

Econ 205

Probability – 2

Course Outline

- **Probability**

- Probability and Inference
- The Discrete Case
- Calculating the Probability of an Event
- Conditional Probability and the Independence of Events
- Laws of Probability
- The Law of Total Probability and Bayes' Rule
- Introduction to Random Variables and their distributions

- **Discrete Random Variables and Their Probability Distributions**

- Discrete Random Variable
- The Probability Distribution for Discrete Random Variable
- The Expected Value of Random Variable
- Moments and Moment-Generating Functions
- The Bernoulli Probability Distribution
- The Binomial Probability Distribution
- The Geometric Probability Distribution
- The Negative Binomial Probability Distribution
- The Hypergeometric Probability Distribution
- The Poisson Probability Distribution
- Probability-Generating Functions

- **Continuous Random Variables and Their Probability Densities**

- Continuous Random Variable
- The Probability Distribution for Continuous Random Variable
- The Expected Value for Continuous Random Variable
- Moments and Moment-Generating Functions
- The Uniform Probability Distribution
- The Normal Probability Distribution
- Exponential Probability Distribution
- Chi-Square Probability Distribution
- Tchebysheff's Theorem

Course Outline (cont'd)

- **Multivariate Probability Distributions**

- Bivariate and Multivariate Probability Distributions
- Independent Random Variables
- Marginal and Conditional Probability Distributions
- The Expected Value of a Function of Random Variables
- Special Theorems
- The Covariance of Two Random Variables
- The Expected Value and Variance of Linear Functions of Random Variables
- The Multinomial Probability Distribution
- The Bivariate Normal Distribution
- Conditional Expectations

- **Functions of Random Variables**

- Finding the Probability Distribution of a Function of Random Variables
- The Method of Distribution Functions
- The Method of Transformations
- The Method of Moment-Generating Functions
- Multivariable Transformations Using Jacobians

Table of Contents of Devore & Berk

2	Probability	50
	Introduction	50
2.1	Sample Spaces and Events	51
2.2	Axioms, Interpretations, and Properties of Probability	56
2.3	Counting Techniques	66
2.4	Conditional Probability	74
2.5	Independence	84
3	Discrete Random Variables and Probability Distributions	96
	Introduction	96
3.1	Random Variables	97
3.2	Probability Distributions for Discrete Random Variables	101
3.3	Expected Values of Discrete Random Variables	112
3.4	Moments and Moment Generating Functions	121
3.5	The Binomial Probability Distribution	128
3.6	Hypergeometric and Negative Binomial Distributions	138
3.7	The Poisson Probability Distribution	146
4	Continuous Random Variables and Probability Distributions	158
	Introduction	158
4.1	Probability Density Functions and Cumulative Distribution Functions	159
4.2	Expected Values and Moment Generating Functions	171
4.3	The Normal Distribution	179
4.4	The Gamma Distribution and Its Relatives	194
4.5	Other Continuous Distributions	202
4.6	Probability Plots	210
4.7	Transformations of a Random Variable	220
5	Joint Probability Distributions	232
	Introduction	232
5.1	Jointly Distributed Random Variables	233
5.2	Expected Values, Covariance, and Correlation	245
5.3	Conditional Distributions	253
5.4	Transformations of Random Variables	265
5.5	Order Statistics	271

Content of this presentation

- Probability (cont'd)
 - Conditional probability
 - The Law of Total Probability
 - Bayes' theorem
 - Independence of events
 - Calculating the probability of an event
 - The Event Composition Method
- Examples

Conditional Probability:

We examine how the information “an event B has occurred” affects the probability assigned to the event A .

Definition:

For any two events A and B with $P(B) > 0$, the **conditional probability of A given that B has occurred** is defined by

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

Corollary:

THE MULTI-
PLICATION
RULE

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Motivation for the definition:

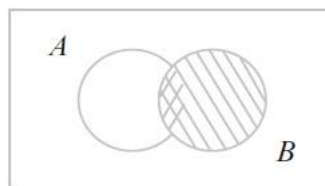


Figure 2.8 Motivating the definition of conditional probability

Given that B has occurred, the relevant sample space is no longer $\mathcal{S}$ but consists of just outcomes in B ; A has occurred if and only if one of the outcomes in the intersection occurred, so the conditional probability of A given B is proportional to $P(A \cap B)$. The proportionality constant $1/P(B)$ is used to ensure that the probability $P(B|B)$ of the new sample space B equals 1.

