Chapter 10

The Probability Calculus

10.1 THE PROBABILITY OPERATOR

Chapter 3 illustrates how truth tables are used to calculate the truth values of complex statements from the truth values of their atomic components and thereby to determine the validity or invalidity of argument forms. It would be useful to have something analogous for inductive reasoning: a procedure that would enable us to calculate probabilities of complex statements from the probabilities of simpler ones and thereby to determine the inductive probabilities of arguments. Unfortunately, no such procedure exists. We cannot always (or even usually) calculate the probability of a statement or the inductive probability of an argument simply from the probabilities of its atomic components.

Yet significant generalizations about probabilistic relationships among statements can be made. Although these do not add up to a general method for calculating inductive probabilities, they shed a great deal of light on the nature of probability and enable us to solve some practical problems. The most important of these generalizations constitute a logical system known as the *probability calculus*.

The probability calculus is a set of formal rules governing expressions of the form '$P(A)$', meaning "the probability of A." These expressions denote numbers. We may write, for example, '$P(A) = \frac{1}{2}$' to indicate that the probability of A is $\frac{1}{2}$. The expression 'A' may stand for a variety of different kinds of objects—sets, events, beliefs, propositions, and so on—depending upon the specific application. For applications of the probability calculus to logic, 'A' generally denotes a proposition, and sometimes an event. These two sorts of entities are formally similar, except that whereas propositions are said to be true or false, events are said to occur or not occur. In particular, events, like propositions, may be combined or modified by the truth-functional operators of propositional logic. Thus, if 'A' denotes an event, '$P(\sim A)$' denotes the probability of its non-occurrence, '$P(A \vee B)$' the probability of the occurrence of either A or B, '$P(A \,\&\, B)$' the probability of the joint occurrence of A and B, and so on.

The operator 'P' can be interpreted in a variety of ways. Under the *subjective interpretation*, '$P(A)$' stands for the degree of belief a particular rational person has in proposition A at a given time. Degree of belief is gauged behaviorally by the person's willingness to accept certain bets on the truth of A.

Under the various *logical interpretations*, '$P(A)$' designates the *logical* or *a priori* probability of A. There are many notions of logical probability, but according to all of them $P(A)$ varies inversely with the information content of A. That is, if A is a weak proposition whose information content is small, $P(A)$ tends to be high, and if A is a strong proposition whose information content is great, then $P(A)$ tends to be low. (Compare Section 9.1.)

Under the *relative frequency interpretation*, A is usually taken to be an event and $P(A)$ is the frequency of occurrence of A relative to some specified reference class of events. This is the interpretation of probability most often used in mathematics and statistics.

The oldest and simplest concept of probability is the *classical interpretation*. Like the relative frequency interpretation, the classical interpretation usually takes the object A to be an event. According to the classical interpretation, probabilities can be defined only when a situation has a finite nonzero number of equally likely possible outcomes, as for example in the toss of a fair die. Here the number of equally likely outcomes is 6, one for each face of the die. The probability of A is defined as the ratio of the number of possible outcomes in which A occurs to the total number of possible outcomes:

$$P(A) = \frac{\text{Number of possible outcomes in which } A \text{ occurs}}{\text{Total number of possible outcomes}}$$

SOLVED PROBLEM

10.1 Consider a situation in which a fair die is tossed once. There are six equally likely possible outcomes: the die will show a one, or a two, or a three, ..., or a six. Let 'A_1' denote the proposition that the die will show a one, 'A_2' denote the proposition that it will show a two, and so on. Calculate the following probabilities according to the classical interpretation:

(*a*) $P(A_1)$
(*b*) $P(A_5)$
(*c*) $P(\sim A_1)$
(*d*) $P(A_1 \vee A_3)$
(*e*) $P(A_1 \,\&\, A_3)$
(*f*) $P(A_1 \,\&\, \sim A_1)$
(*g*) $P(A_1 \vee \sim A_1)$
(*h*) $P(A_1 \vee A_2 \vee A_3 \vee A_4 \vee A_5 \vee A_6)$ [1]

Solution

(*a*) 1/6
(*b*) 1/6
(*c*) 5/6
(*d*) 2/6 = 1/3
(*e*) 0/6 = 0 (The die is tossed only once, so it cannot show two numbers.)
(*f*) 0/6 = 0
(*g*) 6/6 = 1
(*h*) 6/6 = 1

As this example illustrates, logical properties or relations among events or propositions affect the computation of the probability of complex outcomes. These properties or relations can sometimes be verified with the help of a truth table. For instance, a *contradictory* (truth-functionally inconsistent) event such as $A_1 \,\&\, \sim A_1$ can never occur, since every row of the corresponding truth table contains an F; so the probability of this event is zero (case (*f*)). In a similar way, we can sometimes see that some events are *mutually exclusive*, i.e., cannot jointly occur.

