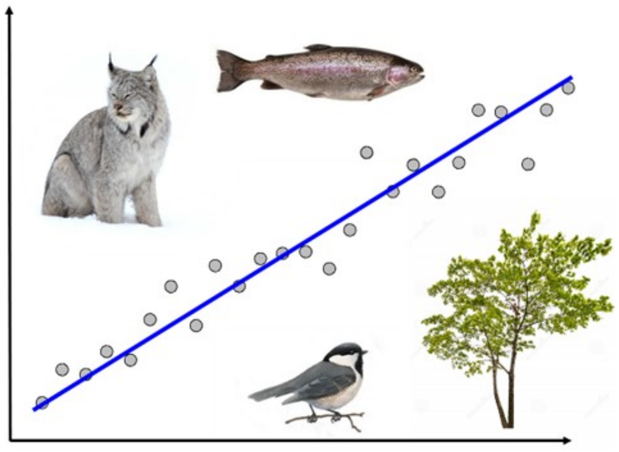
NRC 290b Introduction to Quantitative Ecology

Week 7 – Tests for differences



Meg Graham MacLean, PhD

Department of Environmental Conservation

mgmaclean@umass.edu

2019 - Fall

This week

Tuesday (Monday)

- MAP!
- Share our plots with structured feedback
- Finish reports and plots with edits

Wednesday

- Difference tests!
    - Test for normally distributed data
        - t-test/paired t-test
    - Tests for skewed data
        - U-test
        - Wilcoxon matched-pairs test

Plot feedback

What your group should present (no longer than 3 minutes):

- Your question and your hypotheses
- The year of your data
- What your plot visualizes

Feedback:

- What did we learn from the plot (and what about the plot helped us learn it)?
- What one thing could be improved to better visualize the answer to the question?

Week 7 – Tests for differences

Part II - Wednesday

Statistical testing

Which statistical test is right if you are looking for differences between two samples where you have less than 30 paired samples and the samples are normally distributed?

a) Kruskal-Wallis Test for multiple samples
b) Z-test for matched pairs
c) T-test for matched pairs
d) Mann-Whitney U-Test

Statistical testing

Using a t-test to test the difference between two samples, you calculate a *p*-value of 0.02. If you using a 5% significance (alpha) level – what do you conclude?

a) You reject the null hypothesis

b) You accept the null hypothesis

c) You don't reject the null hypothesis

d) You reject the alternative hypothesis

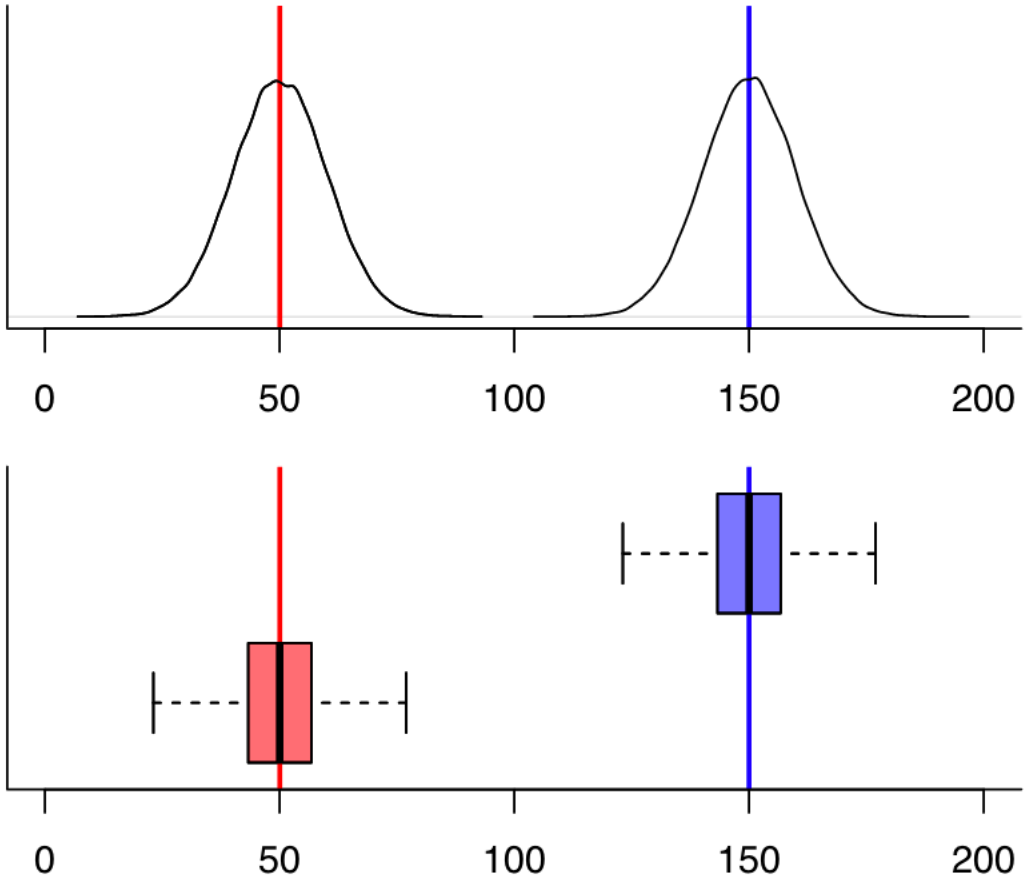
e) You don't reject the alternative hypothesis

Why do we conduct difference tests?

