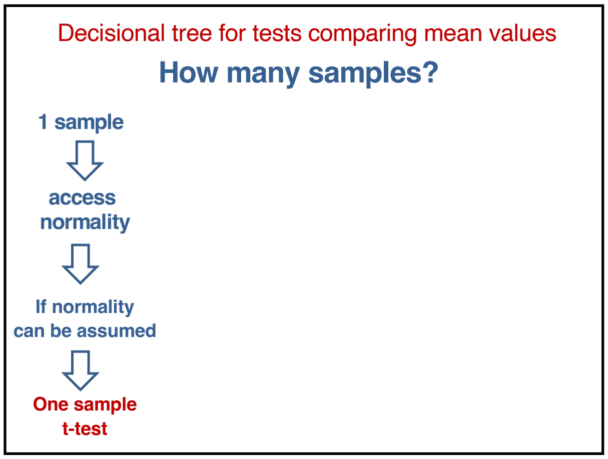
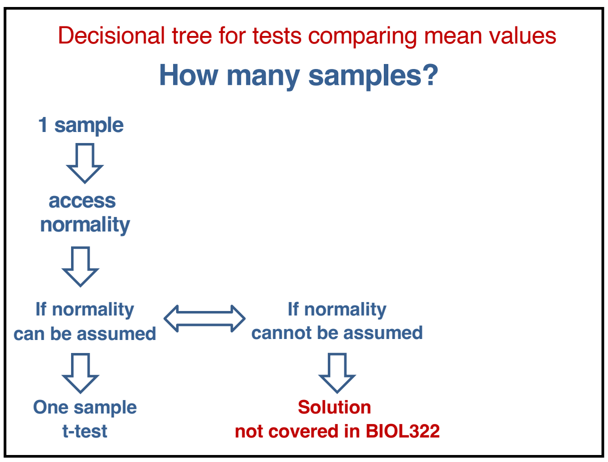
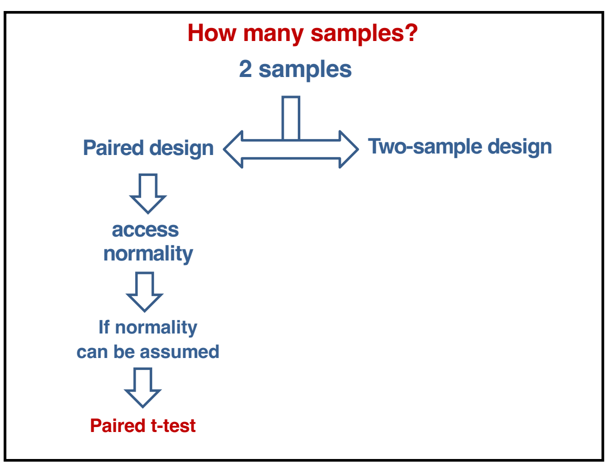

Decisional tree for tests comparing mean values
How many samples?
1 sample
access normality
If normality can be assumed
One sample t-test

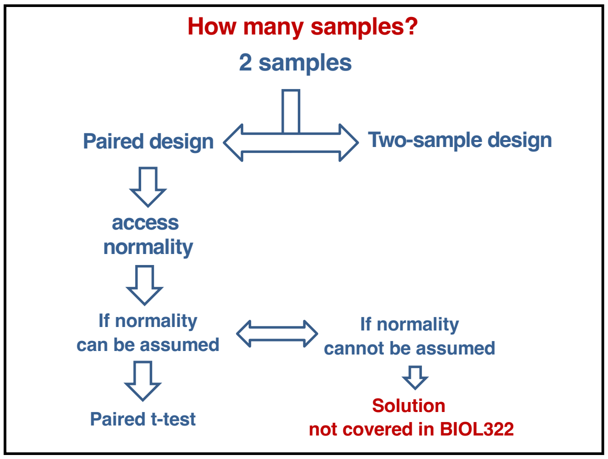
1


Decisional tree for tests comparing mean values
How many samples?
1 sample
access normality
If normality can be assumed
If normality cannot be assumed
One sample t-test
Solution not covered in BIOL322