SOLVED PROBLEM

10.2 Show that any two events of the forms $A \,\&\, B$ and $A \,\&\, \sim B$ are mutually exclusive:

Solution

A	B	$A \,\&\, B$	$A \,\&\, \sim B$
T	T	T	F
T	F	F	T
F	T	F	F
F	F	F	F

[1]Because of the associative law of propositional logic (i.e., the equivalence ASSOC), bracket placement is not crucial when three or more propositions are joined by disjunction alone or by conjunction alone. It is therefore customary to omit brackets in these cases.

Since there is no line of the table on which both $A \& B$ and $A \& \sim B$ receive the value T and the table exhibits all possible situations, it is clear that there is no situation in which the two events can jointly occur.

In some cases, however, truth-functional considerations do not suffice, as an event may be impossible or two events may be mutually exclusive for non-truth-functional reasons. With reference to Problem 10.1(*e*), for instance, the truth table for '$A_1 \& A_3$' will include a line in which both A_1 and A_3 are T. Yet this line represents an impossible case: not for truth-functional reasons, but because of the nature of the die.

Two other logical relations which are important for the probability calculus are *truth-functional consequence* and *truth-functional equivalence*. The former is just the kind of validity detectable by truth tables. In other words, proposition or event A is a truth-functional consequence of proposition or event B if and only if there is no line on their common truth table in which A is false and B is true. Truth-functional equivalence holds when A and B are truth-functional consequences of each other, i.e., when their truth tables are identical. There are many examples of both kinds or relations in Chapter 3 and (in view of the completeness of the propositional calculus) in Chapter 4.

10.2 AXIOMS AND THEOREMS OF THE PROBABILITY CALCULUS

The probability calculus consists of the following three axioms (basic principles), together with their deductive consequences. These axioms are called the Kolmogorov axioms, after their inventor, the twentieth-century Russian mathematician, A. N. Kolmogorov:

AX1 $P(A) \geqslant 0$.
AX2 If A is tautologous, $P(A) = 1$.
AX3 If A and B are mutually exclusive, $P(A \vee B) = P(A) + P(B)$.

Here 'A' and 'B' may be simple or truth-functionally complex (atomic or molecular). We shall use the classical interpretation to illustrate and explain these axioms, though it should be kept in mind that they hold for the other interpretations as well.

AX1 sets the lower bound for probability values at zero. Zero, in other words, is the probability of the least probable things, those which are impossible. AX1 is true under the classical interpretation, since neither the numerator nor the denominator of the ratio which defines classical probability can ever be negative. (Moreover, the denominator is never zero, though the numerator may be.)

It is clear that tautologies ought to have the highest possible probability, since they are certainly true. Thus AX2 says in effect that we take 1 to be the maximum probability. This, too, is in accord with the classical definition, since the numerator of this definition can never exceed its denominator. (See Problem 10.1(*g*).)

AX3 gives the probability of a disjunction as the sum of the probabilities of its disjuncts, provided that these disjuncts are mutually exclusive. We can see that AX3 follows from the classical definition of probability by noting that a disjunctive event occurs just in case one or both of its disjuncts occurs. If the disjuncts are mutully exclusive, then in no possible outcome will they both occur, and so the number of possible outcomes in which the disjunction $A \vee B$ occurs is simply the sum of the number of possible outcomes in which A occurs and the number of possible outcomes in which B occurs. (See Problem 10.1(*d*).) That is, if A and B are mutually exclusive,

$$P(A \vee B) = \frac{\text{Number of possible outcomes in which } A \vee B \text{ occurs}}{\text{Total number of possible outcomes}}$$

$$= \frac{\text{Number of possible outcomes in which } A \text{ occurs}}{\text{Total number of possible outcomes}}$$

10.12 Prove:

If A is a truth-functional consequence of B, then $P(A \,\&\, B) = P(B)$.

Solution

Suppose A is a truth-functional consequence of B. Then $A \,\&\, B$ is truth-functionally equivalent to B, since $A \,\&\, B$ is true on any line of a truth table in which B is true and false on any line of a truth table on which B is false. So by Problem 10.6, $P(A \,\&\, B) = P(B)$.

10.3 CONDITIONAL PROBABILITY

Conditional probability is the probability of one proposition (or event), given that another is true (or has occurred). This is expressed by the notation '$P(A \mid B)$', which means "the probability of A, given B." It is not to be confused with the probability of a conditional statement, $P(B \rightarrow A)$, which plays little role in probability theory. (The difference between $P(A \mid B)$ and $P(B \rightarrow A)$ is illuminated by Problem 10.30 at the end of this section.) Notice that in contrast to the notation for conditional statements, the notation for conditional probabilities lists the consequent first and the antecedent second. This unfortunate convention is fairly standard, though occasionally other conventions are employed.

The notation for conditional probability is introduced into the probability calculus by the following definition:

D1 $$P(A \mid B) =_{\text{df}} \frac{P(A \,\&\, B)}{P(B)}$$

The symbol '$=_{\text{df}}$' means "is by definition." Anywhere the expression '$P(A \mid B)$' occurs, it is to be regarded as mere shorthand for the more complex expression on the right side of D1.

The purport of this definition is easy to understand under the classical interpretation. By the classical definition of probability, we have:

$$\frac{P(A \,\&\, B)}{P(B)} = \frac{\dfrac{\text{Number of possible outcomes in which } A \,\&\, B \text{ occurs}}{\text{Total number of possible outcomes}}}{\dfrac{\text{Number of possible outcomes in which } B \text{ occurs}}{\text{Total number of possible outcomes}}}$$

$$= \frac{\text{Number of possible outcomes in which } A \,\&\, B \text{ occurs}}{\text{Number of possible outcomes in which } B \text{ occurs}}$$

Thus by the classical interpretation, $P(A \mid B)$ is the proportion of possible outcomes in which A occurs among the possible outcomes in which B occurs.

SOLVED PROBLEM

10.13 A single die is tossed once. As in Example 10.6, let 'E' state that the result of the toss is an even number and 'L' state that it is a number less than 5. What is $P(E \mid L)$?

Solution

$E \,\&\, L$ is true in two of the six possible outcomes, and L is true in four of the six. Hence

$$P(E \mid L) = \frac{P(E \,\&\, L)}{P(L)} = \frac{\frac{2}{6}}{\frac{4}{6}} = \frac{2}{4} = \frac{1}{2}$$

Conditional probabilities are central to logic. Inductive probability (Section 2.3 and Chapter 9) is the probability of an argument's conclusion, given the conjunction of its premises. It is therefore a kind of conditional probability. This sort of probability, however, is best understood not by the classical interpretation, but by the subjective interpretation or the logical interpretations, whose technical details are beyond the scope of this discussion.

Note that in the case where $P(B) = 0$, $P(A \mid B)$ has no value, since division by zero is undefined. In general, theorems employing the notation '$P(A \mid B)$' hold only when $P(B) > 0$, but it is cumbersome to repeat this qualification each time a theorem is stated. Therefore we will not repeat it, but it is to be understood implicitly.

SOLVED PROBLEMS

10.14 Prove:

$$P(A \mid A) = 1$$

Solution

$A \,\&\, A$ is truth-functionally equivalent to A. Thus, by Problem 10.6,

$$P(A \mid A) = \frac{P(A \,\&\, A)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

10.15 Prove:

$$P({\sim}A \mid A) = 0$$

Solution

${\sim}A \,\&\, A$ is contradictory, and so by Problem 10.4, $P({\sim}A \,\&\, A) = 0$. Thus, by D1,

$$P({\sim}A \mid A) = \frac{P({\sim}A \,\&\, A)}{P(A)} = \frac{0}{P(A)} = 0$$

10.16 Prove:

If B is tautologous, $P(A \mid B) = P(A)$.

Solution

Suppose B is tautologous. Then by AX2, $P(B) = 1$. Moreover, $A \,\&\, B$ is truth-functionally equivalent to A, so that by Problem 10.6, $P(A \,\&\, B) = P(A)$. Thus

$$P(A \mid B) = \frac{P(A \,\&\, B)}{P(B)} = \frac{P(A)}{1} = P(A)$$

In Problem 10.6 we established that truth-functionally equivalent formulas may validly be substituted for one another in a probability expression. The next two theorems establish this same result for conditional probability expressions.

