

Biology 300
Homework assignment #1
Solutions

Assignment:

Chapter 1, Problems 6, 15
Chapter 2, Problems 6, 8, 9, 12
Chapter 3, Problems 4, 6, 15
Chapter 4, Problem 16

Answers in **bold**.

Chapter 1

6. Identify whether the following variables are numerical or categorical. If numerical, state whether the variable is discrete or continuous. If categorical, state whether the variable is nominal or ordinal.

a) Number of sexual partners in a year by college students

numerical, discrete

b) Petal area of rose flowers

numerical, continuous

c) Key on the musical scale

categorical, nominal

(not ordinal because there is no set starting point for any ordering)

d) Heart beats per minute of Tour de France cyclists

numerical, discrete

e) Stage of fruit ripeness (under ripe, ripe, over ripe)

categorical, ordinal

f) Angle of flower orientation relative to position of sun

numerical, continuous

g) Letter grade on high school report card

categorical, ordinal

h) Tree species

categorical, nominal

i) Year of birth

categorical, ordinal

(not numerical because there is no zero point for the scale)

j) Sex

categorical, nominal

k) Birth weight

numerical, continuous

15. The average age of piñon juniper trees in the coastal range of California was investigated by placing a 10-hectare plot randomly on a distribution map of the tree in California using a computer. Researchers then recorder the location of the random plot, found it in the field, and flagged it using compass and measuring tape. They then proceeded to measure the age of every piñon juniper tree within the 10 hectare plot. The average age within the plot was used to estimate the average age of the whole California population.

a) What is the population of interest in this study?

The whole California population of juniper trees

b) Were the trees sampled randomly from this population? Why or why not?

No. All of the trees selected were very close together, so the sampling of individual trees cannot be considered to be independent - trees are likely to be selected if their neighbors were selected.

Chapter 2

6. The data below are the occurrences of different taxa in the list of endangered and threatened species under the US Endangered Species Act by 2002, separately for each type of organism (U.S. Fish and Wildlife Service 2001). The taxa are listed in no particular order in the table below.

Taxon	Number of species
Birds	92
Clams	70
Reptiles	36
Fish	115
Crustaceans	21
Mammals	74
Snails	32
Plants	745
Amphibians	22
Insects	44
Arachnids	12

a) Rewrite the table, but list the taxa in a more revealing order. Explain your reasons behind the ordering you chose.

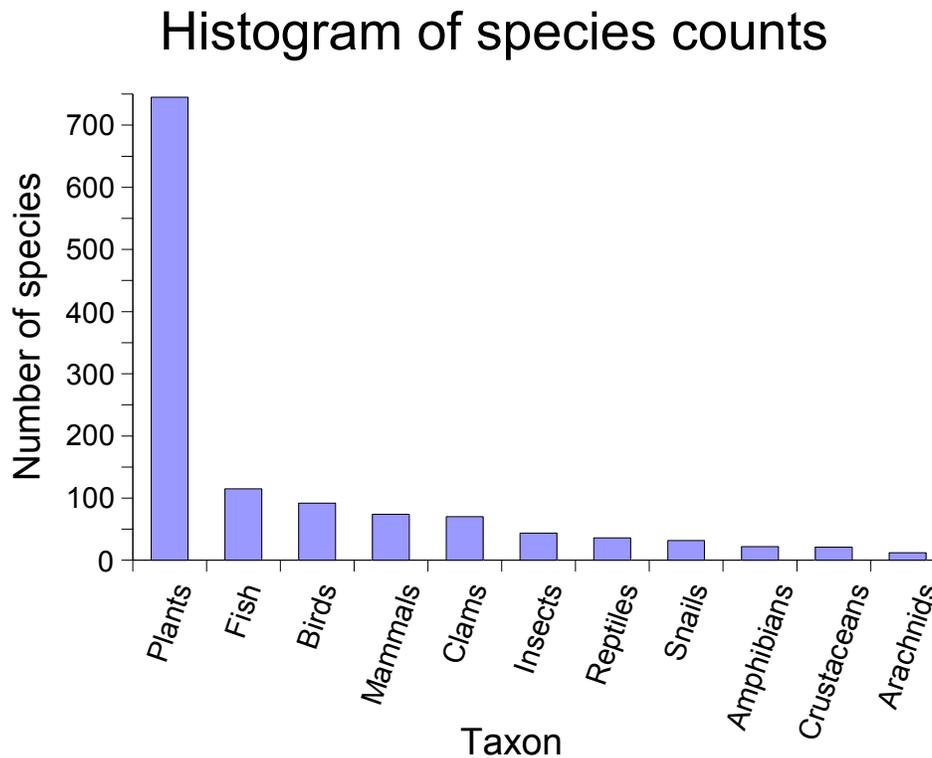
Taxon	Number of species
Plants	745
Fish	115
Birds	92
Mammals	74
Clams	70
Insects	44
Reptiles	36
Snails	32
Amphibians	22
Crustaceans	21
Arachnids	12

