Today: Sections 13.1 to 13.3

ANNOUNCEMENTS:

- We will finish hypothesis testing for the 5 situations today. See **pages 586-587** (end of Chapter 13) for a **summary table**.
- Quiz for week 8 starts Wed, ends *Monday* at noon

HOMEWORK (due Monday, Nov 29):

Chapter 13: #15, 24ac, 25 (partial answer in back)

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.*

Five steps to hypothesis testing – one mean and mean of paired difference: ***Summary Boxes** on **pages 558-559 and 562.***

STEP 1: Determine the null and alternative hypotheses.

One population mean:	Population mean of paired differences
Null hypothesis: H_0: $\mu = \mu_0$	Null hypothesis: H_0: $\mu_d = 0$
Null value is called μ_0	Null value = 0 (*Note special null value*)

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

H_a: $\mu \neq \mu_0$	H_a: $\mu_d \neq 0$
H_a: $\mu > \mu_0$	H_a: $\mu_d > 0$
H_a: $\mu < \mu_0$	H_a: $\mu_d < 0$

Example for testing one population mean:

Is mean human body temperature really 98.6 degrees, or is it lower?

H_0: $\mu = 98.6$ degrees
H_a: $\mu < 98.6$ degrees

n = 101 blood donors at clinic near Seattle, ages 17 to 84

Sample mean = $\bar{x}$ = 97.89 degrees,
Sample standard deviation = s = 0.73 degrees

Standard error = s.e.($\bar{x}$) = $\frac{s}{\sqrt{n}} = \frac{0.73}{\sqrt{101}} = 0.073$

Example for testing population mean of paired differences:

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: $\mu_d = 0$
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 Examples:

Data conditions are met, since both sample sizes are large.

Example for one mean
(Population mean body temperature = 98.6?):

$$t = \frac{\text{sample statistic} - \text{null value}}{\text{(null) standard error}} = \frac{97.89 - 98.6}{\frac{.73}{\sqrt{101}}} = \frac{-.71}{.0726} = -9.77$$

Example for mean of paired differences
(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. 517), 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

Use **Table A.3 on page 729:**

One-Sided *p*-values for Significance Tests Based on a *t*-Statistic

Table will provide a *p*-value *range*, not an exact *p*-value.

Can also use Excel or R Commander.

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
11	.113	.081	.064	.050	.035	.020	.013	.006
12	.112	.080	.062	.049	.034	.019	.012	.006
13	.111	.079	.061	.048	.033	.018	.011	.005
14	.111	.078	.061	.047	.033	.018	.011	.005
15	.110	.077	.060	.046	.032	.017	.010	.004
16	.109	.077	.059	.045	.031	.017	.010	.004
17	.109	.076	.059	.045	.031	.016	.010	.004
18	.108	.075	.058	.044	.030	.016	.009	.004
19	.108	.075	.058	.044	.030	.015	.009	.004
20	.108	.075	.057	.043	.030	.015	.009	.004

Ex: n = 15, df = 14, t = 2.20

$H_a: \mu > \mu_0$

p-value = area above 2.20

Since 2.20 is between 2.00 and 2.33, *p*-value is between .033 and .018:

$.018 < p\text{-value} < .033$

Double it for two-sided:

$H_a: \mu \neq \mu_0$

$.036 < p\text{-value} < .066$

Use with negative values for $H_a: \mu < \mu_0$

Area above 2.20, df = 14

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
11	.113	.081	.064	.050	.035	.020	.013	.006
12	.112	.080	.062	.049	.034	.019	.012	.006
13	.111	.079	.061	.048	.033	.018	.011	.005
14	.111	.078	.061	.047	.033	.018	.011	.005
15	.110	.077	.060	.046	.032	.017	.010	.004
16	.109	.077	.059	.045	.031	.017	.010	.004
17	.109	.076	.059	.045	.031	.016	.010	.004
18	.108	.075	.058	.044	.030	.016	.009	.004
19	.108	.075	.058	.044	.030	.015	.009	.004
20	.108	.075	.057	.043	.030	.015	.009	.004

***p*-value for our two examples**:

Example for one mean (normal body temperature):
H_a: $\mu < 98.6$
$t = -9.77$
p-value = area *below* $t = -9.77$ for df = 100
Best we can do from Table A.3 is *p*-value < .002.
From Excel, *p*-value = 1.6×10^{-16}

Example for paired differences
(weight gain/loss when quitting smoking):
H_a: $\mu_d \neq 0$
$t = 8.07$
p-value = 2 × area above |8.07| for df = 321.
Best we can do from Table A.3 is *p*-value < .004 (take 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.

Examples:
Mean body temperature:
p-value = $1.6 \times 10^{-16} < .05$, so:
- Reject the null hypothesis.
- Accept the alternative hypothesis
- The result is statistically significant

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

For tests involving the *t*-distribution, there is a **Substitute Step 3 and 4**, called the **Rejection Region Approach**.