SOLVED PROBLEMS

10.17 Prove:

If A and B are truth-functionally equivalent, then $P(A \mid C) = P(B \mid C)$.

Solution

Suppose A and B are truth-functionally equivalent. Then so are $A \,\&\, C$ and $B \,\&\, C$. Hence by Problem 10.6, $P(A \,\&\, C) = P(B \,\&\, C)$. But then, by D1,

$$P(A \mid C) = \frac{P(A \,\&\, C)}{P(C)} = \frac{P(B \,\&\, C)}{P(C)} = P(B \mid C)$$

10.18 Prove:

If A and B are truth-functionally equivalent, then $P(C \mid A) = P(C \mid B)$.

Solution

Suppose A and B are truth-functionally equivalent. Then, by Problem 10.6, $P(A) = P(B)$. Moreover, by the reasoning of Problem 10.17, $P(C \& A) = P(C \& B)$. Hence, by D1,

$$P(C \mid A) = \frac{P(C \& A)}{P(A)} = \frac{P(C \& B)}{P(B)} = P(C \mid B)$$

Our next result is of great importance, because it provides a way of calculating the probabilities of conjunctions.

SOLVED PROBLEMS

10.19 Prove:

$$P(A \& B) = P(A) \cdot P(B \mid A)$$

Solution

D1 gives $P(B \mid A) = P(B \& A)/P(A)$. Hence, since $B \& A$ is truth-functionally equivalent to $A \& B$, by Problem 10.6 we have $P(B \mid A) = P(A \& B)/P(A)$. Multiplying both sides of this equation by $P(A)$ yields the theorem.

10.20 Prove:

$$P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

Solution

By Problem 10.19, $P(A \& B) = P(A) \cdot P(B \mid A)$ and also $P(B \& A) = P(B) \cdot P(A \mid B)$. But $A \& B$ is truth-functionally equivalent to $B \& A$. Hence, by Problem 10.6, $P(A \& B) = P(B \& A)$. This proves the theorem, which shows that the order of conjuncts is of no logical importance.

Problems 10.19 and 10.20 give us two ways of expressing the probability of any conjunction. That is, for any conjunction $A \& B$, $P(A \& B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$.

It sometimes happens that $P(A \mid B) = P(A)$, that is, that $P(A)$ is unaffected by the occurrence or nonoccurrence of B. In such a case, we say that A is *independent* of B. For instance, if A is the event of getting a one on the first toss of a single die and B is the event of getting a one on the second toss, then A is independent of B, for $P(A) = P(A \mid B) = \frac{1}{6}$. (The tendency to believe that tosses of dice are not independent is one version of the gambler's fallacy—see Section 8.5.) On the other hand, if A is the event of your living ten more years and B is the event of your having rabies, then clearly $P(A) > P(A \mid B)$, so that A is not independent of B. (Note that it is difficult to make sense of these probabilities under the classical interpretation, since there is no set of equally likely outcomes to appeal to. These probabilities would therefore be best understood by one of the other interpretations mentioned earlier.)

Independence is a concept peculiar to the probability calculus. It cannot be characterized truth-functionally as, for example, tautologousness or truth-functional equivalence can. The next theorem tells us that A is independent of B if and only if B is independent of A; that is, independence is a *symmetrical* relation.

SOLVED PROBLEM

10.21 Prove:

$$P(A \mid B) = P(A) \text{ if and only if } P(B \mid A) = P(B).$$

Solution

By D1, $P(A \mid B) = P(A)$ if and only if $P(A) = P(A \,\&\, B)/P(B)$. Now multiplying both sides of this equation by $P(B)/P(A)$ gives $P(B) = P(A \,\&\, B)/P(A)$, which (since $A \,\&\, B$ is truth-functionally equivalent to $B \,\&\, A$) is true if and only if $P(B) = P(B \,\&\, A)/P(A)$; hence by D1, $P(B) = P(B \mid A)$.

Because of the symmetry of independence, instead of saying "A is independent of B" or "B is independent of A," we may simply say "A and B are independent," without regard to the order of the two terms. Independence is important chiefly because if A and B are independent, then the calculation of $P(A \,\&\, B)$ is very simple. This is shown by the following theorem.

