Appendix B

APPENDIX B (PART 1): Counting Rules

Consider the situation in Example 3.2 (page 88) in which a student takes a pop quiz that consists of three true–false questions. If we consider our experiment to be answering the three questions, each question can be answered correctly or incorrectly. We will let C denote answering a question correctly and I denote answering a question incorrectly. Figure B.1 depicts a tree diagram of the sample space outcomes for the experiment. The diagram portrays the experiment as a three-step process—answering the first question (correctly or incorrectly, that is, C or I), answering the second question (correctly or incorrectly, that is, C or I), and answering the third question (correctly or incorrectly, that is, C or I). The tree diagram has eight different branches, and the eight distinct sample space outcomes are listed at the ends of the branches.

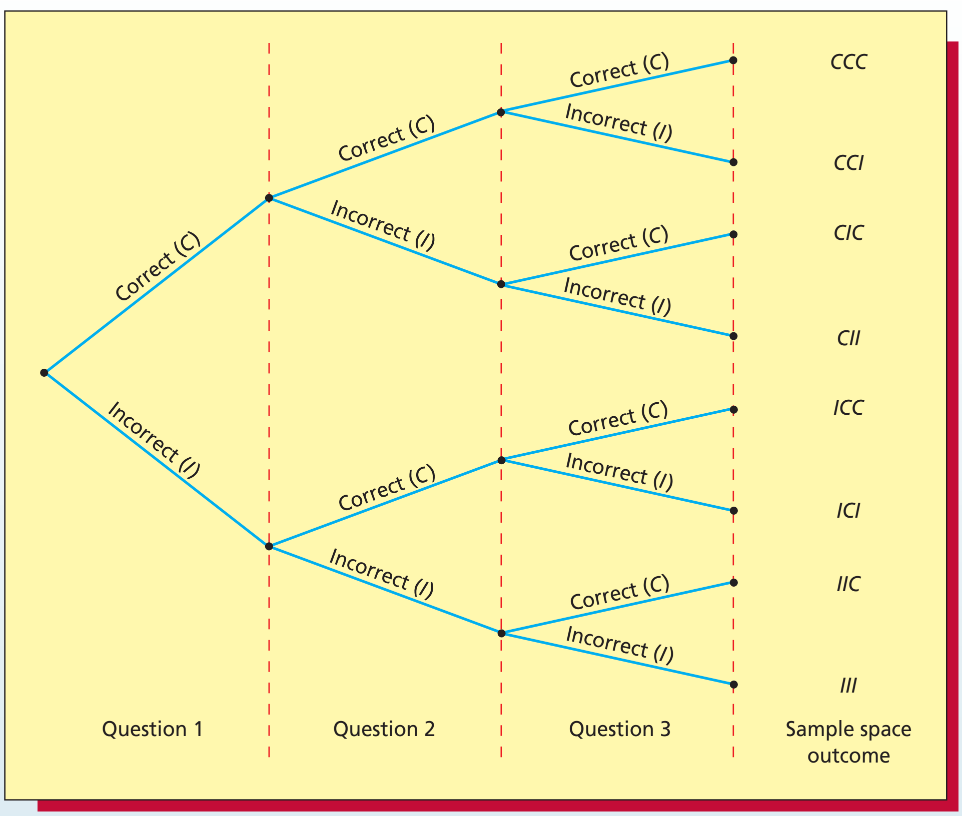
In general, a rule that is helpful in determining the number of experimental outcomes in a multiple-step experiment is as follows:

A Counting Rule for Multiple-Step Experiments

If an experiment can be described as a sequence of k steps in which there are n_1 possible outcomes on the first step, n_2 possible outcomes on the second step, and so on, then the total number of experimental outcomes is given by $(n_1)(n_2) \cdots (n_k)$.

For example, the pop quiz example consists of three steps in which there are $n_1 = 2$ possible outcomes on the first step, $n_2 = 2$ possible outcomes on the second step, and $n_3 = 2$ possible outcomes on the third step. Therefore, the total number of experimental outcomes is $(n_1)(n_2)(n_3) = (2)(2)(2) = 8$, as is shown in Figure B.1. Now suppose the student takes a pop quiz consisting of five true–false questions. Then, there are $(n_1)(n_2)(n_3)(n_4)(n_5) = (2)(2)(2)(2)(2) = 32$ experimental outcomes. If the student is totally unprepared for the quiz and has to blindly guess the answer

FIGURE B.1 A Tree Diagram of Answering Three True–False Questions



to each question, the 32 experimental outcomes might be considered to be equally likely. Therefore, since only one of these outcomes corresponds to all five questions being answered correctly, the probability that the student will answer all five questions correctly is 1/32.

As another example, suppose a bank has three branches; each branch has two departments, and each department has four employees. One employee is to be randomly selected to go to a convention. Since there are $(n_1)(n_2)(n_3) = (3)(2)(4) = 24$ employees, the probability that a particular one will be randomly selected is 1/24.

