Today: Sections 13.1 to 13.3

ANNOUNCEMENTS:

- Discussion Friday is not for credit, review for final and help with this week's homework. You may attend multiple discussions, but first priority goes to those registered for that discussion.

- Quiz for week 8 starts Fri, ends *Monday* at **10am**.

- Final exam Monday – free response on new material only, multiple choice on all material from the quarter, mainly on concepts. Review sheets posted on website. Allowed 4 sheets of notes, both sides.

HOMEWORK (due Friday, March 15): Chapter 13
#14
#26 and 28 (connected, count together for 3 pts)
#54abd (no computation needed for #54)

Finishing what we planned to cover when we started Chapter 9

Five situations we will cover for the rest of this quarter:

Parameter name and description	**Population parameter**	**Sample statistic**
For Categorical Variables:		
One population proportion (or probability)	p	$\hat{p}$
Difference in two population proportions	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$
For Quantitative Variables:		
One population mean	μ	$\bar{x}$
Population mean of paired differences (dependent samples, paired)	μ_d	$\bar{d}$
Difference in two population means (independent samples)	$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$

For each situation will we:

√ Learn about the *sampling distribution* for the sample statistic

√ Learn how to find a *confidence interval* for the true value of the parameter

- *Test hypotheses* about the true value of the parameter
- *For independent samples, will see how to do in R Commander only.*

Hypothesis Testing for One Mean and Mean of Paired Differences (Same process, different notation)

Step 1: Determine the null and alternative hypotheses.

H_0: $\mu = \mu_0$ (Null value is called μ_0)

H_0: $\mu_d = 0$ (Null value is 0.)

Alternative hypothesis is *one* of these, based on context:

H_a: $\mu \neq \mu_0$

H_a: $\mu > \mu_0$

H_a: $\mu < \mu_0$

H_a: $\mu_d \neq 0$

H_a: $\mu_d > 0$

H_a: $\mu_d < 0$

Step 2: Test statistic $t = \dfrac{\text{sample statistic} - \text{null value}}{\text{(null) standard error}}$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s}$$

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{d} - 0)}{s_d}$$

STEP 3: Find the *p*-value. *This is the only thing that's new.*

Finding *p*-value for test statistic *t* (instead of *z*)

Same idea as other situations (see pictures on p. 503), but now we need to use the *t*-distribution with df = n – 1, instead of normal distribution.

Alternative hypothesis:	p-value is:
H_a: $\mu > \mu_0$ (a one-sided hypothesis)	Area above the test statistic *t*
H_a: $\mu < \mu_0$ (a one-sided hypothesis)	Area below the test statistic *t*
H_a: $\mu \neq \mu_0$ (a two-sided hypothesis)	2 × the area above $\|t\|$ = area in tails beyond $-t$ and t

Use **Table A.3 on page 671** (or can use Excel or R Commander)
p-value for test of H_a: $\mu > \mu_0$ based on a *t*-statistic
Table will provide a *p*-value *range*, not an exact *p*-value.
Example: $t = 2.20$, df = 9. Across top, $2.00 < 2.20 < 2.33$
Use df = 9 row, columns for 2.00 and 2.33: **.022 < p-value < .038**

Table A.3 One-Sided *p*-Values for Significance Tests Based on a *t*-Statistic

- A *p*-value in the table is the area to the right of *t*.
- Double the value if the alternative hypothesis is two-sided (not equal).

	Absolute Value of *t*-Statistic							
df	*1.28*	*1.50*	*1.65*	*1.80*	*2.00*	*2.33*	*2.58*	*3.00*
1	.211	.187	.173	.161	.148	.129	.118	.102
2	.164	.136	.120	.107	.092	.073	.062	.048
3	.145	.115	.099	.085	.070	.051	.041	.029
4	.135	.104	.087	.073	.058	.040	.031	.020
5	.128	.097	.080	.066	.051	.034	.025	.015
6	.124	.092	.075	.061	.046	.029	.021	.012
7	.121	.089	.071	.057	.043	.026	.018	.010
8	.118	.086	.069	.055	.040	.024	.016	.009
9	.116	.084	.067	.053	.038	.022	.015	.007
10	.115	.082	.065	.051	.037	.021	.014	.007

Steps for finding the p-value when the test statistic is *positive*:

1. Compute the test statistic t.
2. Find the value(s) in Table A.3 that t falls between (or above/below) and use those columns.
3. Go to the appropriate df row.
4. Read the p-value range from the table.
5. For H_a: parameter > null value, you're done.
6. For H_a: parameter ≠ null value, double the values.
7. For H_a: parameter < null value, you need to subtract from 1.

