Learning Objectives

- **Compare/Contrast** *z*-test and *t*-test
- **Describe** *standardized effect sizes* and compare them to *z*-scores
- **Learn** naming conventions for Cohen's *d,* and three different variants of *d*
- **Calculate** Cohen's $\hat{d}$ (and compare to calculating the *t*-statistic)

t test, binomial, sign test: need to be corrected for w df

t-test

z test is “perfectly normal”, doesn’t need to be corrected for/no df: no estimate for sd, we know it

$$z_{\text{obt}} = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

$$t_{\text{obt}} = \frac{\bar{X} - \mu}{s_{\bar{X}}}$$

sigma = z dist.

Notice: We’re using $s_{\bar{X}}$ instead of $\sigma_{\bar{X}}$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$$

$$s_{\bar{X}} = \frac{s}{\sqrt{N}}$$

t-test estimates σ

- We assume that $s = \sigma$
  - However, s is a *biased* estimator
    - s will systematically underestimate σ

- We correct for bias using ***degrees of freedom***

2 variables: calculate s for x and y, so df is N - 2

number of s calculations = # of “corrections” (N - #)

$$s = \sigma_{\text{est}} = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{\boxed{N-1}}}$$

Determining t_{crit} when $N = 11$, $\alpha = .05$

table D Critical values of Student's t distribution

The values listed in the table are the critical values of t for the specified degrees of freedom (left
(column heading). For two-tailed alpha levels, t_{crit} is both + and –. To be significant, $|t_{obt}| \geq |$

df	Level of Significance for One-Tailed Test, $\alpha_{1\ tai}$			
	.10	.05	.025	.01
	Level of Significance for Two-Tailed Test, $\alpha_{2\ tai}$			
	.20	.10	$\boxed{.05}$	.02
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.986
9	1.383	1.833	2.262	2.821
10	1.372	1.812	$\boxed{2.228}$	2.764

make sure to use df instead of N!

