Lesson 04: Probability

References

- Black, Chapter 4 Probability (pp. 89-115)
- Davies, Chapter 10 Conditions and Loops (pp. 179-208)
- Library Reserves: Teetor, Chapter 8 Probability (pp. 180-185)

Data sets: shoppers.csv

Description: shoppers.csv contains the dollar amounts spent in a store by individual shoppers during one day.

Exercises:

1) Assume the fifty shoppers exit the store individually in random order.

   a) If one shopper is picked at random, what is the probability of picking a shopper who spent $40 or more dollars? What is the probability of picking a shopper who spent less than $10?

   b) If two shoppers are picked at random, what is the probability the pair will include a shopper who spent $40 or more dollars and one who spent less than $10?

Hint : For parts c) and d) it will be necessary to assume sampling without replacement.

   c) If two shoppers are picked at random, what is the probability the pair will include two shoppers who spent no less than $10 and no more than $40?

   d) If four shoppers are picked at random, what is the probability one shopper will have spent less than $10, one shopper will have spent $40 or more dollars and two shoppers will have spent no less than $10 and no more than $40?

   e) If we know a randomly picked shopper has spent more than $30, what is the probability that shopper has spent more than $40?

2) Use R to answer the following questions.

   a) Draw 100 samples with replacement of size 22 from the 365 integers (i.e. 1,2,.,365). Count the number of samples in which one or more of the numbers sampled is duplicated. Divide by 100 to estimate the probability of such duplication occurring. (If 22 people are selected at random, what is the probability of two or more matching birthdays?)

   b) Suppose that 60% of marbles in a bag are black and 40% are white. Generate a random sample of size 20 with replacement using uniform random numbers. For the numbers in each sample, if a random number is 0.6 or less, code it as a 1. If it is not 0.6 or less code it a zero. Add the twenty coded numbers. Do this 50 times and calculate the proportion of times the sum is 11 or greater. What have you estimated? Expand the number of trials to 10,000. The exact binomial estimated probability is 0.755 and the expectation is 12.