Computing Probability

James H. Steiger

October 22, 2003

1 Goals for this Module

In this module, we will

1. Develop a general rule for computing probability, and a special case rule applicable when elementary events are equally likely.
2. Define and discuss
   (a) Joint Events
   (b) *Conditional Probability*
   (c) *Independence*
3. Derive the rule for computing the *probability of a sequence*, and a special case rule applicable when events are independent.

2 Computing Probability

2.1 The General Rule

For any event A in Ω, A is the union of elementary events, which are non-intersecting. Consequently, to compute the probability of A, simply sum the probabilities of the elementary events in A.

Example 1 *You have an unfair die, with probabilities as listed in the following Table. Find the probability of the event* $E = \{2, 4, 6\}$.

X	Pr(X)
6	*.3*
5	*.2*
4	*.1*
3	*.1*
2	*.2*
1	*.1*

Solution. Simply add the probabilities of the events $\{2\}$, $\{4\}$, and $\{6\}$ together, obtaining .6 as the answer.

Example 2 *Given the probabilities in the preceding Table, find the probability that the die will be even and greater than 3.*

Solution. This event is the set $\{4, 6\}$. Adding the probabilities for $\{4\}$, and $\{6\}$, we obtain .4 as the answer.

2.2 Equally Likely Elementary Events

Since the probabilities of elementary events must sum to 1, when elementary events are equally likely each elementary event has probability $1/N_\Omega$, where N_Ω is the total number of elementary events in Ω. Consequently, any event A composed of N_A elementary events must have probability given by

$$\Pr(A) = \frac{N_A}{N_\Omega}$$

Example 3 *Suppose a die is fair. Then each side is equally likely to come up. There are 6 sides, and for their probabilities to add to 1, each side must have a probability of $1/6$. In this case $N_\Omega = 6$. Suppose our event of interest, call it A, is that the number is even and less than 6. To compute the probability of a number being even and less than 6, we simply ask how many elementary events are in A. Since $A = \{2, 4\}$, $N_A = 2$, and*

$$\Pr(A) = \frac{2}{6}$$

3 Joint Events

So far, we have discussing probabilities when the events of interest involve one process. In many cases, however, we are interested in the simultaneous behavior of more than one process. In that case, we are interested in *joint events.*

Definition 4 (Joint Events) ***Joint events*** *are the intersection of outcomes on two different processes.*

Example 5 *You throw a fair coin and toss a fair die. The joint event is the intersection of the outcome on the coin and the outcome on the die. The following table, which is also a Venn diagram, shows the possibilities*

			Die			
Coin	1	2	3	4	5	6
H	$H \cap 1$	$H \cap 2$	$H \cap 3$	$H \cap 4$	$H \cap 5$	$H \cap 6$
T	$T \cap 1$	$T \cap 2$	$T \cap 3$	$T \cap 4$	$T \cap 5$	$T \cap 6$

Definition 6 (Marginal Events) *The events on the margins of the above table are called* ***marginal events****. They are the union of the events in the respective row or column.*

Example 7 *One of the marginal events in the table is the event* $\{3\}$*. Note that the only ways a 3 can occur is if you throw either a Head and a 3 or a Tail and a 3. So* $\{3\} = (H \cap 3) \cup (T \cap 3)$*.Notice in passing that any marginal event is partitioned by the events in its row or column. Consequently, by the 3rd axiom of probability theory, the probabilities of the marginal events in any row or column can be obtained by adding all the entries in the cells of the respective row or column. We examine this notion further on the next slide.*

3.1 Probabilities in a Joint Probability Table

The table below shows probabilities for the joint events in the interior of the table, and probabilities for the marginal events are listed on the margins.

				Die			
		1/6	1/6	1/6	1/6	1/6	1/6
	Coin	**1**	**2**	**3**	**4**	**5**	**6**
1/2	**H**	1/6	0	1/6	0	1/6	0
1/2	**T**	0	1/6	0	1/6	0	1/6

How would you characterize in words the performance of the coin and the die as depicted by the probabilities in the above table? Consider their behavior both separately and together, i.e., their joint and marginal behavior.

4 Conditional Probability

Definition 8 (Conditional Probability) *The conditional probability of* A *given* B*, denoted* $\Pr(A|B)$*, is the probability of* A *within the reduced sample space defined by* B*.*

By way of comment, let me continue by saying that to compute conditional probabilities, you go into the reduced sample space of events where B occurs, and compute the probability of obtaining A within that reduced sample space.

Example 9 *Consider again the table. Find the following*

- $\Pr(1|H)$, the probability of a 1 given a head.
- $\Pr(H|1)$, the probability of a head given a 1.
- $Pr(E|T)$, the probability of an even given a tail.