The rule is easily extended to experiments involving more than two stages. For example,

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_3 | A_1 \cap A_2) \cdot P(A_1 \cap A_2) \\ &= P(A_3 | A_1 \cap A_2) \cdot P(A_2 | A_1) \cdot P(A_1) \end{aligned}$$

where A_1 occurs first, followed by A_2 , and finally A_3 .

Conditional Probability: The Law of Total Probability

Definition:

events $A_1, \dots, A_k$ are **mutually exclusive** if no two have any common outcomes. The events are **exhaustive** if one A_i must occur, so that $A_1 \cup \dots \cup A_k = \mathcal{S}$.

(Remark: The collection of events that are both mutually exclusive and exhaustive is also called a “partition” of the sample space.)

THE LAW OF TOTAL PROBABILITY

Let $A_1, \dots, A_k$ be **mutually exclusive** and **exhaustive** events. Then for any other event B ,

$$\begin{aligned} P(B) &= P(B | A_1) \cdot P(A_1) + \dots + P(B | A_k) \cdot P(A_k) \\ &= \sum_{i=1}^k P(B | A_i) P(A_i) \end{aligned}$$

Proof Because the A_i 's are mutually exclusive and exhaustive, if B occurs it must be in conjunction with exactly one of the A_i 's. That is, $B = (A_1 \text{ and } B) \text{ or } \dots \text{ or } (A_k \text{ and } B) = (A_1 \cap B) \cup \dots \cup (A_k \cap B)$, where the events $(A_i \cap B)$ are mutually exclusive. This “partitioning of B ” is illustrated in Figure 2.11. Thus

$$P(B) = \sum_{i=1}^k P(A_i \cap B) = \sum_{i=1}^k P(B | A_i) P(A_i)$$

as desired.

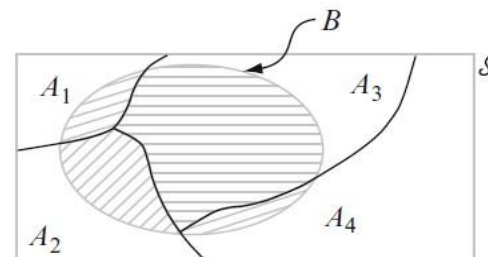


Figure 2.11 Partition of B by mutually exclusive and exhaustive A_i 's

Conditional Probability: Bayes' Theorem

Motivation: The computation of a posterior probability $P(A_j|B)$ from given prior probabilities $P(A_i)$ and conditional probabilities $P(B|A_i)$

BAYES' THEOREM Let $A_1, \dots, A_k$ be a collection of mutually exclusive and exhaustive events with $P(A_i) > 0$ for $i = 1, \dots, k$. Then for any other event B , for which $P(B) > 0$

$$P(A_j | B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B | A_j)P(A_j)}{\sum_{i=1}^k P(B | A_i)P(A_i)} \quad j = 1, \dots, k \quad (2.6)$$

Proof is immediate if the multiplication rule is used in the numerator and the law of total probability is used in the denominator.

Independence of Events

Definition:

Two events A and B are **independent** if $P(A | B) = P(A)$ and are **dependent** otherwise.

Proposition:

The following are equivalent:

$$P(A|B) = P(A),$$

$$P(B|A) = P(B),$$

$$P(A \cap B) = P(A)P(B).$$

Proof follows from
$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

Proposition:

if A and B are independent, then so are the following pairs of events: (1) A' and B , (2) A and B' , and (3) A' and B' .

Proof of (1): Since $P(A' | B) + P(A | B) = 1$,
 $P(A' | B) = 1 - P(A | B) = 1 - P(A) = P(A')$ (where the second equality in the second line of proof is by independence).

Independence of More Than Two Events:

Definition:

Events $A_1, \dots, A_n$ are **mutually independent** if for every k ($k = 2, 3, \dots, n$) and every subset of indices $i_1, i_2, \dots, i_k$,

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$$

Calculating the Probability of an Event: The Event Composition Method

A summary of the steps used in the event-composition method follows:

1. Define the experiment.
2. Visualize the nature of the sample points. Identify a few to clarify your thinking.
3. Write an equation expressing the event of interest—say, A —as a composition of two or more events, using unions, intersections, and/or complements. (Notice that this equates point sets.) Make certain that event A and the event implied by the composition represent the same set of sample points.
4. Apply the additive and multiplicative laws of probability to the compositions obtained in step 3 to find $P(A)$.

