<u>We will look at the following t-tests:</u>

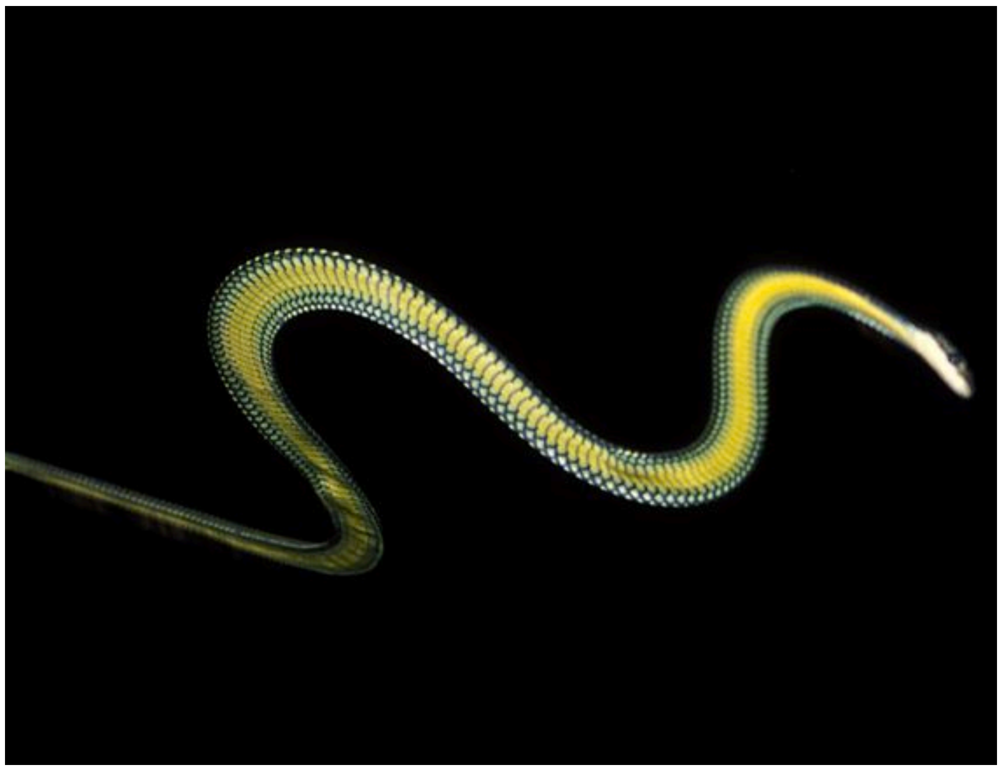
1. Comparing one mean: one-sample t-test
2. Comparing two means:
   a. Paired t-test (which is similar to a one-sample t-test but focused on the difference between two sampels)
   b. Two-sample t-test

Note: All of the above tests have slightly different assumptions which allow our conclusions to be supported. We will investigate what happens when these assumptions are violated and how robust our t-tests are to violations.

