Exercises

6.4.1 Suppose we want to generate a list of pseudorandom numbers $\{x_n : n = 1, \ldots, 1,000\}$ which have values of a, b, or c where a occurs $p_1 \times 100\%$ of the time, b occurs $p_2 \times 100\%$ of the time, and c occurs $p_3 \times 100\%$ of the time where $p_1 + p_2 + p_3 = 1$.

a. Design a spreadsheet which generates this list $\{x_n\}$. You may use the **RAND** function. Make sure the user is able to input the values of a, b, c, p_1, p_2, and p_3.
b. Use the **COUNTIF** function to verify that the list contains the proper percentage of each number a, b, and c.

6.4.2 In each part below, use the **RANDBETWEEN** function to generate a list of equally likely pseudorandom numbers with the given specifications.

a. Even integers between 0 and 20, inclusive.
b. Numbers between 0 and 20, insclusive, with 2 decimal places.
c. Numbers from the set $\{-1, 1\}$. (**Hint**: Use **RANDBETWEEN** to choose between 2 numbers. Use the **IF** function to output a -1 for one of the numbers and a 1 for the other.)
d. Numbers from the set $[-10, -5] \cup [5, 10]$ with 2 decimal places.

6.4.3 Consider the following function for generating a pseudorandom integer x from the set $S = \{a, a+1, \ldots, b\}$ where $a < b$ are integers using the **RAND** function:

$$x = [\mathbf{RAND} \cdot (b - a) + a] \text{ rounded to the nearest integer.}$$

a. Use this function to generate a list of 1000 integers. Allow the user to input the values of a and b.
b. Count the number of times each integer from the set S appears in the list. Do all the integers appear with nearly equal frequency?
c. Suggest a modification to this function so that it generates a list in which the integers from the set S appear with equal frequency. Use your modification to generate another list of 1000 integers. Does the modification generate a list with nearly equal frequencies?

6.4.4 Suppose a pseudorandom number generator gives integers between 0 and 9. One way to test if this generator gives integers with equal frequency is to apply a χ^2 **Goodness–of–fit test**. This can be done by generating a long list of integers (say 500) and then counting the number of times each integer appears in the list. These are called the *observed frequencies*, denoted by O. We then calculate the *expected frequencies*, denoted by E, which is the number of times we expect each integer to appear in the list if they do indeed occur with equal frequencies. If we have 10 different integers in a list of 500, we would expect each one to appear 50 times.

Then we calculate the "test statistic" χ^2 by

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

If this test statistic is "small" (for 10 different integers "small" is less than 16.9) then we can be 95% confident that the generator gives integers with equal frequencies. If it is "large" then we reject the claim that it gives integers with equal frequencies. (For a more detailed description of this test, see any introductory statistics textbook.)

a. The Excel function **RANDBETWEEN** gives pseudorandom integers between two specified values. Use it to generate a list of 500 integers between 0 and 9. Calculate the χ^2 test statistic as described above and press **F9** several times to get several different lists of integers. Does the **RANDBETWEEN** appear to give integers with equal frequencies?
b. Use the linear congruence algorithm with $a = 6$, $b = 9$, $m = 10$, and $x_0 = 2$ to generate a list of 500 integers between 0 and 9. Calculate the χ^2 test statistic. Does this linear congruence algorithm appear to give integers with equal frequencies? Try different values of a, b, and x_0. What do you observe?

6.5 Modeling Random Variables

Simulation is a useful tool for modeling the interaction of **random events**. A random event is an activity where we do not know the outcome until it occurs. Therefore, constructing a simulation involves modeling random events. One of the most important concepts used in modeling random events is the *random variable*.

Definition 6.5.1 (Random Variables) A **Random Variable** is a rule for assigning real numbers to the observations of a random event. A **Discrete Random Variable** can take only certain distinct values (such as integers). A **Continuous Random Variable** can take any value within some interval.

In most cases, defining the random variable involved is rather obvious. For instance, suppose we roll a standard six–sided die, and we define the random variable X as the number on the top face of the die. This is an example of a discrete random variable. In another example, suppose we observe customer arriving at a check–out line at the grocery store. We define the random variable Y as the time between customer arrivals (called the inter-arrival times). This is an example of a continuous random variable.

In a simulation of a dice game, we need to generate values of the roll of the dice. In a simulation of a check–out line, we need to generate values of the inter–arrival times. In other words, simulations involve generating values of random variables. In this section we will discuss how to do this using the **RAND** function.

We will focus on continuous random variables. One of the most important tools use to model continuous random variables is the **density function**. Any function $f(x)$ is a density function if it satisfies the following two properties: