Biostatistics

www.harding.edu/plummer/biostats/biostats.pdf

Spring 2018

Introduction to Statistics	*Descriptive Statistics*	*Graph Construction*	*Inferential Statistics*	*Hypothesis Testing I*	*Hypothesis Testing II*	*Practice Problems*
Will I use this stuff?	*Statistical Tables*	*Faculty Interests*	*Statistical Tests*	*Advanced Stat Tests*	*Protocol Sheet*	*Grades*

Course Description

An introductory computer-based statistics course that includes instruction in SYSTAT. Topics covered include populations and samples, variables, probability distributions, descriptive statistics, statistical inference, and hypothesis testing. Included are selected parametric and non-parametric tests for examining differences in means, variances, and frequencies as well as correlation, regression, and tests of independence.

Emphasis is given to practical matters such as how to choose appropriate analyses and how to interpret results, both statistically and biologically. High school algebra is the only math background you need. Biostats is a practical application course - to learn it, you have to do it. Failing to apply statistical concepts and procedures on a regular basis will diminish your chances of understanding the material and earning the grade you desire.

What we have to learn to do, we have to learn by doing. – Aristotle

Student Learning Outcomes – By the end of the semester you will be able to:

- understand how science and statistics interact
- apply basic statistical procedures using professional statistical software
- read and understand primary biological literature

Textbooks and Software

- Primary text - www.harding.edu/plummer/biostats/biostats.pdf
- Supplementary text - www.khanacademy.org/math/statistics-probability
- Primary software - SYSTAT (provided on computers in S161 and S182).
- Student software - MYSTAT (free student version of SYSTAT; download at www.systat.com)

Evaluation

Exam 1	20%	Exams 1-3 are comprehensive and consist of Content (scantron/short answer ~50%) and Practical (SYSTAT problems/graphing ~50%) sections. An extra point may be earned on each exam if you are present in class when feedback is given on your graded exams. *Exam study guides*
Exam 2	20%	
Exam 3	20%	
Quizzes	20%	~10 announced quizzes and exercises
Final Exam	20%	The final exam is a comprehensive scantron exam taken during the regularly scheduled final exam period. Unlike Exams 1-3, you will not use a computer on the final for any task; this includes SYSTAT. *Exam study guides*

Classroom Policies

- Computer resources that may be viewed during lecture include the course website, SYSTAT, and your M-drive. All other uses (e.g., social notworking sites such as Facebook, Twitter, Instagram, email, blogs, sports news, pictures of your girl/boyfriend, etc.) are off limits during lecture.
- Cell phone use during lecture is prohibited. If you must send or receive a text or call during lecture time, please excuse yourself from the classroom and take it to the hallway.
- Regular class attendance is necessary to do well in this course. Excessive unexcused absences will be handled on an individual basis. An official HU class excuse or prior arrangements with the instructor is necessary to be excused from an exam.
- Cheating in all its forms is inconsistent with Christian faith and practice and will result in sanctions up to and including dismissal from the class with a failing grade. Instances of dishonesty will be handled according to the procedures delineated in the Harding University catalog.
- The visual appearance or use of any unapproved electronic device during an exam will be interpreted as cheating and will result in a zero for that exam.
- In accordance with the official Time Management Policy of the University, you are expected to spend two hours outside of class for each credit hour spent in class each week. That amounts to six additional hours per week, two of which are imposed on you in conjunction with regular class time.
- **THE ONLINE BIOSTATS LECTURE NOTES ARE NOT COMPLETE SOURCES OF INFORMATION FOR EXAMS.** In general, students are responsible for anything discussed in class.

My Responsibilities

Because, as your teacher, I have a substantial responsibility to you and to the Lord (James 3:1), I promise my best effort to you in Biol. 254. I pray that my lectures will be clear, my expectations reasonable, and my exams vigorous, thorough, challenging, and fair. I also pray that your grade will reflect both your ability and your preparation. Finally, I hope that you will learn something substantive in my class regardless of what you think about the subject matter. For further insight into my teaching philosophy, click here - Good luck!

Misc.

- You will need a personal Dropbox account. Data files for the course are available in a shared Dropbox folder called "Student Biostats." You should download these files to your M-drive.
- Statements on academic dishonesty, teaching evolution, and students with disabilities

WILL I EVER USE THIS STUFF?

You may be thinking (and perhaps hoping?) you'll never have to use statistics. The reality is that if you become a professional of any kind, you will very likely use statistics according to at least one of the three objectives of this course. Do yourself a favor and read the unsolicited testimonials from former biostats students.

Introduction to Statistics

Home

The relationship of science and statistics

A. Process of science represents an interplay of ideas and data (**ID**); BC ideas only
   1. Use data to make inferences (decisions) relative to ideas [=analyze data]
   2. Statistics is a tool that assists decision-making
   3. Great increase in use of statistics
      - basic science: American Naturalist
      - applied science: "Evidence-based medicine (EBM) is an approach to medical practice intended to optimize decision-making by emphasizing the use of evidence from well-designed and carefully-conducted research" (Wikipedia); Textbook
   4. Manufacturing-based society → service/information-based society
      - great need for quantitative methods of making decisions using information available
      - e.g., demo quantitative decision-making with scatterplot

B. Levels of organization within biology and the relative use of statistics

Community biology ↑	-Greater need for stats
Population biology	-More uncertainty
Organism biology	-Less understanding
Organ biology	-More variation
Cell biology	-More factors (=more
Molecular biology	complexity)

How do we know what we know? - mechanics of the process of science

5. R.B. Fischer - "Science is what scientists do when they're working."
6. What DO scientists do when they are working?
7. Several processes: HYPOTHETICO-DEDUCTIVE (IF-THEN) MODEL

OBSERVATION → HYPOTHESIS → PREDICTION → DESIGN TEST →
IF → THEN -observation
-experiment

COLLECT DATA → ANALYZE DATA → CONCLUSION → COMMUNICATION
(statistics) (probability) (talks, publications)

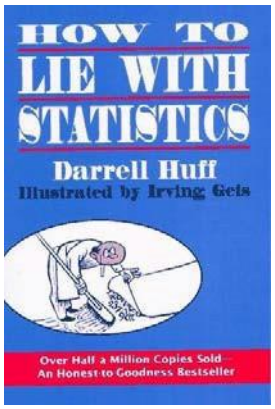
→ REFINE AND GENERALIZE HYPOTHESIS → **THEORY** (= model)

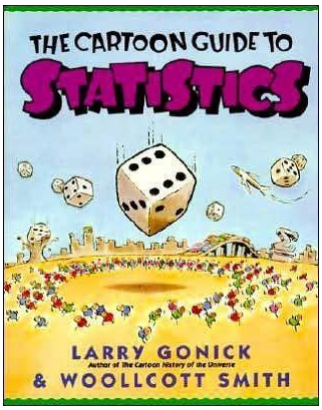
Absolute certainty is a privilege of uneducated minds-and fanatics.
It is, for scientific folk, an unattainable ideal. - C. J. Keyser

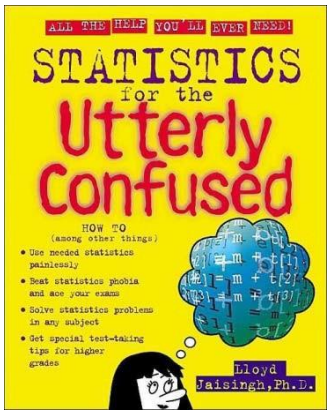
8. Advancement comes by disproving false hypos; "proof" in science means disproving false hypotheses
   - A. Einstein - “No amount of experimentation can ever prove me right; a single experiment can at any time prove me wrong.”
   - D. Hull - “The scientific method does not guarantee that you are right; it guarantees that if you are wrong, someone will find it out.”
   - S. Connery - “Isn’t that what science is all about,...eliminating possibilities?” (video)
9. “proof” is tentative - most models have historically been either discarded or radically modified - no reason to believe that it will be different in the future
   - scientific models are pragmatic (useful) - if model works, use it!
   - scientific “truth” - not necessarily “**TRUTH”**

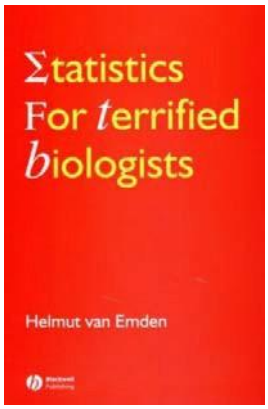
What constitutes the study of “statistics?” (often misunderstood)

-e.g., “***There are three kinds of lies: lies, damn lies, and statistics.”*** -B. Disraeli









What Statistics Is Really About

Population numerical properties = “**parameters**”	***“Error”*** ←	***Sample*** numerical properties = “**statistics**”
The population is what we want to understand.		The sample is what we use to understand the population. -***Descriptive stats***: describe data in the sample -***Inferential stats***: infer from sample to population

Descriptive Statistics

Home

StatisticalBasics

A. Definitions
   1. variable - characteristics that may differ (vary) among individuals
      a. measured
      b. derived (non-measured); derived from measured variables
      c. dependent/response variable vs. independent/predictor variable
   2. data - values of variables for individuals (singular datum)
   3. case/observation - an individual; symbolize: x_1, x_2, ...x_n (n=sample size)

B. Collection of data
   1. population - all individuals of a defined universe (= whatever we say it is!)
   2. sample - subset of population; used to make inferences regarding the population
   3. statistical error - difference between the real population value and the estimates (from sample data) of the population value
   4. randomness - all individuals have equal probability of being sampled
   5. independence - value of one case does not affect the value of other cases

C. Scales of measurement and variable types
   1. Categorical scale (Nominal)
      a. values not quantitative or ranked; no mathematical or value relationship
      b. mutually exclusive categories (e.g., male/female)
      c. 1 variable type: categorical
   2. Ranked scale (Ordinal)
      a. relative differences (e.g., greater than/less than)
      b. no mathematical relationship between values (e.g., small/medium/large; highly active/active/not active)
      c. 1 variable type: ranked
   3. Ratio scale
      a. mathematically defined distance between values; quantitative
      b. absolute zero point (e.g., mass)
      c. 2 variable types:
         - Discrete - may assume only certain values within given range (e.g., 1, 2, 3, 4)
         - Continuous - may assume any value within given range (e.g., 1.0, 2.34, 2.344)
      d. may convert ratio data to ranked/categorical data (but not vice versa)
   4. Interval scale
      a. mathematically defined distance between values; quantitative
      b. arbitrary zero point (e.g., Celsius temperature scale)
      c. 2 variable types:
         - Discrete - may assume only certain values within given range (e.g., 1, 2, 3, 4)
         - Continuous - may assume any value within given range (e.g., 1.0, 2.0, 2.34, 2.344, etc.)
      d. may convert interval data to ranked/categorical data (but not vice versa)

D. Identify variables and measurement scale (variable ID practice)

E. SYSTAT Demo

1. windows (output, data, graph); menus
2. data files
   - columns (variables); numerical vs. string (categorical) variables (e.g., SEX vs. SEX$)
   - rows (values of variables [cases, observations, sample size])
3. creating data files (entering and editing data)
   - raw data file (stacked [=indexed] vs. unstacked data
4. opening existing data files (.SYZ files)
5. graphing frequency distributions (*Graph→Histogram*)
6. creating frequency tables (*Analyze→ One-Way Frequency Tables*; (*Analyze→ Tables→ Two-Way*)
7. calculating an average (*Analyze→Basic Statistics*)
8. selecting cases (*Data→Select Cases*)
9. analyze by groups (*Data→By Groups*); groups = categories
10. transforming data (*Data→Transform→Let* and *Data→Transform→If..,Then Let*

Introduction to SYSTAT

Prepare a SYSTAT data file using the data below. These data are measurements taken from 10 specimens of spiny guanotzits from Arkansas and Missouri. The variables are: collection locality (categorical), length of body (continuous), sex (categorical), weight of body (continuous), amount of pigment on the lower jaw (ranked), and number of scales on the chin (discrete).

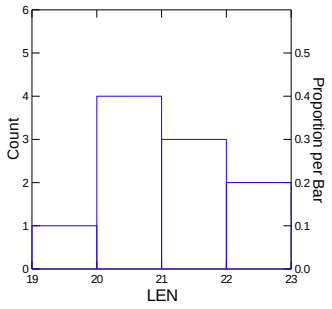
Case	1	2	3	4	5	6	7	8	9	10
Locality	AR	AR	MO	MO	MO	AR	AR	MO	AR	MO
Length (mm)	22.5	21.4	20.8	20.6	19.8	20.1	22.3	21.7	20.4	21.1
Sex	m	m	f	f	f	f	m	f	m	f
Weight (g)	333	298	401	257	21	30	478	400	35	288
Pigment	4	5	5	3	2	1	1	5	4	5
No. scales	23	22	14	26	9	21	17	12	15	12

Name your data file first.syz (the file extension .syz identifies a SYSTAT data file). After you finish entering the data, proofread the file to make sure that the data are correct, edit if necessary, save the file and close it. Reopen the file and use it to learn the following menus and functions:

- File Menu (New, Open, Save, Save As, Print, Exit)
- Edit Menu (Undo, Cut, Copy, Paste, Copy Graph, Delete, Options)
- Data Menu (Variable properties, Transform [Let and If - Then Let], By Groups, Select Cases)
- Graph Menu (Histogram)
- Analyze Menu (One-Way Frequency Tables, Basic Statistics, Tables)

Exercises

1. calculate the average guanotzit weight (254.1g)
2. calculate the average guanotzit weight separately for males and females (m=286.0g; f=232.8g)
3. calculate the average weight for guanotzits from Arkansas (234.8g)
4. draw a histogram of guanotzit lengths
5. transform weight to the common logarithm of weight



(case 1: 333.0 to 2.522)

6. create the new variable USE$ and let its value ("yes" or "no") be determined by a combination of values of the variables SEX$ and LOC$. Example: If SEX$="m" and LOC$="AR"… Then Let USE$="yes." Notice that the variable USE$ is a derived variable, not a measured variable
7. how many quanotzits from Missouri were measured? (n=5)
8. determine the number of guanotzits by scale number and state

LOC$ by NOSCALES

	9	12	14	15	17	21	22	23	26	Total
AR	0	0	0	1	1	1	1	1	0	5
M	1	2	1	0	0	0	0	0	1	5

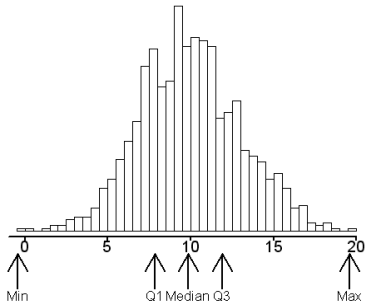
****

Description of Data (from a frequency distribution)

A. Descriptive statistics

1. measures of **central tendency**
   a. mode - most frequent class (of frequency distribution)
   b. median (ordinal or ratio/interval data) - middle class
   c. mean (ratio/interval data) = "average"; $\Sigma x/n$
   d. weighted mean (ratio/interval data) - $\Sigma fx/n$; used when cases have different levels of importance (weights); e.g., grade point average

2. measures of **dispersion** - describe the amount that each observation is likely to vary from the mean/median
   a. maximum, minimum (range): sensitive to extreme values
   b. interquartile range: (quartiles, middle 50% of observations (Q3 – Q1; difference between 25^{th} and 75^{th} percentiles)
   c. sum of squares(SS): $\Sigma(x - \bar{x})^2$
   d. variance: SS/n
   e. standard deviation: √variance



3. symbols for statistics (sample) and parameters (population)

	Parameter	Statistic
Mean	$\mu = \Sigma x/n$	$\bar{x} = \Sigma x/n$ (="x-bar")
Variance	$\sigma^2 = \Sigma(x-\mu)^2/n$	$s^2 = \Sigma(x-\bar{x})^2/n-1$
Standard Deviation	$\sigma = \sqrt{\sigma^2}$	$s = \sqrt{(s^2)}$ (="SD")

4. coefficient of variation (CV)

-expresses SD as a percent of the mean a. **CV = (SD/ $\bar{x}$) 100**

-used to compare relative variation in **one variable** between groups with different means

Example:

	mean	SD	CV	Note that group 2 is relatively more variable despite a greater SD in group 1.
Group 1	14.2	2.5	17.6	
Group 2	7.2	1.8	25.0	

B. Calculating Descriptive Statistics (mean ± SD)

1. calculate descriptive statistics from ***raw data file***; *Analyze→Basic Statistics*

-Use ***CAVESALYS.SYZ*** (Sanders)

QUESTION: What are the descriptive statistics of snout-vent-length for female salamanders collected in Arkansas?

2. calculate descriptive statistics from ***frequency distribution***
   - Step 1: *Data→Case Weighting→By Frequency*
   - Step 2: *Analyze→Basic Statistics*

Value of Variable	No. times observed
0	7
1	24
2	93
3	99
4	24
Total	247

C. How to report sample means (must include a measure of error)
   a. Text (example)
   b. Tables (example)
   c. Graphs (error bars; example)

-how reduce variance? (Σ**(x - $\bar{x}$)/n)**
-what limits n? (availability, money, time)

You now have sufficient knowledge to begin the Graph Construction Exercise on p15.

Probability Distributions (expected probabilities associated with all possible outcomes)

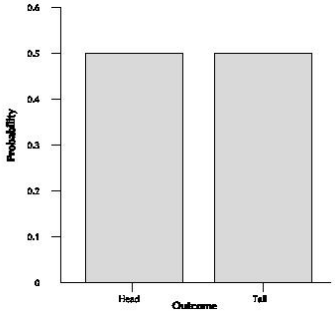
- cannot know if experimental result is due to chance alone unless we know what the expected is (hypothesis testing - basis for much of much of this course!)
- basic question: How well does an observed frequency distribution fit an expected frequency distribution? (goodness of fit - **GOF**)

Discrete probability distribution – Binomial (mutually exclusive categories; either/or); e.g., male/female, red/white, red/not red

Probability Basics

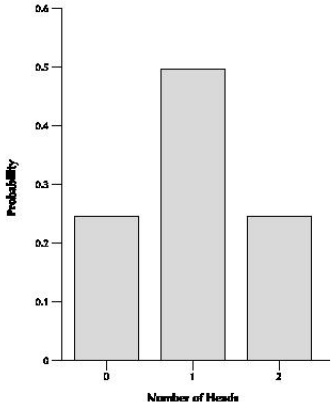
Example: 1 coin toss- possibilities: 1H, 1T

a. probabilities: no. ways an event (H or T) can occur /total no events (2) possible; "**division**" rule; 1H [1/2] = 0.5; 1T [1/2] = 0.5
b. add all possibilities = 1 [0.5 + 0.5 = 1]
c. probability distribution shape



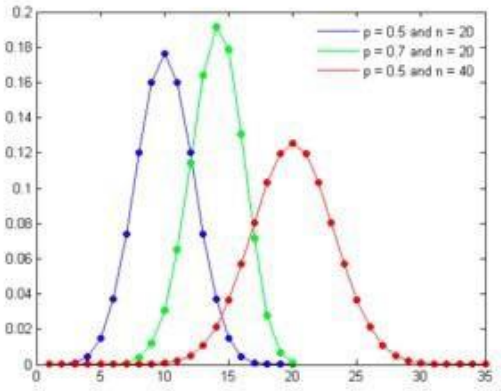
Example: 2 coin toss- possibilities: 2H, 1H1T, 1T1H, 2T (mutually exclusive, independent events)

a. probabilities:
   1) simultaneous events ("**and**" rule, **multiply**): 2H [0.5 **x** 0.5] = 0.25; 2T [0.5 x 0.5] = 0.25
   2) alternative events ("**or**" rule, **add**): 2HT [0.5 x 0.5] **+** [0.5 x 0.5] = 0.5
b. add all probabilities [0.25 + 0.5 + 0.25 = 1]
c. probability distribution shape



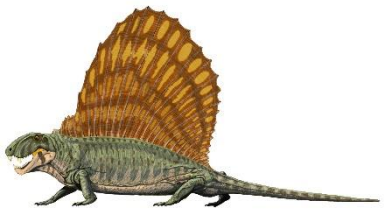
Binomial Distribution

1. formula: $\mathbf{P(x) = (n!/(x!(n-x)!))p^{x}q^{(n-x)}}$
   -no need to memorize the formula but you must be able to recognize the formula and each of its terms
2. terms
   P = probability of the number of occurrences of the event of interest
   p = probability of event of interest = head ("success")
   q = probability of other event (1-p) = not head ("failure")
   n = number of "simultaneous" events (trials)
   x = number of occurrences of the event of interest
3. binomial shape determined by values of n and p



EXAMPLE: A reproductive physiologist counted the number of males in 247 litters of 4 siblings each in a species of *Dimetrodon* (Table). Do these data support the hypothesis of sex being determined by a XX, XY system as occurs in mammals?

No. males Observed	Observed frequency	Expected frequency
0	7	
1	24	
2	93	
3	99	
4	24	
Total	**247**	**247**



Based on the theory of sex determination in mammals (equal chance of being male or female), calculate the expected frequencies for the number of males in these litters.

P(x) = (n!/(x!(n-x)!))$p^x q^{(n-x)}$		Expected proportion		No. litters		Expected number(frequency)
prop (0 males) =	(4!/(0!(4-0)!)) x 0.5^0 x $0.5^{(4-0)}$ =	**0.0625**	x	247 =		**15.438**
prop (1 male) =	(4!/(1!(4-1)!)) x 0.5^1 x $0.5^{(4-1)}$ =	**0.2500**	x	247 =		**61.750**
prop (2 males) =	(4!/(2!(4-2)!)) x 0.5^2 x $0.5^{(4-2)}$ =	**0.3750**	x	247 =		**92.625**
prop (3 males) =	(4!/(3!(4-3)!)) x 0.5^3 x $0.5^{(4-3)}$ =	**0.2500**	x	247 =		**61.750**
prop (4 males) =	(4!/(4!(4-4)!)) x 0.5^4 x $0.5^{(4-4)}$ =	**0.0625**	x	247 =		**15.438**
Total		**1.00**				**247**

SYSTAT calculation **of expected frequencies** (*Utilities→Probability Calculator→Univariate Discrete*)

Question: Is the sex of *Dimetredon* determined by a mechanism similar to that of mammals? Expect 1:1. Compare observed with expected.

No. males Observed	Observed frequency	Expected frequency	Conclusion: because of the large deviations between the expected and observed numbers, we reject the idea of there being equal chances of having equal sexes.
0	7	15.438	
1	24	61.750	
2	93	92.625	
3	99	61.750	
4	24	15.438	
Total	**247**	**247**	

So, what determines sex in Dimetredon?

Importance of sample size for observed data (1 coin example, compare to theoretical)

- IF observed = norm coin, THEN the larger the n, the closer we approximate expected
  conversely, THEN the smaller the n, the more we deviate from expected

Exercise: Binomial Distribution

Assuming that the sex of hatchling turtles is determined by a particular combination of chromosomes as in mammals (i.e., an XX, XY system), fill in the expected frequencies below:

Data are number of male hatchlings emerging from 84 nests of kaw turtles (kaw turtles always lay 6 eggs per nest).

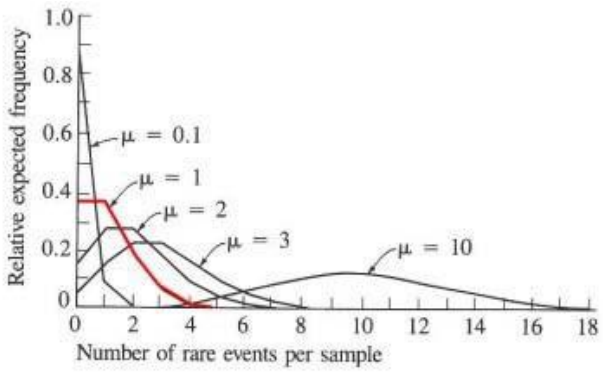
No. Males Observed	Observed No. Nests	Expected No. Nests
0	4	
1	7	
2	15	
3	24	
4	22	
5	7	
6	5	
Total	**84**	**84**

Compare the observed and expected frequencies. Do these data support the hypothesis that sex of hatchlings is genetically determined? (yes or no) Support your conclusion.

ans: exp- 1.310, 7.875, 19.688, 26.250, 19.688, 7.875, 1.310

Discrete probability distribution - Poisson (expected distribution for rare and random events)

1. Poisson: $\mu = \sigma^2$ ($\sigma^2/\mu = 1$) - distribution defined by mean only; low value (rare events; e.g., recapture rates, bacterial viruses infecting bacteria)
2. Poisson formula: **$P(x) = (\bar{x}^x e^{-\bar{x}})/x!$**
   -Students: no need to memorize the formula but you must be able to recognize the formula and each of its terms
3. terms
   -P = probability of the number of occurrences of the event of interest
   - $\bar{x}$ = mean occurrence of event of interest
   - e = mathematical constant (=2.71828)
   - x = number of occurrences of the event of interest
6. Poisson shape determined by $\bar{x}$



Example: An ecologist counted the number of maple seedlings in 100 quadrats

No. Plants	Obs. No. Quadrats	Exp. No. Quadrats
0	35	
1	28	
2	15	
3	10	
4	7	
5	5	
Total	**100**	**100**

Using the mean calculated from the observed frequency distribution of maple seedlings per quadrat in the table ($\bar{x} = 1.41$), calculate the expected frequencies assuming that occurring in a quadrat is a random event.

	Expected proportion	Expected number (frequency)
prop (0 seedlings) = $(1.41^0 e^{-1.41})/0!$ =	**0.244** x 100 =	**24.41**
prop (1 seedling) = $(1.41^1 e^{-1.41})/1!$ =	**0.344** x 100 =	**34.42**
-etc.		

SYSTAT calculation of expected frequencies (*Utilities→Probability Calculator→Univariate Discrete*)

Question: Do seedlings occur randomly in quadrats?

No. Plants	Obs. No. quadrats	Exp. No. Quadrats	Conclusions:
0	35	24.41	1. Is it rare? (mean=1.41)
1	28	34.42	2. Is it random?
2	15	24.27	a. compare obs and exp distributions
3	10	11.41	b. calculate variance/mean ratio (2.18/1.41=1.55)
4	7	4.02	
5	5	1.11	
Total	**100**	**100**	

Exercise: Poisson Distribution

Assuming that being killed by a horse is a rare and random event, fill in the expected frequencies below.

Men killed by being kicked by a horse in the Prussian Army Corps.

No. killed/ yr/corps	Observed Number	Expected Number
0	109	
1	65	
2	22	
3	3	
4	1	
Total	200	200

ans: exp- 108.67, 66.29, 20.22, 4.11, 0.63

$\bar{x}$ = _____(ans: 0.610)

s^2 = _____(ans: 0.611)

$s^2/\bar{x}$ =_____(ans: 1.002)

Compare the observed and expected frequencies.

Do these data support the hypothesis that the chance of being killed by a horse in the Prussian Army Corps is a rare and random event? Support your conclusion.

Exercise: Testing Your Concept of Randomness

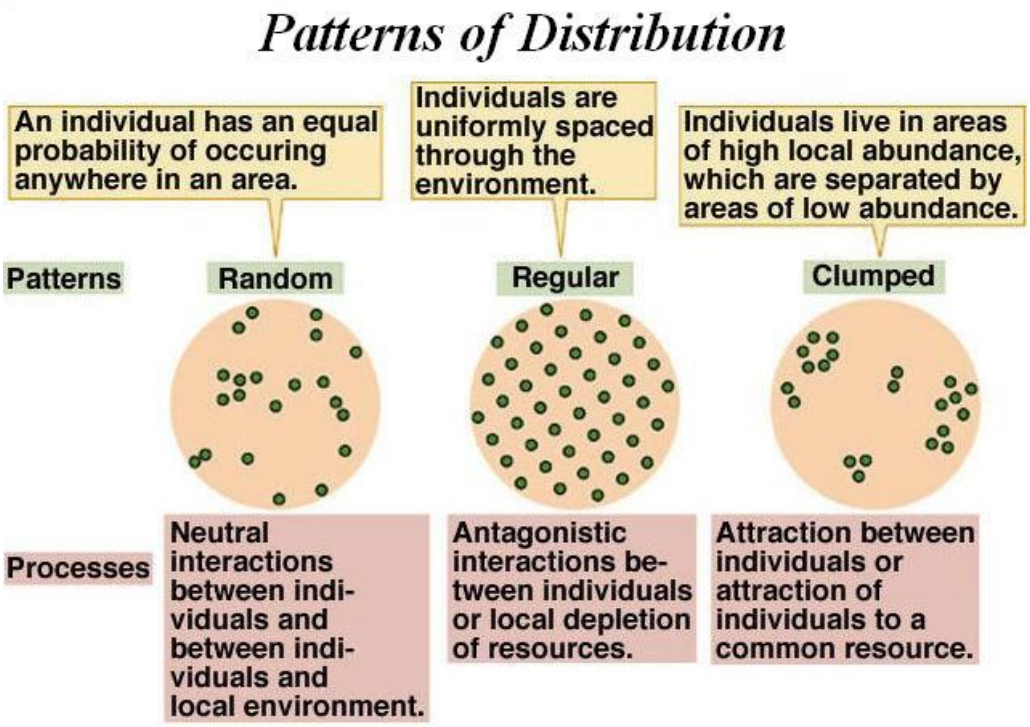
1. draw 100 dots on the 10x10 grid on the next page (keep your eyes open, try to place dots randomly
2. count the number of cells with different numbers of dots
3. create a frequency table of your data
4. calculate the mean and variance of the number of dots per cell

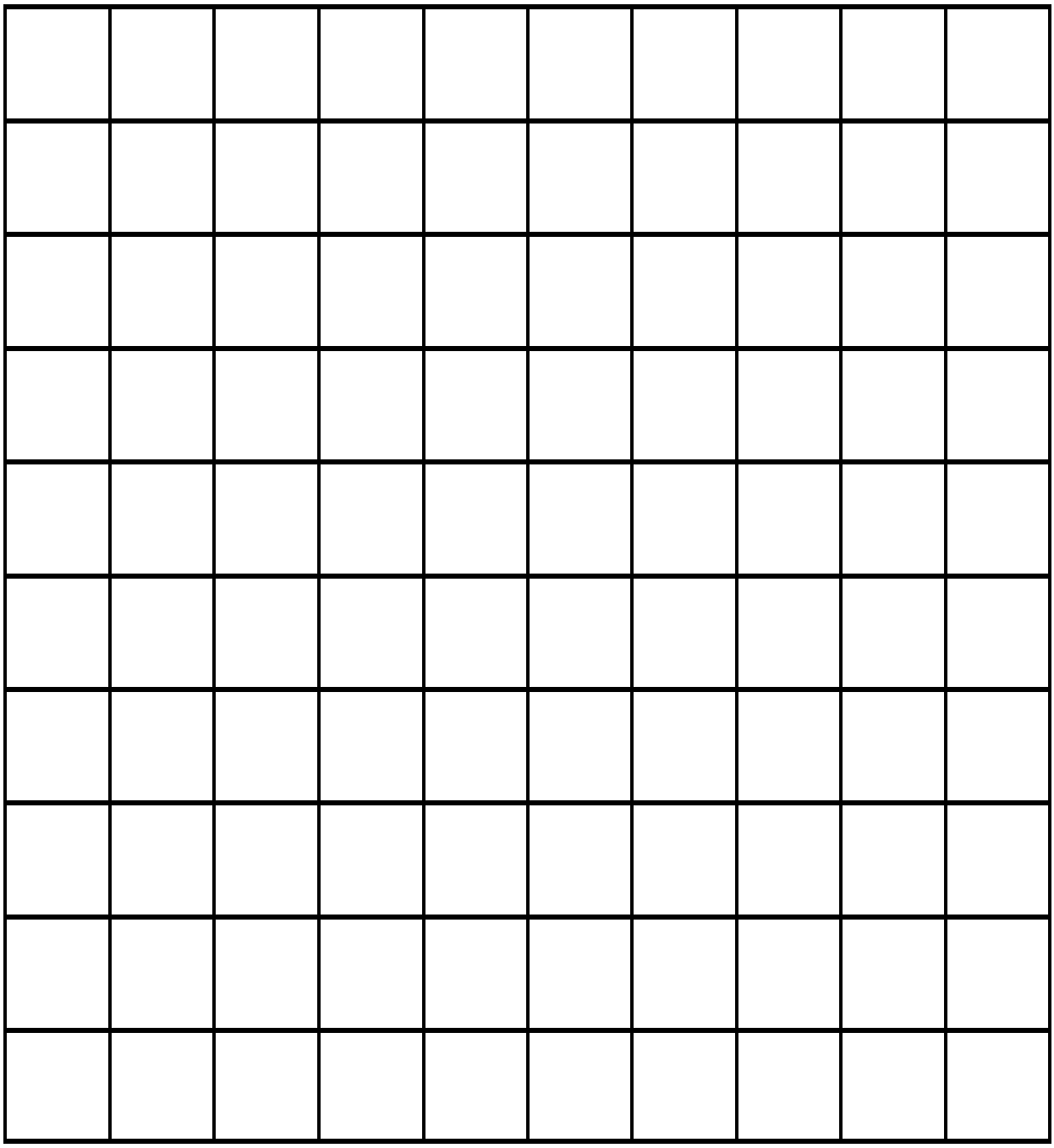
   mean = variance =

5. calculate the variance/mean ratio =

6. interpret: ratio = 1 (random); ratio <1 (evenly spaced); ratio >1 (clumped)

7. Application: patterns of distribution in space reflect biological processes; for example, disease spread and behavioral/ecological interactions





Graph Construction

Home

In this exercise, you will learn to construct five basic graphs used by biologists. The rules for graph construction presented here will **apply to all graphs you construct during the semester**. As you finish graphs, copy and paste each image to a Word file named *graphexercise,* add the caption, and save. There are three parts to the exercise:

1. You will reproduce 5 finished graphs given to you;
2. You will be given data and asked to construct 5 appropriate graphs;
3. You will find an example of each of the 5 graph types in the primary literature.

A. Basic graph types

1. **Histogram** (*Graph→Histogram*) - plots the frequency (counts/proportions/percentages) of occurrence as a bar on the Y-axis against a variable on the X-axis
2. **Bar** (*Graph→Bar*) - plots the mean and error bars of a variable as a bar on the Y-axis against a categorical variable on the X-axis
3. **Dot** (*Graph→Summary Charts→Dot*) - plots the mean and error bars of a variable as a symbol on the Y-axis against a categorical variable on the X-axis
4. **Box Plot** (*Graph→Box Plot*) – plots the median and quartiles of a variable on the Y-axis against a categorical variable on the X-axis
5. **Scatterplot** (*Graph→Scatterplot*) – plots cases of one variable on the Y-axis against cases of another variable on the X-axis

Requirements of all graphs

- The Y variable is always read before the X variable. For example, "plot Y against X", "plot Y by X", and "Y is regressed against X". For this class, X is never plotted against Y.
- Essential graph elements: axes (Y, X), axis labels (with units of measurement, if applicable), ticks, tick labels, caption
- Elements essential for specific graph types: bars, symbols, error bars, data points, line, linear smoother
- Each graph must be self-explanatory and be able to stand alone (figure captions are considered part of the graph). Captions should be descriptive, not interpretative.
- Non-standard abbreviations must be defined.
- Graphs displaying means (Bar, Dot) must portray the mean, error bars, and sample size **for each mean**.

B. **Graph reproduction** - Reproduce each graph (1-5) illustrated below. Read the description of each data file before beginning. Copy and paste your SYSTAT output into a Word file named *graphexercise*, add captions, and save.

1. HISTOGRAM - A SYSTAT Histogram plots the frequency (counts/proportions/percentages) of a single variable. Duplicate the Histogram below. Note axis titles, axis ranges, data plotted, bar fill, etc. The data are in RANDOM.SYZ (Plummer).

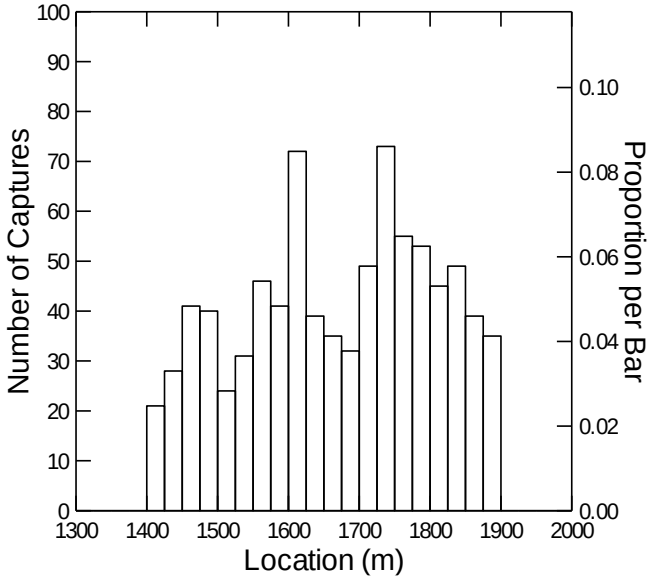


Fig. 1. The distribution of captures of green snakes according to location.

2. BAR - A SYSTAT Bar graph plots the mean of one variable against another variable. Duplicate the BAR graph below. Note bar fill, axis titles, error bars, data plotted, etc. The data are in MOUSEDIET.SYZ (Cooper).

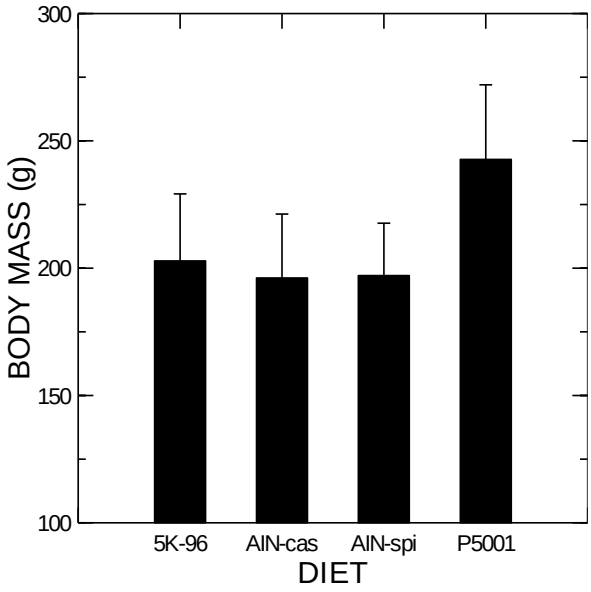


Fig. 2. The relationship of mean body mass and diet in laboratory mice fed different diets. Plotted are mean ± 1 SD. Sample sizes are: 5K-96, n=34; AIN-cas, n=35; AIN-spi, n=32; P5001, n=42.

3. DOT - A SYSTAT Dot graph plots the mean of one variable against a discrete or categorical variable. Duplicate the Dot graph below. Note symbols, error bars, fill, axis titles, axis ranges, data plotted, etc. The data are in WORMSURVIVE.SYZ (JMGoy).

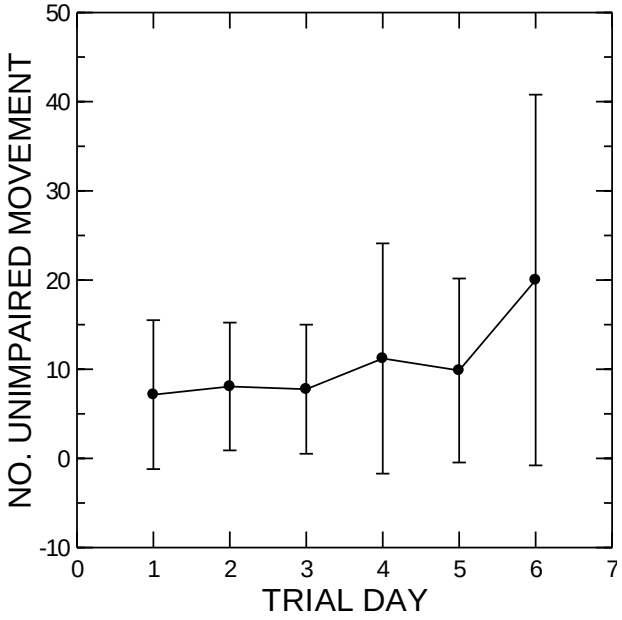


Fig. 3. Mean number of *C. elegans* exhibiting unimpaired movement according to trial day. Plotted are mean ± 1 SD. Sample sizes are day 1, n=48; day 2, n=51; day 3, n=49; day 4, n=15; day 5, n=7; day 6, n=2.

4. BOX – A SYSTAT Box Plot plots the quartiles of one variable against a discrete or categorical variable. Duplicate the Box Plot below. Note symbols, axis titles, axis ranges, selected data plotted, etc. The data are in CAVESALYS.SYZ (Sanders).

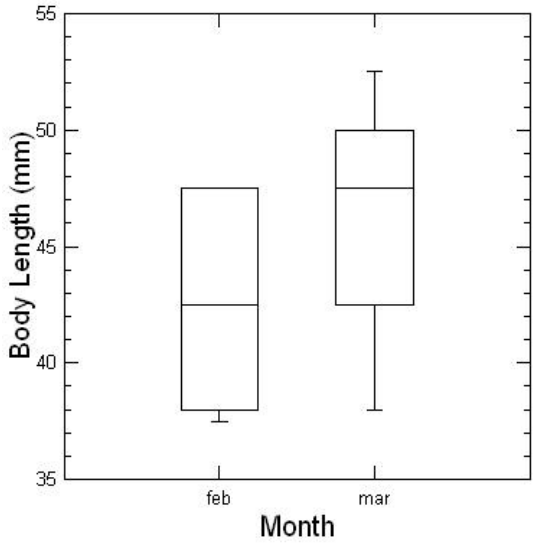


Fig. 4. Box plot of the body lengths of female *Eurycea lucifuga* captured in Arkansas and Kentucky caves in February and March. Plotted are the median (horizontal line), the 25th and 75th quartiles (box) and the maximum and minimum values (whiskers).

5. SCATTERPLOT - A SYSTAT Scatterplot plots individual cases of one variable against another variable. Duplicate the scatterplot below. Note symbols, axis titles, axis ranges, selected data plotted, etc. The data are in LONOKE.SYZ (Plummer).

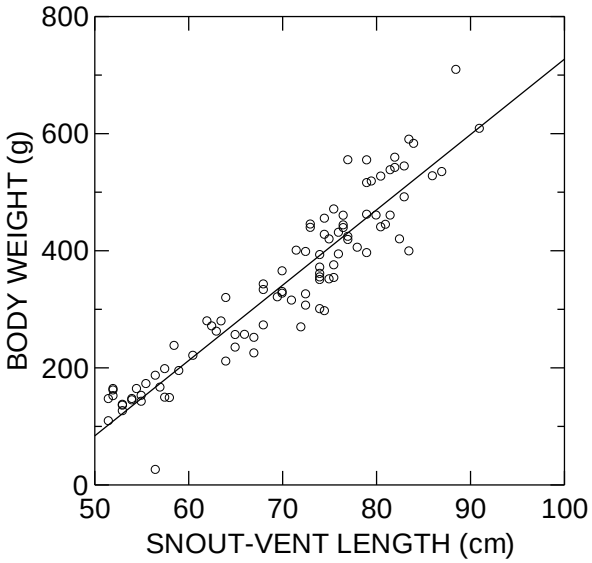


Fig. 5. The relationship of body weight and snout-vent length in 99 adult (=individuals >50 cm SVL) male diamondback water snakes.

C. **Graph construction**: Construct an appropriate graph for each of the following problems and save in your *graphexercise* file.

6. Use the following data on bill lengths (mm) of 42 belted kingfishers to construct a graph (Fig. 6) that plots the median and other quartiles separately for males, females, and the sexes combined (3 groups).

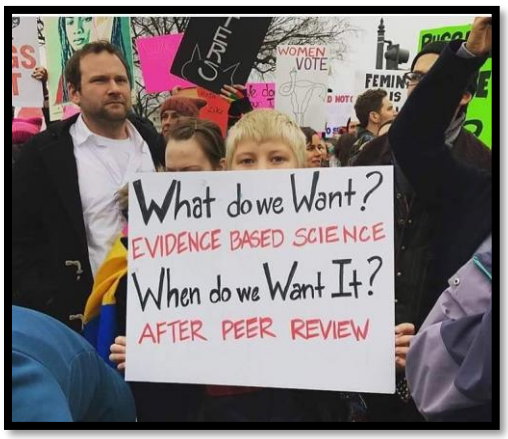
males: 48.1, 47.7, 48.0, 50.6, 50.8, 49.9, 49.3, 50.8, 46.9, 49.9, 48.8, 47.5, 48.2, 51.0, 48.8, 52.0, 51.8, 51.0, 50.1, 47.7, 49.9
females: 53.8, 59.2, 52.3, 59.3, 56.5, 56.2, 55.6, 57.7, 52.5, 47.8, 51.5, 55.8, 57.5, 56.8, 47.0, 50.4, 58.0, 61.2, 56.5, 59.3, 59.2

For graphs 7-10, use the data file LONOKE.SYZ (Plummer).

7. Construct a graph (Fig. 7) that plots cases of weight against length for snakes collected in ponds #53 and #54. Indicate sample size.
8. Construct a graph (Fig. 8) that illustrates the mean body weight for each sex. Restrict cases to snakes ≥30 and ≤90 cm SVL. You can more easily make the X-axis readable by creating a derived variable with this transform: IF sex=1 THEN LET sex$="male"
9. Construct a graph (Fig. 9) that illustrates the frequency of female snakes captured in minnow ponds by snout-vent length. Indicate sample size.
10. Transform variable WGT with common logarithms. Construct a graph (Fig. 10) that plots cases of the transformed variable against SVL. Indicate sample size.

D. **Literature Graphs**: The third part of this exercise consists of finding an example of each of the five graph types in primary literature papers.

What is the Primary Literature?—Journals (evidence-based science; ID)



1. Original research written by the researcher
2. Peer reviewed
3. Publishing process
4. Some useful working categories
   a. First tier—*Science*, *Nature*
      - Broad subject content
      - Publish only the best of the best
      - Papers usually report a major advance in the field

   b. Second tier—*Proceedings of the National Academy of Sciences*, *Ecology*, *Cell*
      - Content frequently has restricted subject areas
      - Publish most of the top papers in that subject area
      - Reject many technically sound papers if they do not advance our knowledge sufficiently

   c. Third tier—*Journal of Herpetology*, *American Midland Naturalist*, *Journal of Immunology*
      - Content limited in subject area and/or geographical coverage
      - Publish the bulk of papers in the subject area
      - Most technically sound papers are accepted even if they do not dramatically advance our knowledge

Structure of a Primary Literature Paper

1. Abstract
   -provides an overview of the paper
2. Introduction
   -provides a theoretical framework for the study
   -provides an overview of what is already known
   -clearly states the question and why it is important
3. **Materials and Methods**
   -provides details of the experimental design
   -provides details about how the data were collected and analyzed (including **statistical analysis**)
4. **Results**
   -provides a textual description of the results of analyses
   -provides **tables** and/or **graphs** showing quantitative and **statistical** results of analyses
5. Discussion
   -compares the results to what was previously reported in the primary literature
   -points out how the results either strengthen or weaken current theoretical models
   -if appropriate, makes suggestions on how theoretical models should be modified
   -highlights questions in need of further research

6. Literature Cited

-contains the full citation for every paper cited in the text. Does not contain citations that are not cited in the text

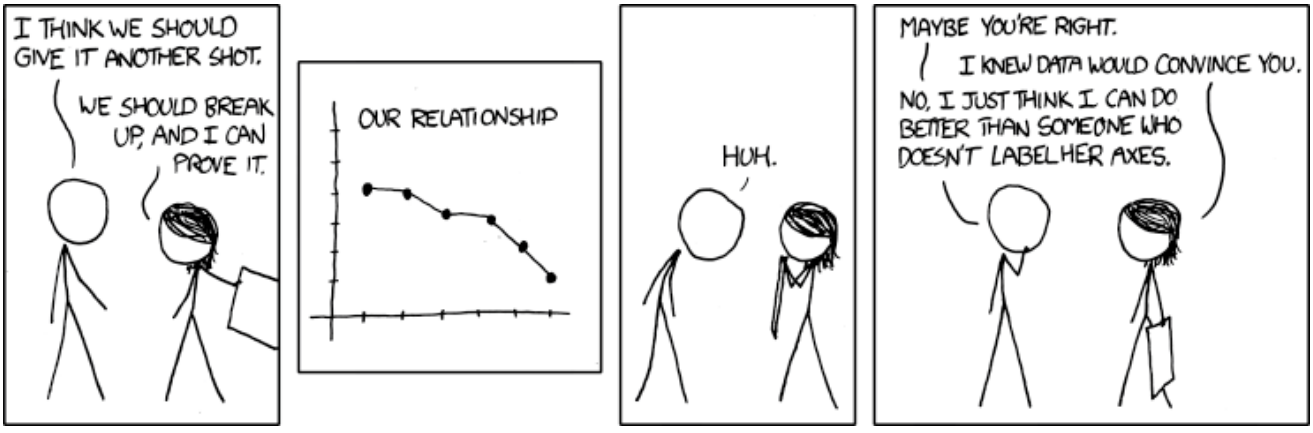
As you locate an example of each graph in the literature, download a digital copy, insert into *graphexercise*, and save in order - Fig. 11 Histogram, Fig. 12 Bar, Fig. 13 Dot, Fig. 14 Box, and Fig. 15 Scatterplot. Make sure to include the caption. Under each graph caption, type the citation of the paper where you found the graph. Proper citation format is: last name, initials, initials, last name, and initials, last name. year. title. journal volume:pages. Here's an example;

Harless, M.L., A.D. Walde, and D.K. Delaney. 2010. Sampling considerations for improving home range estimates of desert tortoises: effects of estimator, sampling regime, and sex. Herpetological Conservation and Biology 5:374-387.

Note: Histogram, Bar, Dot, Box, and Scatterplot are names given to particular graphs by SYSTAT. You may find different names in other statistical software and in the literature; for example, a histogram may be called a frequency distribution or a bar graph. Don't let that confuse you! You should be skilled enough to quickly determine the type of graph just by looking and applying your knowledge. For example, ask yourself what statistic is plotted on the graph; is it frequencies, means, medians, or individual cases?

Turn in a printed copy of *graphexercise* on the due date. Print two graphs per page. Do not separate the graphs from their respective captions.

How to Search Primary Literature (Google Scholar; Library)



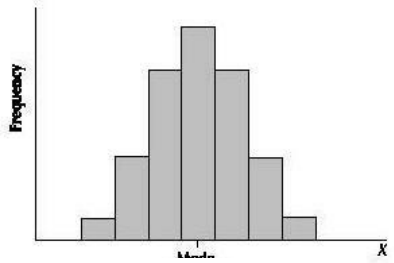
Inferential Statistics

Home

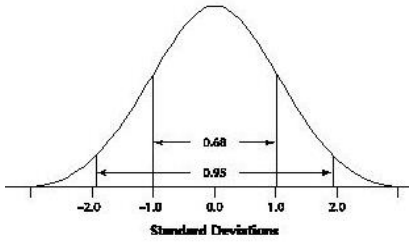
The Normal distribution

-very important frequency distribution for 2 reasons:

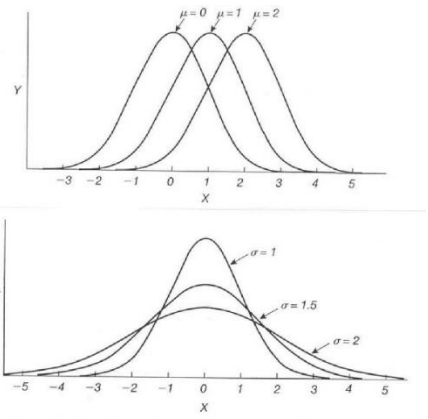
A. Data that are influenced by many small and unrelated random effects are approximately normally distributed (math: Fuzzy Central Limit Theorem); extremely widespread and common in nature
B. Forms the conceptual basis of a large number of statistical procedures - one of the most important theoretical distributions in statistics
C. Properties
   1. formula: $\mathbf{1/(\sigma\sqrt{2\pi})exp(-(x-\mu)^2/2\sigma^2)}$
   2. students – no need to memorize the formula but you must be able to recognize it
   3. shape determined by mean and SD
   4. symetrical around the mean (mean=mode=median)



   5. $\bar{x}\pm1SD$ = approx. 68% of cases; $\pm2SD$ = approx 95%



D. Standard normal distribution
   1. many different "normal" distributions
   2. standardize any normal distribution (directly compare)
   3. express individual cases in terms of SND; $z = (x - \bar{x})/s$; "z-score"
   4. z-score = distance from mean in standard deviation units; e.g., z = 1 (=1SD greater than the mean)
   5. Areas of normal curve (Tables)



E. Testing observed data for normality; SYSTAT output (TREAT.SYZ, EGGWGT)
   1. qualitative: Probability plot (*Graph→Distribution Plots→Probability Plot*): **DEMO**
   2. quantitative: Kolmogorov-Smirnov Test: **DEMO**
   3. SYSTAT path*: Analyze→Nonparametric Tests→One-sample KS (Enter selected variable and Lilliefors distribution*)
      - hypothesis: frequency distribution of EGGWGT is normally distributed
      - test statistic, probability

- if probability ≤0.05, reject the hypothesis; conclusion: EGGWGT distribution is not normally distributed (="skewed")
- if probability >0.05, cannot reject the hypothesis; conclusion: EGGWGT distribution is normally distributed

---

Exercise: practice SYSTAT Probability Plot and One-sample KS Test using the variable H2OOUT from file DLWMEANS.SYZ. Note that H2OOUT is not normally distributed (skewed)

-Data transformation has the potential to normalize non-normal data)

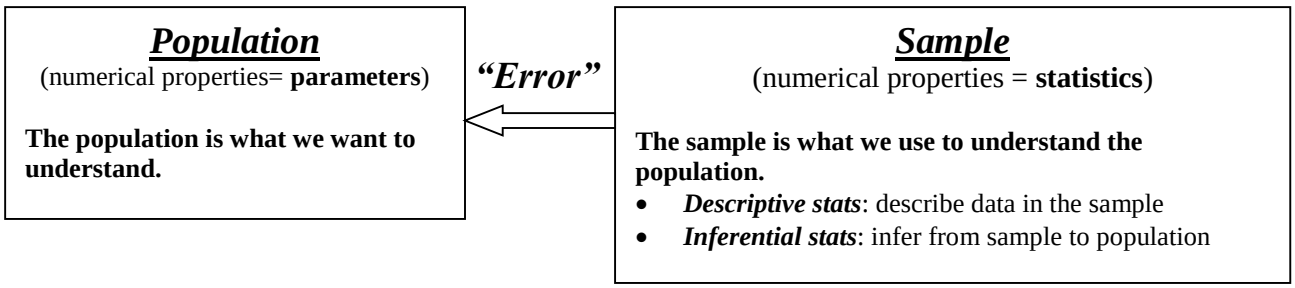
1. Data transformations - many procedures in statistics assume that data are normally distributed. If data are not normally distributed, one can transform the data to another measurement scale in an effort to normalize them. Deciding which transformation to use is entirely practical, i.e., the "right" transformation is whatever makes the data normally distributed. Trial-and-error applications of various transformations may be necessary to determine which will work. However, some transformations work better in some situations than in others. Examples of transformations commonly used in biology are the logarithmic, arcsine, and square-root transformations.

   - the logarithmic transformation is useful in a wide variety of situations and is by far the most commonly used transformation in biology
   - the arcsine (inverse sine) transformation is used specifically when data are in the form of proportions or percentages
   - the square-root transformation is used specifically when data are in the form of counts

2. Transform the variable H2OOUT with common logarithms and retest for normality with both Probability Plot and KS. Note that the SYSTAT designation for common logs is **L10** (always use common logs in Biol. 254). After transformation, the new variable **L10H2OOUT** should now be normal

   Always create a NEW variable name for the transformed variable!

Statistical inference - draw conclusions regarding populations based on analysis of samples from those populations



1. Two major categories of statistical inference
   a. **Estimate parameters** (e.g., μ, σ)
   b. **Test hypotheses** (infer population from sample)
2. The foundation for both concepts is the **Sampling Distributioin**
   a. take repeated samples from population
   b. examine distribution of sample means
3. Two major predictions of the **Central Limit Theorem** regarding sampling distributions
   a. **Means of samples from a normally distributed population will be normally distributed**
      - mean of means = $\Sigma\ \bar{x}/n$
      - SD of means (=standard error of mean, SE or SEM); SE = SD/$\sqrt{n}$
   b. **Means of samples from a non-normally distributed population will be normally distributed if n is sufficiently large** (required n is proportional to amount of variation)

Simulation: Rice University Virtual Stats Lab

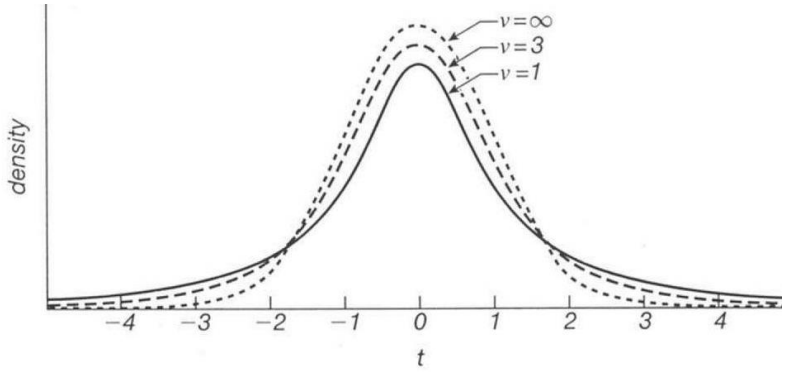
Estimation of parameters

1. How well does the sample mean ($\bar{x}$) estimate the population mean (μ)?
   a. in a normally distributed population, 95% of the **cases** lie between $\bar{x}$ - 1.96 SD and $\bar{x}$ + 1.96 SD
   b. in a normal sampling distribution, 95% of the **means** lie between $\bar{x}$ - 1.96 SE and $\bar{x}$ + 1.96 SE
   c. interpretation: 95% chance that population mean is enclosed within these limits (**95% confidence limits**)

> *Absolute certainty is a privilege of uneducated minds-and fanatics. It is, for scientific folk, an unattainable ideal. - C. J. Keyser*

   d. problem: sampling distributions of means may depart from normality if sample size is small (central limit theorem)
   e. solution: use distribution that adjusts for sample size - Student's t-distribution (shape determined by 3 characteristics):

- mean, SD, df
- areas of curve that exclude a given proportion of the distribution vary with n (Tables)
- at infinity df, $t_{0.05}$ = 1.96 *as in normal distribution*

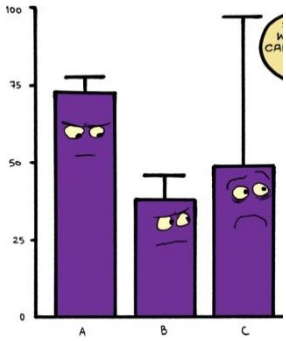


f. to calculate 95% CLs using a t-distribution, replace 1.96 with value from t-table
- UL: mean + ($t_{[0.05,\ n-1]}$) x SE
- LL: mean - ($t_{[0.05,\ n-1]}$) x SE

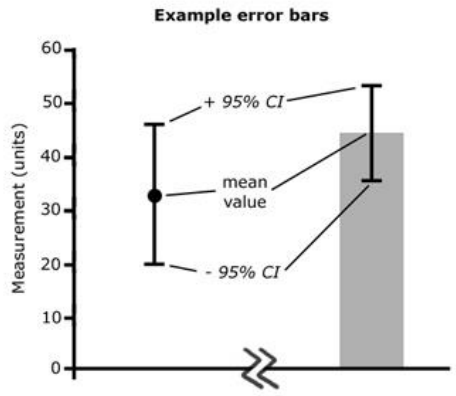
g. examples: calculate 95% CLs for these sample means:
- $\bar{x}$ = 4.7, SD = 0.27, N = 25 95% CI = **4.58 – 4.81** (higher n; narrower CLs)
- $\bar{x}$ = 4.7, SD = 0.27, N = 7 95% CI = **4.45 – 4.95** (lower n; broader CLs)

2. 95% CL in the public media: GPS accuracy, political polls, church surveys

3. How to report sample means
- $\bar{x}$ ± SD - provides idea of how much variation there is in the data but does not provide information on how well statistic $\bar{x}$ estimates parameter μ
- $\bar{x}$ (95% CLs) provides information on how well $\bar{x}$ estimates μ and if two means are significantly different from each other
- **$\bar{x}$ ± 1SE** (**most common way of reporting means in text, tables, and graphs**)



Differences in means: graphic methods for 'informed guessing' whether means are statistically different



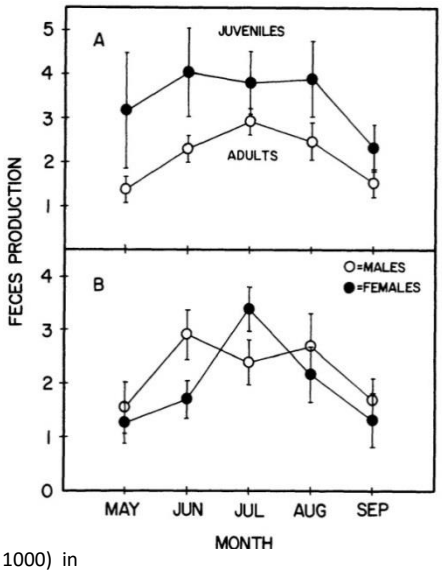


Fig. 1. (A) Feces production (x 1000) in juvenile and adult green snakes by month. (B) Feces production (x1000) in adult male and female green snakes by month.

Plotted are means ± 2 SE.

Caution! The most common way of reporting descriptive statistics in the literature is mean ± 1 SE; Proper interpretation requires that you visually double the value of the SE to get 95% CLs.

VS.

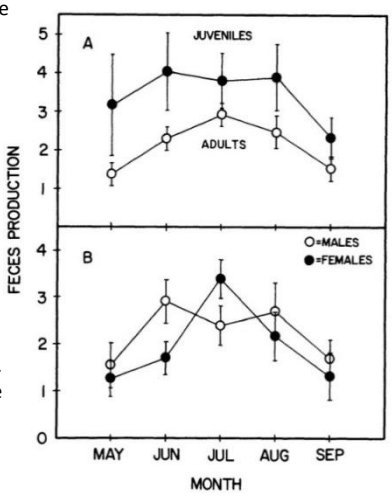


Fig. 1. (A) Feces production (x 1000) in juvenile and adult green snakes by month. (B) Feces production (x1000) in adult male and female green snakes by month.

Plotted are means ± 1 SE.

To properly interpret graphs displaying descriptive statistics, you must know what the error bars represent! (info found in the figure caption or in the M&M)

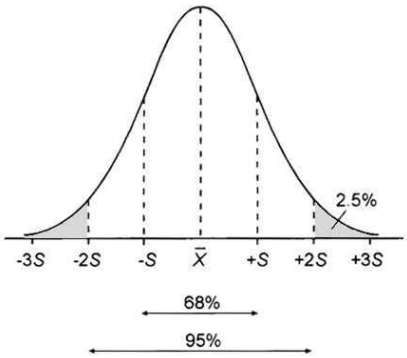
II. Hypothesis testing

A. Scientific hypothesis testing (sci_method)
   1. Scientific method (ID; science begins when we try to explain observations (hypothesis)
   2. Primary attributes of a good hypothesis
      a. if it is correct, then it will explain what has been observed (**consistent with observations**)
      b. if it is false, it can be shown to be false (**falsifiable**)
   3. Cannot prove a true hypothesis; science advances by disproving false hypotheses
   4. Process of hypothesis testing
      a. if-then logic (**IF** the hypothesis is true, **THEN** this should be the result); MP?

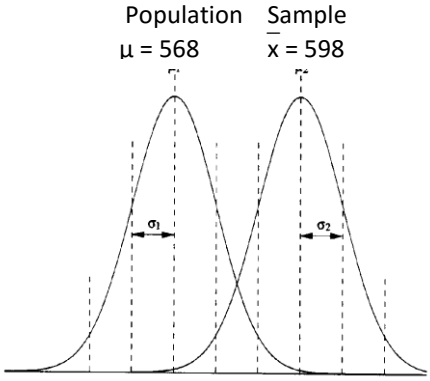
b. if testing results in something other than expected outcome, we reject hypothesis and look for a better explanation

B. Statistical hypothesis testing - similar procedure

1. State hypothesis such that there are only 2 possible outcomes, e.g.,
   a. H_A: A ≠ B (cannot test directly) = **research** [alternative] hypothesis
   b. H_0: A = B (if false; assume H_A by default) = **null** hypothesis



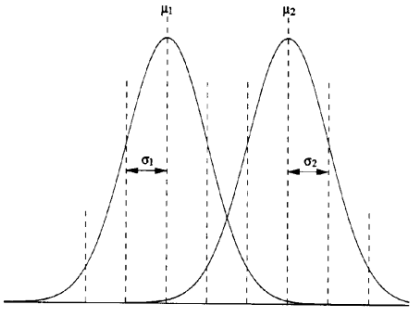
2. Example 1: compare case with known population
   H_0: case is from population
   H_A: case is not from population
   What is the probability that the null hypothesis is true?
   -if low, the research hypothesis is more likely true



3. Example 2: compare sample mean with known population
   - *SYSTAT: Analyze->Hypothesis Testing->Mean->one-sample t-test*

   a. SYSTAT (onesamplet.syz): $\bar{x}$ = 598; SD = 70.3; n = 30
   b. assume population mean is known [μ= 568]
   c. H_0: $\bar{x} = \mu$; H_A: $\bar{x} \neq \mu$
   d. calculate (SYSTAT); one-sample t-test; test statistic, t_{calc} = 2.31
   e. determine probability by comparing t_{calc} to t_{tab} (tabled value; df=29; Tables); P = between 0.02 and 0.05
   f. at **P=0.05** (alpha level); t_{tab} = **2.045** (critical value)
   g. t_{calc} (2.31) is greater than t_{tab} (2.045), therefore P<0.05
   h. two explanations for obtaining a high t value (2.31)
      - null hypothesis is true; sample mean differed by chance alone (unlikely)
      - null hypothesis is false (more likely)
   i. 1-sample t-test: rarely done in science… Why?

4. Example 3: compare two sample means (populations unknown - common question in many areas of biology)
   - *Hypothesis Testing 1* (next lecture section)



C. Writing null hypotheses for parametric difference tests and their nonparametric counterparts (does not include tests of frequencies or tests of relationships): **required components**

1. indicator (H_0)
2. parameter (e.g., μ, σ^2)
3. variable (e.g., length, mass)
4. group (e.g., sex, color); for questions of differences between independent data only (no grouping variable for dependent data)
5. relational operator (e.g., =, ≥, ≤)

-groups are designated by being enclosed in parentheses
-examples: independent: H_0: μlength(males) = μlength(females)
dependent: H_0: μbeforelength = μafterlength

D. Two-tailed vs. one-tailed hypotheses

1. two-tailed research hypothesis: $\mathbf{H_A: \mu_A \neq \mu_B}$ (non-directional)
   -null hypothesis (opposite of H_A:): $\mathbf{H_0: \mu_A = \mu_B}$

2. one-tailed research hypothesis: $\mathbf{H_A: \mu_A < \mu_B}$ (directional)
   -null hypothesis (opposite of H_A:): $\mathbf{H_0: \mu_A \geq \mu_B}$

3. one-tail: use only 1/2 of distribution (divide probability by 2)
4. how know if one-tail or two-tail? read question carefully

III. Statistical decision-making

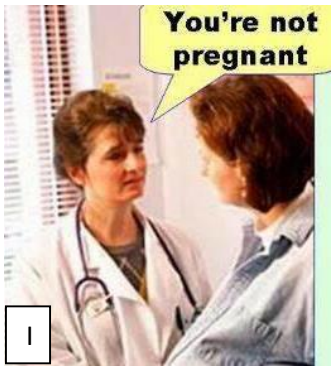
1. researchers set alpha level before statistical test is performed (usually 0.05)
2. onesample.syz example: what would happen if you changed alpha to 0.01 after the test was done? (Tables; t_{calc} (2.31) < t_{tab} (2.756; P>0.01)
3. possible to reject or not reject null hypothesis with the same set of data! Which one is "true?" (two types of errors)

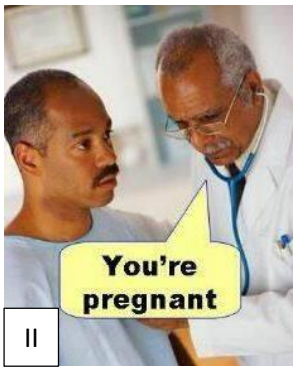
TRUTH TABLE	The real world; H_0 is actually:		
		TRUE	**FALSE**
Your analysis; you say that H_0 is:	true	Correct	Type II error
	false	Type I error	Correct

- type I error (rejecting a true null hypothesis); fixed value set by scientific community (P=0.05); make mistake 1 out of 20 times
- type II error (failure to reject a false null hypothesis); can be minimized by:
  1. increasing sample size
  2. choosing the most powerful test (power = probability of rejecting a false null hypothesis); minimum power of 80% generally necessary for an acceptable biological conclusion when you cannot reject the null hypothesis
- Why not reduce probability of type I error? – increases probability of type II error
- Alpha set at 0.05 because it represents a compromise between making type I and type II errors
- **SYSTAT** - how to calculate power or to determine minimum sample size needed for a specific power level (*Utilities->Power Analysis->specific test*)

4. Medical application of Truth Table – **Diagnostic Testing Outcomes**

	Disease present (pregnant)	**Disease absent** (not pregnant)
Test positive (cannot reject H_0; pregnant)	True positive	**False positive** (Type II error)
Test negative (reject H_0; not pregnant)	**False negative** (Type I error)	True negative





Reporting significance levels (**definition of "significant" = H_0 has been rejected**)

a. conventional method (non-exact probability from statistical table)
   - nonsignificant = $P>0.05$ = ns
   - significant = $P\leq0.05$ = *
   - highly significant = $P\leq0.01$ = **
   - very highly significant = $P\leq0.001$ = ***

Statistical decisions are always made at the P≤0.05 level.

Absolute certainty is a privilege of uneducated minds-and fanatics. It is, for scientific folk, an unattainable ideal. - C. J. Keyser

b. modern method (exact probability from computer calculation)
c. both methods are correct, so students may use either method in Biol. 254
d. "Statistically significant" is one of those phrases scientists would love to have a chance to take back and rename. "Significant" suggests importance; but the test of statistical significance, developed by the British statistician R.A. Fisher, **doesn't measure the importance or size of an effect**; only whether we are able to distinguish it, using our keenest statistical tools, from zero. "Statistically noticeable" or statistically discernable" would be much better." -Mathematician Jordan Ellenberg
e. if you are talking science, avoid using the non-qualified term "significant" in a non-statistical context

5. Why is it incorrect to "accept" a null hypothesis?
   a. it implies that the null hypothesis has been proven true (NO!); the null hypothesis is only assumed true
   b. legal analogy: defendant is assumed innocent until proven guilty (jury decisions: "guilty" or "not guilty")

c. modern experimental design was developed by Ronald Fisher (1930s). "…it should be noted that the null hypothesis is never proved or established, but is possibly disproved in the course of experimentation."

IV. Statistical Software (usually found toward the end of M&M in primary literature papers)

- SAS (no. 1 statistical software for scientists); high learning curve
- SYSTAT
- Minitab
- SPSS
- many others (http://en.wikipedia.org/wiki/Comparison_of_statistical_packages)
- Excel is not recommended for inferential statistical analysis.