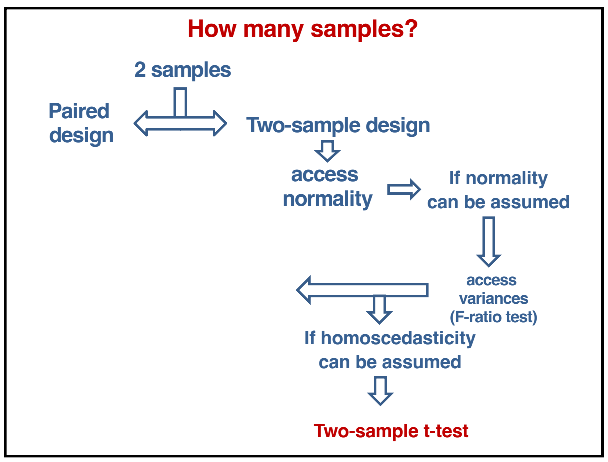
2


How many samples?
2 samples
Paired design
Two-sample design
access normality
If normality can be assumed
Paired t-test

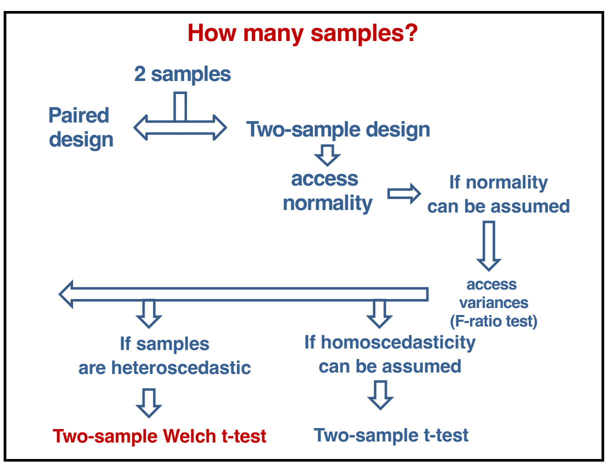
3


How many samples?
2 samples
Paired design
Two-sample design
access
normality
If normality
can be assumed
If normality
cannot be assumed
Paired t-test
Solution
not covered in BIOL322

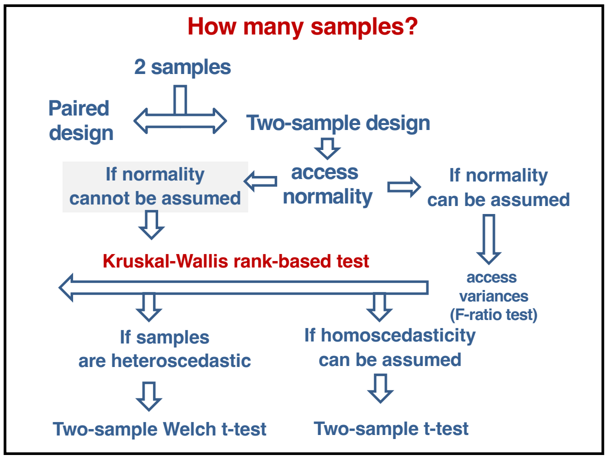
4


How many samples?
2 samples
Paired
design
Two-sample design
access
normality
If normality
can be assumed
access
variances
(F-ratio test)
If homoscedasticity
can be assumed
Two-sample t-test