Recall the additive law :

For any events A and B ,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Recall the multiplicative law
(multiplication rule) :

THE MULTI-
PLICATION
RULE

$$P(A \cap B) = P(A|B) \cdot P(B)$$

EXAMPLES

Examples 1-2: Conditional Probability

Example 1:

Complex components are assembled in a plant that uses two different assembly lines, A and A' . Line A uses older equipment than A' , so it is somewhat slower and less reliable. Suppose on a given day line A has assembled 8 components, of which 2 have been identified as defective (B) and 6 as nondefective (B'), whereas A' has produced 1 defective and 9 nondefective components. This information is summarized in the accompanying table.

Line	Condition	
	B	B'
A	2	6
A'	1	9

Unaware of this information, the sales manager randomly selects 1 of these 18 components for a demonstration. Prior to the demonstration

$$P(\text{line } A \text{ component selected}) = P(A) = \frac{N(A)}{N} = \frac{8}{18} = .444$$

However, if the chosen component turns out to be defective, then the event B has occurred, so the component must have been 1 of the 3 in the B column of the table. Since these 3 components are equally likely among themselves after B has occurred,

$$P(A|B) = \frac{2}{3} = \frac{2/18}{3/18} = \frac{P(A \cap B)}{P(B)}$$

Example 2:

A news magazine includes three columns entitled “Art” (A), “Books” (B), and “Cinema” (C). Reading habits of a randomly selected reader with respect to these columns are

Read regularly	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
Probability	.14	.23	.37	.08	.09	.13	.05

(See Figure 2.9.)

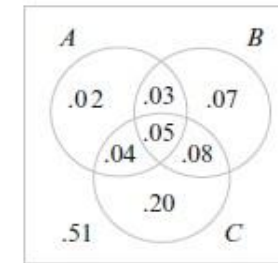


Figure 2.9 Venn diagram for Example 2.26

We thus have

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{.08}{.23} = .348 \\ P(A|B \cup C) &= \frac{P(A \cap (B \cup C))}{P(B \cup C)} = \frac{.04 + .05 + .03}{.47} = \frac{.12}{.47} = .255 \\ P(A|\text{reads at least one}) &= P(A|A \cup B \cup C) = \frac{P(A \cap (A \cup B \cup C))}{P(A \cup B \cup C)} \\ &= \frac{P(A)}{P(A \cup B \cup C)} = \frac{.14}{.49} = .286 \end{aligned}$$

and

$$P(A \cup B|C) = \frac{P((A \cup B) \cap C)}{P(C)} = \frac{.04 + .05 + .08}{.37} = .459$$

Example 3: Conditional Probability

Example 3: The Multiplication Rule

Four individuals have responded to a request by a blood bank for blood donations. None of them has donated before, so their blood types are unknown. Suppose only type O+ is desired and only one of the four actually has this type. If the potential donors are selected in random order for typing, what is the probability that at least three individuals must be typed to obtain the desired type?

Making the identification $B = \{\text{first type not O+}\}$ and $A = \{\text{second type not O+}\}$, $P(B) = 3/4$. Given that the first type is not O+, two of the three individuals left are not O+, so $P(A|B) = 2/3$. The multiplication rule now gives

$$\begin{aligned} P(\text{at least three individuals are typed}) &= P(A \cap B) \\ &= P(A|B) \cdot P(B) \\ &= \frac{2}{3} \cdot \frac{3}{4} = \frac{6}{12} \\ &= .5 \end{aligned}$$

What is the probability that the third type is O+?

$$\begin{aligned} P(\text{third type is O+}) &= P(\text{third is | first isn't} \cap \text{second isn't}) \\ &\quad \cdot P(\text{second isn't | first isn't}) \cdot P(\text{first isn't}) \\ &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{4} = .25 \end{aligned}$$



Example 4: Conditional Probability: Bayes' Theorem

INCIDENCE OF A RARE DISEASE Only 1 in 1000 adults is afflicted with a rare disease for which a diagnostic test has been developed. The test is such that when an individual actually has the disease, a positive result will occur 99% of the time, whereas an individual without the disease will show a positive test result only 2% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease?