<table>
<tr><th colspan="2">STATISTICAL TESTS
Home</th><th>Parametric
-more power, more assumptions</th><th>Nonparametric
-less power, fewer assumptions</th><td rowspan="5"><u>Assumptions of parametric tests</u>
• Data are randomly sampled and independent (except dependent designed tests(= repeated measures)
• Data are measured on ratio or interval scale
• Data (or residuals in ANOVA and regression) are normally distributed <u>for each group</u>
• For questions regarding means, the variances among groups (or residuals in ANOVA and regression) are homogeneous

<u>Assumptions of non-parametric tests</u>
• Data are randomly sampled and independent (except dependent designed tests)

_____

Tests covered on Exam II</td></tr>
<tr><td rowspan="3">Differences</td><td>Frequencies</td><td>-----</td><td>Goodness-of-fit (GOF)
• Chi-square
• Kolmogorov-Smirnov (KS)</td></tr>
<tr><td>Variances</td><td>Bartlett's</td><td>Levene's</td></tr>
<tr><td>2 Means</td><td>t- tests
• Independent samples t
• Paired samples t (assumes data are dependent)</td><td>Mann-Whitney

Wilcoxon (assumes data are dependent)</td></tr>
<tr><td rowspan="2">Relationships</td><td>>2 Means</td><td>Analysis of Variance
• One-way ANOVA
• Post-hoc pairwise comparisons (Tukey)
• Two-way ANOVA</td><td>Kruskal-Wallis
Post-hoc pairwise comparisons (Dwass-Steel-Critchlow-Fligner; DSCF)</td></tr>
<tr><td>Frequencies</td><td>-----</td><td>Test of Independence (contingency table analysis)
• Chi-square
• Fisher Exact Test</td></tr>
<tr><td rowspan="2">Variables/ Cases</td><td>Pearson correlation</td><td>Spearman correlation</td><td rowspan="2"></td></tr>
<tr><td>Linear Regression</td><td>-----</td></tr>
</table>

<u>How does one know which test is appropriate?</u>

- Read question carefully; make sure you understand what the question is asking
- Look for key words in the question: difference, differ, same as, more/less than, relationship, association, correlation, linked
- A "v" word, (vary, variance, variation) will be present in the question for differences in variances
- If a "v" word does not appear in a difference question and question does not concern frequencies, assume question concerns means
- "Affect" and "effect" can be used in both difference and relationship questions. You must understand their use in context; for example, it likely is a difference question if there is a grouping variable present.

Protocol for hypothesis testing *- fill in each blank; write "NA" for questions that are not applicable.*

Home

A. Justify test used [2].

1. What are the variables? [.2] _____
2. What is the respective measurement scale of each variable? [.2] _____
3. Is the question about differences or relationships? [.2] _____
   a. If a difference question, does it concern means, variances, or frequencies? [.2] _____
   b. If a relationship question, does it concern variables or frequencies? [.2] _____
4. To determine if a parametric test can be used, ask these questions:
   a. ***Means***: If you think the appropriate test is a parametric test of differences in means

      -are the data independent or dependent? [.2] _____

      -is each group/variable normally distributed? [.2] Y/N_____; probs _____

      -are the variances homogeneous? [.2] Y/N_____; prob**.** _____
   b. ***Variances***: If you think the appropriate test is a parametric test of differences in variances,

      -is each group normally distributed? [.2] Y/N_____; probs**.**_____
   c. ***Variables***: If you think the appropriate test is a parametric test of relationships between variables,

      -are the residuals or each variable normally distributed? [.2] Y/N_____; probs_____

B. State research hypothesis(es) [0]. H_A: _____

C. State null hypothesis(es) [2]. H_0: _____
(**variables must match answers in A1**)

D. What is the most appropriate test? [1] _____
(an incorrect answer limits further points)

E. Execute test(s) and identify and state value of **each** test statistic [2]. _____
(an incorrect answer limits further points)

F. State probability of **each** test statistic [1]. _____

G. State reject or cannot reject for **each** null hypothesis [1]. _____

H. Concisely state a biological conclusion for **each** test [1].

Hypothesis Testing 1

Home

Frequencies: Goodness-of-Fit StatTests

1. Test whether an observed frequency distribution fits an expected frequency distribution
2. One variable, mutually exclusive categories, each frequency occurs in one category, no cell has an expected frequency <5 (must pool categories if violated), no proportions or percentages
3. Null hypothesis: **H_0: Ovar = Evar**
4. Test statistic (χ^2) and probability source: Calculator/Statistical Table

   -calculation: $\chi^2 = \Sigma((O-E)^2/E)$; reading a chi-square table (Tables)

5. Probability models used for determining expected frequencies
   - The equal probability model occurs if all categories are equally likely. The expected number of outcomes for each category is n / no. categories.
   - The unequal probability model occurs if there are several categories with unequal probabilities. The expected number of outcomes for each category is np1, np2, ..etc.
   - The binomial distribution model occurs if there are two possible outcomes for any item, with a constant probability of success with repeated independent encounters of subjects. To calculate the expected number of outcomes in n experiments, multiply the binomial probabilities by n.
   - The Poisson distribution model is used as a probability model for events that occur randomly. To calculate the expected number of outcomes in n experiments, multiply the Poisson probabilities by n.

8. df: extrinsic hypothesis (theoretical): df = no. categories – 1
   df: intrinsic hypothesis (empirical; e.g., estimating the mean from the data): df = no. categories – 2
9. Examples:
   - **Question 1**: Is the sex ratio of Wood Ducks skewed? (equal probability model; extrinsic)
   - **Question 2**: Do Rough Green Snakes prefer a particular kind of tree when sleeping? (unequal probability model; intrinsic); **pic**)
   - **Question 3:** Do the sample data fit a binomial distribution? (Binomial model; extrinsic; PP#36)
   - **Question 4**: Are seedlings randomly distributed among quadrats? (Poisson model; intrinsic; PP #60)

Example problems

1. Two purple-flowered pea plants, both heterozygous for flower color, were crossed, resulting in 78 purple-flowered offspring and 22 white-flowered offspring. **Question**: Does this outcome differ from the expected 3:1 ratio of purple-flowered to white-flowered offspring? (*Protocol link*)

2. The data below are number of juvenile manatees killed by boats in Florida. **Question**: Are males and females equally susceptible to being killed by boats? (*Protocol link*)
   no. males killed (1985-1995): 206
   no. females killed (1985-1995): 127

Frequencies: Test of Independence (=test of association) StatTests

1. Test whether the frequencies of two categorical variables are independent (unrelated)
2. Two categorical variables, each frequency occurs in multiple mutually exclusive categories, no proportions or percentages, no cell has an expected frequency of <5 (Systat will inform you of violations)
3. Null hypothesis: **H_0: row var independent of column var**
4. Test statistic (**X^2**) and probability source: Systat/Systat
5. SYSTAT path: *Analyze→Tables→Two-Way (enter row and column variables)*

6. **Question**: Is habitat dependent on (related to) sex?

SYSTAT output: (GINMOVE.SYZ; Plummer)

Frequencies
HAB$ (rows) by SEX$ (columns)

	F	M	Total
P	480	420	900
R	2	25	27
Total	482	445	927

Test statistic	Value	DF	Prob
Pearson Chi-square	**22.1511**	1.0000	**0.000**

6. Frequency table data - start with table (no raw data)
   a. example 1 – **Question**: Is there an association between the hemoglobin S allele and resistance to malaria?

	Did not contract malaria	Contracted malaria
Heterozygotes	1	14
Homozygotes	13	2

SYSTAT output:
Frequencies
MALARIA$ (rows) by GENES$ (columns)

	het	hom	Total
n	1	13	14
y	14	2	16
Total	15	15	30

Test statistic	Value	DF	Prob
Pearson Chi-square	19.286	1.000	**0.000**

b. example 2 – **Question**: Is the frequency of breaking bones independent of taking calcium supplements? (supplements)

Example problems

1. The following data are frequency of rabies in skunks collected from three geographic areas. **Question**: Is the incidence of rabies dependent on geographic area? (*Protocol link*)

Area	With Rabies	Without Rabies
Ozarks	14	29
Ouachitas	12	38
Delta	11	35

2. The following data are frequency of individuals with different hair colors according to sex. **Question**: Is human hair color dependent on sex? (*Protocol link*)

sex	black	brown	blond	red
male	32	43	16	9
female	55	65	64	16

Frequencies: Fisher Exact Test StatTests

1. Test whether the frequencies of two categorical variables are independent; **2 x 2 table only**
2. Two categorical variables, each frequency occurs in multiple mutually exclusive categories, no proportions or percentages; **no minimum expected cell frequency**
3. Null hypothesis: **H_0: row var independent of column var**
4. Calculates probability directly; no intermediate test statistic
5. SYSTAT path: *Analyze→Tables→Two-Way (check Fisher's Exact Test in Measures, enter row and column variables)*

6. **Question**: Is phenotype independent of genotype?

Measures of Association for genetics$ and malaria$

genetics(rows) by malaria(columns)	n	y	Total
het	14	1	15
hom	2	13	15
Total	**16**	**14**	**30**

Test Statistic	Value	df	p-Value
Fisher Exact Test (Two-Tail)			0.0000

Variances: Bartlett's and Levene's Tests StatTests

1. Test whether sample variances are from the same population (=homogeneous)
2. Bartlett's is sensitive to departures from normality (not robust)
3. Null hypothesis: **H_0: σ^2var(group a) = σ^2var(group b) = σ^2var(group c), etc.**
4. Test statistic for Bartlett's test (χ^2) and Levene's test (**F**) and probability source: Systat/Systat
5. SYSTAT path: *Analyze→Hypothesis Testing→Variance→Equality of Several Variances (enter dependent and grouping variables)*

6. **Question**: Does variation in total absorbance differ between concentrations?

SYSTAT output: (ABSORBANCE.SYZ; Moore)
-Equality of Several Variances

Variable	CONC	N	Mean	Variance	Median
ABSORB_TOT	8	6.000	0.476	0.020	0.467
	16	6.000	0.412	0.052	0.449

Bartlett's Test

Variable	Chi-Square	df	p-Value
ABSORB_TOT	1.004	1.000	0.316

Levene's Test - *For Levene's, use the F-ratio based on the median.

Variable		F-Ratio	df	p-Value
ABSORB_TOT	Based on Mean	1.173	1, 10	0.304
	Based on Median	1.045	1, 10	0.331

Example problems

1. The following data are systolic blood pressure in two breeds of domestic cats. **Question**: Does variation in pressure (mm/Hg) differ between Siamese and Mynx cats? (*Protocol link*)
   Siamese: 122, 138, 129, 152, 149, 166, 110, 114, 155, 136, 189, 145, 129, 115, 144, 134
   Mynx: 129, 128, 109, 115, 108, 116, 125, 124, 117, 132, 111, 113, 127

2. Three different methods were used to determine the dissolved oxygen content of lake water. Each of the three methods was applied to a sample of water six times, with the following results. **Question**: Do the three methods yielded equally variable results? (*Protocol link*)

method 1	method 2	method 3
10.96	10.88	10.73
10.77	10.71	10.79
10.90	10.88	10.78
10.69	10.86	10.82
10.87	10.70	10.88
10.60	10.89	10.81

3. The following data are growth rate (g/d) in newborn rats fed four different diets. **Question**: Is growth rate equally variable among diets? (*Protocol link*)
   diet A: 1.6, 1.9, 0.9, 1.1, 1.5, 1.0, 1.8, 1.6
   diet B: 2.5, 2.0, 2.8, 2.6, 2.6, 2.9, 1.9, 2.1
   diet C: 0.8, 0.9, 0.5, 0.6, 0.7, 0.5, 0.9, 0.8
   diet D: 1.0, 1.1, 0.7, 0.8, 0.9, 0.7, 1.1, 1.0

4. The following data are number of moths caught during the night by four different trap types.
   Question: Is there a difference in the variance of trap effectiveness? (*Protocol link*)
   Trap type 1: 41, 34, 33, 36, 40, 25, 31, 37, 34, 30, 38
   Trap type 2: 52, 55, 62, 56, 64, 56, 56, 55
   Trap type 3: 25, 33, 34, 37, 41, 34, 40, 36
   Trap type 4: 36, 41, 33, 28, 34, 40, 27, 37

REVIEW

Graphic methods for 'informed guessing' whether means are statistically different (not a substitute for a formal statistical test)

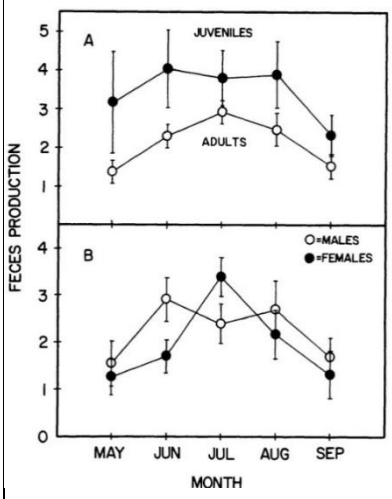
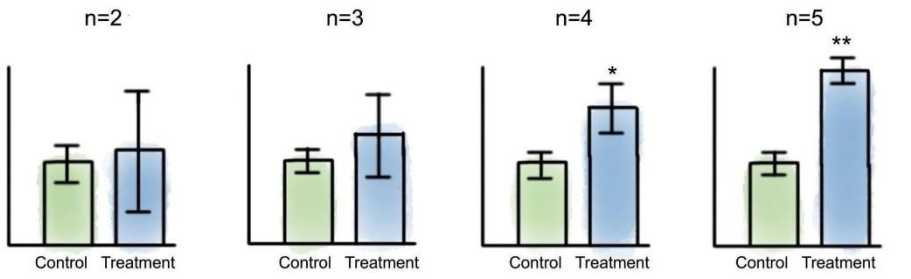


Fig. 1. (A) Feces production (x 1000) in juvenile and adult green snakes by month. (B) Feces production (x1000) in adult male and female green snakes by month. **Plotted are means ± 2 SE.**

VS.

Fig. 1. (A) Feces production (x 1000) in juvenile and adult green snakes by month. (B) Feces production (x1000) in adult male and female green snakes by month. **Plotted are means ± 1 SE.**



What is the message of this image? Is there anything wrong with how it is portrayed?

Means: Independent samples t-test StatTests

1. Test whether two sample means are from the same population
2. Powerful, robust (in literature ="Students" t-test"; William Gossett 1904)
3. Null hypothesis: **H_0: μvar (group a) = μvar (group b)**
4. Test statistic (**t**; absolute value) and probability source: Systat/Systat
5. SYSTAT path: *Analyze→Hypothesis Testing→Mean→Two-Sample t-test (enter dependent and grouping variables)*
6. Calculate power if you cannot reject H_0

Question: Do IAA levels differ between the wild type and triple mutants in the 4D germination treatments?

a. H_0: μ IAA(WS) = μ IAA(ILR/IAR/ILL)
b. pooled variance t ("regular" t-test - assumes homogeneous variances); use this one
c. separate variance t ("approximate" t-test - does not assume homogeneous variances)

SYSTAT output: (DRAMPEY.SYZ; Rampey)

Variable	PLANT$	N	Mean	Standard Deviation
IAA	ILR/IAR/ILL	3.000	11.133	3.350
	WS	3.000	21.000	3.378

Separate Variance

Variable	PLANT$	Mean Difference	95.00% Confidence Interval		t	df	p-Value
			Lower Limit	Upper Limit			
IAA	ILR/IAR/ILL	-9.867	-17.493	-2.240	-3.592	4.000	0.023
	WS						

Pooled Variance

Variable	PLANT$	Mean Difference	95.00% Confidence Interval		t	df	p-Value
			Lower Limit	Upper Limit			
IAA	ILR/IAR/ILL	-9.867	-17.493	-2.241	-3.592	4.000	0.023
	WS						

:1

Example problems

1. The effect of copper sulfate on the mucus cells in the gill filaments of a species of fish was investigated. The number of mucus cells per square micron in the gill filaments of untreated fish and in fish exposed for 24 hours to copper sulfate (mg/l) was as follows. **Question**: Does exposure to copper sulfate affect the number of mucus cells in these fish? (*Protocol link*)

   untreated: 16, 17, 12, 18, 11, 18, 12, 15, 16, 14, 18, 12
   exposed: 8, 10, 12, 13, 14, 6, 5, 7, 10, 11, 9, 8

2. A species of bacterium was grown with either glucose or sucrose as a carbon source. After a period of incubation, the number of cells (X 10^6) was determined. **Question**: Is there a difference in growth rate of the bacterium between the two carbon sources? (*Protocol link*)

   glucose: 6.3, 5.7, 6.8, 6.1, 5.2
   sucrose: 5.8, 6.2, 6.0, 5.1, 5.8

Means: Mann-Whitney StatTests

1. Test whether two sample means are from the same population
2. Null hypothesis: **H_0: μvar (group a) = μvar (group b)** (technically testing differences in medians)
3. Test statistic (**U**) and probability source: Systat/Systat (if provided an outside answer, may need to convert test statistic ($U'=n_1n_2-U$)
4. SYSTAT path: *Analyze→Nonparametric Tests→Kruskal-Wallis (enter dependent and grouping variables)*
5.

6. **Question**: Does weight differ between the sexes?

SYSTAT output: (LONOKE.SYZ; Plummer)
Mann-Whitney U Test for female length within range as male <50) length.

The categorical values encountered during processing are

Variables	Levels	
SEX (2 levels)	1.000	2.000

Dependent Variable	WGT
Grouping Variable	SEX

Group	Count	Rank Sum
1	27	912.000
2	44	1,644.000

Mann-Whitney U Test Statistic	:	534.000
p-Value	:	0.477
Chi-Square Approximation	:	0.505
df	:	1

Example problems

1. Twenty people were randomly assigned to two groups of ten each. One group viewed a hairy spider, and the other group viewed a similar but nonhairy spider. Each person was asked to score the spider she or he viewed on a ranked scariness scale from 1 to 10 (10 being the most scary). The results are below. **Question**: Do people find hairy spiders scarier than nonhairy spiders? (*Protocol link*).

   hairy: 10, 8, 7, 9, 9, 10, 9, 9, 5, 8
   nonhairy: 7, 6, 8, 6, 1, 5, 4, 5, 6, 3

2. The mass (g) of random samples of adult male tuatara from two localities in New Zealand are given below. **Question**: Do animals from locality A differ in mean mass from locality B? (*Protocol link*)

   loc A: 510, 773, 840, 505, 765, 780, 235, 790, 440, 435, 815, 460, 690
   loc B: 650, 600, 600, 575, 452, 320, 660

Means: Paired samples t-test StatTests

1. Test whether two sample means are from the same population
2. Each individual is measured twice or selected pairs are matched ("repeated measures"); more powerful than independent t-test (reduced error variance); robust; Exercise in Twins, NASA Twins
3. Data must be in an unstacked format
4. Null hypothesis: **H_0: μvar1 = μvar2** (no grouping variable)
5. Test statistic (**t**) and probability source: Systat/Systat
6. SYSTAT path: *Analyze→Hypothesis Testing→Mean→Paired t-test (enter paired variables)*
7. Calculate power if you cannot reject H_0

Question: Does early field metabolic rate differ from late field metabolic rate?