Next, consider the population of the percentage returns for last year of six high-risk stocks. This population consists of the percentage returns −36, −15, 3, 15, 33, and 54 (which we have arranged in increasing order). Now consider randomly selecting without replacement a sample of $n = 3$ stock returns from the population of six stock returns. Below we list the 20 distinct samples of $n = 3$ returns that can be obtained:

Sample	$n = 3$ Returns in Sample	Sample	$n = 3$ Returns in Sample
1	−36, −15, 3	11	−15, 3, 15
2	−36, −15, 15	12	−15, 3, 33
3	−36, −15, 33	13	−15, 3, 54
4	−36, −15, 54	14	−15, 15, 33
5	−36, 3, 15	15	−15, 15, 54
6	−36, 3, 33	16	−15, 33, 54
7	−36, 3, 54	17	3, 15, 33
8	−36, 15, 33	18	3, 15, 54
9	−36, 15, 54	19	3, 33, 54
10	−36, 33, 54	20	15, 33, 54

Since each sample is specified only with respect to what returns are contained in the sample, and therefore not with respect to the different orders in which the returns can be randomly selected, each sample is called a **combination of $n = 3$ stock returns selected from $N = 6$ stock returns.** In general, the following result can be proven:

A Counting Rule for Combinations

The number of combinations of n items that can be selected from N items is

$$\binom{N}{n} = \frac{N!}{n!\,(N-n)!}$$

where

$$N! = N(N-1)(N-2)\cdots 1$$

$$n! = n(n-1)(n-2)\cdots 1$$

Note: 0! is defined to be 1.

For example, the number of combinations of $n = 3$ stock returns that can be selected from the six previously discussed stock returns is

$$\binom{6}{3} = \frac{6!}{3!(6-3)!} = \frac{6!}{3!\,3!} = \frac{6 \cdot 5 \cdot 4\,\cancel{(3 \cdot 2 \cdot 1)}}{(3 \cdot 2 \cdot 1)\,\cancel{(3 \cdot 2 \cdot 1)}} = 20$$

The 20 combinations are listed above. As another example, the Ohio lottery system uses the random selection of 6 numbers from a group of 47 numbers to determine each week's lottery winner. There are

$$\binom{47}{6} = \frac{47!}{6!\,(47-6)!} = \frac{47 \cdot 46 \cdot 45 \cdot 44 \cdot 43 \cdot 42\,\cancel{(41!)}}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1\,\cancel{(41!)}} = 10{,}737{,}573$$

combinations of 6 numbers that can be selected from 47 numbers. Therefore, if you buy a lottery ticket and pick six numbers, the probability that this ticket will win the lottery is 1/10,737,573.

Exercises for Appendix B (Part 1)

B.1 A credit union has two branches; each branch has two departments, and each department has four employees. How many total people does the credit union employ? If you work for the credit union, and one employee is randomly selected to go to a convention, what is the probability that you will be chosen?

B.2 How many combinations of two high-risk stocks could you randomly select from eight high-risk stocks? If you did this, what is the probability that you would obtain the two highest-returning stocks?

APPENDIX B (PART 2): The Hypergeometric Distribution

Suppose a population consists of N items (for example, stocks); r of these items are "successes" (for example, stocks with positive returns for a year), and $(N - r)$ of these items are "failures" (for example, stocks with zero or negative returns for the year). If we randomly select n of the N items without replacement, it can be shown that the probability that x of the n randomly selected items will be successes is given by the **hypergeometric distribution**

$$p(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

For example, recall from Appendix B (Part 1) that $N = 6$ high-risk stocks gave the following returns last year: -36, -15, 3, 15, 33, and 54. Note that $r = 4$ of these returns are positive, and $N - r = 2$ of these returns are negative. Suppose you randomly selected $n = 3$ of the six stocks to invest in at the beginning of last year, and let x denote the number of the three stocks that gave positive returns. It follows that the probability that x is at least 2 (that is, the probability that x is either 2 or 3) is

$$p(2) + p(3) = \frac{\binom{4}{2}\binom{2}{1}}{\binom{6}{3}} + \frac{\binom{4}{3}\binom{2}{0}}{\binom{6}{3}} = \frac{\left(\frac{4!}{2!\,2!}\right)\left(\frac{2!}{1!\,1!}\right)}{\frac{6!}{3!\,3!}} + \frac{\left(\frac{4!}{3!\,1!}\right)\left(\frac{2!}{0!\,2!}\right)}{\frac{6!}{3!\,3!}}$$

$$= \frac{(6)(2)}{20} + \frac{(4)(1)}{20} = .8$$

Note that, on the first random selection from the population of N items, the probability of a success is r/N. Since we are making selections **without replacement,** the probability of a success changes as we continue to make selections. However, if the population size N is "much larger" than the sample size n (say, at least 20 times as large), then making the selections will not substantially change the probability of a success. In this case, we can assume that the probability of a success stays essentially constant from selection to selection, and the different selections are essentially independent of each other. Therefore, we can approximate the hypergeometric distribution by the easier-to-compute binomial distribution:

$$p(x) = \frac{n!}{x!\,(n-x)!}p^x(1-p)^{n-x} = \frac{n!}{x!\,(n-x)!}\left(\frac{r}{N}\right)^x\left(1-\frac{r}{N}\right)^{n-x}$$

The reader will use this approximation in Exercise B.4.

Exercises for Appendix B (Part 2)

B.3 Suppose that you purchase (randomly select) 3 TV sets from a production run of 10 TV sets. Of the 10 TV sets, 9 are destined to last at least five years without needing a single repair. What is the probability that all three of your TV sets will last at least five years without needing a single repair?

B.4 Suppose that you own an electronics store and purchase (randomly select) 15 TV sets from a production run of 500 TV sets. Of the 500 TV sets, 450 are destined to last at least five years without needing a single repair. Set up an expression using the hypergeometric distribution for the probability that at least 14 of your 15 TV sets will last at least five years without needing a single repair. Then, using the binomial tables, approximate this probability by using the binomial distribution. What justifies the approximation? Hint: $p = r/N = 450/500 = .9$.