Steps for finding the p-value when the test statistic is *negative*:

1. Do Steps 2 to 4 above using $|t|$
2. For H_a: parameter < null value, you're done.
3. For H_a: parameter ≠ null value, double the values.
4. For H_a: parameter > null value, you need to subtract from 1.

Using the portion of the table below, do these examples:

H_a: $\mu > \mu_0$, $t = 1.90$, df = 7. Result: $.043 < p < .057$

H_a: $\mu \neq \mu_0$, $t = 1.90$, df = 7. Result: $.086 < p < .114$

H_a: $\mu < \mu_0$, $t = 1.90$, df = 7. Result: $.943 < p < .957$

H_a: $\mu < \mu_0$, $t = -1.90$, df = 7. Result: $.043 < p < .057$

	Absolute Value of *t*-Statistic							
df	*1.28*	*1.50*	*1.65*	*1.80*	*2.00*	*2.33*	*2.58*	*3.00*
1	.211	.187	.173	.161	.148	.129	.118	.102
2	.164	.136	.120	.107	.092	.073	.062	.048
3	.145	.115	.099	.085	.070	.051	.041	.029
4	.135	.104	.087	.073	.058	.040	.031	.020
5	.128	.097	.080	.066	.051	.034	.025	.015
6	.124	.092	.075	.061	.046	.029	.021	.012
7	.121	.089	.071	.057	.043	.026	.018	.010

The Rejection Region Approach (Easier!)

Rejection region is the set of test statistic values that will lead us to *reject* the null hypothesis. Refer to the bottom of **Table A.2**:

80	1.29	1.66	1.99	2.37
90	1.29	1.66	1.99	2.37
100	1.29	1.66	1.98	2.36
1000	1.282	1.646	1.962	2.330
Infinite	1.282	1.645	1.960	2.326
Two-tailed α	.20	.10	.05	.02
One-tailed α	.10	.05	.025	.01

Alternative hypothesis	Column of Table A.2	Rejection region
H_a: $\mu \neq \mu_0$	Two-tailed α	$\|t\| \geq t^*$
H_a: $\mu > \mu_0$	One-tailed α	$t \geq t^*$
H_a: $\mu < \mu_0$	One-tailed α	$t \leq -t^*$

Rejection Region Approach, continued

Use **Table A.2**. (See full table on next page.)

- Find one-tailed or two-tailed α at the *bottom* of the table.
- Find appropriate df row.
- Read t^* to form the rejection region (based on which H_a)

Examples based on segment of table below:

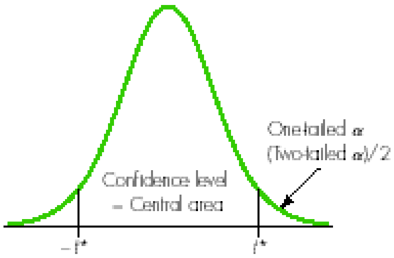
H_a: $\mu > \mu_0$, $\alpha = .05$, df = 100. Reject H_0 if test statistic $t \geq 1.66$.