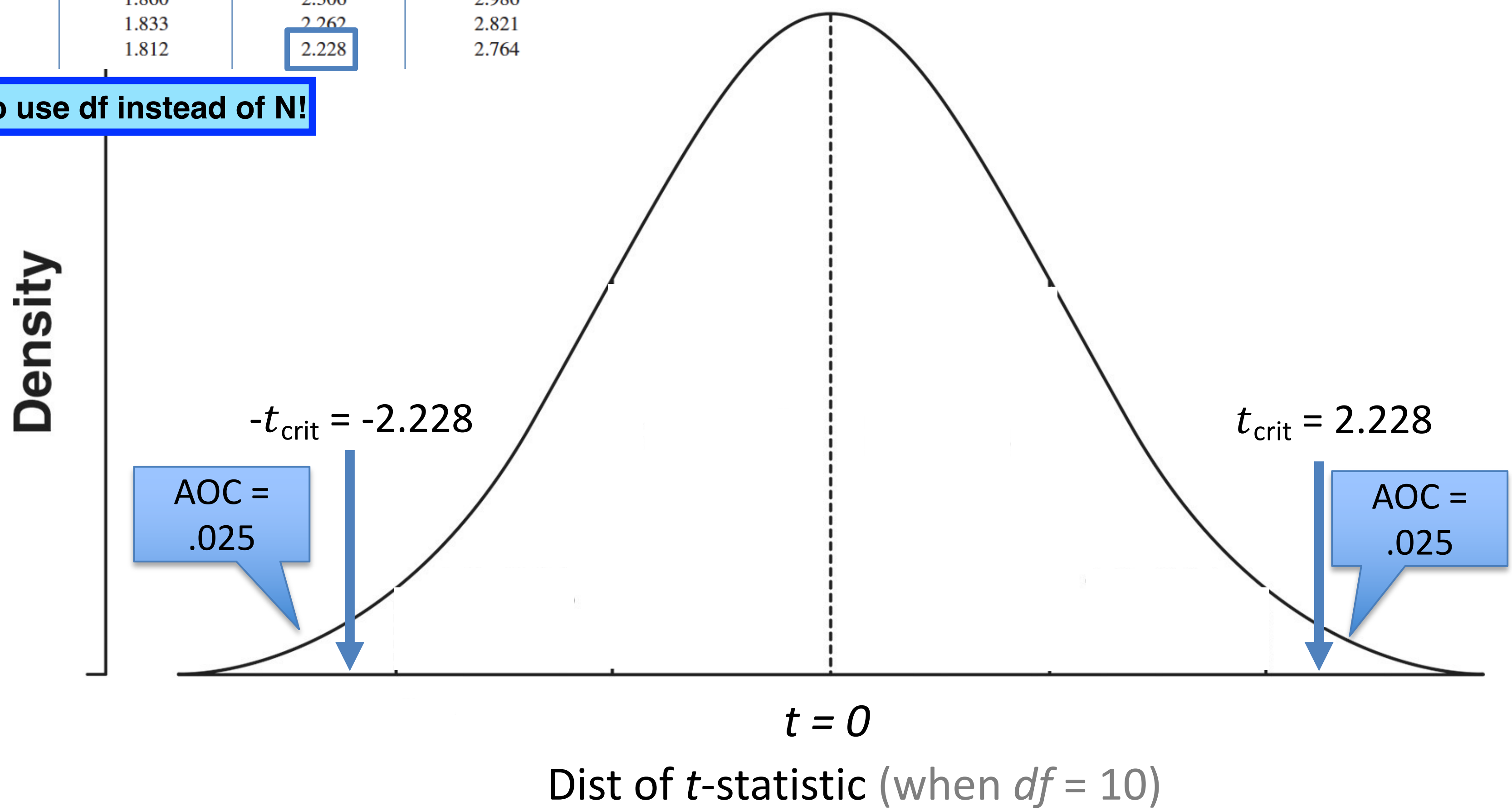


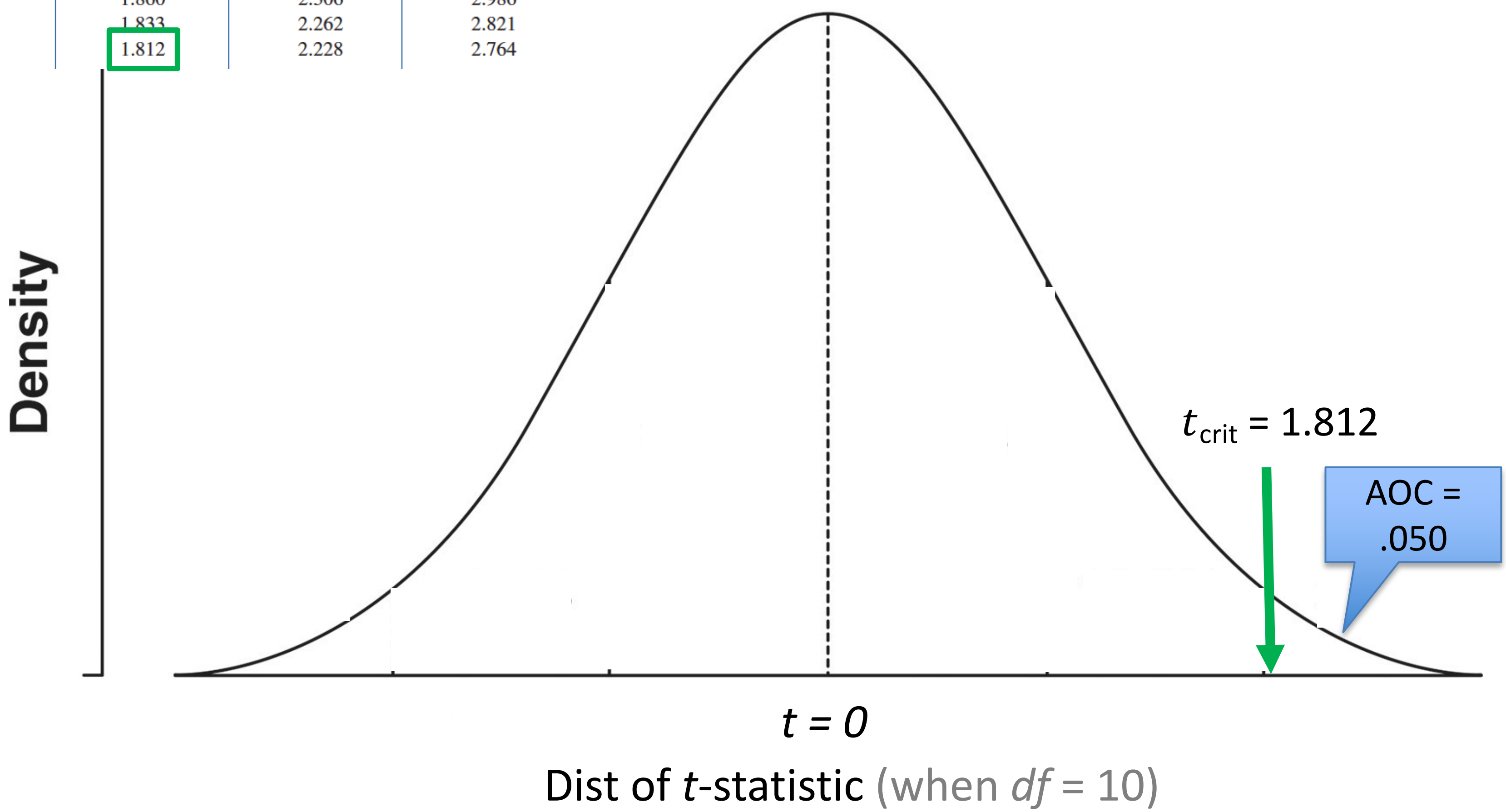
table D Critical values of Student's *t* distribution

The values listed in the table are the critical values of t for the specified degrees of freedom (left
(column heading). For two-tailed alpha levels, t_{crit} is both + and –. To be significant, $|t_{obt}| \geq |$

df	Level of Significance for One-Tailed Test, $\alpha_{1\ tai}$			
	.10	$\boxed{.05}$	.025	.01
	Level of Significance for Two-Tailed Test, $\alpha_{2\ tai}$			
	.20	.10	.05	.02
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.986
9	1.383	1.833	2.262	2.821
10	1.372	$\boxed{1.812}$	2.228	2.764

Determining t_{crit} when N = 11, α = .05

how close is t to z?