SYSTAT output: (DLWMEANS.SYZ; Plummer)

Paired samples t test on EARLYFMR vs LATEFMR with 6 cases

Mean EARLYFMR	=	0.1552		
Mean LATEFMR	=	0.1268		
Mean Difference	=	0.0283	95.00% CI = -0.0580 to 0.1147	
SD Difference =		0.0823	t = **0.8437**	
		DF = 5	Prob = **0.4373**	

How to stack dependent data files for testing equality of variances

1. manual stacking (create grouping variable)
2. SYSTAT stacking (*Data→Reshape→Stack*)

Example problems

1. *Brucella abortus* antibody titers (pfc/10^6 cells) in 15 turkeys were measured before and after a period of stress. **Question**: Did stress decrease antibody titer in these turkeys? (*Protocol link*)

turkey no.:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
before stress:	20	18	19	18	17	14	17	10	13	16	20	17	16	19	8
after stress:	17	14	16	19	14	18	8	10	12	15	8	6	17	5	3

2. Male hoop snakes, upon encountering one another, may engage in a protracted ritualized combat behavior until one establishes himself as dominant over the other. Six males were tested in the presence of a female and again in the absence of a female. Whether each male was tested first with or without a female was randomly determined. The results in interaction time (min.) are below. **Question**: Do these encounters last longer in the presence of a female? (*Protocol link*)

snake no.:	1	2	3	4	5	6
w/o female:	10	15	8	30	1	80
w/ female:	59	35	70	65	43	90

Means: Wilcoxon (Chap. 9) StatTests

1. Test whether two sample means are from the same population
2. Each individual is measured twice or selected pairs are matched (repeated measures)
3. Data must be in an unstacked format
4. Null hypothesis: **H_0: μvar1 = μvar2** (no grouping variable)
5. Test statistic (**Z**) and probability source: Systat/Systat
6. SYSTAT path: *Analyze→Nonparametric Tests→Wilcoxon (enter paired variables)*

7. **Question**: Does field metabolic rate differ between early and late measurements?

SYSTAT output: (DLWMEANS.SYZ; Plummer)

Wilcoxon Signed Ranks Test Results

Counts of differences (row variable greater than column)

	EARLYFMR	LATEFMR
EARLYFMR	0	4
LATEFMR	2	0

Z = (Sum of signed ranks)/square root(sum of squared ranks)

	EARLYFMR	LATEFMR
EARLYFMR	0.0	
LATEFMR	**-0.3145**	0.0

Two-sided probabilities using normal approximation

	EARLYFMR	LATEFMR
EARLYFMR	1.0000	
LATEFMR	**0.7532**	1.0000

Example problems

1. The wattle thickness (mm) of 10 randomly selected chickens was measured before and after treatment with PHA. **Question**: Does treatment with PHA affect wattle thickness? (*Protocol link*)

Chicken no.	1	2	3	4	5	6	7	8	9	10
pretreatment	1.05	1.01	0.78	0.98	0.81	0.95	1.00	0.83	0.78	1.05
posttreatment	3.48	5.02	5.37	5.45	5.37	3.92	6.54	3.42	3.72	3.25

2. Ten young men were asked to rate their feeling of well-being on a scale of 1 (worst) to 10 (best) before and after taking an experimental drug. **Question**: Does the drug increase a person's sense of well-being? (*Protocol link*)

individual no.:	1	2	3	4	5	6	7	8	9	10
before drug:	5	8	2	7	5	2	9	3	9	6
after drug:	7	9	1	9	5	9	9	9	10	7

You are responsible for knowing how to work all the Practice Problems concerning differences in frequencies, association of frequencies, and differences in variances and two means (Goodness-of-Fit, Test of Independence, Fisher's Exact Test,, Bartlett's, Levene's, Independent Samples t-test, Paired Samples t-test, Mann-Whitney, Wilcoxon). Exam problems will be taken directly or modified from Example and Practice Problems.

Hypothesis Testing 2

Home

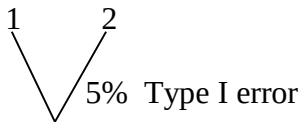
Analysis of Variance **StatTests**

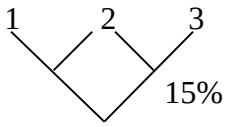
1. ANOVA – important part of experimental design (Fisher 1935); extremely common in the literature
2. Goal is to partition the sources of natural variability for any given system
   -total variability = source1 + source2 + source3, etc. (additive)
3. Also permits measurement of interaction (e.g., drug interaction); source1 x source2 (not additive)
4. Many different ANOVA models; e.g.,

- One-way ANOVA (1 dependent variable, 1 independent variable)
- Two-way ANOVA (1 dependent variable, 2 independent variables)
- Analysis of Covariance; ANCOVA (1 dep, 1 indep, 1 covariate) – at end of course if enough time

One-way ANOVA

1. Test whether sample means are from the same population
2. Powerful and robust
3. Null hypothesis: **H_0: $\mu_{var(group1)} = \mu_{var(group2)} = \mu_{var(group3)}$, etc.**
4. Why not use multiple t-tests? – *"The problem of multiple comparisons"*





4 means = 30%
5 means = 50%

5. Partition total variation into between-group and within-group ("error") variation
   - between group: variation due to being part of a certain group (treatment)
   - error variation: all variation not due to being in that group
6. Calculate ratio of between-groups variance/within-groups variance (F-ratio; test statistic)
   - F-ratio relatively large when treatment accounts for significant variation
7. Determine probability; compare F-ratio with F-distribution (shape determined by 2 separate dfs)
   - numerator (no. treatments – 1)
   - denominator (no. observations in all groups - no. groups)
8. Test statistic (**F**) and probability source: Systat/Systat

REVIEW: Required components of a null hypothesis for questions of differences in means or variances.

1. Indicator (H_0)
2. Parameter (e.g., μ, σ^2)
3. Variable (e.g., length, mass)
4. Group (e.g., sex, color); for questions of differences between independent data only (no grouping variable for dependent data). Groups are designated by being enclosed in parentheses.
5. Relational operator (e.g., $=, \geq, \leq$)

Examples:
-independent: H_0: $\mu_{length(males)} = \mu_{length(females)}$
-dependent: H_0: $\mu_{beforelength} = \mu_{afterlength}$

9. SYSTAT path: *Analyze→ANOVA→Estimate Model (enter dependent and grouping[=factor] variables)→Options (KS, Levene)*
10. Calculate power if you cannot reject H_0

SYSTAT output: (TREAT.SYZ; Plummer); **treat.ppt**

Categorical values encountered during processing are: CLUTNO (24 levels)
1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,24,25

Dep Var: EGGWGT N: 245 Multiple R: **0.7090** Squared multiple R: **0.5027**

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
CLUTNO	334.2372	23	14.5321	**9.7115**	**0.000**
Error	330.6987	221	1.4964		

Normality and homogeneity assumptions are tested after ANOVA with the residuals (=difference between observed value and value predicted by the model)

Example Problems

1. Random samples of a certain species of zooplankton were collected from five lakes and their selenium content (ppm) was determined. Was there a difference among lakes with respect to selenium content? (*Protocol link*)

   lake A: 23, 30, 28, 32, 35, 27, 30, 32
   lake B: 34, 42, 39, 40, 38, 41, 40, 39
   lake C: 15, 18, 12, 10, 8, 16, 20, 19
   lake D: 18, 15, 9, 12, 10, 17, 10, 12
   lake E: 25, 20, 22, 18, 30, 22, 20, 19

2. The following data are amount of food (kg) consumed per day by adult deer at different times of the year. Test the null hypothesis that food consumption was the same for all the months tested. (*Protocol link*)

February	May	August	November
4.7	4.6	4.8	4.9
4.9	4.4	4.7	5.2
5.0	4.3	4.6	5.4
4.8	4.4	4.4	5.1
4.7	4.1	4.7	5.6
4.2	4.8		

After significant ANOVA: Which means are different from which other means?

Post hoc pairwise tests counteract the problem of maintaining an alpha level of 0.05 for multiple comparisons; many different post hoc tests

1. Example: Tukey test
2. Test statistic (Difference) and probability source: Systat/Systat
3. SYSTAT path: *Analyze→ANOVA→Pairwise comparisons→Tukey (add group)*
4. MYSTAT path: *not available*

SYSTAT output: (TREAT.SYZ; Plummer, select clutno<6); **treat.ppt**

Categorical values encountered during processing are:
CLUTNO (5 levels) - 1, 2, 3, 4, 5

Dep Var: EGGWGT N: 63 Multiple R: 0.5851 Squared multiple R: 0.3423

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
CLUTNO	166.0769	4	41.5192	**7.5471**	**0.0001**
Error	319.0783	58	5.5014		

Post Hoc Test of EGGWGT
Using least squares means.

Tukey's Honestly-Significant-Difference Test

CLUTNO(i)	CLUTNO(j)	Difference	p-Value	95% Confidence Interval	
				Lower	Upper
1	2	-0.4850	0.9911	-3.4991	2.5291
1	3	1.2625	0.7430	-1.6643	4.1893
1	4	-0.3904	0.9957	-3.3171	2.5364
1	5	3.6288	0.0070	0.7378	6.5199
2	3	1.7475	0.3321	-0.8504	4.3454
2	4	0.0946	1.0000	-2.5032	2.6925
2	5	4.1138	0.0003	1.5562	6.6714
3	4	-1.6529	0.3478	-4.1488	0.8431
3	5	2.3663	0.0639	-0.0877	4.8203
4	5	4.0192	0.0002	1.5652	6.4732

Note that each Tukey comparison in the matrix, e.g., 1 vs. 5, 2 vs. 5, etc., is a separate statistical test. Each test requires its own null hypothesis, test statistic, probability, and conclusion.

Example Problems

1. In a study of snake hibernation, fifteen pythons of similar size and age were randomly assigned to three groups. One group was treated with drug A, one group with drug B, and the third group was not treated. Their systolic blood pressure (mmHg) was measured 24 hours after administration of the treatments. Do the drugs affect blood pressure? If so, do they have similar effects? (*Protocol link*)

   control: 130, 135, 132, 128, 130
   drug A: 118, 120, 125, 119, 121
   drug B: 105, 110, 98, 106, 105

2. Fourteen hucksters were assigned at random to one of three experimental groups and fed a different diet for six months. Use the following data on huckster mass (kg) at the end of the experiment to determine if diet affected body size. Which diet produced the heaviest hucksters? (*Protocol link*)

diet 1	diet 2	diet 3
60.8	68.7	102.6
57.0	67.7	102.1
65.0	74.0	100.2
58.6	66.3	96.5
61.7	69.8	

Kruskal_Wallis Test

1. Test whether three or more sample means are from the same population
2. Non-parametric counterpart to one-way ANOVA
3. Null hypothesis: **H_0: μvar (group1) = μvar(group2) = μvar(group3), etc.**
4. Test statistic (**H**) and probability source: Systat/Systat
5. SYSTAT path: *Analyze→Nonparametric tests→Kruskal-Wallis (enter dependent and grouping (=factor) variables)*

SYSTAT output: (TREAT.SYZ; Plummer, select clutno<8); **treat.ppt**

Categorical values encountered during processing are:
CLUTNO (7 levels)
1, 2, 3, 4, 5, 6, 7

Kruskal-Wallis One-Way Analysis of Variance for 89 cases
Dependent variable is EGGWGT
Grouping variable is CLUTNO

Group	Count	Rank Sum
1	8	374.0000
2	12	731.5000
3	14	245.0000
4	14	833.5000
5	15	490.0000
6	9	720.0000
7	17	611.0000

Kruskal-Wallis Test Statistic [H] = **46.9358**
Probability is **0.0000** assuming Chi-square distribution with 6 DF

Dwass-Steel-Chritchlow-Fligner Test for All Pairwise Comparisons

Group(i)	Group(j)	Statistic	p-Value
1	2	7.8558	0.0000
1	3	1.2552	0.9745
1	4	9.8964	0.0000
1	5	5.8438	0.0007
1	6	6.1237	0.0003
1	7	6.5521	0.0001
2	3	-4.1468	0.0524
2	4	0.9094	0.9954
2	5	0.8282	0.9972
etc.			

- *For post hoc pairwise comparisons after significant KW*

Dwass-Steel-Critchlow-Fligner Test *(DSCF)*

Example Problems

1. Twenty-four freshwater clams were randomly assigned to four groups of six each. One group was placed in deionized water, one group was placed in a solution of 0.5 mM sodium sulfate, and one group was placed in a solution of 0.74 mM sodium chloride. At the end of a specified time period, blood potassium levels (μM K^+) were determined. Did treatment affect blood potassium levels? *(Protocol link)*

   pond water: 0.518, 0.523, 0.499, 0.502, 0.520, 0.507
   deionized water: 0.308, 0.385, 0.301, 0.390, 0.307, 0.371
   sodium sulfate: 0.393, 0.415, 0.351, 0.390, 0.385, 0.397
   sodium chloride: 0.383, 0.405, 0.398, 0.352, 0.381, 0.407

2. An entomologist interested in the vertical distribution of a fly species collected the following data on numbers of flies (no. flies/m^3) from each of tree different vegetation layers. Use these data to test the hypothesis that fly abundance was the same in all three vegetation layers. *(Protocol link)*

herbs	shrubs	trees
14.0	8.4	6.9
12.1	5.1	7.3
5.6	5.5	5.8
6.2	6.6	4.1
12.2	6.3	5.4

Two-way ANOVA - factorial design; 2 independent variables (=factors)

1. Test whether sample means are from the same population; access interaction between independent variables
2. Powerful and robust; test assumptions with residuals
3. Test statistic (**F**) and probability source: Systat/Systat
4. Null hypotheses:
   - **H_0: μvar(group1) = μvar(group2) = μvar(group3), etc.** (for each main effect)
   - **H_0: no interaction among factors (**interaction = the extent to which the effects of one factor differ according to the levels of another factor; synergism or antagonism)
5. SYSTAT path: *Analyze→ANOVA→Estimate Model (enter dependent variable and >1 grouping variable)*

Example 1:

SYSTAT output; MOUSEDIET.SYZ; Cooper

Variables	Levels			
DIET$ (4 levels)	5K-96	AIN-cas	AIN-spi	P5001
NPDOSE$ (2 levels)	0	2000		

Dependent Variable	BODWGT
N	143
Multiple R	0.783
Squared Multiple R	0.614

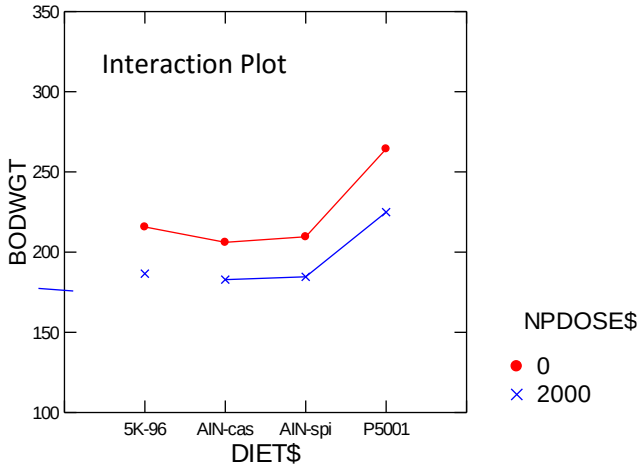
Analysis of Variance

Source	Type III SS	df	Mean Squares	F-ratio	p-value
DIET$	66249.445	3	22083.148	47.793	0.000
NPDOSE$	29869.989	1	29869.989	64.645	0.000
DIET$*NPDOSE$	1538.354	3	512.785	1.110	0.347
Error	62378.033	135	462.060		

*Note there are 3 separate hypotheses tested

Conclusions

- Diet explains a significant amount of variation in body weight. Body weight is greater in mice with the P 5001 diet.
- NPdose explains a significant amount of variation in body weight. Body weight is greater in mice not receiving NPdose.
- There is no interaction between diet and NPdose. Body weight responds the same to diet and NPdose.



Example 2: effect of diet and stress on weight gain in mice

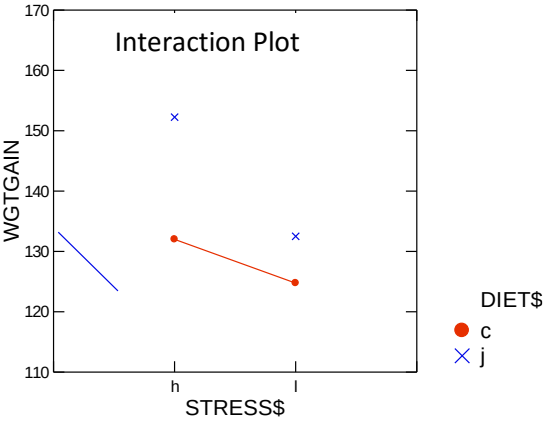
SYSTAT output: (dietstress.syz)

Variables	Levels
DIET$ (2 levels)	c j
STRESS$ (2 levels)	h l

Dependent Variable	WGTGAIN
N	32
Multiple R	0.844
Squared Multiple R	0.712

Analysis of Variance

Source	Type III SS	df	Mean Squares	F-ratio	p-value
DIET$	1568.000	1	1568.000	32.449	0.000
STRESS$	1458.000	1	1458.000	30.173	0.000
DIET$*STRESS$	312.500	1	312.500	6.467	0.017
Error	1353.000	28	48.321		



How affect? Conclusions

- Diet explains a significant amount of variation in weight gain. Mice with junk food diets gain more weight than mice with regular diets.
- Stress explains a significant amount of variation in weight gain. Mice experiencing high stress gain more weight than mice experiencing low stress.
- The interaction between diet and stress explains significant variation in weight gain. Weight gain caused by a junk food diet is exacerbated (i.e. made worse) by high stress. Or stated from another perspective, the weight gain caused by high stress is exacerbated by a junk food diet.