Tests for differences

Ask the questions:

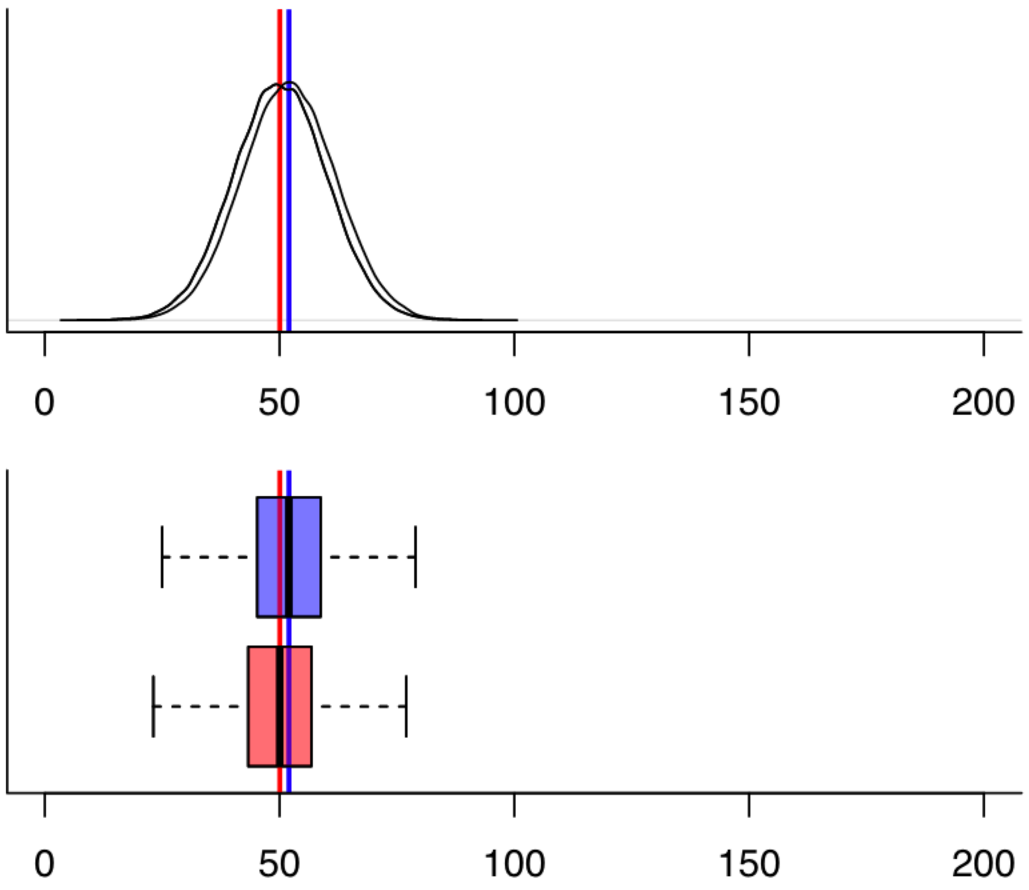
- Are the sample means different?
- Are the sample medians different?
- Are those differences statistically significant?



Tests for differences

Ask the questions:

- Are the sample means different?
- Are the sample medians different?
- Are those differences statistically significant?



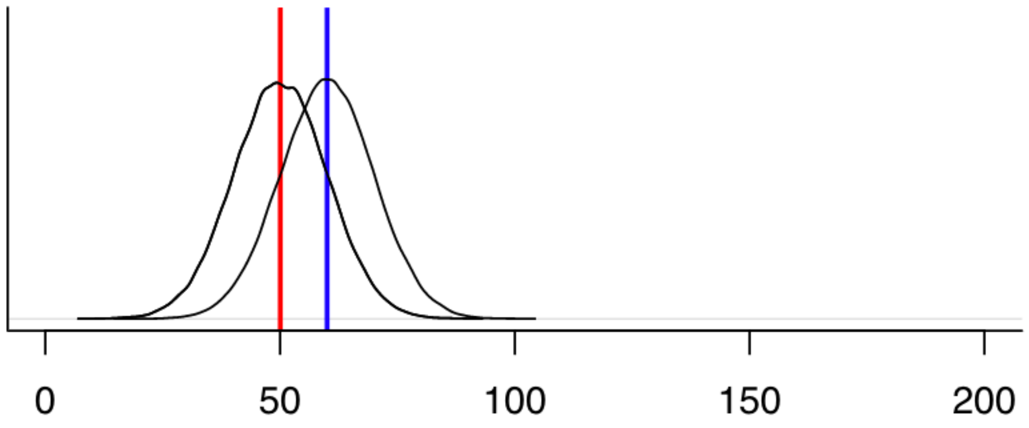
Tests for differences

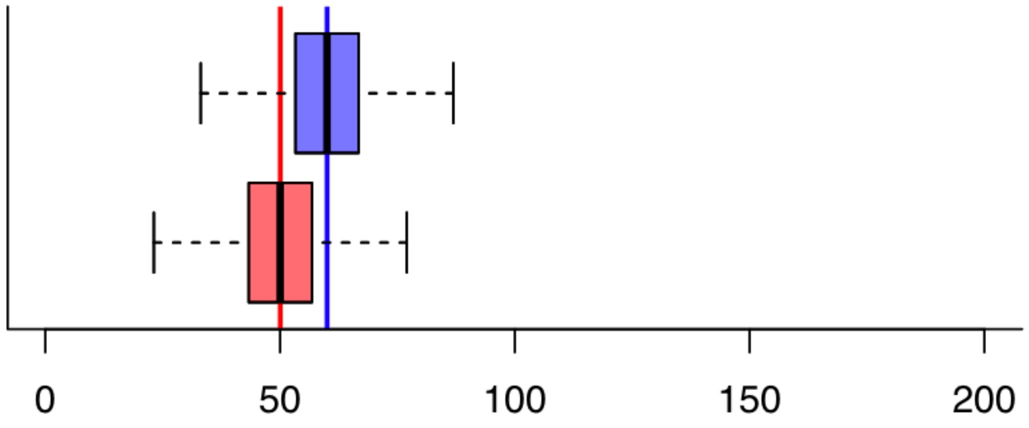
Ask the questions:

- Are the sample means different?
- Are the sample medians different?
- Are those differences statistically significant?

Statistical tests:

- T-test (normally distributed data)*
- U-test (skewed data)





Student’s t-test

Compares the means of two samples where H_0: $\bar{x}_a = \bar{x}_b$

$$t = \frac{|\bar{x}_a - \bar{x}_b|}{\sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}}}$$

Assumptions:

- Both samples are randomly collected, normally distributed, and independent
- Both samples have equal variances (i.e. homogeneity of variance)
- BUT the t-test is robust to slight deviations from these assumptions!

T-test

Answer the two questions using the t-test equation:

$$t = \frac{|\bar{x}_a - \bar{x}_b|}{\sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}}}$$

If the absolute value of the difference in means is large, then *t* is…

If the sum of the variances is large, then *t* is…

a) Large, large
b) Large, small
c) Small, large
d) Small, small

T-test

- Critical value of t is determined by your defined alpha (significance value) and degrees of freedom

$$df = (n_a - 1) + (n_b - 1)$$

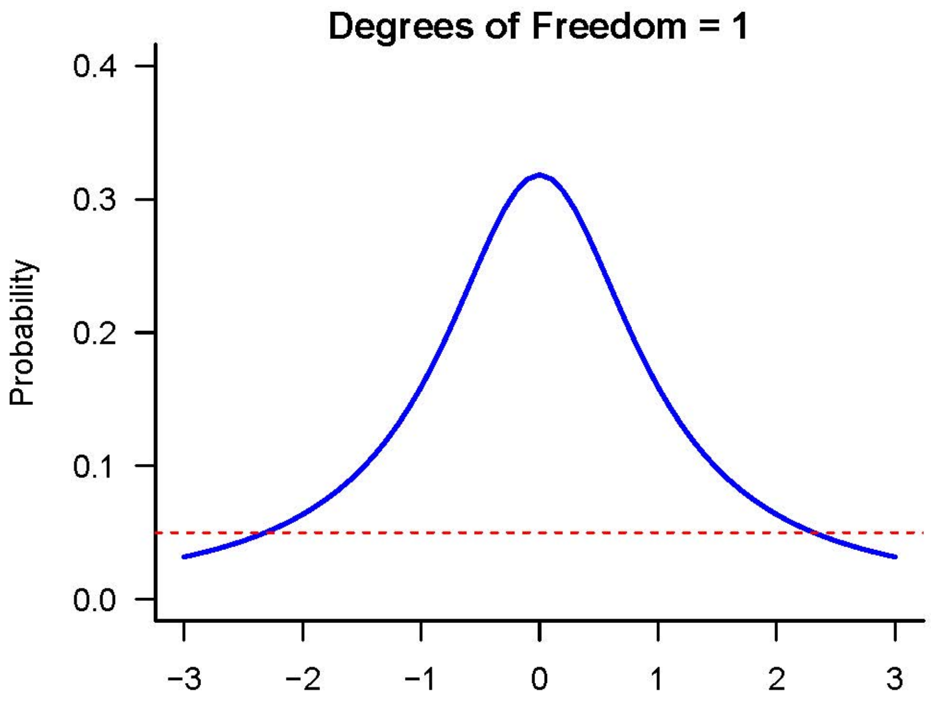


Table A-2 The *t*-Distribution

$df \backslash \alpha$	0.10	0.05	0.025	0.01	0.005
	0.20	0.10	0.05	0.02	0.01
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
11	1.36	1.80	2.20	2.72	3.11
12	1.36	1.78	2.18	2.68	3.05
13	1.35	1.77	2.16	2.65	3.01
14	1.35	1.76	2.14	2.62	2.98
15	1.34	1.75	2.13	2.60	2.95
16	1.34	1.75	2.12	2.58	2.92
17	1.33	1.74	2.11	2.57	2.90
18	1.33	1.73	2.10	2.55	2.88
19	1.33	1.73	2.09	2.54	2.86
20	1.33	1.72	2.09	2.53	2.85
21	1.32	1.72	2.08	2.52	2.83
22	1.32	1.72	2.07	2.51	2.82
23	1.32	1.71	2.07	2.50	2.81
24	1.32	1.71	2.06	2.49	2.80
25	1.32	1.71	2.06	2.49	2.79
26	1.31	1.71	2.06	2.48	2.78
27	1.31	1.70	2.05	2.47	2.77
28	1.31	1.70	2.05	2.47	2.76
29	1.31	1.70	2.05	2.46	2.76
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
60	1.30	1.67	2.00	2.39	2.66
120	1.29	1.66	1.98	2.36	2.62
∞	1.28	1.64	1.96	2.33	2.58

T-test

- Critical value of t is determined by your defined alpha (significance value) and degrees of freedom

$$df = (n_a - 1) + (n_b - 1)$$

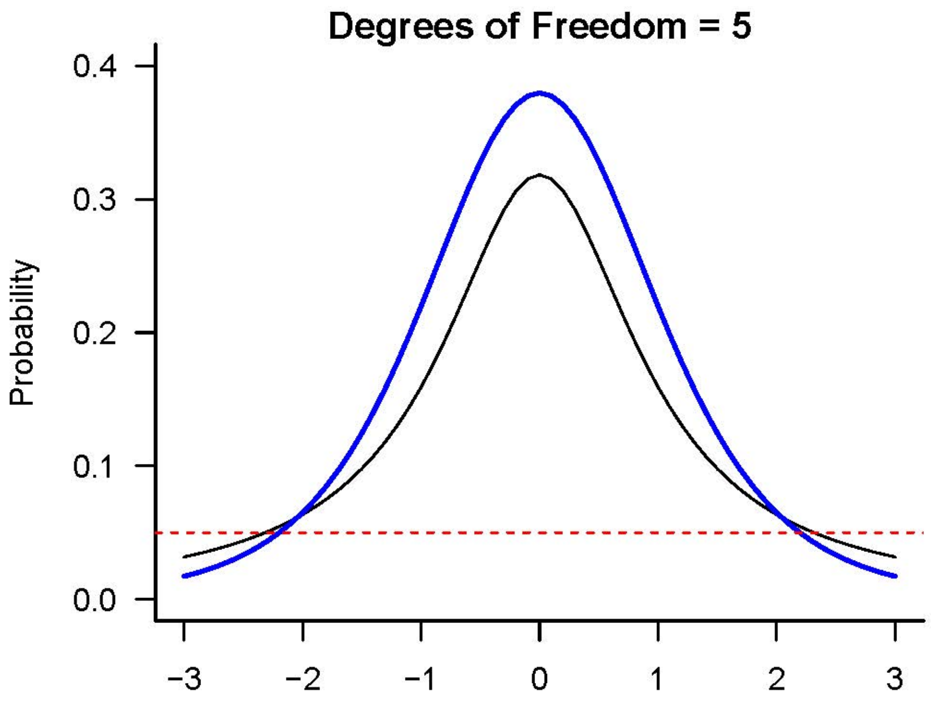


Table A-2 The *t*-Distribution

$df\backslash\alpha$	0.10	0.05	0.025	0.01	0.005
	0.20	0.10	0.05	0.02	0.01
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
11	1.36	1.80	2.20	2.72	3.11
12	1.36	1.78	2.18	2.68	3.05
13	1.35	1.77	2.16	2.65	3.01
14	1.35	1.76	2.14	2.62	2.98
15	1.34	1.75	2.13	2.60	2.95
16	1.34	1.75	2.12	2.58	2.92
17	1.33	1.74	2.11	2.57	2.90
18	1.33	1.73	2.10	2.55	2.88
19	1.33	1.73	2.09	2.54	2.86
20	1.33	1.72	2.09	2.53	2.85
21	1.32	1.72	2.08	2.52	2.83
22	1.32	1.72	2.07	2.51	2.82
23	1.32	1.71	2.07	2.50	2.81
24	1.32	1.71	2.06	2.49	2.80
25	1.32	1.71	2.06	2.49	2.79
26	1.31	1.71	2.06	2.48	2.78
27	1.31	1.70	2.05	2.47	2.77
28	1.31	1.70	2.05	2.47	2.76
29	1.31	1.70	2.05	2.46	2.76
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
60	1.30	1.67	2.00	2.39	2.66
120	1.29	1.66	1.98	2.36	2.62
∞	1.28	1.64	1.96	2.33	2.58