I chose this ordering because then the relative ranking of species numbers can be easily seen.

b) What kind of table is this?

A frequency table

c) Choosing the most appropriate graphical method, display the number of species listed in each taxon of organism in 2002. What kind of graph did you choose? Why?



I created a histogram, because that's a good way to show a frequency distribution.

d) Should the baseline for number of species in your graph in (c) be 0 or 21, the smallest number in the data set for 2002? Why?

Zero, so that the relative areas of the bars is proportional to the species counts - this is the most accurate way to visually display this information.

8. Can environmental factors influence the incidence of schizophrenia? A recent project measured incidence of the disease among children born in a region of eastern China. 192 of 13,748 babies born in the midst of a severe famine in the region in 1960 later developed schizophrenia. This compared with 483 schizophrenics out of 59,088 births in 1956, before the famine, and 695 out of 83,536 births in 1965, after the famine (St Claire et al. (2005).

a) What two variables are compared in this example?

year and incidence of schizophrenia

b) Are the variables numerical or categorical? If numerical, are they continuous or discrete; if categorical, are they nominal or ordinal?

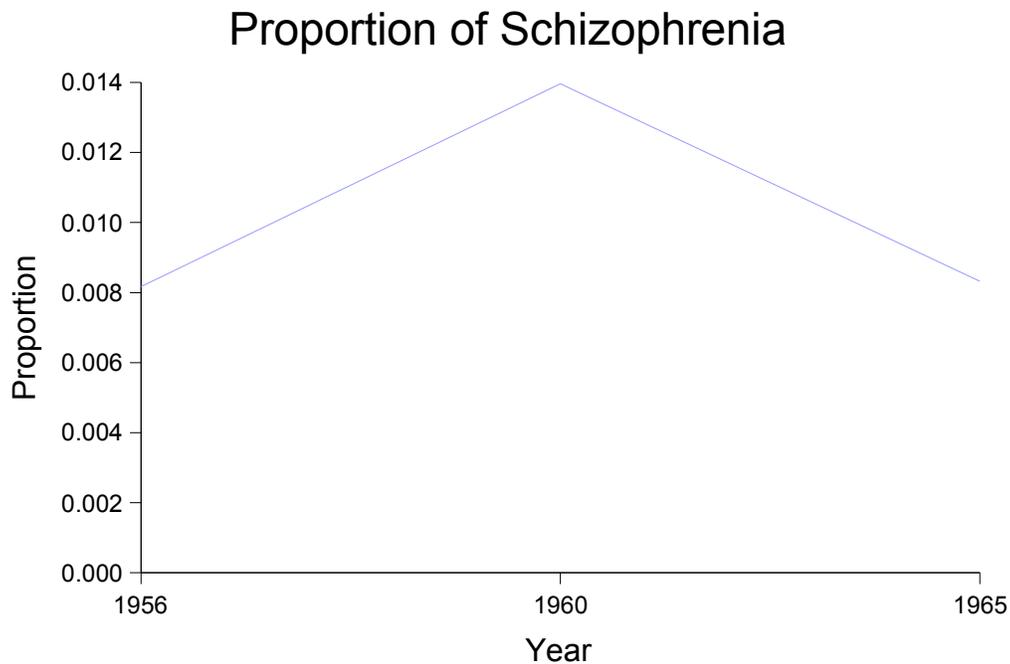
both are categorical; year is ordinal and schizophrenia is nominal

c) Effectively display the findings in a table. What kind of table did you use?