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

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

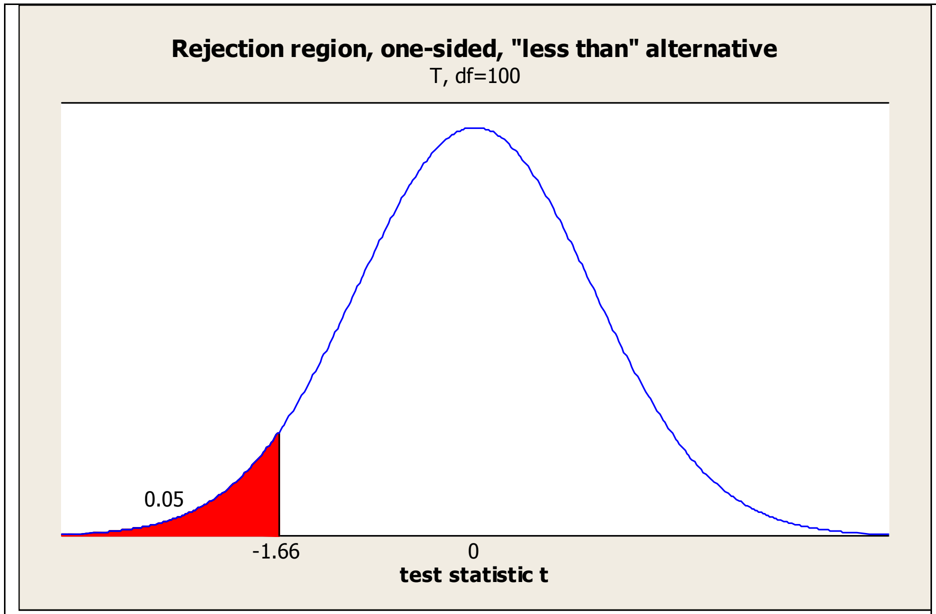
80	1.29	1.66	1.99	2.37	2.64	3.20	3.42
90	1.29	1.66	1.99	2.37	2.63	3.18	3.40
100	1.29	1.66	1.98	2.36	2.63	3.17	3.39
1000	1.282	1.646	1.962	2.330	2.581	3.098	3.300
Infinite	1.281	1.645	1.960	2.326	2.576	3.090	3.291
Two-tailed α	.20	.10	.05	.02	.01	.002	.001
One-tailed α	.10	.05	.025	.01	.005	.001	.0005

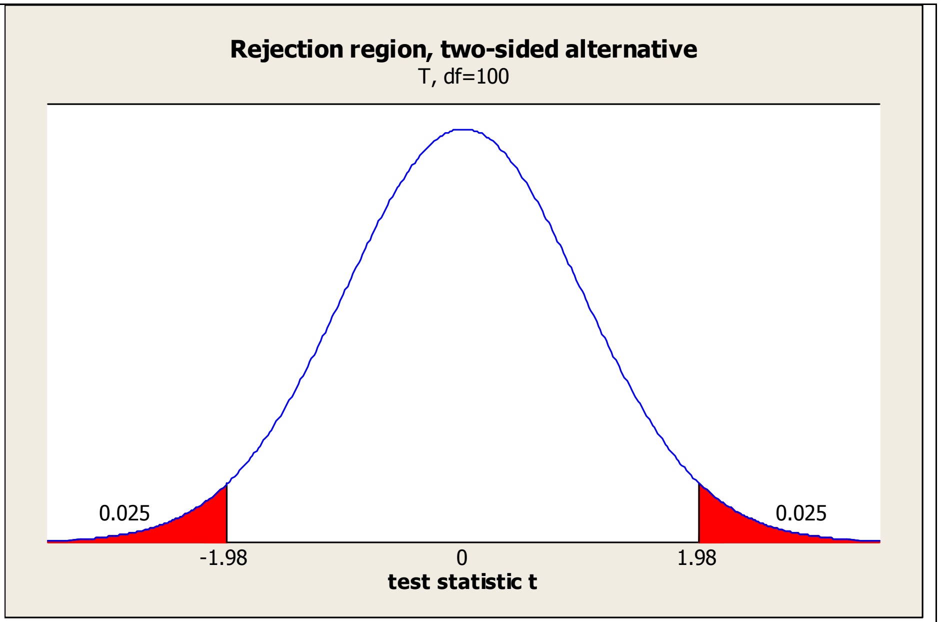
Note that the *t*-distribution with infinite df is the standard normal distribution.

Examples (Use $\alpha = .05$):

Mean body temperature, n = 101, df = 100
One-sided test H_a: $\mu < 98.6$, Rejection region is $t \leq -1.66$

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 both examples, the test statistic is definitely in the rejection region, so we reject the null hypothesis.

Step 5: Report the conclusion in the context of the situation.

Example 1:
The mean body temperature for healthy human adults is less than 98.6 degrees.

Note: We found a 95% confidence interval for this in an earlier lecture. It was 97.75 to 98.03 degrees.

Example 2:
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.

Hypothesis test for difference in two 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 Exercise 11.51: Two-sample t-test to compare pulse for those who do and don't exercise

- Data → New data set – give name, enter data
- One column for Exercise (Y,N) and one column for pulse
- Statistics → Means →Independent samples t-test
- Choose the alternative (≠,>,<) and conf. level

```
data:  Pulse by Exercises
t = 1.7156, df = 13.014, p-value = 0.05496
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.727387 15.060720
sample estimates:
mean in group N mean in group Y
       72.00000        65.33333
```

New Example: 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”

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.

95% confidence interval is $9.5 \pm 2.26(1.14)$ or <u>6.9 to 12.1</u> mins.