T-test

- Critical value of t is determined by your defined alpha (significance value) and degrees of freedom

$$df = (n_a - 1) + (n_b - 1)$$

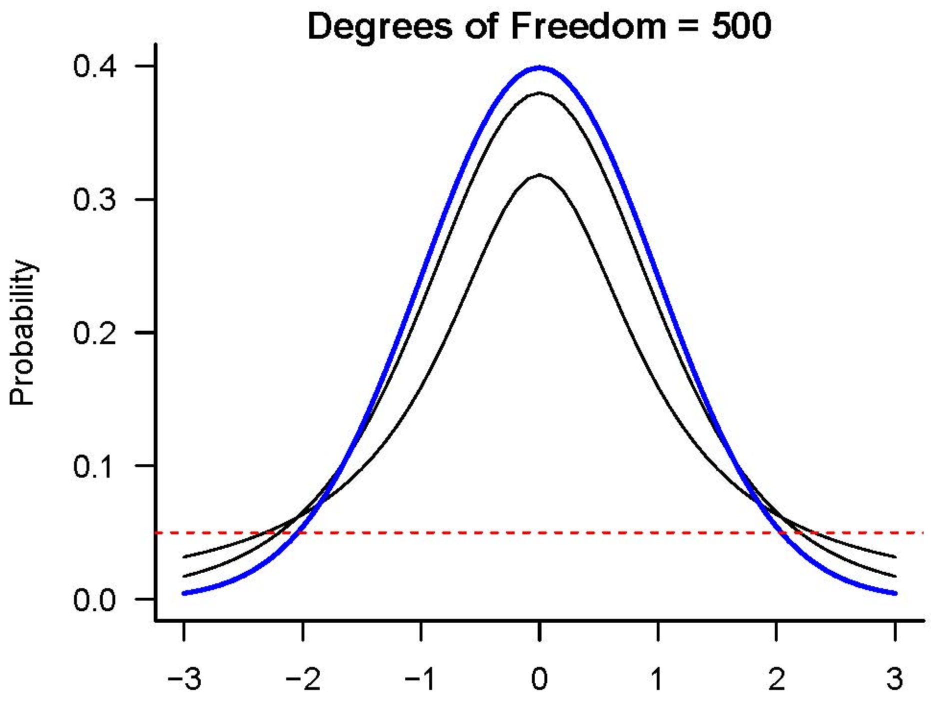


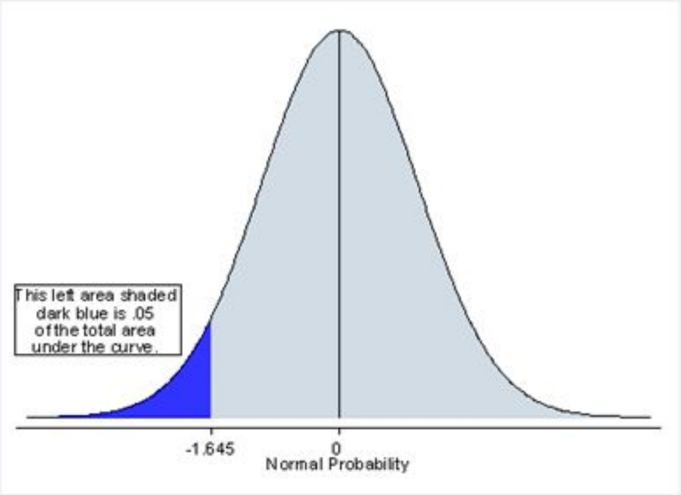
Table A-2 The *t*-Distribution

$df \setminus \alpha$	0.10	0.05	0.025	0.01	0.005
	0.20	0.10	0.05	0.02	0.01
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
11	1.36	1.80	2.20	2.72	3.11
12	1.36	1.78	2.18	2.68	3.05
13	1.35	1.77	2.16	2.65	3.01
14	1.35	1.76	2.14	2.62	2.98
15	1.34	1.75	2.13	2.60	2.95
16	1.34	1.75	2.12	2.58	2.92
17	1.33	1.74	2.11	2.57	2.90
18	1.33	1.73	2.10	2.55	2.88
19	1.33	1.73	2.09	2.54	2.86
20	1.33	1.72	2.09	2.53	2.85
21	1.32	1.72	2.08	2.52	2.83
22	1.32	1.72	2.07	2.51	2.82
23	1.32	1.71	2.07	2.50	2.81
24	1.32	1.71	2.06	2.49	2.80
25	1.32	1.71	2.06	2.49	2.79
26	1.31	1.71	2.06	2.48	2.78
27	1.31	1.70	2.05	2.47	2.77
28	1.31	1.70	2.05	2.47	2.76
29	1.31	1.70	2.05	2.46	2.76
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
60	1.30	1.67	2.00	2.39	2.66
120	1.29	1.66	1.98	2.36	2.62
∞	1.28	1.64	1.96	2.33	2.58

