**AnSc 5403**
*Biometry*

Lecture Notes 1

I. What is statistics?

A. "Statistics is the study of the methods and procedures for **collecting**, **classifying**, **summarizing**, and **analyzing** data and for making scientific **inferences** from such data." – Remington and Schork. 1970. *Statistics with Applications to the Biological and Health Sciences*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

1. Two categories

a. **Descriptive statistics** –

(a) Frequency, average, variability, and so on

(b) Not concerned with implications or conclusions that can be drawn from data sets – but methods used will affect conclusions that can be drawn and could introduce bias

b. **Inferential statistics** –

(a) Biological populations of interest are often too large to sample all members of the population

(i) Inferential statistics involves sampling a population so as to draw conclusions (make inferences) about the population

B. Both descriptive and inferential statistics deal primarily with numerical data –

1. Measurements might be **discrete** or **continuous**

a. **Discrete** –

b. **Continuous –**

(a) Values reflect the precision of the measuring instrument

2. Measurements might be **direct** or **subjective**

a. Example –

b. As noted for continuous variables, measurements are not typically error-free

(a) Continuous variables are recorded in discrete form, with the size of the gap between measurements reflecting the precision of the measurement

c. Subjective measurements are likewise not error free and depend on the observer who assigns the measurement

d. Use common sense to evaluate the measurements taken to determine whether they should be subjected to statistical analyses

(a) Case 1 – The difference between two treatments in average daily gain (ADG) for a 28-d period = 0.2 lb, but the scale used to measure body weight had a readability of ± 5 lb. How meaningful is the difference in ADG?

(b) Case 2 – Condition score of beef cows were assigned by two people and then averaged. Is agreement of the scores by the two observers important in interpreting the results?

II. Presentation of Statistical Data – Graphs and Tables

A. Graphs and tables should be self-explanatory

1. Table titles and figure captions should be descriptive enough to stand alone

B. For graphs:

1. Vertical and horizontal scales should be clearly labeled and units identified
2. Keep graphs as simple as possible – avoid too many bars or lines – two or three lines are appropriate – four or more are probably too many
3. Graphs are designed to provide a "snapshot" of the results – use tables for details
4. Avoid presentation of numbers in the body of a graph

C. For tables:

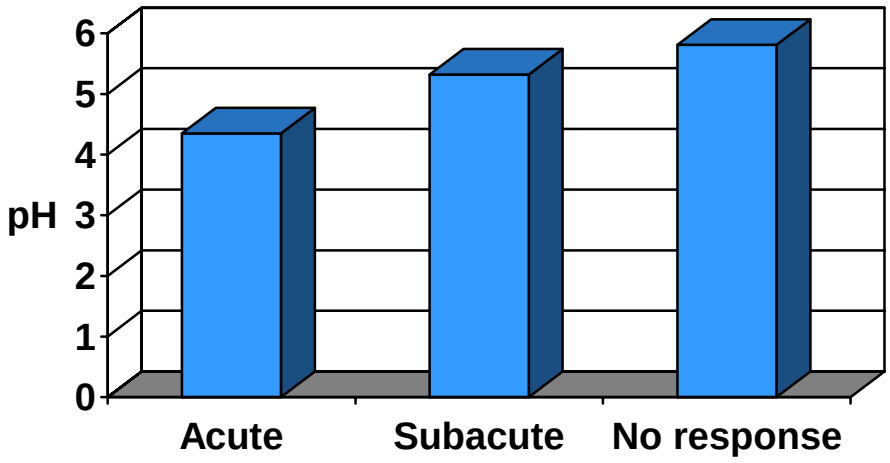
1. Make sure that units of measurement are clearly indicated for each variable
2. Avoid horizontal lines within tables, except at the start and end of tabular entries
3. Avoid putting too much information in a single table. Several small tables are often better than one large table
4. Use dashes to indicate table entries that have been left blank, but use a numerical entry for an observation that is zero (e.g., 0 or 0.0).
5. Report all entries for a given variable with the same precision (same number of decimal places).

D. Avoiding misleading graphs

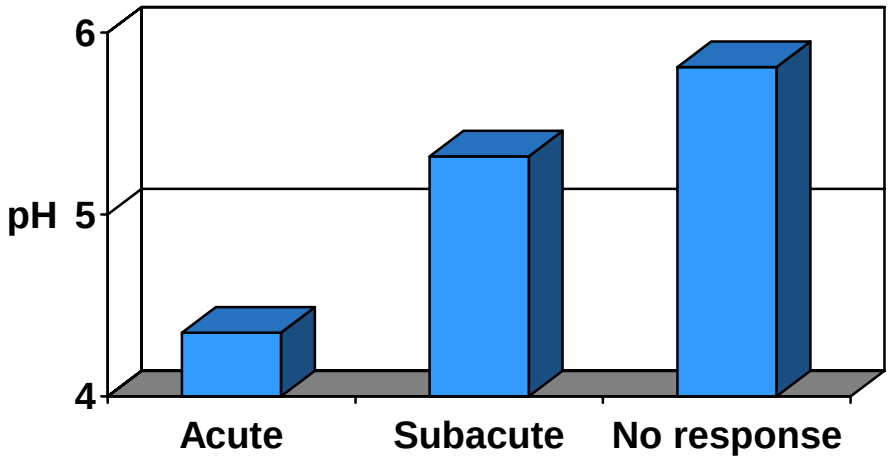
1. Graphs give an overall impression of the results as well as the facts of the results
2. Scaling of graphs can be used to give a misleading impression of the data

3. Example

   a.




   b.




   c. Data in these two graphs are identical, but scaling of the Y-axis changes the impression of the magnitude of differences

      (a) The problem can be avoided by indicating that a scale break has been used

4. Carefully examine graphs to ensure they do not bias presentation of the results