To use Bayes' theorem, let $A_1 = \{\text{individual has the disease}\}$, $A_2 = \{\text{individual does not have the disease}\}$, and $B = \{\text{positive test result}\}$. Then $P(A_1) = .001$, $P(A_2) = .999$, $P(B|A_1) = .99$, and $P(B|A_2) = .02$. The tree diagram for this problem is in Figure 2.12.

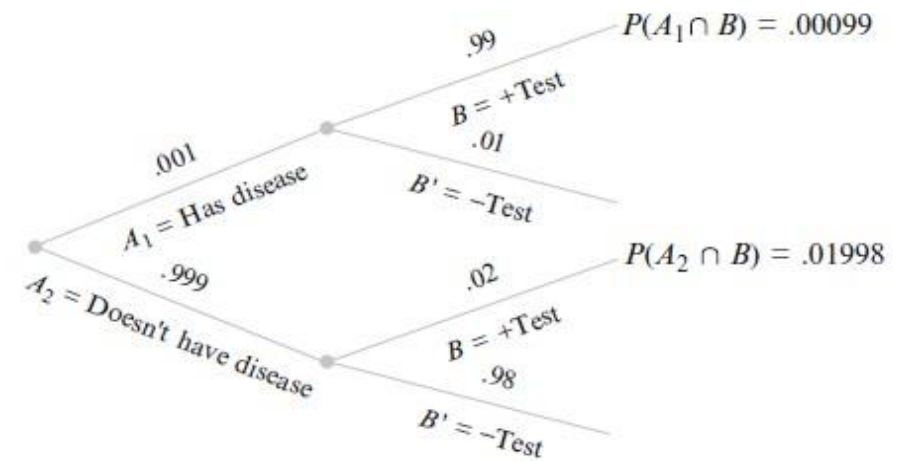


Figure 2.12 Tree diagram for the rare-disease problem

Next to each branch corresponding to a positive test result, the multiplication rule yields the recorded probabilities. Therefore, $P(B) = .00099 + .01998 = .02097$, from which we have

$$P(A_1 | B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{.00099}{.02097} = .047$$

This result seems counterintuitive; because the diagnostic test appears so accurate, we expect someone with a positive test result to be highly likely to have the disease, whereas the computed conditional probability is only .047. However, because the disease is rare and the test only moderately reliable, most positive test results arise from errors rather than from diseased individuals.

Examples 5-6: Independence of Events

Example 5:

Consider the following events in the toss of a single die:

- A : Observe an odd number.
- B : Observe an even number.
- C : Observe a 1 or 2.

- a Are A and B independent events?
- b Are A and C independent events?

Recall:

Two events A and B are **independent** if $P(A|B) = P(A)$ and are **dependent** otherwise.

Solution

- a To decide whether A and B are independent, we must see whether they satisfy the conditions of Definition 2.10. In this example, $P(A) = 1/2$, $P(B) = 1/2$, and $P(C) = 1/3$. Because $A \cap B = \emptyset$, $P(A|B) = 0$, and it is clear that $P(A|B) \neq P(A)$. Events A and B are dependent events.
- b Are A and C independent? Note that $P(A|C) = 1/2$ and, as before, $P(A) = 1/2$. Therefore, $P(A|C) = P(A)$, and A and C are independent. ■

Example 6:

Let A and B be any two mutually exclusive events with $P(A) > 0$. For example, for a randomly chosen automobile, let $A = \{\text{car is blue}\}$ and $B = \{\text{car is red}\}$. Since the events are mutually exclusive, if B occurs, then A cannot possibly have occurred, so $P(A|B) = 0 \neq P(A)$. The message here is that *if two events are mutually exclusive, they cannot be independent*. When A and B are mutually exclusive, the information that A occurred says something about B (it cannot have occurred), so independence is precluded. ■

Example 7: Independence of Events, The Event-Composition method

Each day, Monday through Friday, a batch of components sent by a first supplier arrives at a certain inspection facility. Two days a week, a batch also arrives from a second supplier. Eighty percent of all supplier 1's batches pass inspection, and 90% of supplier 2's do likewise. What is the probability that, on a randomly selected day, two batches pass inspection? We will answer this assuming that on days when two batches are tested, whether the first batch passes is independent of whether the second batch does so. Figure 2.13 displays the relevant information.