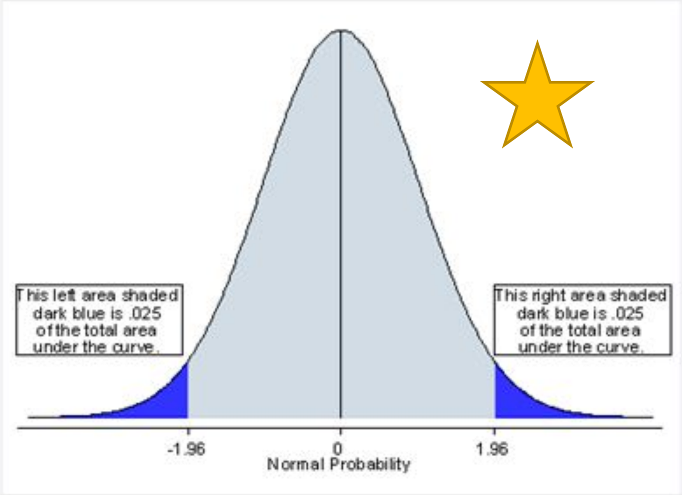
T-test

- Critical value of t is determined by your defined alpha (significance value) and degrees of freedom

$$df = (n_a - 1) + (n_b - 1)$$



A one-tailed test will test either if the mean is significantly greater than x or if the mean is significantly less than x, but not both. The one-tailed test provides more power to detect an effect in one direction by not testing the effect in the other direction.



A two-tailed test will test both if the mean is significantly greater than x and if the mean significantly less than x. The mean is considered significantly different from x if the test statistic is in the top 2.5% or bottom 2.5% of its probability distribution, resulting in a p-value less than 0.05.

Table A-2 The t-Distribution

df \ α	0.10	0.05	0.025	0.01	0.005
	0.20	0.10	0.05	0.02	0.01
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
11	1.36	1.80	2.20	2.72	3.11
12	1.36	1.78	2.18	2.68	3.05
13	1.35	1.77	2.16	2.65	3.01
14	1.35	1.76	2.14	2.62	2.98
15	1.34	1.75	2.13	2.60	2.95
16	1.34	1.75	2.12	2.58	2.92
17	1.33	1.74	2.11	2.57	2.90
18	1.33	1.73	2.10	2.55	2.88
19	1.33	1.73	2.09	2.54	2.86
20	1.33	1.72	2.09	2.53	2.85
21	1.32	1.72	2.08	2.52	2.83
22	1.32	1.72	2.07	2.51	2.82
23	1.32	1.71	2.07	2.50	2.81
24	1.32	1.71	2.06	2.49	2.80
25	1.32	1.71	2.06	2.49	2.79
26	1.31	1.71	2.06	2.48	2.78
27	1.31	1.70	2.05	2.47	2.77
28	1.31	1.70	2.05	2.47	2.76
29	1.31	1.70	2.05	2.46	2.76
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
60	1.30	1.67	2.00	2.39	2.66
120	1.29	1.66	1.98	2.36	2.62
∞	1.28	1.64	1.96	2.33	2.58

T-test

- Critical value of *t* is determined by your defined alpha (significance value) and degrees of freedom

$$df = (n_a - 1) + (n_b - 1)$$

- If *t* is **higher** than the given value in the table at your alpha and df, you **reject** the null hypothesis!

Special case: when you have a large sample size (n > 30) you use a very similar test, the z-test! Not discussed in this book.

Table A-2 The *t*-Distribution

$df \backslash \alpha$	*0.10*	*0.05*	*0.025*	*0.01*	*0.005*
	0.20	*0.10*	*0.05*	*0.02*	*0.01*
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
11	1.36	1.80	2.20	2.72	3.11
12	1.36	1.78	2.18	2.68	3.05
13	1.35	1.77	2.16	2.65	3.01
14	1.35	1.76	2.14	2.62	2.98
15	1.34	1.75	2.13	2.60	2.95
16	1.34	1.75	2.12	2.58	2.92
17	1.33	1.74	2.11	2.57	2.90
18	1.33	1.73	2.10	2.55	2.88
19	1.33	1.73	2.09	2.54	2.86
20	1.33	1.72	2.09	2.53	2.85
21	1.32	1.72	2.08	2.52	2.83
22	1.32	1.72	2.07	2.51	2.82
23	1.32	1.71	2.07	2.50	2.81
24	1.32	1.71	2.06	2.49	2.80
25	1.32	1.71	2.06	2.49	2.79
26	1.31	1.71	2.06	2.48	2.78
27	1.31	1.70	2.05	2.47	2.77
28	1.31	1.70	2.05	2.47	2.76
29	1.31	1.70	2.05	2.46	2.76
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
60	1.30	1.67	2.00	2.39	2.66
120	1.29	1.66	1.98	2.36	2.62
∞	1.28	1.64	1.96	2.33	2.58

Paired t-test

What if your samples are paired?