Example Problems

1. Use USOPHEO.SYZ; Plummer to determine if body size is affected by sex and/or location. Read the description of the data file before proceeding. (*Protocol link*)

2. Qualime epithelial cancer is hypothesized to result from either genotype or several environmental factors that vary by season. To address this hypothesis, use the data below on QSA level (µg/g; the diagnostic test indicator of qualime cancer) that were collected on 20 individuals in different seasons. (*Protocol link*)

QSA	Genotype	Season	QSA	Genotype	Season	QSA	Genotype	Season	QSA	Genotype	Season
478	ZZ	Winter	425	ZW	Summer	428	ZZ	Summer	466	ZW	Winter
538	ZZ	Winter	467	ZW	Summer	478	ZZ	Summer	522	ZW	Winter
502	ZZ	Winter	444	ZW	Summer	455	ZZ	Summer	489	ZW	Winter
496	ZZ	Winter	438	ZW	Summer	446	ZZ	Summer	475	ZW	Winter
483	ZZ	Winter	431	ZW	Summer	432	ZZ	Summer	501	ZW	Winter

3. Work practice problem #56. Why is it a one-way rather than a two-way ANOVA? You will have to create a derived variable to work the problem. There are two ways to do this: (1) enter the derived variable directly on the SYSTAT data sheet or (2) enter all of the data shown and use *TRANSFORM →If.., Then Let* to create the derived variable. You likely will need to review how to create derived variables.

Correlation

- correlation analysis is a test of association that makes no assumption about a cause-and-effect relationship (i.e., there is no dependent and independent variable)
- addresses two questions
  - does an association exist between two variables?
  - if the association exits, what is its strength (effect)?
- requires that both variables be normally distributed random variables

Pearson correlation StatTests

1. Test whether the cases of two variables are correlated (positive or negative)
2. Linear relationships only
3. Null hypothesis: assume no relationship; **H_0: ρvar1,var2 = 0** (Note there is no grouping variable, just two ratio or interval variables)
4. Test statistic (correlation coefficient, **r** (varies from -1 to +1; measure of strength) and probability source: Systat/Systat
5. r^2 (coefficient of determination) - proportion of variation in one variable that is explained by variation in the other variable (r^2 is **not** a test statistic)
6. SYSTAT path: *Analyze→Correlation→Simple (enter variables; Continuous Data)*
7. Calculate power if you cannot reject H_0

SYSTAT output (AMPHIBIANS.SYS; Mills); assume normality for purposes of demonstration only

Number of Non-Missing Cases: 40

Means

BUFO	SPECIES
1.5750	2.5000

Pearson Correlation Matrix

	BUFO	SPECIES
BUFO	1.0000	
SPECIES	0.6198	1.0000

Matrix of Bonferroni Probabilities

	BUFO	SPECIES
BUFO	0.0000	
SPECIES	0.0000	0.0000

Bonferroni probability correction **(**counteracts the "*The problem of multiple comparisons*"); reduces chances of making a Type 1 error (= "false negative" in the medical literature)

SYSTAT output: (AMPHIBIANS.SYZ; Mills); assume normality for purposes of demonstration only

Number of Non-Missing Cases: 40

Means

BUFO	RASP	HYLA	INDIVIDUALS	SPECIES
1.5750	1.7750	0.7500	5.1750	2.5000

Pearson Correlation Matrix

	BUFO	RASP	HYLA	INDIVIDUALS	SPECIES
BUFO	1.0000				
RASP	0.2408	1.0000			
HYLA	0.1034	0.2000	1.0000		
INDIVIDUALS	0.7103	0.7239	0.5245	1.0000	
SPECIES	0.6198	0.5630	0.4854	0.8761	1.0000

Matrix of Bonferroni Probabilities

	BUFO	RASP	HYLA	INDIVIDUALS	SPECIES
BUFO	0.0000				
RASP	1.0000	0.0000			
HYLA	1.0000	1.0000	0.0000		
INDIVIDUALS	0.0000	0.0000	0.0051	0.0000	
SPECIES	0.0002	0.0016	0.0150	0.0000	0.0000

Example problems

1. Use the following data on wing length (cm) and tail length (cm) in cowbirds to determine if there is a relationship between the two variables. (*Protocol link*)

Wing	10.4	10.8	11.1	10.2	10.3	10.2	10.7	10.45	10.8	11.2	10.6
Tail	7.4	7.6	7.9	7.2	7.4	7.1	7.4	7.2	7.8	7.7	7.8

2. Use the following data taken from crabs to determine if there is a relationship between weight of gills (g) and weight of body (g) and between weight of thoracic shield (g) and weight of body. (*Protocol link*)

Body	159	179	100	45	384	230	100	320	80	220	320
Gill	14.4	15.2	11.3	2.5	22.7	14.9	11.4	15.81	4.19	15.39	17.25
Thorax	80.5	85.2	49.9	21.1	195.3	111.5	56.6	156.1	39.0	108.91	160.1

Spearman correlation StatTests

1. Test whether the cases of two variables are correlated
2. Linear relationships only
3. Null hypothesis: **H_0: $\rho_s$var1, var2 = 0** (Note there is no grouping variable, just two ratio, interval, or ranked variables)

4. Test statistic (r_s) and probability source: Systat/Statistical Table
5. SYSTAT path: *Analyze→Correlation→Simple (enter variables; Rank Order Data)*

SYSTAT output: (AMPHIBIANS.SYZ; Mills)

Number of Non-Missing Cases: 40

Spearman Correlation Matrix

	BUFO	RASP	HYLA	GACA	NOVI	INDIVIDUALS	SPECIES
BUFO	1.0000						
RASP	0.3113	1.0000					
HYLA	0.2886	0.3879	1.0000				
GACA	0.3407	0.3682	0.1901	1.0000			
NOVI	0.2314	0.0436	0.3467	-0.0526	1.0000		
INDIVIDUALS	0.7264	0.7678	0.5804	0.3506	0.2044	1.0000	
SPECIES	0.7512	0.6482	0.6438	0.3226	0.3001	0.9173	1.0000

Spearman probabilities are not available in SYSTAT; must get probabilities from a Spearman Table

Example problems

1. The following data are ranked scores for ten students who took both a math and a biology aptitude examination. Is there a relationship between math and biology aptitude scores for these students? (*Protocol link*)

Math	53	45	72	78	53	63	86	98	59	71
Biology	83	37	41	84	56	85	77	87	70	59

2. Test the following data to determine if there is a relationship between the total length of aphid stem mothers and the mean thorax length of their parthenogenetic offspring. (*Protocol link*)

Mother	8.7	8.5	9.4	10.0	6.3	7.8	11.9	6.5	6.6	10.6
offspring	5.95	5.65	6.00	5.70	4.40	5.53	6.00	4.18	6.15	5.93

Correlation vs. causation

1. Earlier: alcoholics in FL vs HU grads; spurious correlations
2. sometimes results from a common correlation with 3rd variable (e.g., B correlated with C because both B&C are functionally correlated with A)**;** Cause and effect

Regression analysis is a test of association that

- assumes a cause-and-effect relationship between an independent and dependent variable
- is used to address the same basic questions as correlation analysis (with one important additional question), but from the perspective of cause-and-effect
  - does the independent variable explain significant variation in the dependent variable?
  - how strong is the explanatory power of the independent variable?
  - what is the mathematical relationship between the variables? (i.e., what is the mathematical equation that describes the relationship?)
- requires that the dependent variable be a normally distributed random variable. The independent variable may be controlled or selected and thus may not be a normally distributed random variable.

Regression (Chap. 14) StatTests

1. Test whether the cases of one variable are functionally (mathematically) related to the cases of another variable (i.e., can be predicted from)
2. Linear relationships only
3. Normality assumptions are analyzed with residuals after the regression analysis; robust
4. Null hypothesis: **H_0: βyvar, xvar = 0** (Note there is no grouping variable, just two ratio or interval variables)
5. Test statistic (**F-ratio**) and probability source: Systat/Systat
6. SYSTAT path: *Analyze→Regression→Linear→Least Squares (enter dependent and independent variables; enter KS on options tab)*

Procedure

a. Fit regression line (least squares method; minimize Σ(residuals^2)
b. Test for significance of slope
c. Write the **regression equation** (general form **Y = a (intercept) + b (slope) X**
   -do NOT use math format (y = mx + b)
d. **Add regression statistics and variable names**

SYSTAT output: (SHRIMP.SYZ; Goy)

Output format

- Regression statistics: intercept (=constant), slope (=regression coefficient); standard error
- ANOVA table (test statistic, probability)
- KS test of assumptions

Dependent Variable	EGGNO
N	68
Multiple R	0.7763
Squared Multiple R	0.6027
Adjusted Squared Multiple R	0.5967
Standard Error of Estimate	1142.1881

Regression Coefficients B = (X'X)$^{-1}$X'Y

Effect	Coefficient	Standard Error	Std. Coefficient	Tolerance	t	p-Value
CONSTANT (intercept)	-4914.5822	683.9281	0.0000	.	-7.1858	0.0000
FEMLEN (slope)	561.5867	56.1225	0.7763	1.0000	10.0065	0.0000

Analysis of Variance

Source	SS	df	Mean Squares	F-Ratio	p-Value
Regression	1.3063E+008	1	1.3063E+008	**100.1291**	**0.0000**
Residual	8.6103E+007	66	1304593.7089		

← Test statistic and probability

Test for Normality

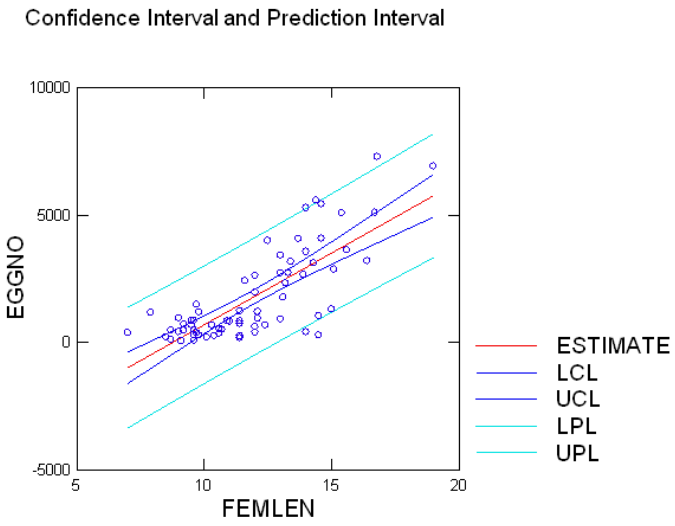
	Test Statistic	p-Value
K-S Test (Lilliefors)	0.0775	0.3660

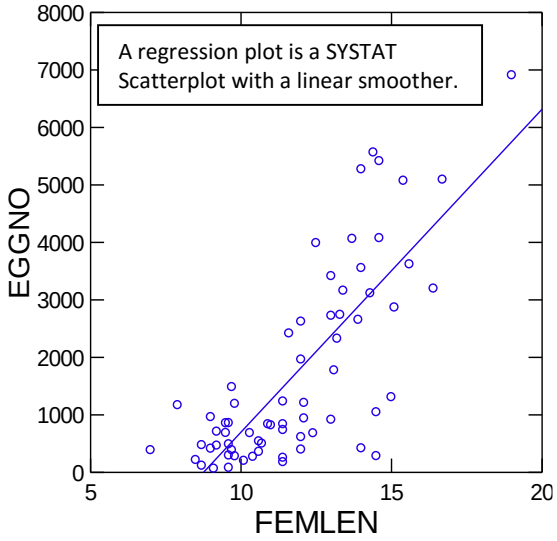
← KS test of normality assumption for residuals

The regression equation from the above analysis and represented on the graph is:

EGGNO = -4914.6 + 561.6 FEMLEN

In the regression equation, note that 'X' and 'Y' are replaced with the specific variables in question, i.e., FEMLEN and EGGNO. Also note that the dependent variable, EGGNO, is plotted on the Y axis, and the independent variable, FEMLEN, is plotted on the X axis. Another way of stating this is, "EGGNO is plotted against FEMLEN", or "EGGNO is regressed on FEMLEN."





Example problems

1. The following data are rate of oxygen consumption (ml/g/hr) in crows at different temperatures (°C). Does temperature affect oxygen consumption in crows? Determine the equation for predicting oxygen consumption from temperature. (Protocol link)

temp	-18	-15	-10	-5	0	5	10	19
oxygen	5.2	4.7	4.5	3.6	3.4	3.1	2.7	1.8

2. Use the following data on mean adult body weight (mg) and larval density (no./mm^3) of fruit flies to determine if there is a functional relationship between adult body mass and the density at which it was reared. Determine the equation for predicting body weight from larval density. (Protocol link)

density	1	3	5	6	10	20	40
weight	1.356	1.356	1.284	1.252	0.989	0.664	0.475

Extrapolation: linear regressions are statistically valid only within limits of the data (independent variable, X); beyond data - do not know if relationship is linear

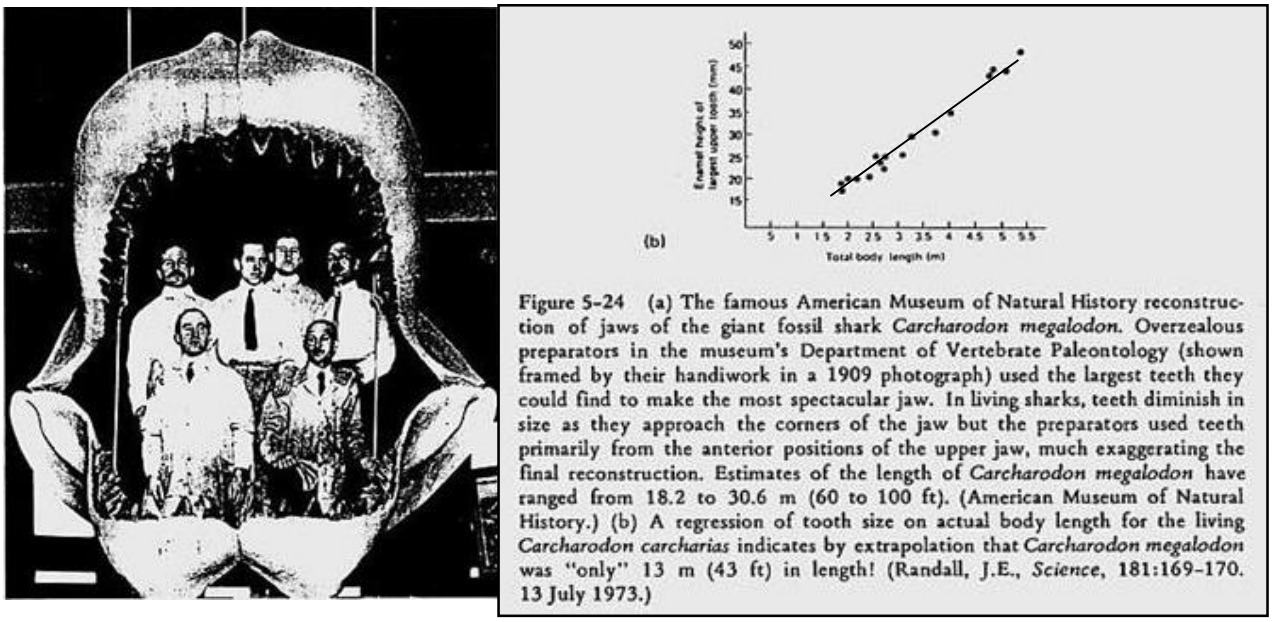


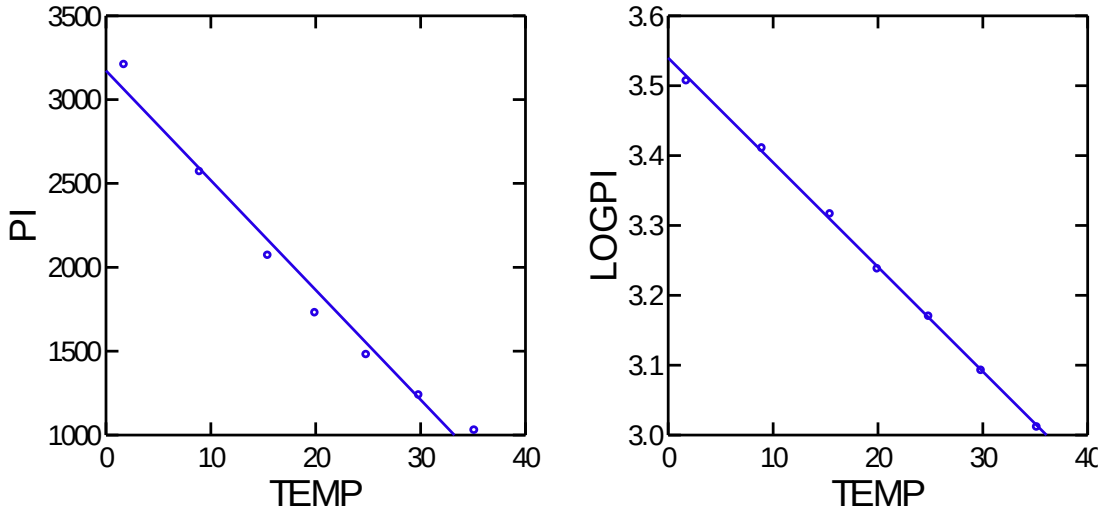
Figure 5-24 (a) The famous American Museum of Natural History reconstruction of jaws of the giant fossil shark *Carcharodon megalodon*. Overzealous preparators in the museum's Department of Vertebrate Paleontology (shown framed by their handiwork in a 1909 photograph) used the largest teeth they could find to make the most spectacular jaw. In living sharks, teeth diminish in size as they approach the corners of the jaw but the preparators used teeth primarily from the anterior positions of the upper jaw, much exaggerating the final reconstruction. Estimates of the length of *Carcharodon megalodon* have ranged from 18.2 to 30.6 m (60 to 100 ft). (American Museum of Natural History.) (b) A regression of tooth size on actual body length for the living *Carcharodon carcharias* indicates by extrapolation that *Carcharodon megalodon* was "only" 13 m (43 ft) in length! (Randall, J.E., *Science*, 181:169–170. 13 July 1973.)