	Year			
	1956	1960	1965	Row totals
not schizophrenic	58,605	13,556	83,841	155,002
schizophrenic	483	192	695	1,370
column totals	59,088	13,748	83,536	156,372

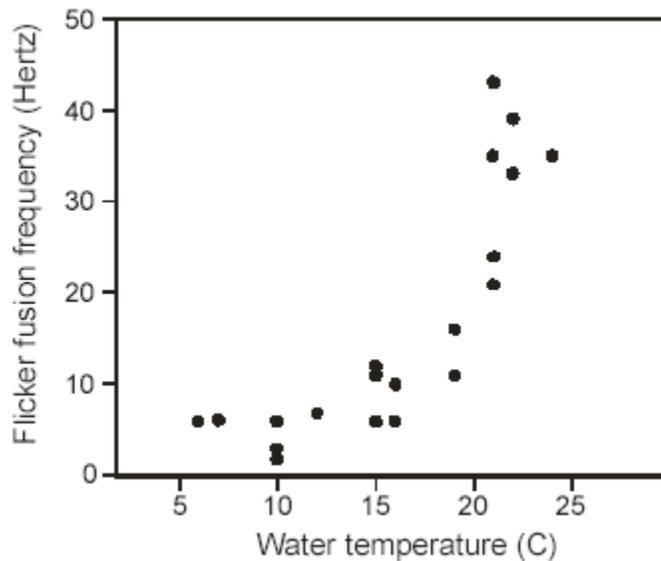
This is a contingency table.

d) Calculate the relative frequency (proportion) of children born in each of the three years that later developed schizophrenia. Plot these proportions in a line graph. What pattern is revealed?



There is a higher proportion of schizophrenia in the famine year.

9. Swordfish have a unique "heater organ" that maintains elevated eye and brain temperatures when hunting in deep cold water, but its function is unclear. The graph below illustrates the results of a study by Fritsches et al. (2005) that measured how the ability of swordfish retinas to detect rapid motion, measured by the flicker fusion frequency, changes with eye temperature.



a) What types of variables are displayed?

Both are numerical and continuous

b) What type of graph is this?

A scatter plot.

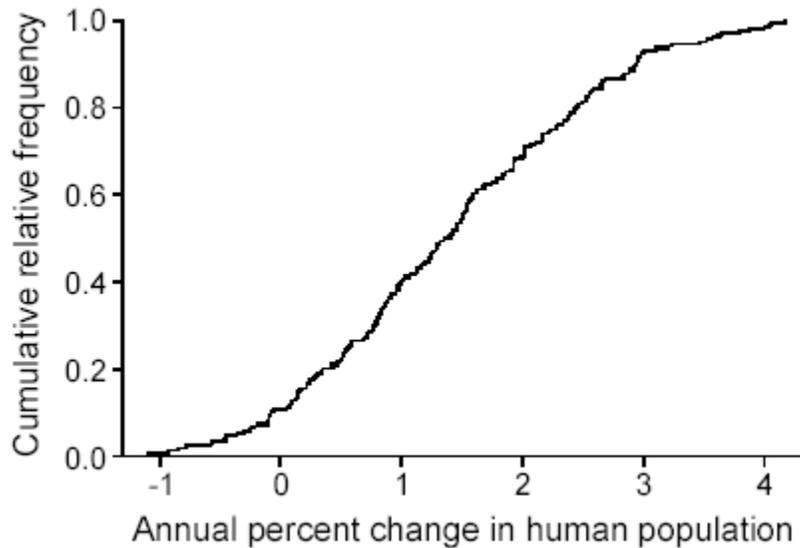
c) Describe the association between the two variables. Is the relationship between flicker fusion frequency and temperature positive or negative? Is the relationship linear or nonlinear?

