

Ma 322: Biostatistics

Homework Assignment 1

Prof. Wickerhauser

Due Monday, February 1st, 2021

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. Any version number 2.2 or greater should work. You may then also download R Studio for its convenience.

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. Generate 100 samples from a standard normal density using the following R commands:

```
set.seed(63130); x <- rnorm(100)
```

- (a) Plot the histogram of \mathbf{x} using R defaults.
 - (b) Plot the histogram of \mathbf{x} using integer breakpoints $-4, -3, \dots, 2, 3$ with bins that include their left endpoint but not their right endpoint. HINT: this is not the default for `hist()` in R. It will be necessary to set both the `right=` and `include.lowest=` parameters to non-default values. Read the documentation page!
 - (c) Calculate the mean and median of the samples.
 - (d) Calculate the standard deviation, variance, and mean absolute deviation of the samples.
2. Reuse the population \mathbf{x} generated in Problem 1, keeping the original ordering. Set the random number generator seed to 63130 before each sampling (that is, before each `rnorm()` and each `sample()`) to get reproducible results.
 - (a) Pick a random subsample of 10 values, without replacement, and compute the sample mean and sample median for those 10 values.
 - (b) Pick a random subsample of 10 values with replacement, then compute the sample mean and sample median for those 10 values.
 3. Consider the following table of tree species in a complete count from a section of 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.
4. 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?
5. A DNA modeling kit contains 13 base units: 4 A's, 2 C's, 4 G's, and 3 T's.
- (a) How many distinct sequences of length 2 can be formed from this kit?
- (b) How many distinct sequences of length 13 can be formed from this kit?
- (c) How many distinct sequences of length 3 can be formed from this kit?
- (d) How many distinct sequences of length 6 can be formed from this kit? Of length 9? Of length 12? (Hint: use `deduct()`.)
6. Subsets A, B, C, D, E satisfy $B \subset A$, $C \subset B$, $D \subset B$, $C \cap D = \emptyset$, and $A \cap E = \emptyset$.
- a. Depict the sets using a Venn diagram.
- b. Is $C \cap E = \emptyset$?
7. A standard set of 52 playing cards is divided into 4 suits of 13 ranks each: ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, and king. The suits are called clubs, diamonds, hearts and spades, with clubs and spades being “black” and hearts and diamonds being “red.”
- (a) Taking 1 card at random, what is the probability of drawing a king of spades? a black king? a face card (jack, queen, or king)?
- (b) Taking 2 cards at random without replacement, what is the probability of drawing a pair of kings? a pair of clubs? a pair of black cards? a pair of cards of different ranks and suits?
- (c) Taking 5 cards at random without replacement, what is the probability of drawing a “full house,” namely 3 cards of one rank and 2 cards of a second rank?