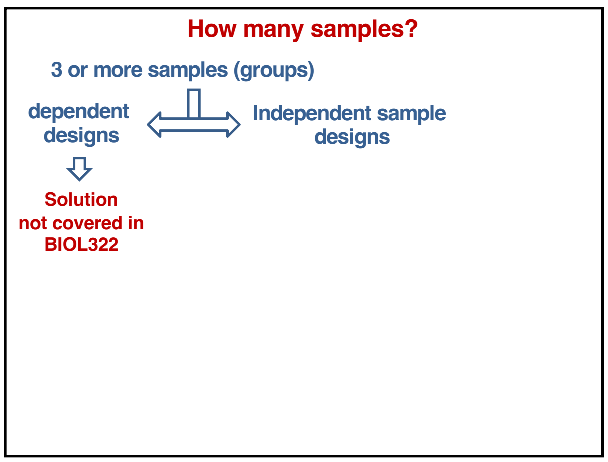
5


How many samples?
2 samples
Paired
design
Two-sample design
access
normality
If normality
can be assumed
access
variances
(F-ratio test)
If samples
are heteroscedastic
If homoscedasticity
can be assumed
Two-sample Welch t-test
Two-sample t-test

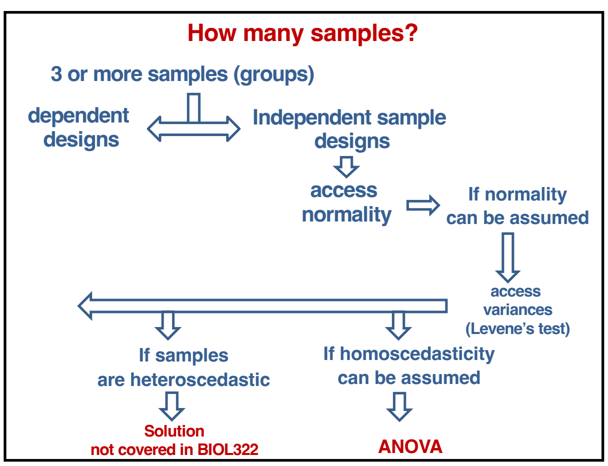
6


How many samples?
2 samples
Paired design
Two-sample design
If normality cannot be assumed
access normality
If normality can be assumed
Kruskal-Wallis rank-based test
access variances (F-ratio test)
If samples are heteroscedastic
If homoscedasticity can be assumed
Two-sample Welch t-test
Two-sample t-test

7

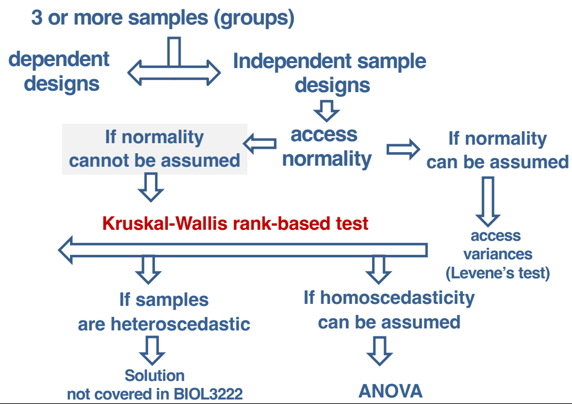

How many samples?
3 or more samples (groups)
dependent designs
Independent sample designs
Solution not covered in BIOL322

8


How many samples?
3 or more samples (groups)
dependent designs
Independent sample designs
access normality
If normality can be assumed
access variances (Levene's test)
If samples are heteroscedastic
If homoscedasticity can be assumed
Solution not covered in BIOL322
ANOVA

9

How many samples?



10

Inferential (statistical hypothesis testing) approaches for testing means and variances covered in BIOL322:

One sample t-test
Two sample t-test
Welch's t-test
Paired t-test
ANOVAs
Levene's test
F-ratio

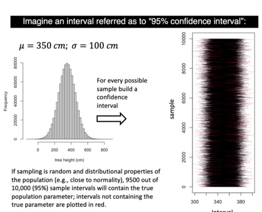
They can be conducted by regression in which response and predictors are modified according to the test in question.

11

1) A soybean farmer took a random sample of 30 plants after growing for 5 weeks and measured their size (length of plant in cm). They went back to the same plants one week later and measured the plant sizes again. The mean difference between the two samples was 10.0 cm and standard deviation of the difference was 1.2 cm. The resulting 95% confidence interval for the mean difference was 8 cm - 12 cm. Which of the following statements is a correct interpretation of the 95% confidence interval?