A regression of tooth size on actual body length for the living *Carcharodon carcharias* indicates by extrapolation (assuming continued linearity) that *C. megalodon* was "only" 13 m (43 ft) in length!

Model building in regression (goal is to build a better model by increasing r^2; results in more accurate prediction)

Data transformation

1. SYSTAT e.g.: calibrate transmitters; **DEMO**
2. Linear vs. log10 data regressions - note increase in r^2 and linearity with log transformation



Dep Var: **PI** N: 7 Multiple R: 0.989 Squared multiple R: **0.978**

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P
CONSTANT	3172.273	97.857	0.000	.	32.417	0.000
TEMP	-65.363	4.390	-0.989	1.000	-14.888	0.000

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	3518606.514	1	3518606.514	221.638	0.000
Residual	79377.200	5	15875.440		

Dep Var: **LPI** N: 7 Multiple R: 1.000 Squared multiple R: **0.999**

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P
CONSTANT	3.540	0.004	0.000	.	834.735	0.000
TEMP	-0.015	0.000	-1.000	1.000	-78.645	0.000

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	0.184	1	0.184	6185.068	0.000
Residual	0.000	5	0.000		

Predicting dependent variable Y from independent variable X

A. ***Linear (Y, X) equations:*** **Y = a + bX**

Example 1: using the regression equation Y = 14.5 + 2.56X, predict Y when X = 63

Y = 14.5 + 2.56(63) = 175.78

Example 2: **inverse prediction** (predict X from Y); Y = 14.5 + 2.56X; by algebraic manipulation Y-14.5 = 2.56X; **(Y-14.5)/2.56 = X**

predict X when Y = 175.78:

X = (175.78-14.5)/2.56 = 63

B. ***Semilog (logY, X) equations:*** **log Y = log a + bX** (must take the inverse log of log Y to get final answer on linear scale)

Example: using the regression equation log Y = 1.42234 +0.047560X, predict Y when X = 12.1

logY = 1.42234 + 0.047560(12.1) = 1.99782 (calculate regression coefficients and

answer to at least **5** decimal places); inverse log 1.99782 = **99.49**

Note that the intercept (1.42234) is a log value (i.e., log a = 1.42234). You must not take the log of this value when calculating log Y; that would be the equivalent of taking the log of a log!

C. ***Log-log (logY, logX) and exponential equations***: **log Y = log a + b(log X); $Y = aX^b$**

Example 1 (logarithmic form): using the regression equation log Y = 2.53403 + 0.72000(log X), predict Y when X = 1.98

log Y = 2.53403 + 0.72000 (log 1.98) = 2.74763 (calculate regression coefficients and

answer to at least **5** decimal places)

inverse log 2.74763 = **559.28**

The most common form of the log-log regression equation, and one that is much easier to use is the exponential form:

log Y = log a + b(log X) = log a + log x^b; take inverse logs: **$Y = aX^b$** (exponential form)

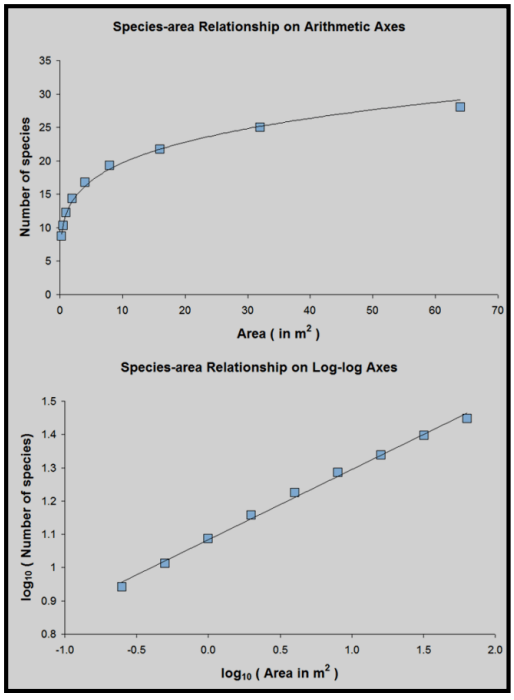
- **Example 2 (exponential form)**: using the regression equation $Y = 342X^{0.720}$, predict Y when

X = 1.98; ***note that 342 = the inverse log of 2.53403**

$Y = 342(1.98^{0.720}) = 559.28$

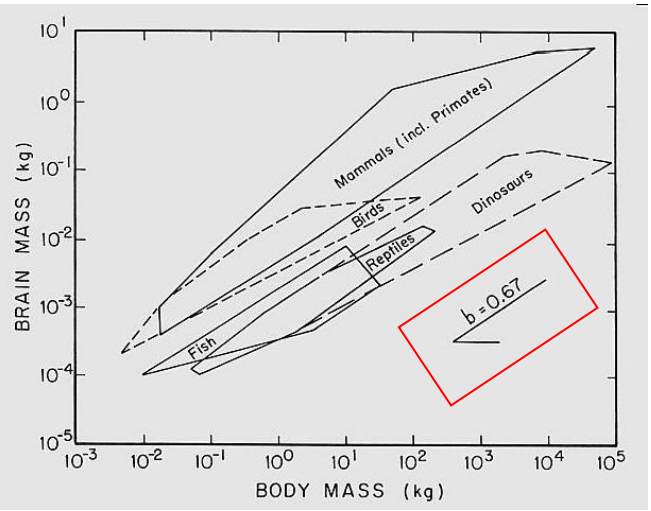
Examples of important uses of exponential regressions in biology

1. **Ecology**: species-area curves (Isle Biogeography Theory)

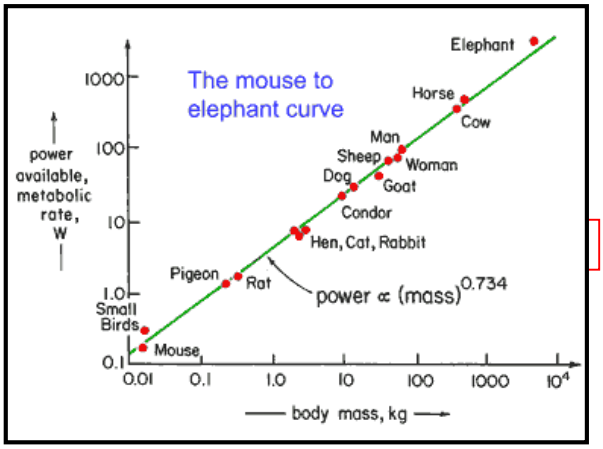


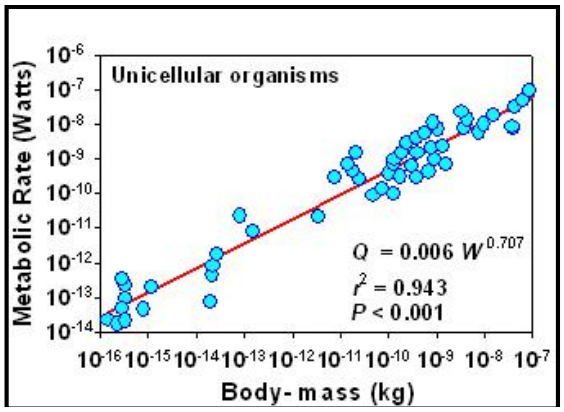
Common slope in some (~0.3)
-West Indian snakes: $S = 1.19A^{0.33}$
-Galapagos land plants: $S = 28.6A^{0.32}$
-Sierra Nevada mammals: $S = 1.18A^{0.32}$

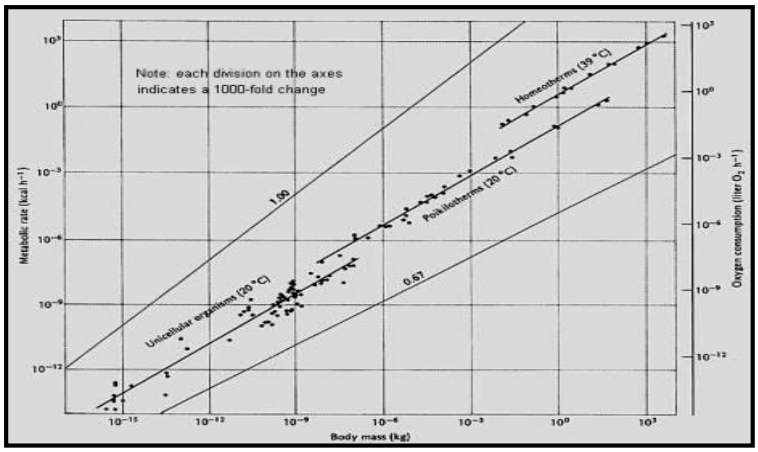
2. **Morphology**: effects of scaling; e.g., brain size



Physiology: effects of scaling; e.g., metabolic rate and body mass







Advanced statistical procedures commonly seen in the literature

Home

1. Analysis of Covariance **(ANCOVA)**

- Method of comparing regression lines: eg,
  -marsupials: MR = 0.409 $M^{0.75}$
  -eutherians MR = 0.676 $M^{0.75}$ (>60% higher)
- Detect differences among means of two or more groups when the dependent variable is affected by a third (continuous) variable (=covariate)
- A covariate adds unwanted variability to the dependent variable. ANCOVA removes that variability and yields least squares means (means adjusted for the covariate effect)
- ANCOVA combines the use of both ANOVA and regression methods

Example1: A common belief is that men are stronger than women. Is this belief due to men being bigger or are men actually stronger when compared to women of similar body size? Test this question on data from a sample of healthy young adults (**stronger.syz**). The variables are sex, lean body mass, and a measure of strength called "slow, right extensor knee peak torque."

Add Example problems

2. Circular statistics (Raleigh Test) – techniques for data measured on an angular scale. Angular scales are circular in nature, have no designated zero, and the designation of high and low values is arbitrary. For example, 0° and 360° point to the same direction.
3. Principal component analysis (PCA) - variable reduction technique that describe variability among multiple observed variables in terms of a lower number of non-measured derived variables

4. MANOVA (multivariate analysis of variance) – a generalized form of ANOVA in which there are two or more independent and/or two or more dependent variables. MANOVA assesses main effects and possible interactions among the dependent variables and among the independent variables
5. Repeated measures ANOVA – each individual is measured ≥ two times
6. Logistic regression – regression with a binary dependent variable (e.g., presence/absence
7. Non-linear regression
8. Multiple regression – regression with >1 independent variable (Fig. 2)

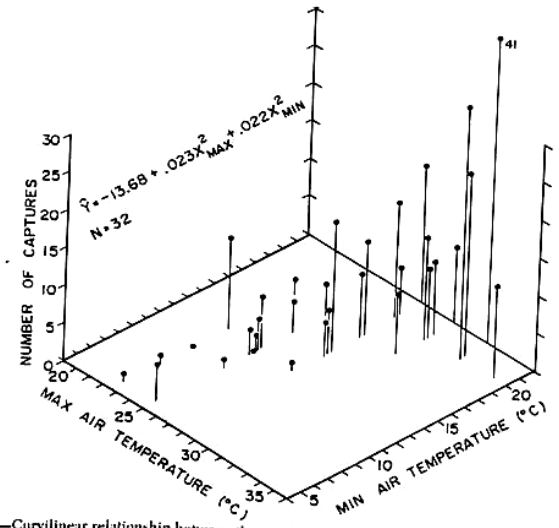


FIG. 2.—Curvilinear relationship between the number of *Tantilla coronata* captured per week and the mean minimum and maximum air temperatures per week. Data are pooled from all study areas; sample size indicates the number of weekly totals.

Statistical Tables

Home

TABLE A.1 AREAS OF THE NORMAL DISTRIBUTION

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997

Table A.3 Critical values of the chi-square distribution

	α				
df	**0.05**	**0.02**	**0.01**	**0.001**	**0.0001**
1	3.84	5.41	6.63	10.83	15.14
2	5.99	7.82	9.21	13.82	18.42
3	7.81	9.84	11.34	16.27	21.11
4	9.49	11.67	13.28	18.47	23.51
5	11.07	13.39	15.09	20.51	25.74
6	12.59	15.03	16.81	22.46	27.86
7	14.07	16.62	18.48	24.32	29.88
8	15.51	18.17	20.09	26.12	31.83
9	16.92	19.68	21.67	27.88	33.72
10	18.31	21.16	23.21	29.59	35.56
11	19.68	22.62	24.72	31.26	37.37
12	21.03	24.05	26.22	32.91	39.13
13	22.36	25.47	27.69	34.53	40.87
14	23.68	26.87	29.14	36.12	42.58
15	25.00	28.26	30.58	37.70	44.26
16	26.30	29.63	32.00	39.25	45.92
17	27.59	31.00	33.41	40.79	47.57
18	28.87	32.35	34.81	42.31	49.19
19	30.14	33.69	36.19	43.82	50.80
20	31.41	35.02	37.57	45.31	52.39
21	32.67	36.34	38.93	46.80	53.96
22	33.92	37.66	40.29	48.27	55.52
23	35.17	38.97	41.64	49.73	57.08
24	36.42	40.27	42.98	51.18	58.61
25	37.65	41.57	44.31	52.62	60.14
26	38.89	42.86	45.64	54.05	61.66
27	40.11	44.14	46.96	55.48	63.16
28	41.34	45.42	48.28	56.89	64.66
29	42.56	46.69	49.59	58.30	66.15
30	43.77	47.96	50.89	59.70	67.63

~~Table 9.~~ **Critical Values of Spearman's Rank Correlation Coefficient.**

~~*The* α *values correspond to a one-tailed test of* $H_0: \rho_s = 0$. *The value should be doubled for two-tailed tests.*~~

n	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$	n	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
5	.900				18	.399	.476	.564	.625
6	.829	.886	.943		19	.388	.462	.549	.608
7	.714	.786	.893		20	.377	.450	.534	.591
8	.643	.738	.833	.881	21	.368	.438	.521	.576
9	.600	.683	.783	.833	22	.359	.428	.508	.562
10	.564	.648	.745	.794	23	.351	.418	.496	.549
11	.523	.623	.736	.818	24	.343	.409	.485	.537
12	.497	.591	.703	.780	25	.336	.400	.475	.526
13	.475	.566	.673	.745	26	.329	.392	.465	.515
14	.457	.545	.646	.716	27	.323	.385	.456	.505
15	.441	.525	.623	.689	28	.317	.377	.448	.496
16	.425	.507	.601	.666	29	.311	.370	.440	.487
17	.412	.490	.582	.645	30	.305	.364	.432	478

Table A.2 Critical values of the *t* distribution

	α (Two-Tailed)						
df	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**	**0.001**	**0.0001**
1	3.078	6.314	12.706	31.821	63.657	636.619	6366.198
2	1.886	2.920	4.303	6.695	9.925	31.598	99.992
3	1.638	2.353	3.182	4.541	5.841	12.924	28.000
4	1.533	2.132	2.776	3.747	4.604	8.610	15.544
5	1.476	2.015	2.571	3.365	4.032	6.869	11.178
6	1.44	1.943	2.447	3.143	3.707	5.959	9.082
7	1.415	1.895	2.365	2.998	3.499	5.408	7.885
8	1.397	1.860	2.306	2.896	3.355	5.041	7.120
9	1.383	1.833	2.262	2.821	3.250	4.781	6.594
10	1.372	1.812	2.228	2.764	3.169	4.587	6.211
11	1.363	1.796	2.201	2.718	3.106	4.437	5.921
12	1.356	1.782	2.179	2.681	3.055	4.318	5.694
13	1.35	1.771	2.160	2.650	3.012	4.221	5.513
14	1.345	1.761	2.145	2.624	2.977	4.140	5.363
15	1.341	1.753	2.131	2.602	2.947	4.073	5.239
16	1.337	1.746	2.120	2.583	2.921	4.015	5.134
17	1.333	1.740	2.110	2.567	2.898	3.965	5.044
18	1.33	1.734	2.101	2.552	2.878	3.922	4.966
19	1.328	1.729	2.093	2.539	2.861	3.883	4.897
20	1.325	1.725	2.086	2.528	2.845	3.850	4.837
21	1.323	1.721	2.080	2.518	2.831	3.819	4.784
22	1.321	1.717	2.074	2.508	2.819	3.792	4.736
23	1.319	1.714	2.069	2.500	2.807	3.767	4.693
24	1.318	1.711	2.064	2.492	2.797	3.745	4.654
25	1.316	1.708	2.060	2.485	2.787	3.725	4.619
26	1.315	1.706	2.056	2.479	2.779	3.707	4.587
27	1.314	1.703	2.052	2.473	2.771	3.690	4.558
28	1.313	1.701	2.048	2.467	2.763	3.674	4.530
29	1.311	1.699	2.045	2.462	2.756	3.659	4.506
30	1.31	1.697	2.042	2.457	2.750	3.646	4.482
40	1.303	1.684	2.021	2.423	2.704	3.551	4.321
60	1.296	1.671	2.000	2.390	2.660	3.460	4.169
100	1.292	1.660	1.984	2.364	2.626	3.390	4.053
∞	1.282	1.645	1.960	2.326	2.576	3.291	3.750

Table A.6 Critical values of the *F* distribution ($\alpha = 0.05$; df_1 = treatment degrees of freedom, df_2 = error degrees of freedom). Part A: $\alpha = 0.05$

	df_1											
df_2	1	2	3	4	5	6	7	8	9	10	11	12
2	18.5	19.0	19.2	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.71	4.68
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
7	5.59	4.74	4.35	4.12	3.97	3.87	3.77	3.73	3.68	3.64	3.60	3.57
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.21	2.16
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.04
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83