Compares the means of two samples that are NOT independent (e.g., before/after study)

$$t = \frac{\bar{D}}{\sqrt{\frac{s_D^2}{n}}}$$

$\bar{D}$ is the mean of the differences

s_D is the standard deviation of the differences

n is the number of paired samples

Assumptions:

- Both samples are normally distributed
- Both samples have equal variances/sd (i.e. homogeneity of variance)
- Sample sizes must be exactly the same!

Paired t-test

Let's say you have a group of tree frogs you gave a vaccine to and you are trying to see if there is a difference in mucus production of those frogs before and after the vaccine… however, one of your treated frogs was eaten by a predator between your first measurement and your second – what do you do with the initial mucus value you recorded for this frog when conducting your paired t-test?

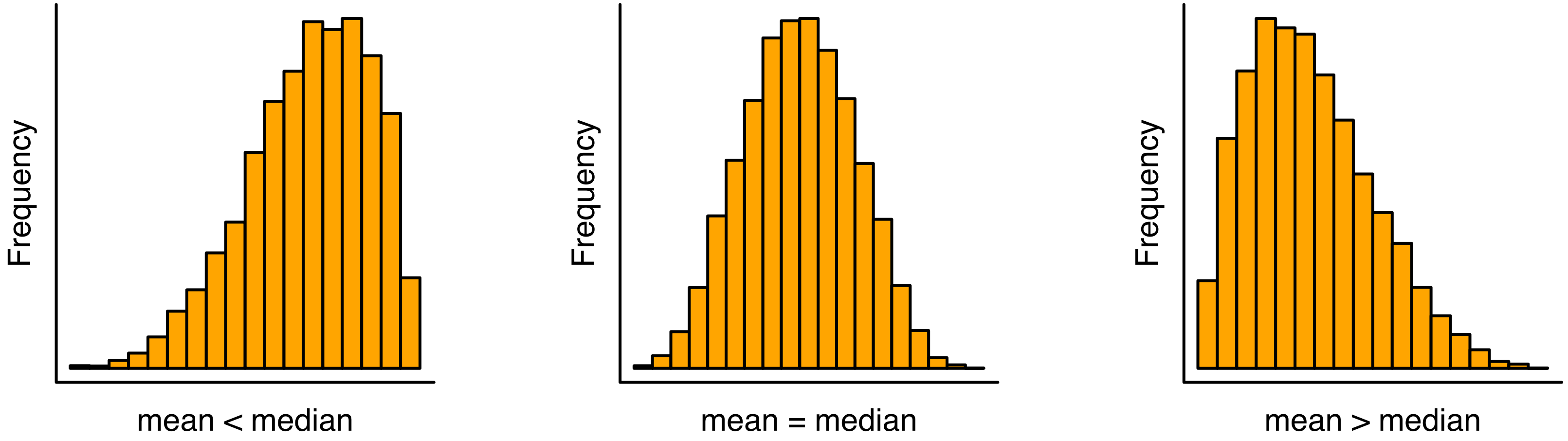
a) Get rid of it

b) Use it, no problem there

c) Use it and duplicate another similar frogs' second reading

d) Use it for both before and after treatment

What happens when our data aren’t normally distributed?



Difference tests for skewed data

- U-test (for unpaired samples)
  - Like the t-test but uses ranks, median, and range to determine how much overlap there is between samples
  - LOWER U-values are more significant (opposite from t-test)
- Wilcoxon matched-pairs test (for paired samples)
  - Based on ranked differences (calculated differences first, then rank them)
  - Again, lower W-values are more likely to be significant