A) We can extrapolate to the entire farm that the plants grew between 8 cm and 12 cm on 95% of the days that passed between the two samples.
B) There is a 0.95 probability that the true mean amount of plant growth for the entire farm in that one-week period is between 8 cm and 12 cm.
C) We can't speak of probabilities. What we can say is that if we were to repeat the process of sampling multiple times, 95% of the intervals would contain the true growth rate for the entire farm.
D) We can't speak of probabilities. What we can say is that if we were to repeat the process of sampling multiple times, 95% of the intervals would be between 8 cm and 12 cm.
E) None of these answers are appropriate.

ANSWER: C

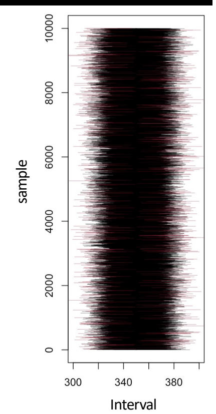


12

Very important!

For any given sample confidence interval, we can state that "we are 95% confident that the true population mean lies between the lower and upper limits of the interval".

We cannot say that "there is a 95% probability that the true population mean lies within the confidence interval". Either the parameter is within the interval or not! So, no probability attached to this condition.



13

Which of the following best describes samples that exhibit high **ACCURACY** and low **PRECISION** when estimating the variance of a population?

A) The values tend to be like one another and, in average, different from the true population value.

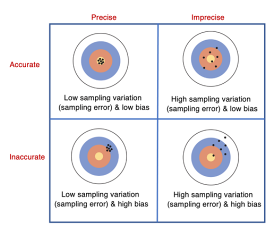
B) The values tend to be like one another and, in average, and to be similar to the true population value.

C) The values tend to be quite different from one another and, in average, different from the true population value.

D) The values tend to be quite different from one another and, in average, they tend to be similar to the true population value.

E) None of the answers.

ANSWER: D



14

Which of the following DOES NOT describe a random sample?

A) Whether one individual is selected has no bearing on whether another individual is selected as well.

B) Each individual observation's chance of being selected is independent of the other individuals selected.

C) Each individual observation in the population is equally likely to be chosen.

D) Each **numerical value** for an observation in the population is equally likely to be chosen.

E) None of the answers.

ANSWER: D

1,2,2,2,2,2,2,2,2,2,2,3,3,4

15

Which of the following is the criterion for classifying a study as an experimental study?

A) Individuals are assigned to different treatments based on criteria out of the researcher's control.

B) Individuals are assigned to treatments based on a measurable trait (e.g., body size).

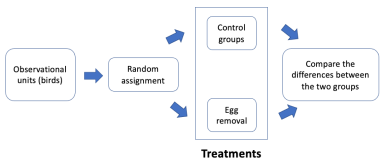
C) Individuals are observed and measured in assigned laboratory settings.

D) Individuals are randomly assigned to different treatments.

E) All these are expected to happen under experimental studies

ANSWER: D

Experimental study - Researcher randomly assigns observational units (birds) to different groups (often called treatments), i.e., they control the treatments.



16

A good experiment is designed to do ALL the following except:

A) minimize bias in estimating treatment effects.
B) minimize bias is testing treatment effects.
C) minimize P-values.
D) minimize sampling error.
E) All experiments should try their best to achieve all these goals.

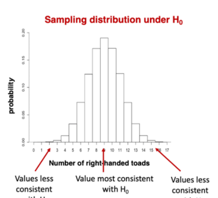
ANSWER: C

17

A P-value indicates:

A) A measure of how compatible the observed data is with the null hypothesis.
B) The probability that the alternative hypothesis is true.
C) The probability of committing a false positive.
D) A statement about the truth about a hypothesis.
E) None of these statements.

ANSWER: A



So, the sampling distribution underlying the null hypothesis is the realm of values for the test statistics that are consistent with the null hypothesis.