H_a: $\mu < \mu_0$, $\alpha = .05$, df = 100. Rejection region is $t \leq -1.66$

H_a: $\mu \neq \mu_0$, $\alpha = .05$, df = 100. Rejection region is $t \leq -1.98$ and $t \geq 1.98$

80	1.29	1.66	1.99	2.37
90	1.29	1.66	1.99	2.37
100	1.29	1.66	1.98	2.36
1000	1.282	1.646	1.962	2.330
Infinite	1.282	1.645	1.960	2.326
Two-tailed α	.20	.10	.05	.02
One-tailed α	.10	.05	.025	.01

Table A.2 t^* Multipliers for Confidence Intervals and Rejection Region Critical Values



	Confidence Level						
df	*.80*	*.90*	*.95*	*.98*	*.99*	*.998*	*.999*
1	3.08	6.31	12.71	31.82	63.66	318.31	636.62
2	1.89	2.92	4.30	6.96	9.92	22.33	31.60
3	1.64	2.35	3.18	4.54	5.84	10.21	12.92
4	1.53	2.13	2.78	3.75	4.60	7.17	8.61
5	1.48	2.02	2.57	3.36	4.03	5.89	6.87
6	1.44	1.94	2.45	3.14	3.71	5.21	5.96
7	1.41	1.89	2.36	3.00	3.50	4.79	5.41
8	1.40	1.86	2.31	2.90	3.36	4.50	5.04
9	1.38	1.83	2.26	2.82	3.25	4.30	4.78
10	1.37	1.81	2.23	2.76	3.17	4.14	4.59
11	1.36	1.80	2.20	2.72	3.11	4.02	4.44
12	1.36	1.78	2.18	2.68	3.05	3.93	4.32
13	1.35	1.77	2.16	2.65	3.01	3.85	4.22
14	1.35	1.76	2.14	2.62	2.98	3.79	4.14
15	1.34	1.75	2.13	2.60	2.95	3.73	4.07
16	1.34	1.75	2.12	2.58	2.92	3.69	4.01
17	1.33	1.74	2.11	2.57	2.90	3.65	3.97
18	1.33	1.73	2.10	2.55	2.88	3.61	3.92
19	1.33	1.73	2.09	2.54	2.86	3.58	3.88
20	1.33	1.72	2.09	2.53	2.85	3.55	3.85
21	1.32	1.72	2.08	2.52	2.83	3.53	3.82
22	1.32	1.72	2.07	2.51	2.82	3.50	3.79
23	1.32	1.71	2.07	2.50	2.81	3.48	3.77
24	1.32	1.71	2.06	2.49	2.80	3.47	3.75
25	1.32	1.71	2.06	2.49	2.79	3.45	3.73
26	1.31	1.71	2.06	2.48	2.78	3.43	3.71
27	1.31	1.70	2.05	2.47	2.77	3.42	3.69
28	1.31	1.70	2.05	2.47	2.76	3.41	3.67
29	1.31	1.70	2.05	2.46	2.76	3.40	3.66
30	1.31	1.70	2.04	2.46	2.75	3.39	3.65
40	1.30	1.68	2.02	2.42	2.70	3.31	3.55
50	1.30	1.68	2.01	2.40	2.68	3.26	3.50
60	1.30	1.67	2.00	2.39	2.66	3.23	3.46
70	1.29	1.67	1.99	2.38	2.65	3.21	3.44
80	1.29	1.66	1.99	2.37	2.64	3.20	3.42
90	1.29	1.66	1.99	2.37	2.63	3.18	3.40
100	1.29	1.66	1.98	2.36	2.63	3.17	3.39
1000	1.282	1.646	1.962	2.330	2.581	3.098	3.300
Infinite	1.282	1.645	1.960	2.326	2.576	3.090	3.291
Two-tailed α	.20	.10	.05	.02	.01	.002	.001
One-tailed α	.10	.05	.025	.01	.005	.001	.0005

Note that the t-distribution with infinite df is the standard normal distribution.

Summary of Notation and Tables for Student's t-distribution

Confidence intervals:

- Notation for multiplier is ***t****
- Use **Table A.2** (page 670 or inside back cover) to find ***t****

Hypothesis tests:

- Notation for the test statistic is ***t*** (computed from data)
- *p-value approach*: After you compute ***t*** from the data, use **Table A.3** (page 671) to find the *p*-value range
- *Rejection region approach*
  - Compute ***t*** from the data
  - Find the rejection region of form $t \geq t^*$, etc. (form depends on H_a); use **Table A.2** to find ***t****
  - See if the *computed* test statistic ***t*** is in the rejection region

Testing One Population Mean or Paired Differences
Go through steps with an example

Example: Do people gain or lose weight when they quit smoking?
American Journal of Public Health, 1983, pgs 1303-05.

