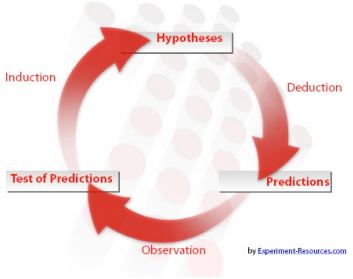
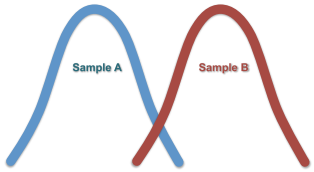
Inferential Statistics
Experiment design & analysis



What is the probability that two samples are significantly different?
(I.e., represent different populations.)

Inferential Statistics
Experiment design & analysis

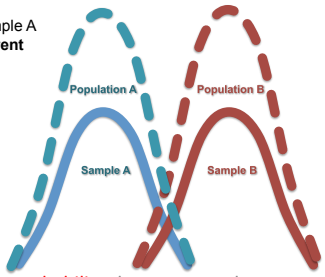
- Compare data set **A** *vs.* data set **B**
  - Ecosystem X *vs.* Ecosystem Y
  - Experimental *vs.* Control
  - Etc.



What is the probability that two samples are significantly different?
(I.e., represent different populations.)

Inferential Statistics
Experiment design & analysis

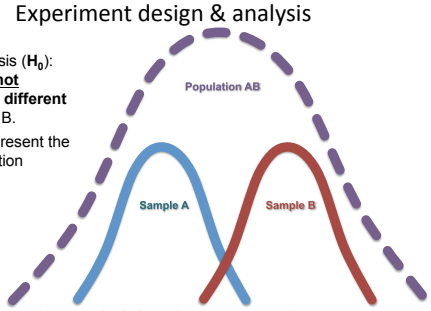
- Hypothesis (H_A): Sample A is **significantly different** from Sample B.
- I.e., They represent different populations
  A≠B



What is the probability that two samples are significantly different?
(I.e., represent different populations.)

Inferential Statistics
Experiment design & analysis

- Null Hypothesis (H_0): Sample A is **not significantly different** from Sample B.
- I.e., They represent the same population
  A⇔B



What is the probability that two samples are significantly different?
(I.e., represent different populations.)

Testing the Null Hypothesis (A⇔B)

Possible conclusion from observations and/or experiments

- **Accept null hypothesis** [Samples A & B are **not different**.] A⇔B [Reject alternate hypothesis tested]
- **Reject null hypothesis** [Samples A & B **are different**.] A≠B [Accept alternate hypothesis tested]

Reality

- Samples A & B represent the **same population** (A⇔B) [Null hypothesis is correct. Alternate hypothesis incorrect.]
- Samples A & B represent **different populations** (A≠B) [Alternate hypothesis is correct. Null hypothesis incorrect.]

Testing the Null Hypothesis (A⇔B)

* **Type I Error**: Rejecting true null hypothesis / Accepting false alternate hypothesis

Possible conclusion from observations and/or experiments

- **Accept null hypothesis** [Samples A & B are **not different**.] A⇔B [Reject alternate hypothesis tested]
- **Reject null hypothesis** [Samples A & B **are different**.] A≠B [Accept alternate hypothesis tested]

Wrong conclusion *

Correct conclusion

Reality

- Samples A & B represent the **same population** (A⇔B) [Null hypothesis is correct. Alternate hypothesis incorrect.]
- Samples A & B represent **different populations** (A≠B) [Alternate hypothesis is correct. Null hypothesis incorrect.]

Testing the Null Hypothesis (A⇔B)

* **Type II Error**: Accepting false null hypothesis / Rejecting true alternate hypothesis

Possible conclusion from observations and/or experiments

- **Accept null hypothesis** [Samples A & B are **not different**.] A⇔B [Reject alternate hypothesis tested]
- Reject null hypothesis [Samples A & B **are different**.] A≠B [Accept alternate hypothesis tested]

Correct conclusion

Wrong conclusion *

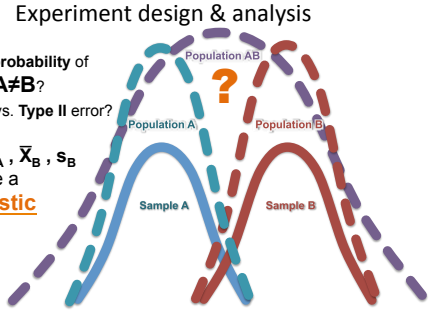
Reality

- Samples A & B represent the **same population** (A⇔B) [Null hypothesis is correct. Alternate hypothesis incorrect.]
- Samples A & B represent **different populations** (A≠B) [Alternate hypothesis is correct. Null hypothesis incorrect.]

Inferential Statistics

Experiment design & analysis

- What is the **probability** of **A⇔B** vs. **A≠B**?
- **Type I** error vs. **Type II** error?
- Use $\bar{X}_A$, s_A , $\bar{X}_B$, s_B to calculate a **test statistic**



•**Type I Error**: Concluding A and B are different when they really are not.
•**Type II Error**: Concluding A and B are not different when they really are.

Remember: The Empirical Rule

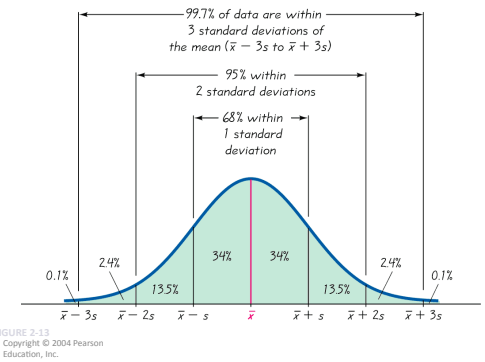


FIGURE 2-13

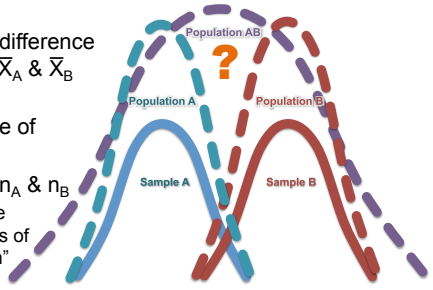
Terms

confidence interval: The range of values we can be reasonably certain includes the true value.
- We can be 95% confident that the "true" population mean is within 2 standard deviations from the sample mean.
- But there is still a 5% probability, the "true" mean is >2s from the sample mean.

p value: the probability that an observed *apparent* difference could have occurred by chance.
- i.e., the probability the **null** hypothesis is true.
- In biology, a **p value** ≤0.05 is usual recognized as a significant difference
  - i.e., ≥95% probability H_A is correct; ≤5 % probability H_0 is correct.

Increasing *Power* of the test statistic

- Increase difference between $\bar{X}_A$ & $\bar{X}_B$
- Minimize magnitude of s_A & s_B
- Increase n_A & n_B
  - Increase "degrees of freedom"



*Cannot alter *natural* variation.
***Can** (and should) reduce variation from *sampling error* and *bias*!

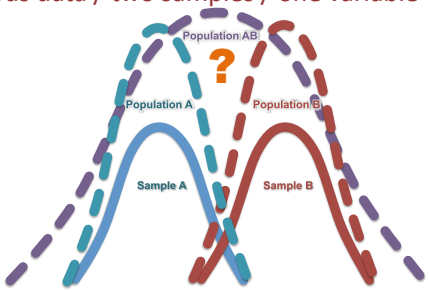
How to choose the best statistical test?

- Comparing means, proportions, and relationships
- Four major factors:

1. What type of data?
2. How many variables are being compared?
3. How many samples (populations)?
4. What is the purpose of the test?

Example:

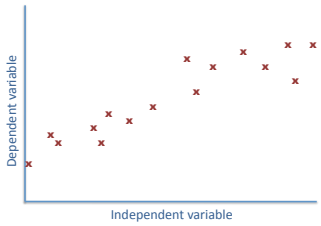
continuous data / two samples / one variable



- Test: comparison of two independent means:
  - Students t-test

Example:

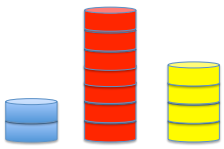
continuous data / one sample / two variables



- Test: comparison of two variables:
  - Correlation coefficient / regression analysis

Example:

nominal data / multiple samples / one variable



- Test: comparison of frequency of occurrence:
  - Chi-square test