Effect Sizes

- Effect sizes are meant to capture *magnitude* rather than *reliability*
  - *ES* are descriptive stats, not inferential stats

- Examples of raw *ES*:
  - $(\bar{X}_{\text{obt}} - \mu)$
  - $(\bar{X}_1 - \bar{X}_2)$
  - $(\bar{X}_{\text{pre}} - \bar{X}_{\text{post}})$

- Raw effect sizes are expressed in original units of measurement (e.g., IQ points, percentage grade, likert scale ratings)

Standardized Effect Sizes

- **Standardized** ES is like standardized scores
  - They're put on the same scale
    - dividing by standard deviation!

comparable bc they are divided by s

mu or sample mean; can be either

$$z_1 = \frac{X_1 - \bar{X}}{s}$$

Standardized

Score

$$\hat{d} = \frac{\bar{X}_{\mathrm{obt}} - \mu}{s}$$

Standardized

Difference or ES

Standardized Effect Sizes

- Easily comparable across different studies:
  - We could compare:
    - *z* scores to each other
    - *r* values to each other
    - ***d* values to each other**
      - How different are these patterns, in units of standard deviations?

- Naming conventions for *d* (just like *r*):
    - *d* = 0, “no effect”
    - *d* = .20, “small effect” **antihistamines have d of .20**
    - *d* = .50, “medium effect”
    - *d* = .80+, “large effect”

Three Cohen's *d*'s

There are different calculations for *d*, depending on the research design:

- Comparing a sample mean to a population mean:
  (single sample *t*-test)

  $$\hat{d}$$

- Comparing two independent sample means to each other:
  (Student's or independent *t*-test)

  $$d_s$$

- Comparing two dependent means to each other:
  (dependent *t*-test)

  $$d_z$$

Each have slightly different formulas, but the idea is the same!

Calculating Cohen's $\hat{d}$

best for: single sample t test

- An instructor tested a new teaching technique on a *sample* of 500 students last year. The sample scored 10% better on the final exam, 75%, compared to her *population* of past students. The standard deviation of her sample was 14%.

What was the effect size for this teaching technique?

First, what Cohen's *d* should we use?

$$\hat{d} = \frac{\bar{X} - \mu}{s}$$

$$\hat{d} = \frac{75 - 65}{14} = 0.71$$

Calculating Cohen's $\hat{d}$

- An instructor tested a new teaching technique on a *sample* of **20** students last year. The sample scored **10% better** on the final exam, **75%**, compared to her *population* of past students. The standard deviation of her sample was **14%**.

What was the effect size for this teaching technique?

$$\hat{d} = \frac{\bar{X} - \mu}{s}$$

$$\hat{d} = \frac{\mathbf{75} - \mathbf{65}}{\mathbf{14}} = 0.71$$

Calculating *t*- & *p*-values

- An instructor tested a new teaching technique on a *sample* of **20** students last year. The sample scored 10% better on the final exam compared to her *population* of past students. The standard deviation of her sample was 14%.

Is this effect statistically reliable?

$$\boxed{s_{\bar{X}} = \frac{s}{\sqrt{N}}}$$

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s_{\bar{X}}}$$

Calculating *t*- & *p*-values

- An instructor tested a new teaching technique on a *sample* of **20** students last year. The sample scored **10%** better on the final exam compared to her *population* of past students. The standard deviation of her sample was **14%**.

Is this effect statistically reliable?

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{14}{\sqrt{20}} = 3.1304$$

$$t_{\mathrm{obt}} = \frac{\bar{X}_{\mathrm{obt}} - \mu}{s_{\bar{X}}}$$

Calculating *t*- & *p*-values

- An instructor tested a new teaching technique on a *sample* of **<u>20</u>** students last year. The sample scored 10% better on the final exam compared to her *population* of past students. The standard deviation of her sample was 14%.

Is this effect statistically reliable?

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{14}{\sqrt{20}} = 3.1304$$

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s_{\bar{X}}} = \frac{75 - 65}{3.1304} = 3.194$$

Lookup t_{crit} in Table D ($\alpha_{2-\text{tail}}$ = .05, $df = N - 1 = 19$)

$t_{\text{crit}} = \pm 2.093$

Compare t_{obt} to t_{crit} **passes, reject H_0**