

Ma 322: Biostatistics

Homework Assignment 1

Prof. Wickerhauser

Due Friday, January 27th, 2012

Begin by obtaining access to the R software package, either by downloading a copy onto your computer or else by finding a computer with a working installation. The current version number is 2.12.

Read Chapter 6, pages 60–79, of our e-text to review basic principles of probability. Consult Chapters 1-5 as needed to find function names and syntax to solve the computation problems below.

1. Let $\{1.3, 1.6, 1.8, 3.1, 3.2, 2.2, 3.5, 3.8\}$ be the values of a variable X over an 8-member population.
 - (a) Find the frequency distribution table for groupings into intervals $\{[0, 1), [1, 2), [2, 3), [3, 4)\}$. Plot the resulting frequency bar graph (histogram).
 - (b) Find the cumulative frequency distribution table and the relative cumulative frequency distribution table for this variable and population. Plot the results on a single frequency bar graph.
 - (c) Calculate the mean of the population.
 - (d) Estimate the mean of the population using the frequency distribution table from part b, with the midpoints of the intervals as the representative values.
 - (e) Calculate the median of the population using the 8 values.
 - (f) Estimate the median of the population using interpolation with the cumulative frequency distribution table from part b.
 - (g) Find the mode of the grouped values of the population using the frequency distribution table from part b. Use the midpoint of each interval as the representative value.
2. Pick a random sample of 3 values, without replacement, from the population in Problem 1, keeping the original ordering. Compute the sample mean and sample median for those 3 values. Then perform a similar sampling with replacement. Set the random number generator seed to 1 029 384 before each sampling to get reproducible results.

3. Consider the following five sample observations: $X = \{63.2, 72.5, 65.7, 61.7, 68.3\}$.
- (a) Compute the range, mean deviation, variance, standard deviation and coefficient of variation of X .
- (b) Suppose one additional observation, with value 70.0, is appended to X . Compute the range, mean deviation, variance, standard deviation and coefficient of variation of the six samples.

4. Consider the following table of tree species in a random sample from a forest:

Species	Frequency
White Oak	45
Red Oak	3
Shagbark hickory	27
Black walnut	12
Basswood	2
Slippery Elm	8

- (a) Use the Shannon index to express the tree species diversity. Compute the maximum Shannon diversity possible for this number of species, and then calculate the Shannon evenness for this table.
- (b) Compute the Brillouin diversity index for the frequency table in the previous problem. Find the maximum Brillouin diversity, then calculate the Brillouin evenness.
5. A tennis team has 6 boys and 7 girls.
- (a) How many distinct mixed doubles pairs (one boy and one girl) can be formed using members of the team?
- (b) How many distinct practice matchups of two mixed doubles pairs can be formed using members of the team?
6. A DNA modeling kit contains 15 base units: 4 A's, 4 C's, 4 G's, and 3 T's.
- (a) How many distinct sequences of length 15 can be formed from this kit?
- (b) How many distinct sequences of length 3 can be formed from this kit?
- (c*) How many distinct sequences of length 6, 9, or 12 can be formed from this kit?