There is a positive nonlinear relationship.

d) The 20 points in the graph were obtained from measurements of 6 swordfish. Can we treat the 20 measurements as a random sample? Why or why not?

No, because more than one measurement was taken for at least some of the fish. Repeated measurements taken from the same individual are not independent - these data are pseudoreplicated.

12. The following graph shows the population growth rates of 204 countries recognized by the United Nations. Growth rate is measured as the average annual percent change in the total human population between 2000 and 2004 (United Nations Statistics Division 2004).



a) Identify the type of graph depicted.

A cumulative frequency distribution

b) Explain the quantity along the Y-axis.

Y is cumulative relative frequency. It is the proportion of the sample that is less than a given X value. In this case, the y-axis depicts the proportion of countries with a growth rate less than X.

c) Approximately what percentage of countries had a negative change in population?

Approximately 10%. (Read the value off the y axis above 0 on the X axis.)

d) Identify by eye the 0.10, 0.50, and 0.90 quantiles of change in population size.

The 0.1 quantile is approximately 0.

The 0.50 quantile is approximately 1.3.

The 0.90 quantile is approximately 3.

(You get these by finding 0.1, 0.5, or 0.9 on the y-axis and finding the X value that corresponds.)

e) Identify by eye the 60th percentile of change in population size.

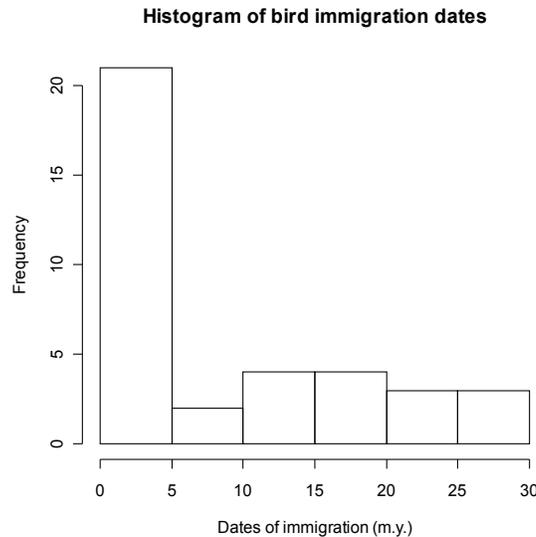
The 60th percentile is approximately 1.9.

Chapter 3

4. Birds of the Caribbean islands of the Lesser Antilles are descended from immigrants originating from larger islands and the nearby mainland. Here are the approximate dates of immigration, in millions of years, of each of 37 bird species now present on the Lesser Antilles (Ricklefs and Bermingham 2001). The dates were calculated from the difference in mitochondrial DNA sequence between each of the species and its closest living relative on larger islands or the mainland.

0.00, 0.00, 0.04, 0.21, 0.29, 0.54, 0.63, 0.88, 0.96, 1.25, 1.67, 1.75, 1.84,
 1.96, 2.01, 2.51, 2.72, 3.30, 3.51, 4.05, 4.85, 6.94, 8.73, 10.57, 11.11, 12.45,
 14.00, 17.30, 17.92, 18.05, 18.43, 22.48, 22.48, 23.48, 26.32, 26.45, 28.87

a) Plot the data in a histogram and describe the shape of the frequency distribution.



These data are skewed to the right.

b) By viewing the graph alone, approximate the mean and median of the distribution. Which is greater? Explain your reasoning. (Here and forever after, provide units with your answer.)

Estimated mean: 10 m.y.

Estimated median: 4 m.y.

The mean will be greater because it is more strongly affected by outliers - that is, very old immigration times.

c) Calculate the mean and median. Was your intuition in (b) correct?

Mean: 8.67 m.y.

Median: 3.51 m.y.

Yes, my intuition was correct - the mean is much larger than the median.

d) Calculate the first and third quartiles and the interquartile range.