Applications of one sample t-test

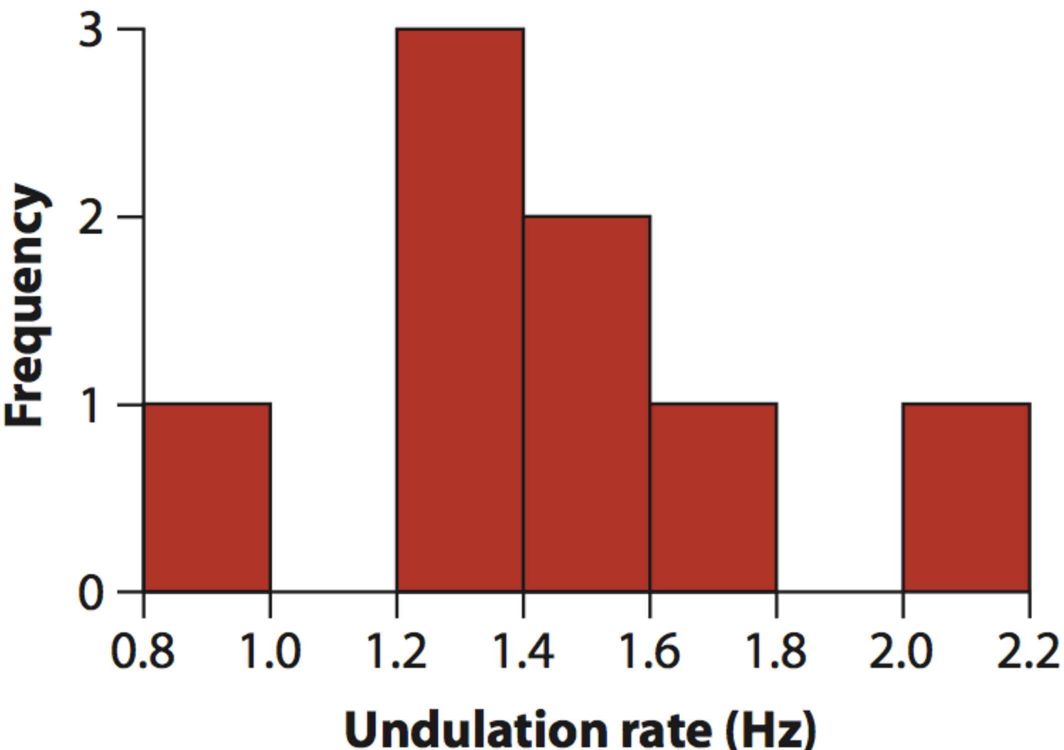
Example: What is the 95% confidence interval for the mean undulation rate of the paradise flying snakes. The rate of undulation (hz) from 8 snakes:

0.9,1.4,1.2,1.2,1.3,2.0,1.4,1.6



Example: What is the 95% confidence interval for the mean undulation rate of the paradise flying snakes. The rate of undulation (hz) from 8 snakes:

0.9,1.4,1.2,1.2,1.3,2.0,1.4,1.6



$$\bar{Y} = 1.375$$

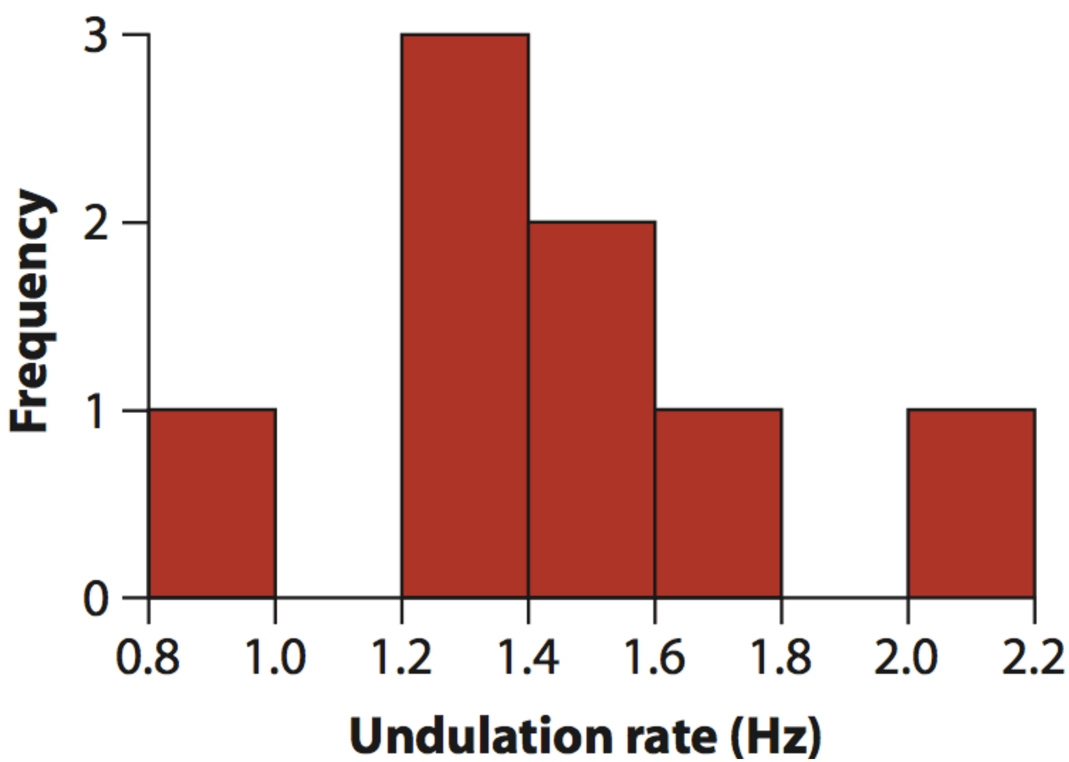
$$s = 0.324$$

$$n = 8$$

<u>Example</u>: What is the 95% confidence interval for the mean undulation rate of the paradise flying snakes. The rate of undulation (hz) from 8 snakes:

0.9,1.4,1.2,1.2,1.3,2.0,1.4,1.6

____



dof = n - 1 = 7

$$t_{\alpha(2),df} = t_{0.05(2),7} = 2.36$$

Inference for a Normal Population

<u>Example</u>: What is the 95% confidence interval for the mean undulation rate of the paradise flying snakes. The rate of undulation (hz) from 8 snakes:

0.9,1.4,1.2,1.2,1.3,2.0,1.4,1.6

	$\alpha(2)$:	0.2	0.10	0.05	0.02	0.01	0.001	0.0001
df	$\alpha(1)$:	0.1	0.05	0.025	0.01	0.005	0.0005	0.00005
1		3.08	6.31	12.71	31.82	63.66	636.62	6366.20
2		1.89	2.92	4.30	6.96	9.92	31.60	99.99
3		1.64	2.35	3.18	4.54	5.84	12.92	28.00
4		1.53	2.13	2.78	3.75	4.60	8.61	15.54
5		1.48	2.02	2.57	3.36	4.03	6.87	11.18
6		1.44	1.94	2.45	3.14	3.71	5.96	9.08
7		1.41	1.89	2.36	3.00	3.50	5.41	7.88
8		1.40	1.86	2.31	2.90	3.36	5.04	7.12
9		1.38	1.83	2.26	2.82	3.25	4.78	6.59

Example: What is the 95% confidence interval for the mean undulation rate of the paradise flying snakes. The rate of undulation (hz) from 8 snakes:

0.9,1.4,1.2,1.2,1.3,2.0,1.4,1.6

Answer:

$$\bar{Y} \pm t_{0.05(2),7} SE_{\bar{Y}} = 1.375 \pm 0.115(2.36)$$

$\boxed{1.10 < \mu < 1.65}$ (95% Confidence Interval)

$$\bar{Y} \pm t_{0.01(2),7} SE_{\bar{Y}} = 1.375 \pm 0.115(3.50)$$

$\boxed{0.97 < \mu < 1.78}$ (99% Confidence Interval)

The one-sample t test:

Compares the mean of a random sample from a normal population with the population mean proposed in a null hypothesis

- H_0: True mean equals μ_0
  H_A: True mean *does not* equal μ_0

- Assumptions:
  - The variable is normally distributed
  - The sample is a random sample

HOW WOULD YOU TEST THE FOLLOWING:

Do Rochester high school seniors who attend a summer math camp score above the state mean on the math subtest of the state's standardized achievement test ? (n=15)

a. Test one mean against a hypothesized constant.

b. Test the difference between two means (independent samples).

c. Test the difference in means between two paired or dependent samples.

d. Use a chi-squared test of association.

The one-sample t test:

- H_0: True mean equals μ_0
  H_A: True mean *does not* equal μ_0

- Test Statistic:

$$\boxed{t = \frac{\bar{Y} - \mu_0}{SE_{\bar{Y}}} = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}}}$$

Example: The consequences of mutation on fitness after 100 generations in $\phi 6$ using 5 lines:

0.063, -0.062, 0.064, -0.043, 1.34

Has the mean fitness changed?