				Die			
		1/6	1/6	1/6	1/6	1/6	1/6
	Coin	**1**	**2**	**3**	**4**	**5**	**6**
1/2	**H**	1/6	0	1/6	0	1/6	0
1/2	**T**	0	1/6	0	1/6	0	1/6

Notice that, to compute conditional probability, you went to a group of cells characterized by a row or a column, and "restandardized" the joint probabilities in that row or column so that they added up to 1. Perhaps without realizing it, you performed this standardization by dividing each entry in a row or column by the sum of the entries in that row or column. For example, to compute $\Pr(1|H)$, you went into the row labeled "H" and noted that in that reduced sample space, only 3 values ever occur for the die, and they occur equally often. So the conditional probability of a 1 given a head must be $1/3$. In other words, you divided $\Pr(1 \cap H)$ by $\Pr(H)$, obtaining $(1/6)/(1/2) = 1/3$. This type of operation leads to the following formula for conditional probability.

Definition 10 (Conditional Probability Formula)

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

5 Independence

Probabilistic independence is a key idea in probability theory. Notice that, in the example of the preceding table, knowing the outcome on the coin provides information about the outcome on the die, and vice versa, in the sense that the conditional probabilities were different from the marginal probabilities. For example, before knowing the outcome on the coin, the probability of an odd numer on the die is 1/2. On the other hand, once you know that the coin is a head, the probability of an odd becomes zero. Two marginal events are independent if the status of one does not provide information about the other in the sense of changing the probability distribution of the other.

This leads to the following definition

Definition 11 (Probabilistic Independence) *Two events are independent if*

$$\Pr(A|B) = \Pr(A)$$

or, alternatively, if

$$\Pr(A \cap B) = \Pr(A)\Pr(B)$$

6 Sequences

Many interesting problems in probability involve sequences of events. For example, if you are playing 5 card stud poker, you might be interested in the probability of the sequence A_1A_2, i.e., drawing an ace on the first card and an ace on the second card. In general, sequences may be viewed as intersections of events in time, so that the probability we are interested in is actually $\Pr(A_1 \cap A_2)$. If we look carefully at the definition of conditional probability, we can see a general rule for the probability of a sequence. We begin by deriving a general rule for the probability of the intersection of two events.

Theorem 12 (Probability of an Intersection) *The probability of the intersection of two events A and B is given by*

$$\Pr(A \cap B) = \Pr(A)\Pr(B|A)$$

Proof. *Consider the definition of conditional probability, but reverse the roles of A and B. One obtains*

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

and the result immediately follows. ■

The preceding theorem leads immediately to a rule for calculating the probability of a sequence.

Corollary 13 (Probability of a Sequence) *The probability of the sequence of events $A_1A_2A_3\cdots A_N$ is the product of the probabilities of events at each point in the sequence conditional on everything that happened previously, i.e.,*

$$\begin{aligned}\Pr(A_1A_2A_3\cdots A_N) &= \Pr(A_1)\Pr(A_2|A_1)\\ &\quad \times \Pr(A_3|A_1A_2)\\ &\quad \times \Pr(A_4|A_1A_2A_3)\\ &\quad \times \cdots \Pr(A_N|A_1A_2\cdots A_{N-1})\end{aligned}$$

The result is simpler when the events are independent, because conditionalizing has no effect when events are independent.

Corollary 14 (Product Rule) *Consider a sequence of* ***independent*** *events $A_1A_2A_3\cdots A_N$. The probability of the sequence is given by*

$$\Pr(A_1A_2\cdots A_N) = \Pr(A_1)\Pr(A_2)\cdots\Pr(A_N)$$

Example 15 (Drawing Two Aces) *We wish to compute the probability of drawing two consecutive aces of the top of a perfectly shuffled deck in stud poker. In poker there are 52 cards, and 4 of them are aces. Poker involves* ***sampling without replacement****, because what happens on one draw affects the probabilities for subsequent draws. Suppose cards are dealt completely at random. What is the probability of drawing an ace on the first card?*

Solution 16 *We need* $\Pr(A_1 A_2)$*, which can be calculated as* $\Pr(A_1)\Pr(A_2|A_1)$*. Since there are* 52 *cards in the sample space, and* 4 *of them are aces, and elementary events are assumed to be equally likely,* $\Pr(A_1)$ *is* $4/52 = 1/13$*. What is* $\Pr(A_2|A_1)$*? Once the first card has been drawn, there are* 51 *cards in the deck and* 3 *are aces. So* $\Pr(A_2|A_1) = 3/51 = 1/17$*. Conseqently,* $\Pr(A_1 A_2) = (1/13)(1/17) = 1/221$*.*

Example 17 (Throwing Two Sixes) *If you throw 2 independent fair dice, what is the probability of throwing two sixes?*

Solution 18 *If we can assume the dice are fair, the probability of a six on either die is* $1/6$*. If we can assume independence, the product rule holds, and the probability of two sixes is the product of the individual probabilities, i.e,* $1/36$*.*