Learning Objectives

- **Describe** how the *t*-distribution differs from the *z*-distribution

- **Practice** finding **critical values** of the *t*-statistic

- **Conduct** a *single-sample t*-test

Comparing *z* and *t*

- When using *z*-test, we know:
  - <u>Know</u> population mean (μ)
  - <u>Know</u> population standard deviation (σ)

But in the real world, we usually don't know one (or both) of these!

- When using *t*-test, we:
  - <u>Set</u> hypothetical population mean
    - Either we know μ, or we know what value to test against
  - <u>Estimate</u> σ using sample standard deviation (*s*)

t-test

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

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

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 <u>estimates</u> σ

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

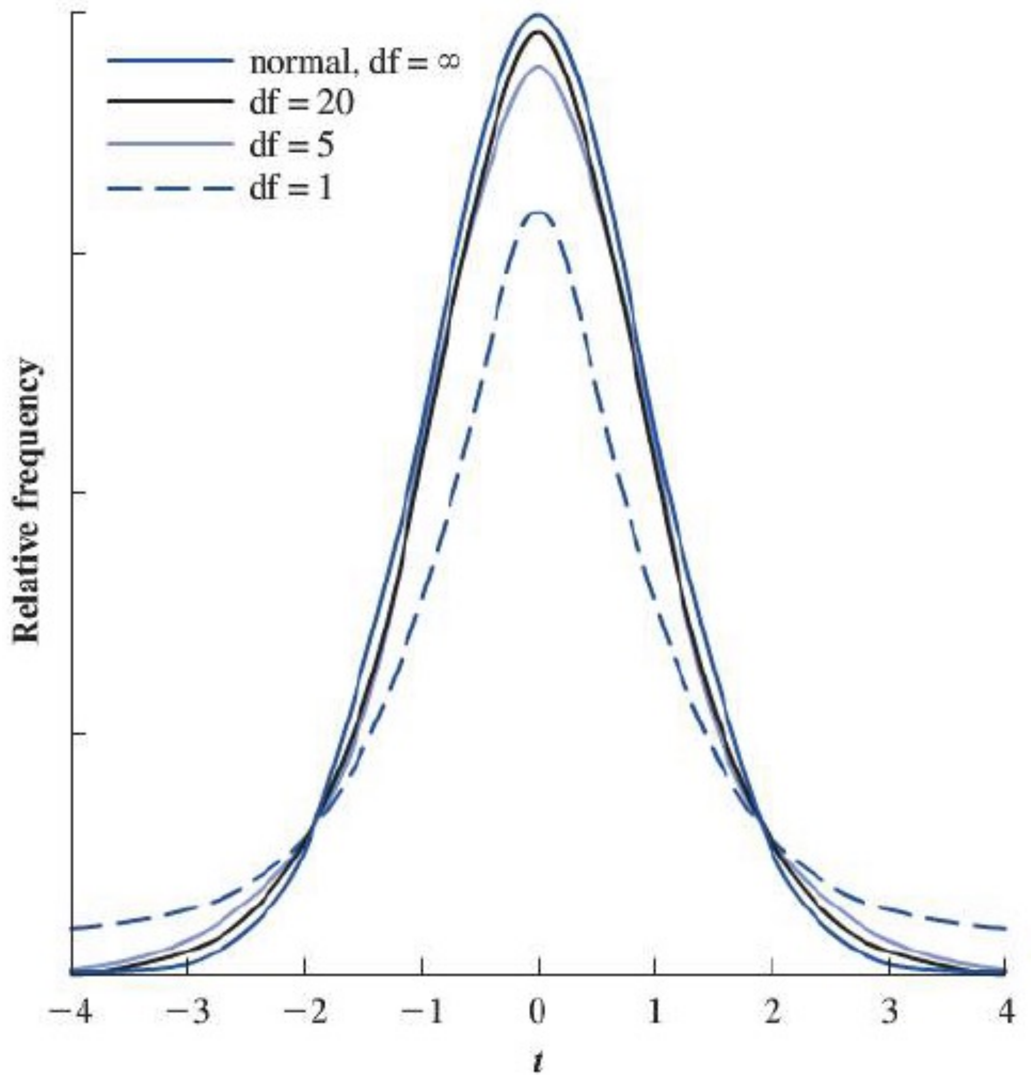
- We correct for bias using ***degrees of freedom***
  - ***Remember class 4 when we learned about standard deviation?***

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

Comparing *z*-test *vs.* *t*-test

z-test	t-test
Standard deviation of population: $\boldsymbol{\sigma}$	*Estimated* standard deviation of Ho population: **s**
Standard error of the mean: $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}}$	*Estimated* standard error of the mean: $s_{\overline{X}} = \frac{s}{\sqrt{N}}$
z-score: $z_{obt} = \frac{\overline{X}_{obt} - \mu}{\sigma_{\overline{X}}}$	t-score: $t_{obt} = \frac{\overline{X}_{obt} - \mu}{s_{\overline{X}}}$

t-distribution <u>approximates</u> *z*-distribution as *N* increases



Pagano, p. 331

z-test vs. *t*-test

1. For both tests, we either know μ or we know what μ will be if H_0 is true
2. *t*-test ***estimates*** population standard error from sample standard error
   – Both $\sigma_{\bar{X}}$ & $s_{\bar{X}}$ depend on sample size
3. *t*-distribution approaches normality as *N* increases; *z*-distribution is always normal
   – We choose which *t*-distribution to use based on our sample size (just like binomial)

Two methods for *t*-test

<u>*p*-value method</u>

- Set H_0/H_1, α
- Calculate $s_{\bar{X}}$
- Calculate t_{obt}
- Determine *p*-value
  - We will not do this…
- Compare *p* & α

<u>t_{crit} method</u>

- Set H_0/H_1, α
- Calculate $s_{\bar{X}}$
- Calculate t_{obt}
- Look up t_{crit} (pp. 604)
  - We will do this!
- Compare t_{obt} & t_{crit}

Try it! Look up t_{crit}

- We hypothesize that students in our class have a *higher* IQ than the population
  - We will conduct a one-tailed hypothesis test
  - We set $\alpha = .05$
  - We sample from 30 students

- What is t_{crit}? Look on pp. 604
  - Answer: $t_{crit} = 1.699$
    - Remember *df* = (N - 1) for single-sample *t*-test

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) and the alpha level (column heading). For two-tailed alpha levels, t_{crit} *is both + and –. To be significant,* $|t_{obt}| \geq |t_{crit}|$.

df	Level of Significance for One-Tailed Test, $\alpha_{1\ tail}$					
	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-Tailed Test, $\alpha_{2\ tail}$					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.986	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

Example *t*-test

- Young people check their phones an average of 150 times per day
  - *Do UBC students check their phones at a different rate than young people generally?*

- You observe 24 students from class and find:

$$\bar{X} = 136$$

$$s = 10$$

Do UBC students use their phones differently?

$\mu = 150$
$\bar{X}_{\mathrm{UBC}} = 136$
$s = 10$
$N = 24$

$s_{\bar{X}} = 2.04$
$t_{\mathrm{obt}} = -6.8587$
$p < .0001$

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{10}{\sqrt{24}} = 2.0412$$

$$t_{\mathrm{obt}} = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{136 - 150}{2.0412} = -6.8587$$

Do UBC students use their phones differently?

$\mu = 150$
$\bar{X}_{\mathrm{UBC}} = 136$
$s = 10$
$N = 24$

$s_{\bar{X}} = 2.04$
$t_{\mathrm{obt}} = -6.8587$
t_{crit} = pp. 604

$t_{\mathrm{obt}} = -6.8587$

$t_{\mathrm{crit}} = \pm 2.069$

Is $t_{\mathrm{obt}} > t_{\mathrm{crit}}$, or is $-t_{\mathrm{obt}} < -t_{\mathrm{crit}}$?
If yes, then reject H_0

Formatted Conclusion

- "Students at UBC checked their cell phones reliably less than the population on average, $t(23) = -6.858$, $p < .001$."