- <u>Example:</u>The consequences of mutation on fitness after 100 generations in ϕ6 using 5 lines:

  0.063, -0.062, 0.064, -0.043, 1.34

  Has the mean fitness changed?

H_0: No change in mean fitness, $\mu = 0$

H_A: Change in mean fitness after 100 gens, $\mu \neq 0$

Example: The consequences of mutation on fitness after 100 generations in $\phi6$ using 5 lines:

0.063, -0.062, 0.064, -0.043, 1.34

Has the mean fitness changed?

H_0: No change in mean fitness, $\mu = 0$

H_A: Change in mean fitness after 100 gen, $\mu \neq 0$

$\bar{Y} = 0.2724$

s = 0.600

n = 5

dof = 4

$$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}} = \frac{0.2724 - 0}{0.600/\sqrt{5}} = 1.02$$

One Sample t Test

<u>Example:</u> The consequences of mutation on fitness after 100 generations in ϕ6 using 5 lines:

0.063, -0.062, 0.064, -0.043, 1.34

Has the mean fitness changed?

H_0: No change in mean fitness, $\mu = 0$

H_A: Change in mean fitness after 100 gen, $\mu \neq 0$

$$\bar{Y} = 0.2724$$
$$s = 0.600$$
$$n = 5$$
$$dof = 4$$

$$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}} = \frac{0.2724 - 0}{0.600/\sqrt{5}} = 1.02$$

One Sample t Test

Example (PP#4): The consequences of mutation on fitness after 100 generations in $\phi 6$ using 5 lines:

0.063, -0.062, 0.064, -0.043, 1.34

Has the mean fitness changed?

H_0: No change in mean fitness, $\mu = 0$

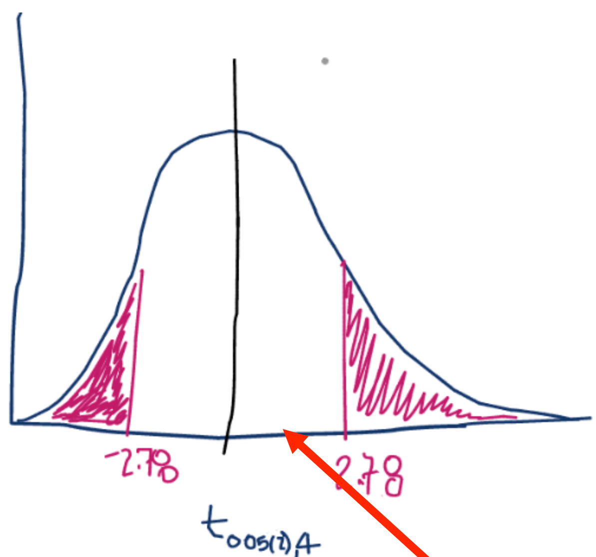
H_A: Change in mean fitness after 100 gen, $\mu \neq 0$

$\bar{Y} = 0.2724$

$s = 0.600$

$n = 5$

$dof = 4$



$$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}} = \frac{0.2724 - 0}{0.600/\sqrt{5}} = 1.02$$

How unusual is this data?

One Sample t Test

<u>Example (PP#4):</u> The consequences of mutation on fitness after 100 generations in $\phi 6$ using 5 lines:

0.063, -0.062, 0.064, -0.043, 1.34

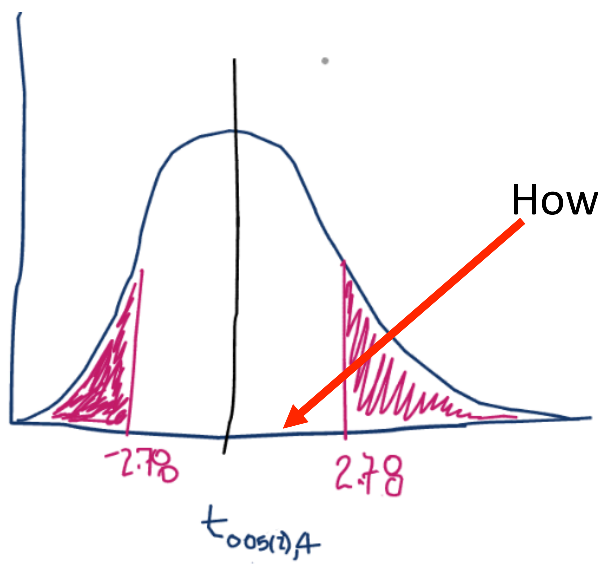
Has the mean fitness changed?

H_0: No change in mean fitness, $\mu = 0$

H_A: Change in mean fitness after 100 gen, $\mu \neq 0$

$\overline{Y} = 0.2724$
s = 0.600
n = 5
dof = 4

$$t = \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} = \frac{0.2724 - 0}{0.600/\sqrt{5}} = 1.02$$



How unusual is this data?

Answer: Not very. Based on this sample of mean fitnesses, we fail to reject H_0.

We could also calculate the confidence interval for mean fitnesses after 100 generations…

$$\boxed{\bar{Y} - t_{\alpha(2),df} SE_{\bar{Y}} < \mu < \bar{Y} + t_{\alpha(2),df} SE_{\bar{Y}}}$$

For 95%:

$-0.47 < \mu < 1.02$

Assumptions:

- Random sample
- Normally (ish) distributed variable

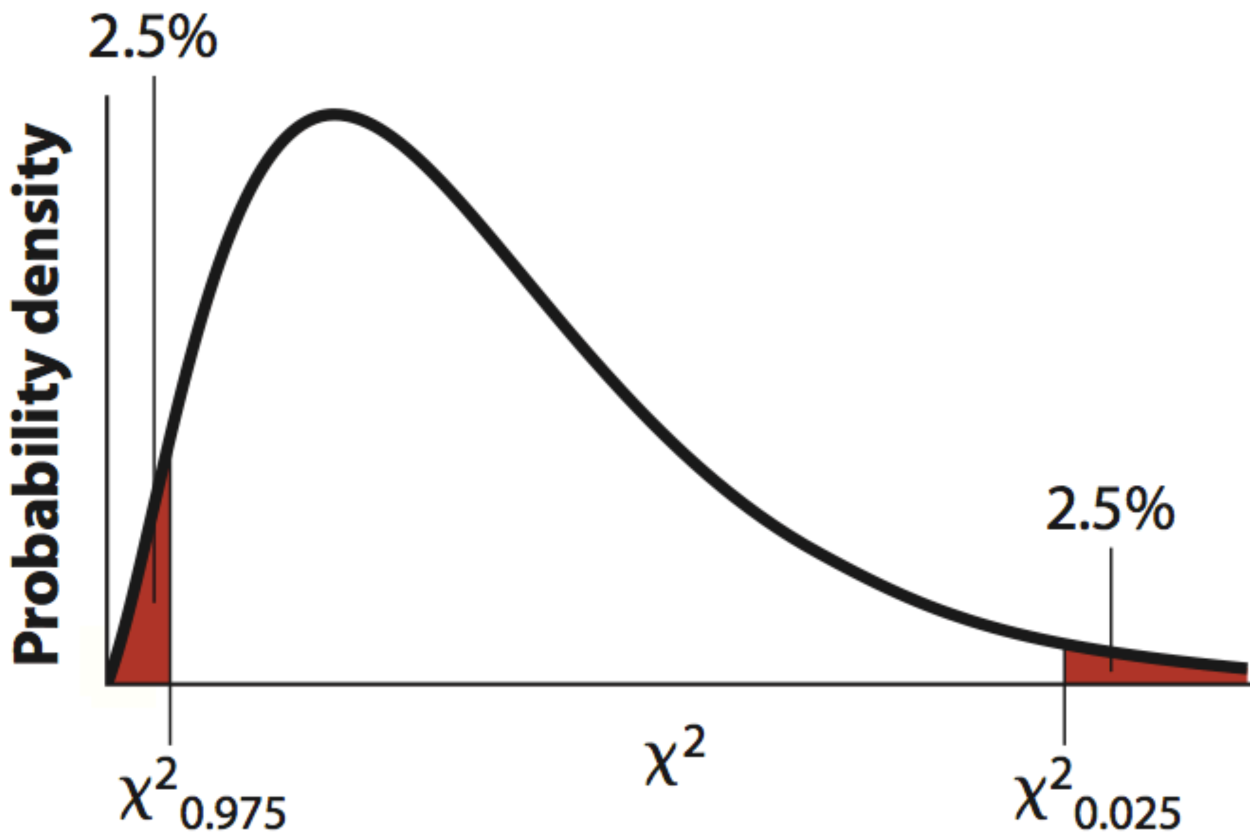
Estimating the Standard Deviation and Variance of a Normal Population:

- Y is a normally distributed variable then the sampling distribution of:

$$(n-1)\frac{s^2}{\sigma^2} \sim \chi^2_{n-1}$$

is the χ^2 distribution with dof = n-1

Estimating the Standard Deviation and Variance of a Normal Population:



$$(n-1)\frac{s^2}{\sigma^2} \sim \chi^2_{n-1}$$

Estimating the Standard Deviation and Variance of a Normal Population:

$$(n-1)\frac{s^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\boxed{\frac{df * s^2}{\chi^2_{\frac{\alpha}{2},df}} < \sigma^2 < \frac{df * s^2}{\chi^2_{1-\frac{\alpha}{2},df}}}$$

Confidence Interval: $$\sqrt{\frac{df * s^2}{\chi^2_{\frac{\alpha}{2},df}}} < \sigma < \sqrt{\frac{df * s^2}{\chi^2_{1-\frac{\alpha}{2},df}}}$$

Example: What is the 95% confidence interval for the mean undulation rate of the paradise flying snakes. The rate of undulation (hz) from 8 snakes:

0.9,1.4,1.2,1.2,1.3,2.0,1.4,1.6

Answer:

$$\bar{Y} \pm t_{0.05(2),7} SE_{\bar{Y}} = 1.375 \pm 0.115(2.36)$$

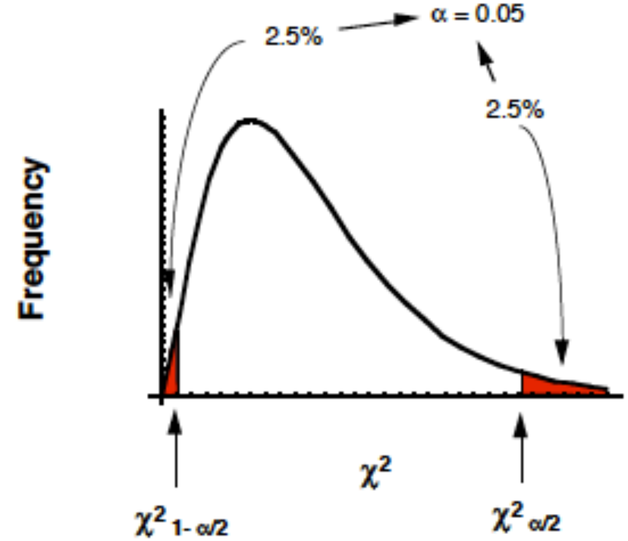
$\boxed{1.10 < \mu < 1.65}$ (95% Confidence Interval)

$$\bar{Y} \pm t_{0.01(2),7} SE_{\bar{Y}} = 1.375 \pm 0.115(3.50)$$

$\boxed{0.97 < \mu < 1.78}$ (99% Confidence Interval)

95% confidence interval for the variance of the flying snake undulation rate:

$$\boxed{\frac{df * s^2}{\chi^2_{\frac{\alpha}{2},df}} < \sigma^2 < \frac{df * s^2}{\chi^2_{1-\frac{\alpha}{2},df}}}$$



df=n-1=7

$s^2 = (0.324)^2 = 0.105$

$$\chi^2_{\frac{\alpha}{2},df} = \chi^2_{0.025,7} = 16.01$$

$$\chi^2_{1-\frac{\alpha}{2},df} = \chi^2_{0.975,7} = 1.69$$

df \ X	0.999	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005	0.001
1	1.6 E-6	3.9E-5	0.00016	0.00098	0.00393	3.84	5.02	6.63	7.88	10.83
2	0	0.01	0.02	0.05	0.1	5.99	7.38	9.21	10.6	13.82
3	0.02	0.07	0.11	0.22	0.35	7.81	9.35	11.34	12.84	16.27
4	0.09	0.21	0.3	0.48	0.71	9.49	11.14	13.28	14.86	18.47
5	0.21	0.41	0.55	0.83	1.15	11.07	12.83	15.09	16.75	20.52
6	0.38	0.68	0.87	1.24	1.64	12.59	14.45	16.81	18.55	22.46
7	0.6	0.99	1.24	1.69	2.17	14.07	16.01	18.48	20.28	24.32
8	0.86	1.34	1.65	2.18	2.73	15.51	17.53	20.09	21.95	26.12

95% confidence interval for the variance of the flying snake undulation rate:

$$\frac{df * s^2}{\chi^2_{\frac{\alpha}{2},df}} < \sigma^2 < \frac{df * s^2}{\chi^2_{1-\frac{\alpha}{2},df}}$$

$$\frac{7(0.324)^2}{16.01} < \sigma^2 < \frac{7(0.324)^2}{1.69}$$

$$0.0459 < \sigma^2 < 0.435$$

95% confidence interval for the **Standard deviation** of the flying snake undulation rate:

$$\sqrt{\frac{df * s^2}{\chi^2_{\frac{\alpha}{2},df}}} < \sigma < \sqrt{\frac{df * s^2}{\chi^2_{1-\frac{\alpha}{2},df}}}$$

$$\sqrt{\frac{7(0.324)^2}{16.01}} < \sigma < \sqrt{\frac{7(0.324)^2}{1.69}}$$

$$\sqrt{0.0459} < \sigma < \sqrt{0.435}$$

$$0.21 < \sigma < 0.66$$

Two species of net-casting spiders, ***deinopsis*** and ***menneus***, co-exist in eastern Australia, a place that – in my opinion - produces a disproportionate number of deadly organisms. The following summary statistics about the size of the prey of these two species were obtained:

	deinopsis	***menneus***
n	**10**	**10**
E(x)	10.26 mm	9.02
s^2_X	$(2.51)^2$	$(1.90)^2$

With 95% confidence: Are the two population variances the same?

a. FTR the null hypothesis

b. Reject the null hypothesis

Example A professor wants to test if her introductory class has a good grasp of basic concepts. 6 students are randomly chosen and given a proficiency test. The professor wants the class to be able to score above 70 on the test.

62, 92, 75, 68, 83, 95

Can the professor be at least **90%** certain that the mean score for the class on the test would be 70%?

A. Yes

B. No

Example A professor wants to test if her introductory class has a good grasp of basic concepts. 6 students are randomly chosen and given proficiency test. The professor wants the class to be able to score above 70 on the test.

62, 92, 75, 68, 83, 95

Can the professor be at least 90% certain that the mean score for the class on the test would be 70%?

Step 1: H_o: $\mu \leq 70$; H_A: $\mu > 70$

Step 2: Use a one sample t test. Calculate the relevant values from our data: sample mean = 79.17

sample stand. dev. = 13.17

$$\textbf{t-value: } t = \frac{79.17-70}{13.17/\sqrt{6}} = \frac{9.17}{5.38} = 1.71$$

Step 3: compare to the critical value of **t** that has **5 df** and alpha = 0.10 (one-tailed so you don't divide alpha by 2); $t_{0.10,5}=1.476$

Step 4: Since 1.71 > 1.476, we reject Ho and with 90% confidence state that the true class mean on the math test would be at least 70%. The 90% Confidence Interval:

$$79.17 - 1.476*5.38 < \mu < 79.17 + 1.476*5.38$$

71.23 < <**μ** < 87.11 (note that 70 is not included)