For each person, d_i = difference in weight (after – before) for people who quit smoking for 1 year. (Positive = weight *gain*)

μ_d = *population* mean weight gain in 1 year for smokers who quit.

H_0: $\mu_d = 0$ $\qquad\qquad$ H_a: $\mu_d \neq 0$

n = 322, *Sample mean* = $\bar{d}$ = 5.15 pounds,
Sample standard deviation = s_d = 11.45 pounds

$$\textit{Standard error of } \bar{d} = \frac{s_d}{\sqrt{n}} = \frac{11.45}{\sqrt{322}} = .6381$$

STEP 2:
Verify data conditions. If met, summarize data into test statistic.

Data conditions:
Bell-shaped data (no extreme outliers or skewness) or large sample.

Test statistic (remember, use *t* for means):

$$t = \frac{\text{sample statistic} - \text{null value}}{\text{(null) standard error}}$$

One population mean:

Sample statistic $= \bar{x}$
Null value $= \mu_0$

Null standard error $= \frac{s}{\sqrt{n}}$

Mean of paired differences:

Sample statistic $= \bar{d}$
Null value $= 0$

Null standard error $= \frac{s_d}{\sqrt{n}}$

Note that the word "null" is unnecessary in std. error involving means.

Step 2 for the Example:

Data conditions are met, since sample size is large (n = 322).

Population mean weight loss after quitting smoking = 0?

$$t = \frac{\text{sample statistic} - \text{null value}}{\text{(null) standard error}} = \frac{5.15 - 0}{\frac{11.45}{\sqrt{322}}} = \frac{5.15}{.6381} = 8.07$$

STEP 3:
Assuming the null hypothesis is true, find the p-value.

General: *p*-value = the *conditional* probability of a test statistic as extreme as the one observed or more so, in the direction of H_a, *if* the null hypothesis is true.

Same idea as other situations (see pictures on p. 503), but now we need to use the *t*-distribution with df = n – 1, instead of normal distribution.

Alternative hypothesis (similar for μ_d):	p-value is:
H_a: $\mu > \mu_0$ (a one-sided hypothesis)	Area above the test statistic ***t***
H_a: $\mu < \mu_0$ (a one-sided hypothesis)	Area below the test statistic ***t***
H_a: $\mu \neq \mu_0$ (a two-sided hypothesis)	2 × the area above $|t|$ = area in tails beyond –***t*** and ***t***

***p*-value for our example of paired differences**:

Weight gain/loss when quitting smoking:

- H_0: $\mu_d = 0$
  H_a: $\mu_d \neq 0$

- $t = 8.07$

- *p*-value =
  2 × P(sample mean ≥ 5.15 pounds | population mean = 0)
  2 × area above |8.07| for *t*-distribution with df = 321.
  Best we can do from **Table A.3** is *p*-value < .004 (=2 × .002)

- From Excel, *p*-value = 1.4×10^{-14}

STEP 4 – using *p*-values: Decide whether or not the result is statistically significant based on the p-value.

Example:

Paired difference, mean weight gain/loss after quitting smoking:
p-value = $1.4 \times 10^{-14} < .05$, so:

- Reject the null hypothesis.
- Accept the alternative hypothesis
- The result is statistically significant

STEP 5: Report the conclusion in the context of the situation.
Example : The mean change in weight for one year after quitting smoking is significantly different from 0.

Note: A 95% confidence interval for the mean change is weight is:

$$5.15 \pm 1.97(.638) \quad \text{or} \quad 3.89 \text{ to } 6.41 \text{ pounds.}$$

Possible problem: No control group! People gain weight as they age.

Rejection Region Method (Substitute Steps 3 and 4):

Use **Table A.2** to find rejection region
H_a: $\mu_d \neq 0$, $\alpha = .05$, df = 321. Best we can do is use df = 100.