SOLVED PROBLEM

10.22 Prove:

If A and B are independent, then $P(A \,\&\, B) = P(A) \cdot P(B)$.

Solution

This follows immediately from Problem 10.19 and the definition of independence.

The next result also plays a central role in probability theory. It was first proved by Thomas Bayes (1702–1761), one of the founders of probability theory. Bayes' theorem enables us to calculate conditional probabilities, given converse conditional probabilities together with some nonconditional probabilities (the so-called priors). It has a number of important practical applications. We shall state it first in a simplified version (Problem 10.23) and then in a more fully articulated form (Problem 10.25). Here is the simple version:

SOLVED PROBLEM

10.23 Prove:

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B)}$$

Solution

This follows immediately from Problem 10.20.

This version of the theorem allows us to calculate $P(A \mid B)$ if we know the converse conditional probability $P(B \mid A)$ and the prior probabilities $P(A)$ and $P(B)$.

To state the full version of Bayes' theorem, we need two additional concepts: the concept of an exhaustive series of propositions or events, and the concept of a pairwise mutually exclusive series. A series of propositions or events $A_1, A_2, \ldots, A_n$ is *exhaustive* if $P(A_1 \vee A_2 \vee \ldots \vee A_n) = 1$. For instance, the series $A_1, A_2, \ldots, A_6$, representing the six possible outcomes of a single toss of a single die, is exhaustive, under the classical interpretation. Also, any series of the form $A \,\&\, B$, $A \,\&\, {\sim}B$, ${\sim}A \,\&\, B$, ${\sim}A \,\&\, {\sim}B$ is exhaustive, as can be seen by observing that the disjunction of these four forms is a tautology.

The second crucial concept for the general version of Bayes' theorem is that of a *pairwise mutually exclusive* series. We have already discussed mutual exclusivity for pairs of propositions or events. Now in effect we simply extend this idea to a whole series of them. Thus a series of propositions or events $A_1, A_2, \ldots, A_n$ is pairwise mutually exclusive if for each pair A_i, A_j of its members, $P(A_i \& A_j) = 0$. For instance, the exhaustive series $A_1, A_2, \ldots, A_6$ mentioned above is pairwise mutually exclusive, since only one outcome may result from a single toss. So is the other exhaustive series mentioned above, but here the reason is truth-functional: on no line of a truth table are any two of these statements both true.

Now we observe an important fact about pairwise mutually exclusive and exhaustive series:

SOLVED PROBLEM

10.24 Prove:

If $A_1 A_2, \ldots, A_n$ is a pairwise mutually exclusive and exhaustive series, then $P(B) = P(A_1 \& B) + P(A_2 \& B) + \ldots + P(A_n \& B)$.

Solution

Let $A_1, A_2, \ldots, A_n$ be a pairwise mutually exclusive and exhaustive series. Then $P(A_1 \vee A_2 \vee \ldots \vee A_n) = 1$, and so, by Problem 10.11,

(*a*) $P((A_1 \vee A_2 \vee \ldots \vee A_n) \& B) = P(B)$.

Now as can be seen by repeated application of the distributive law of propositional logic, $(A_1 \vee A_2 \vee \ldots \vee A_n) \& B$ is truth-functionally equivalent to $(A_1 \& B) \vee (A_2 \& B) \vee \ldots \vee (A_n \& B)$. Thus, applying Problem 10.6 to item (*a*) we get:

(*b*) $P(B) = P((A_1 \& B) \vee (A_2 \& B) \vee \ldots \vee (A_n \& B))$.

Moreover,

(*c*) The series $(A_1 \& B), (A_2 \& B), \ldots, (A_n \& B)$ is pairwise mutually exclusive.

For consider any two members $(A_i \& B)$, $(A_j \& B)$ of this series. Since the series $A_1, A_2, \ldots, A_n$ is pairwise mutually exclusive, we know that $P(A_i \& A_j) = 0$. Hence, by Problem 10.8 and AX1, $P(A_i \& A_j \& B) = 0$. So, by Problem 10.6, $P((A_i \& B) \& (A_j \& B)) = 0$.

Now from item (*c*) it follows that for each i such that $1 \leq i < n$,

(*d*) $P(((A_1 \& B) \vee (A_2 \& B) \vee \ldots \vee (A_i \& B)) \& (A_{i+1} \& B)) = 0$.