I'll use the method outlined in the book:

first quartile: $(0.96+1.25)/2 = 1.105$

third quartile: $(17.30+17.92)/2 = 17.61$

Interquartile range: $17.61 - 1.105 = 16.505$

6) The data below are from an ecological study of the entire rainforest community at El Verde in Puerto Rico (Waide and Reagan 1996). Diet breadth is the number of types of prey eaten by the

consumer species present. The number of species having each diet breadth is shown in the second column. The total number of consumer species is $n = 127$.

Diet breadth (No. prey types eaten)	Frequency
1	21
2	8
3	9
4	10
5	8
6	3
7	4
8	8
9	4
10	4
11	4
12	2
13	5
14	2
15	1
16	1
17	2
18	1
19	3
20	2
More than 20	25
Total	127

a) Calculate the median number of prey types of species in the community.

The median is 8.

b) What is the interquartile range in the number of prey types? Use the simple method outlines in Section 3.2 to calculate the quartiles.

First quartile: consider just the first 63 values, and find their median (the 32nd number)

First quartile = 3

Third quartile: similarly, consider the last 63 values, and find their median

Third quartile = 17

Interquartile range = 17-3 = 14

c) Can you calculate the mean number of prey types in the diet?

No, because you don't know how many prey types are represented by the last category, more than 20.

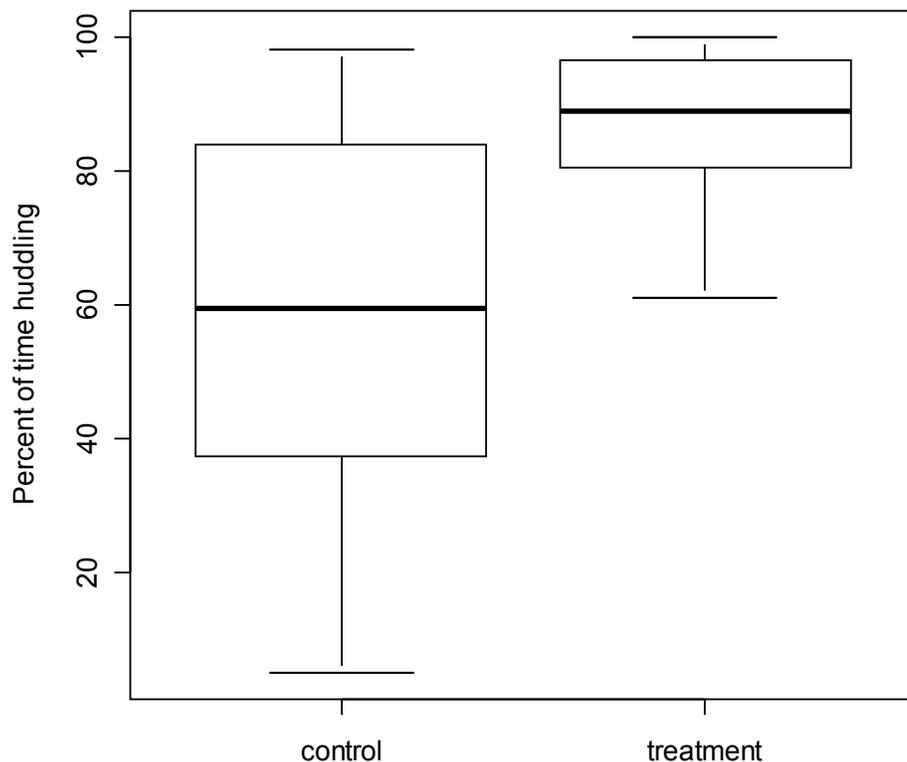
15) The gene for the vasopressin receptor *V1a* is expressed at higher levels in the forebrain of monogamous vole species than promiscuous vole species? Can expression of this one gene itself influence monogamy? To test this, Lim et al. (2004) experimentally enhanced *V1a* expression in the forebrain of 11 males of the meadow vole, a solitary promiscuous species. The percent of time each male spent huddling with the female provided to him was recorded. The same measurements were taken in 20 control males left untreated:

Control males: 98, 96, 94, 88, 86, 82, 77, 74, 70, 60, 59, 52, 50, 47, 40, 35, 29, 13, 6, 5

