom and whether the test is one- or two-tailed.

	One-Tailed Tests			Two-Tailed Tests		
df	0.10	0.05	0.01	0.10	0.05	0.01
1	3.078	6.314	31.821	6.314	12.706	63.657
2	1.886	2.920	6.965	2.920	4.303	9.925
3	1.638	2.353	4.541	2.353	3.182	5.841
4	1.533	2.132	3.747	2.132	2.776	4.604
5	1.476	2.015	3.365	2.015	2.571	4.032

Stop and think. Which is more conservative: one-tailed or two-tailed tests? Why?

Hypothesis Testing: Step 5

> The *t* found statistic

$$t = \frac{(M - \mu_M)}{S_M}$$

Calculating in *R*

- To get *t*
  - Mean – mu / se
- If you have the calculated numbers (i.e. you are given M, u, SD, N), then you can fill in the formulas.
- If you are given the raw numbers, then we can calculate with the t.test() function.

Calculating in *R*

> t.test(*column of y data*,

mu = #,

alternative = “less” OR “greater” OR
“two.sided”,

conf.level = .95 OR .99)

Note: alternative and conf.level options.

Hypothesis Testing: Step 6

> Decide to reject or fail to reject the null hypothesis.

> Beyond hypothesis testing:

- Calculate effect size
- Calculate confidence interval

Calculating Effect size

$$d = \frac{(M - \mu)}{s}$$

Confidence Interval

- We are usually discussion two tailed confidence intervals
  - You can do one tailed confidence intervals but they are not very common.
- They are calculated in the same way as z-tests but with a t-critical instead of z-critical.

Confidence Interval

> Lower = $M_{sample} - t_{critical}$*SE
> Upper = $M_{sample} + t_{critical}$*SE

> A quicker way!
- Use t.test() with a TWO tailed test to get the two tailed confidence interval.

Interpretation of Confidence Interval

If we were to sample N students from the same population over and over, the 95% confidence interval would include the population mean 95% of the time.

t, Effect Size, and Confidence Interval Cheat!

> Use the effsize and calculate code Dr. B has written.

- Not required, but if you want to check to make sure you doing it correctly.
- <u>Definitely</u> much easier to use when calculating *d* for other types of t-tests.

t, Effect Size, and Confidence Interval Cheat!

- Make sure the functions appear in your window or you won't be able to use them!