80	1.29	1.66	1.99	2.37
90	1.29	1.66	1.99	2.37
100	1.29	1.66	1.98	2.36
1000	1.282	1.646	1.962	2.330
Infinite	1.282	1.645	1.960	2.326
Two-tailed α	.20	.10	.05	.02
One-tailed α	.10	.05	.025	.01

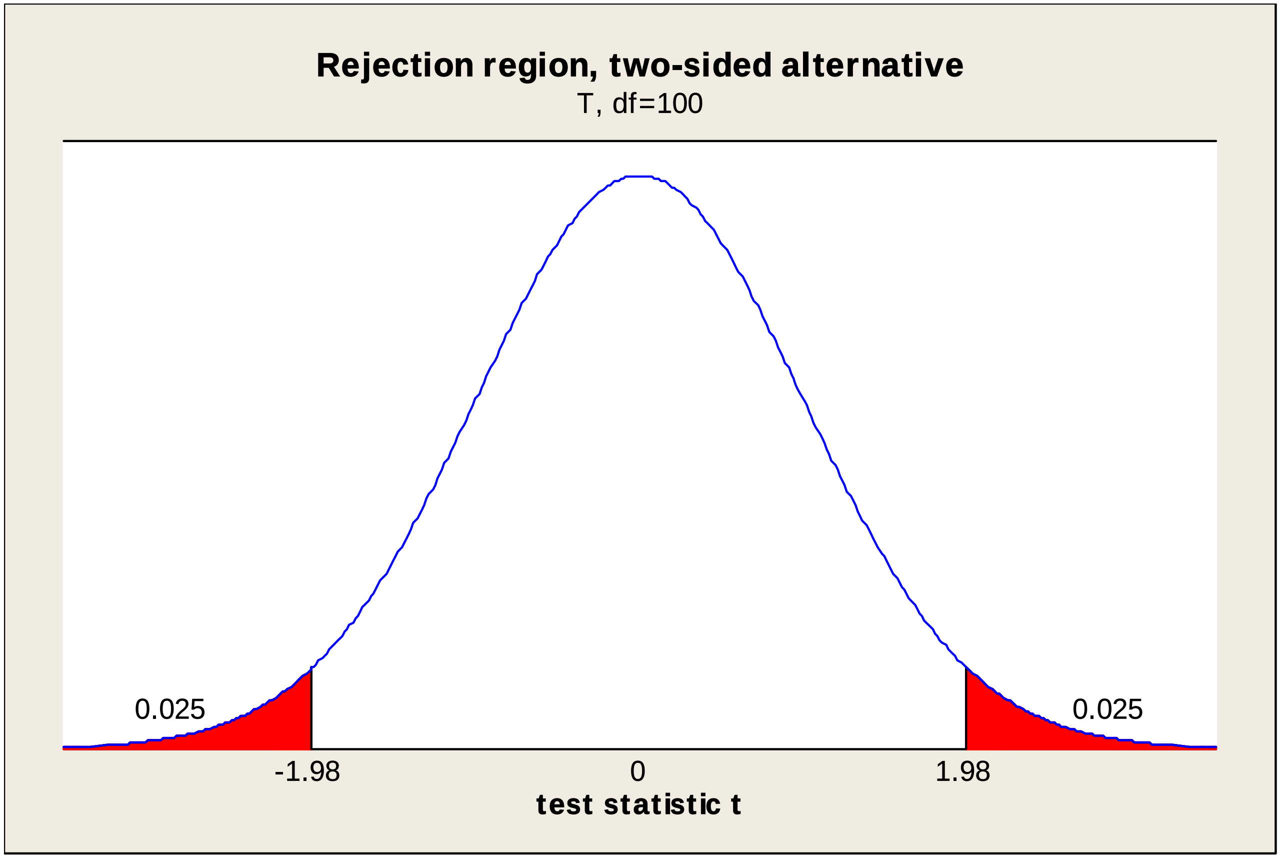
Rejection region is $t \leq -1.98$ and $t \geq 1.98$

Test statistic $t = 8.07$ *is* clearly in the rejection region!
So reject H_0.