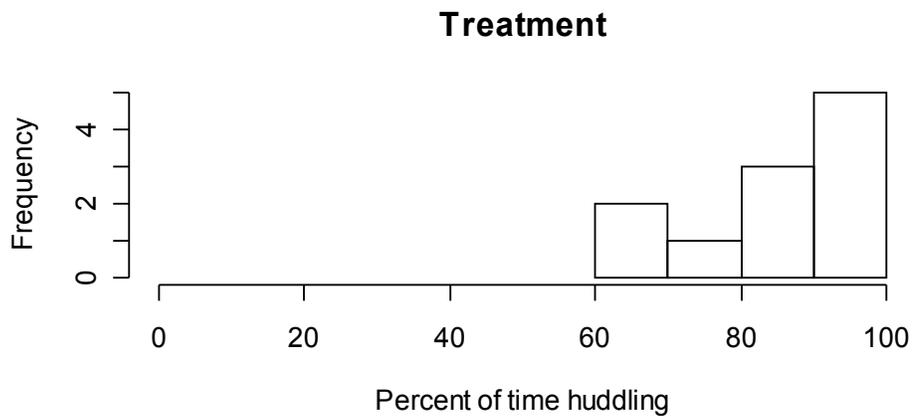
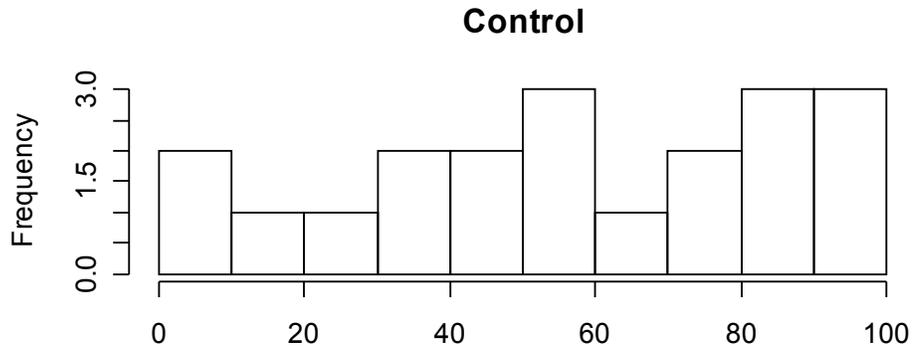
V1a enhanced males: 100, 97, 96, 97, 93, 89, 88, 84, 77, 67, 61

a) Display these data in a graph.

Two possibilities: a box plot



or stacked histograms:



b) Which group has the higher mean percent time spent huddling with females?

The treatment group: mean = 86.3% versus 58.1% for the controls.

c) Which group has the higher standard deviation in percent time spent huddling with females?

The control group: s = 29.8 versus s = 12.9 for the controls.

Chapter 4

16) Examine the times to rigor mortis of 114 corpses displayed in Problem 12 of Chapter 3.

Hours	Number of bodies
1	0
2	2
3	14
4	31
5	14
6	20

Hours	Number of bodies
7	11
8	7
9	4
10	7
11	1
12	1
13	2
Total	114

a) What is the standard error of the mean time to rigor mortis?

First, calculate the mean:

$$x = (0*1 + 2*2 + 14*3 + 31*4 + \dots + 2*13)/114 = 648/114 = 5.68$$

Next, calculate the standard deviation:

$$s^2 = [0*(1-5.68)^2 + 2*(2-5.68)^2 + 14*(3-5.68)^2 + 31*(4-5.68)^2 + \dots + 2*(13-5.68)^2]/(114-1)$$

$$s^2 = 630.63/113 = 5.58$$

$$s = \sqrt{5.58} = 2.36$$

$$\text{standard error} = s / \sqrt{n} = 2.36/\sqrt{114} = 0.22$$

b) The standard error calculated in (a) measures the width of what frequency distribution?

The sampling distribution of the mean.

c) What assumption does your calculation in (a) require?

That the cadavers used represent a random sample of the population.