Even in a world where H_0 is true, some values for the test statistic of interest are more **consistent** and some values are **less consistent** with H_0.

What is **consistency**? compatible or in agreement with something (here H_0).

18

A P-value indicates:

A) A measure of how compatible the observed data is with the null hypothesis.
B) The probability that the alternative hypothesis is true.
C) The probability of committing a false positive.
D) A statement about the truth about a hypothesis.
E) None of these statements.

ANSWER: A

H_0: Right-handed and left-handed toads are equally frequent in the population.

H_A: Right-handed and left-handed toads are NOT equally frequent in the population.

The test statistic that we will use here is the number of right-handed frogs.

Remember that the test statistic is a number calculated from the data that is used to evaluate how compatible the observed (sample) data are with the result expected under random sampling from a statistical population in which the null hypothesis is true (i.e., the sampling distribution under H_0).

19

In a paired design study, if the variances of the two samples are heteroscedastic, what is the best course of analysis to compare their means?

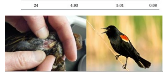
A) Conduct the Welch's t-test.
B) Conduct the standard t-test.
C) Conduct a paired t-test.
D) Conduct an F-ratio test.
E) Conduct a Levene's test.

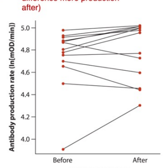
Are males with high testosterone paying a cost for this extra mating success in other ways (trade-offs)?

Antibody production rates measure optically
ln[mOD/min] = log optical density per minute

Male identification number	Before implant: Antibody production (ln[mOD/min])	After implant: Antibody production (ln[mOD/min])	d
1	4.65	4.44	−0.21
4	3.91	4.30	0.39
5	4.91	4.98	0.07
6	4.50	4.45	−0.05
9	4.80	5.00	0.20
10	4.88	5.00	0.12
15	4.88	5.01	0.13
16	4.78	4.96	0.18
17	4.98	5.02	0.04
19	4.87	4.73	−0.14
20	4.75	4.77	0.02
23	4.70	4.60	−0.10
24	4.93	5.01	0.08

d is the difference between treatments (positive difference more production after)





Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

20

Which of the following does NOT increase the power of a statistical test?

A) Larger sample size.

B) Larger discrepancies from null hypothesis expectations.

C) Lower significance level thresholds.

D) Lower variability in the population.

E) Impossible to tell as it will depend on the data.

ANSWER: C

21

Which of the following IS NOT a method to generate data appropriate for a paired t-test?

A) Compare the left and right sides of individuals given different treatments to each arm.

B) Measure the same thing in twins divided so one is in each treatment.

C) Measure the same thing in individuals before and after an intervention.

D) Place individuals randomly into the treatments.

E) All options are appropriate for a paired t-test

22

Which of the following IS NOT one of the main questions to consider when choosing which statistical test to use?

A) Are the variables categorical or numerical?

B) Are the data values paired in some way?

C) Does the sample size allow estimation of P-values?

D) Does our data meet the assumptions of the test we're considering?

E) All of them are important questions.

ANSWER: C

23

STATISTICAL LITERACY: THINKING CRITICALLY ABOUT STATISTICS

Milo Schield, Augsburg College

Department of Business & MIS Minneapolis, MN

Statistical literacy is the ability to read and interpret data: the ability to use statistics as evidence in arguments.

Statistical literacy is a competency: the ability to think critically about statistics.

24

Statistical literacy: one shouldn't simply trust statistics without having a good knowledge of the problem at hands. use statistics as evidence in arguments!

25

Statistical literacy is a basic skill: the ability to think critically about arguments using statistics as evidence.

Consider the story of two photographers being chased by a bear. [Adapted from David Friedman, (1996)] The first says, "It's hopeless! This bear can run twice as fast as we can." The second, realizing the flaw in the argument says, "No, it's not hopeless! I don't have to outrun this bear. I just have to outrun you!" The truth of this statistic ("twice as fast") does not give strong support for this conclusion ("it's hopeless").

The second photographer was statistically literate; the first photographer wasn't.

26