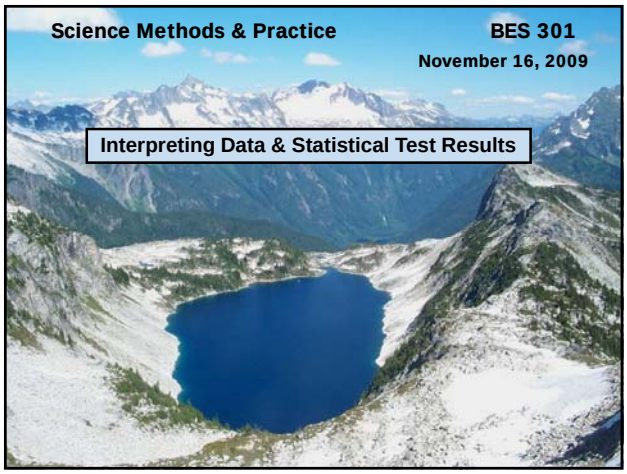
Science Methods & Practice BES 301

November 16, 2009



Interpreting Data & Statistical Test Results

Interpreting Scientific Data

Reed Canary Grass (*Phalaris arundinacae*) - RCG

Invasive non-native species of grass that takes over freshwater wetland areas in our region.

Interpreting Scientific Data

Reed Canary Grass (*Phalaris arundinacae*)

What factors influence the rate of RCG invasion?

Perhaps watershed development makes a difference?

Let's compare RCG abundance in different watersheds with differing degrees of development:

Bear Creek vs. North Creek

Interpreting Scientific Data

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
1	14.0	32.6
2	20.6	84.5
3	18.3	54.2
4	19.6	22.1
5	15.2	77.9

What do we do first with these data?

Conclusions ?

Interpreting Scientific Data

Let's try a different data set

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
1	55.9	32.6
2	54.1	84.5
3	51.4	54.2
4	47.6	22.1
5	62.3	77.9

Analysis ?

Conclusions ?

Interpreting Scientific Data

Let's try YET ANOTHER different data set

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
1	52.9	32.6
2	75.2	84.5
3	51.4	54.2
4	32.1	22.1
5	28.6	77.9

Analysis ?

Conclusions ?

Interpreting Scientific Data

In order to conduct any kind of test the study QUESTION and HYPOTHESES **must be** clearly defined.

Question: Is RCG invasion greater in North Creek than Bear Creek?

Study Hypothesis: RCG invasion is greater in North Creek than Bear Creek.

Stating Statistical Hypotheses for comparisons

"Null Hypothesis": There is no difference in RCG invasion between North Creek and Bear Creek.

"Alternate Hypothesis": There is a difference in RCG invasion between North Creek and Bear Creek.

Interpreting Scientific Data

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
1	52.9	32.6
2	75.2	84.5
3	51.4	54.2
4	32.1	22.1
5	28.6	77.9
Mean	48.0	54.3
SD	18.7	27.3

What can you conclude?

What kind of errors can we make in our conclusions?

Interpreting Scientific Data

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
Mean	48.0	54.3
SD	18.7	27.3

What kind of errors can we make in our conclusions?

TYPE I Error:

TYPE II Error:

Interpreting Scientific Data

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
Mean	48.0	54.3
SD	18.7	27.3

We test our "alternate hypothesis": "There is a difference"

Result of statistical test: P = 0.69

"P" indicates "Probability" (fractional)

Interpreting Scientific Data

% Cover of RCG in wetland

Wetland Sampled	Bear Creek	North Creek
Mean	48.0	54.3
SD	18.7	27.3

We test our "alternate hypothesis": There is a difference

Result of statistical test: P = 0.69

What can we conclude?

Interpreting Scientific Data

P = 0.69

	Bear Creek	North Creek
Mean	48.0	54.3
SD	18.7	27.3

There is a 69% chance of committing a type I error:

Stating that there IS a difference between the creeks when there really is not.

In other words, there is a 69% chance that the alternate hypothesis is WRONG

Given that probability, what do we say about these results?

Interpreting Scientific Data

Let's go back to the very first data set

P = 0.017

	Bear Creek	North Creek
Mean	17.5	54.3
SD	2.8	27.3

There is a 1.7% chance of committing a type I error:

Stating that there IS a difference between the creeks when there really is not.

Interpreting Scientific Data

What % chance of making a type I error are we willing to accept?

1%? 5%? 10%? 20%? 50%?

by general convention

There is less than a 5% chance of committing a type I error:

Stating that there IS a difference (between the creeks) when there really is not.

Why can't we be 100% confident about our conclusions?

Interpreting Scientific Data

Is P<.05 (5%) a magical cut-off value?

- P-value used for significance MUST be clearly stated
- Often it is best to report the actual P value to allow readers to draw their own conclusions
- P cutoffs of 0.1 are not uncommon in field studies

Interpreting Scientific Data

Designating Statistical Results

Ecological Characteristic	Bear Creek	North Creek
RCG Invasion (% cover)	33.0 ± 3.8 a	36.2 ± 1.2 b
% Tree Cover	22.1 ± 0.02 a	18.4 ± 0.1 b
Stream pH	6.6 ± 0.1 a	6.4 ± 0.1 b

Interpreting Scientific Data

Designating Statistical Results

Ecological Characteristic	Bear Creek	North Creek
RCG Invasion (% cover)	33.0 ± 3.8 x	36.2 ± 1.2 y
% Tree Cover	22.1 ± 0.02 &	18.4 ± 0.1 $
Stream pH	6.6 ± 0.1 a	6.4 ± 0.1 b

Any symbols or letter combinations can be used

Don't get lost in the fog of Statistical Significance!

There is a difference between

STATISTICAL SIGNIFICANCE

&

ECOLOGICAL / BIOLOGICAL SIGNIFICANCE

Ecological Characteristic	Bear Creek	North Creek
RCG Invasion (% cover)	33.0 ± 3.8 a	36.2 ± 1.2 b
% Tree Cover	22.1 ± 0.02 a	18.4 ± 0.1 b
Stream pH	6.6 ± 0.1 a	6.4 ± 0.1 b