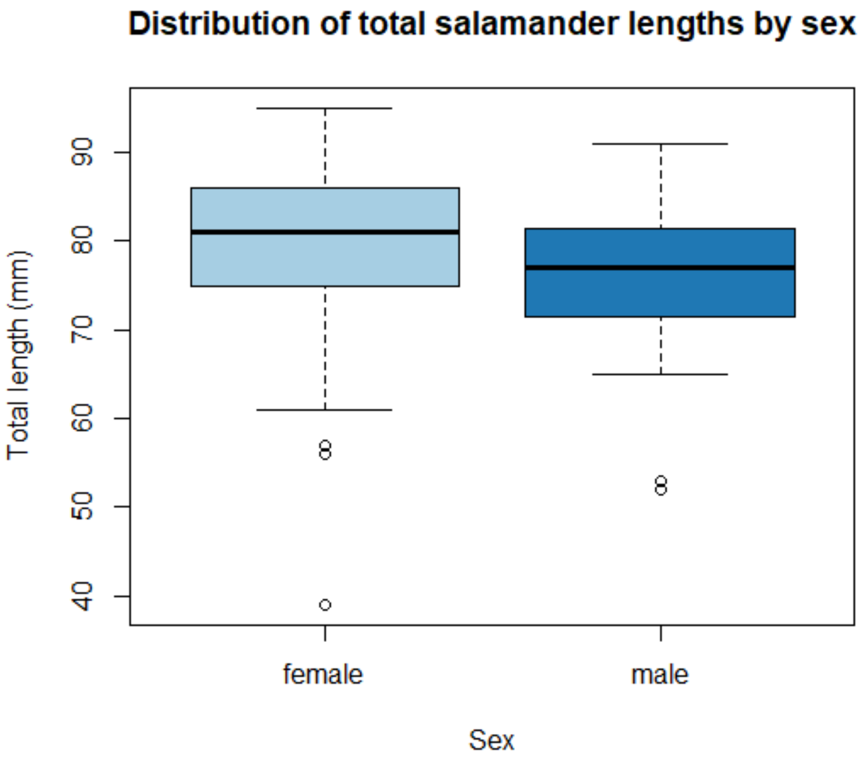
Today's Exercise

Back to the redbacked salamander (*Plethodon cinereus)* data! We're trying to answer the question:

Is the average length of salamanders significantly different between salamanders identified as male or female?

To do so:

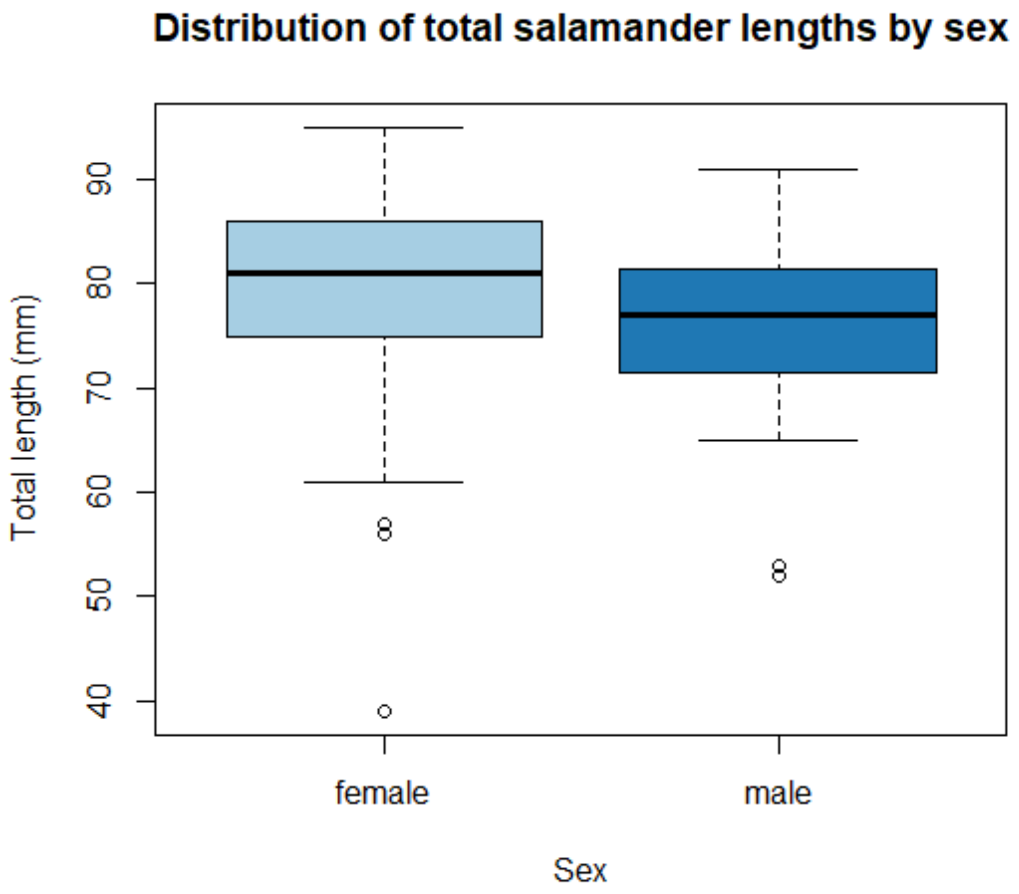
1. Download my code from moodle labeled mander_sex_differences.R
2. Go through the code and add comments each place you see a hashtag and code where it needs code. The comments should tell me:
   1. What each section of of code does
   2. Answer the questions outlined in the code
3. Upload your ***individual*** code to moodle



Quick aside – outliers?

In R (and many other places), data points that are >1.5x the interquartile range (IQR) from the .25 and .75 quartiles are identified as *possible* outliers (something to explore)

- Plotted here as the open circles
- Identifying data points as possible outliers doesn't mean they are "bad data"!



For Monday:

1) Read Ch 10 (Skip 10.2) in Gardner (2017) – Differences between more than 2 samples
2) Answer the individual evaluation questions on moodle

All before 11:55pm on Sunday