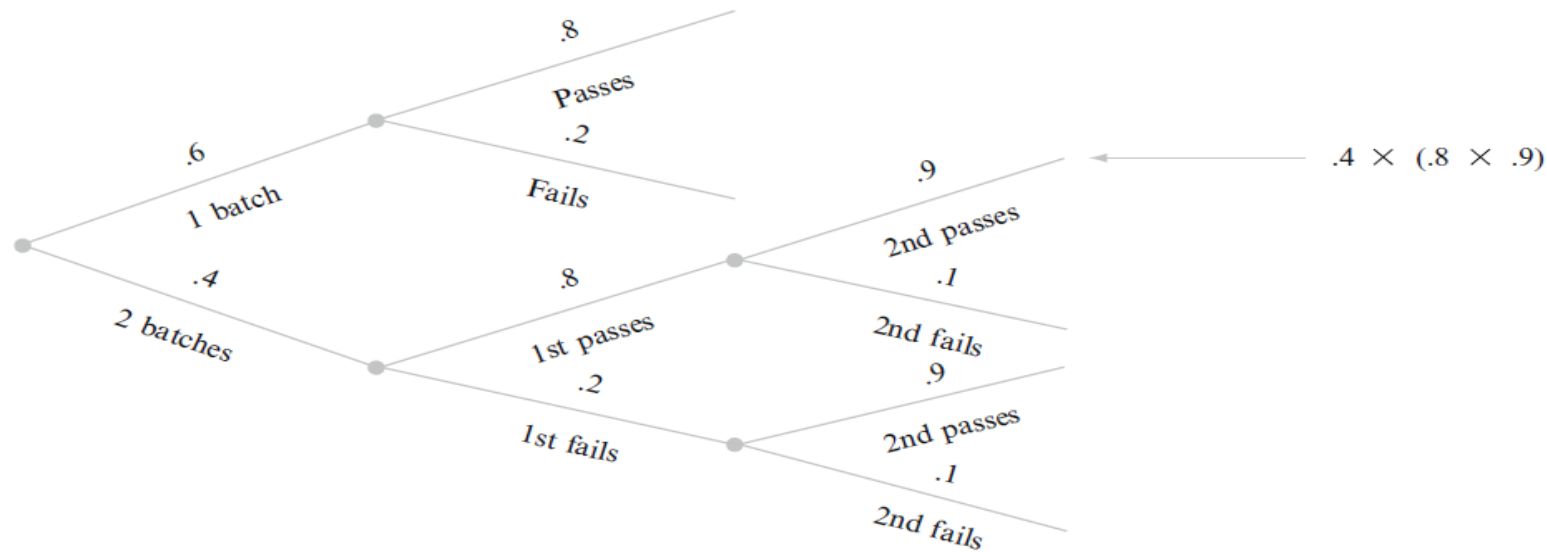


Figure 2.13 Tree diagram for Example 2.34

$$\begin{aligned} P(\text{two pass}) &= P(\text{two received} \cap \text{both pass}) \\ &= P(\text{both pass} \mid \text{two received}) \cdot P(\text{two received}) \\ &= [(.8)(.9)](.4) = .288 \end{aligned}$$



Example 8: Independence of Events, The Event-Composition Method

Example 8: Independence of More Than 2 Events

The article “Reliability Evaluation of Solar Photovoltaic Arrays” (*Solar Energy*, 2002: 129–141) presents various configurations of solar photovoltaic arrays consisting of crystalline silicon solar cells. Consider first the system illustrated in Figure 2.14a. There are two subsystems connected in parallel, each one containing three cells. In order for the system to function, at least one of the two parallel subsystems must work. Within each subsystem, the three cells are connected in series, so a subsystem will work only if all cells in the subsystem work. Consider a particular lifetime value t_0 , and suppose we want to determine the probability that the system lifetime exceeds t_0 . Let A_i denote the event that the lifetime of cell i exceeds t_0 ($i = 1, 2, \dots, 6$). We assume that the A_i 's are independent events (whether any particular cell lasts more than t_0 hours has no bearing on whether any other cell does) and that $P(A_i) = .9$ for every i since the cells are identical. Then

$$\begin{aligned} P(\text{system lifetime exceeds } t_0) &= P[(A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_5 \cap A_6)] \\ &= P(A_1 \cap A_2 \cap A_3) + P(A_4 \cap A_5 \cap A_6) \\ &\quad - P[(A_1 \cap A_2 \cap A_3) \cap (A_4 \cap A_5 \cap A_6)] \\ &= (.9)(.9)(.9) + (.9)(.9)(.9) \\ &\quad - (.9)(.9)(.9)(.9)(.9)(.9) = .927 \end{aligned}$$

Alternatively,

$$\begin{aligned} P(\text{system lifetime exceeds } t_0) &= 1 - P(\text{both subsystem lives are } \leq t_0) \\ &= 1 - [P(\text{subsystem life is } \leq t_0)]^2 \\ &= 1 - [1 - P(\text{subsystem life is } > t_0)]^2 \\ &= 1 - [1 - (.9)^3]^2 = .927 \end{aligned}$$

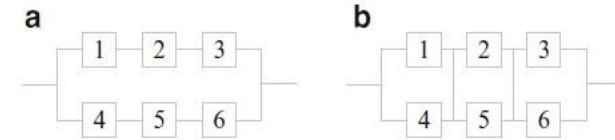


Figure 2.14 System configurations for Example 2.35: (a) series-parallel; (b) total-cross-tied

Next consider the total-cross-tied system shown in Figure 2.14b, obtained from the series-parallel array by connecting ties across each column of junctions. Now the system fails as soon as an entire column fails, and system lifetime exceeds t_0 only if the life of every column does so. For this configuration,

$$\begin{aligned} P(\text{system lifetime exceeds } t_0) &= [P(\text{column lifetime exceeds } t_0)]^3 \\ &= [1 - P(\text{column lifetime is } \leq t_0)]^3 \\ &= [1 - P(\text{both cells in a column have lifetime } \leq t_0)]^3 \\ &= 1 - [1 - (.9)^2]^3 = .970 \end{aligned}$$

There is a mistake here
(on p.88 of Devore & Berk). It should be:

$$= (1 - [1 - (0.9)(0.9) - 2(0.9)(0.1)])^3 = 0.970$$

Example 9: Independence of Events, The Event-Composition Method

EXAMPLE Observation of a waiting line at a medical clinic indicates the probability that a new arrival will be an emergency case is $p = 1/6$. Find the probability that the r th patient is the first emergency case. (Assume that conditions of arriving patients represent independent events.)

Solution The experiment consists of watching patient arrivals until the first emergency case appears. Then the sample points for the experiment are

E_i : The i th patient is the first emergency case, for $i = 1, 2, \dots$

Because only one sample point falls in the event of interest,

$$P(\text{rth patient is the first emergency case}) = P(E_r).$$

Now define A_i to denote the event that the i th arrival is not an emergency case. Then we can represent E_r as the intersection

$$E_r = A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{r-1} \cap \bar{A}_r.$$

Applying the multiplicative law, we have

$$P(E_r) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(\bar{A}_r|A_1 \cap \dots \cap A_{r-1}),$$

and because the events $A_1, A_2, \dots, A_{r-1}$, and $\bar{A}_r$ are independent, it follows that

$$\begin{aligned} P(E_r) &= P(A_1)P(A_2) \dots P(A_{r-1})P(\bar{A}_r) = (1 - p)^{r-1}p \\ &= (5/6)^{r-1}(1/6), \quad r = 1, 2, 3, \dots \end{aligned}$$

Notice that

$$\begin{aligned} P(S) &= P(E_1) + P(E_2) + P(E_3) + \dots + P(E_i) + \dots \\ &= (1/6) + (5/6)(1/6) + (5/6)^2(1/6) + \dots + (5/6)^{i-1}(1/6) + \dots \\ &= \frac{1}{6} \sum_{i=0}^{\infty} \left(\frac{5}{6}\right)^i = \frac{1/6}{1 - (5/6)} = 1. \end{aligned}$$

This result follows from the formula for the sum of a *geometric series* given in Appendix A1.11. This formula, which states that if $|r| < 1$, $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$, is useful in many simple probability problems. ■

References

- Devore, J.L. and Berk, K.N. (2012) Modern Mathematical Statistics with Applications, 2nd Edition, New York: Springer.
- Wackerly, D.D., Mendenhall, W. and Scheaffer, R.L., (2008) Mathematical Statistics With Applications, 7th Edition, Pacific Grove: Duxbury.