Picture for the Example (Use $\alpha = .05$):

Weight gain or loss one year after quitting smoking, df = 321

Two-sided test H_a: $\mu_d \neq 0$, Rejection region is $|t| \geq 1.98$ (use df = 100)



Substitute Step 4: Rejection Region Approach

If the test statistic is *not* in the rejection region:

- Do not reject the null hypothesis.
- There is not enough evidence to accept the alternative hypothesis
- The result is not statistically significant

If the test statistic *is* in the rejection region:

- Reject the null hypothesis.
- Accept the alternative hypothesis
- The result is statistically significant

For our example, the test statistic is definitely in the rejection region, so we reject the null hypothesis.

Hypothesis test for difference in means, independent samples

Called a “two-sample t-test” or “independent samples t-test.” You already learned how to do this with R Commander.

Example from last time:
Two-sample t-test to compare study hours for those who drink and those who don’t. (Don’t confuse terms “two-sample” and “two-sided.”) Makes sense to do *one-sided* test.

Population 1:
Students who have or will take Stat 7, who don’t drink alcohol
μ_1 = mean hours of Stat 7 study for this *population*

Population 2:
Students who have or will take Stat7, who drink alcohol
μ_2 = mean hours of Stat 7 study for this *population*

Research question: Is the mean hours of study per week for Stat 7 higher in the *population* who don't drink alcohol than in the *population* who drink? (We *know* it's higher in the sample.)

$H_0: \mu_1 = \mu_2$ (i.e $H_0: \mu_1 - \mu_2 = 0$), $H_a: \mu_1 > \mu_2$ (i.e. $H_a: \mu_1 - \mu_2 > 0$)

- One column for Drink or Don't drink, one for Study hours
- Statistics → Means →Independent samples t-test
- Choose the alternative (>) and confidence level

```
data:  HoursStudy by Drink
t = 1.9923, df = 70.556, p-value = 0.02511
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.2817279       Inf
sample estimates:
mean in group Don't drink       mean in group Drink
                 6.500000                  4.775735
```

[Note one-sided confidence interval for difference]

p-value = .02511 < .05, so reject H_0 and conclude drinkers study fewer hours on average. C.I. gives lower bound as 0.28 hours.

New Example (CS 13.1): Work through from start to finish

Research question: Can drinking an ice slushie increase endurance when exercising in hot weather?

Australia study published in *Medicine and Science in Sports and Exercise*, 2010

- 10 Male volunteers, average age 28
- Two treatments administered to all 10 men:
  - Drink fruit-flavored ice slushie
  - Drink fruit-flavored cold water
- Then run on treadmill in 93 degree room until exhausted
- Response variable = time until exhaustion
- Order randomized, administered a few weeks apart
- Did some practice runs to eliminate "learning effect"

Possible Errors for the Example

Type 1 error:
Slushies *do not* increase endurance, but this research study concludes that they do.

Consequence: People consume slushies to increase endurance, when cold water would have worked just as well.

Type 2 error:
Slushies *do* increase endurance, but this research study fails to discover that fact.

Consequence: People miss out on finding out about something that could help them increase endurance.

Parameter of interest:
μ_d = mean difference in exhaustion times if everyone in the population were to run under both conditions.

Hypotheses:
H_0: $\mu_d = 0$ (Slushie and water have same effect on endurance)
Ha: $\mu_d > 0$ (Slushie improves endurance)

Data and Test Statistic:

$$\bar{d} = 9.5 \text{ minutes, } s_d = 3.6 \text{ minutes, so s.e.}(\bar{d}) = \frac{3.6}{\sqrt{10}} = 1.14$$

$$t = \frac{9.5 - 0}{1.14} = 8.3, \text{ df} = 9, \text{ p-value } \approx 0.$$

Reject H_0, conclude ice slushie *does* increase endurance compared to drinking cold water.
(Two-sided) 95% confidence interval: Multiplier ***t** = 2.26**
9.5 ± 2.26(1.14) or <u>6.9 to 12.1</u> minutes more, on average.