For again, by repeated application of the distributive law, $((A_1 \& B) \vee (A_2 \& B) \vee \ldots \vee (A_i \& B)) \& (A_{i+1} \& B)$ is truth-functionally equivalent to $((A_1 \& B) \& (A_{i+1} \& B)) \vee ((A_2 \& B) \& (A_{i+1} \& B)) \vee \ldots \vee ((A_i \& B) \& (A_{i+1} \& B))$. But by item (*c*), the probability of each of the disjuncts of this latter formula is zero; and repeated application of Problem 10.10 implies that any disjunction whose disjuncts all have probability zero must itself have probability zero. Hence item (*d*) follows by Problem 10.6.

The theorem then follows from formula (*b*) by repeated application of formula (*d*) and Problem 10.7. To see this, suppose for the sake of concreteness that $n = 3$. Then formula (*b*) will read:

(*b'*) $P(B) = P((A_1 \& B) \vee (A_2 \& B) \vee (A_3 \& B))$

and we will have these two instances of formula (*d*):

(*d'*) $P(((A_1 \& B) \vee (A_2 \& B)) \& (A_3 \& B)) = 0$.

(*d''*) $P((A_1 \& B) \& (A_2 \& B)) = 0$.

From Problem 10.7 we obtain:

(*e*) $P((A_1 \& B) \vee (A_2 \& B) \vee (A_3 \& B))$
$= P((A_1 \& B) \vee (A_2 \& B)) + P(A_3 \& B) - P(((A_1 \& B) \vee (A_2 \& B)) \& (A_3 \& B))$

which reduces by (d') to:

(f) $P((A_1 \,\&\, B) \vee (A_2 \,\&\, B) \vee (A_3 \,\&\, B))$
$= P((A_1 \,\&\, B) \vee (A_3 \,\&\, B)) + P(A_3 \,\&\, B)$

Now, again, by Problem 10.7:

(g) $P((A_1 \,\&\, B) \vee (\mathrm{A}_2 \,\&\, B))$
$= P(A_1 \,\&\, B) + P(A_2 \,\&\, B) - P((A_1 \,\&\, B) \,\&\, (A_2 \,\&\, B))$

which reduces by (d'') to:

(h) $P((A_1 \,\&\, B) \vee (A_2 \,\&\, B)) = P(A_1 \,\&\, B) + P(A_2 \,\&\, B)$

Combining (f) and (h), we get:

(i) $P((A_1 \,\&\, B) \vee (A_2 \,\&\, B) \vee (A_3 \,\&\, B)) = P(A_1 \,\&\, B) + P(A_2 \,\&\, B) + P(A_3 \,\&\, B)$

which, by (b'), gives:

$$P(B) = P(A_1 \,\&\, B) + P(A_1 \,\&\, B) + P(A_3 \,\&\, B)$$

which proves the theorem for $n = 3$. The reasoning for other values of n is similar.

Using the result of Problem 10.24, we now establish the general version of Bayes' theorem:

SOLVED PROBLEM

10.25 Prove:

If $A_1, A_2, \ldots, A_n$ is a pairwise mutually exclusive and exhaustive series and A_i is some member of this series, then

$$P(A_i \mid B) = \frac{P(A_i) \cdot P(B \mid A_i)}{P(A_1) \cdot P(B \mid A_1) + P(A_2) \cdot P(B \mid A_2) + \ldots + P(A_n) \cdot P(B \mid A_n)}$$

Solution

Assume $A_1, A_2, \ldots, A_n$ is a pairwise mutually exclusive and exhaustive series. Now by Problem 10.23 we know that

(a) $P(A_i \mid B) = \dfrac{P(A_i) \cdot P(B \mid A_i)}{P(B)}$

and by Problem 10.24 we have:

(b) $P(B) = P(A_1 \,\&\, B) + P(A_2 \,\&\, B) + \ldots + P(A_n \,\&\, B)$

Applying Problem 10.19 to formula (b) we get

(c) $P(B) = P(A_1) \cdot P(B \mid A_1) + P(A_2) \cdot P(B \mid A_2) + \ldots + P(A_n) \cdot P(B \mid A_n)$

Then applying formula (c) to formula (a) proves the theorem.

The general version of Bayes' theorem is often expressed in the more compact summation notation:

$$P(A_i \mid B) = \frac{P(A_i) \cdot P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j) \cdot P(B \mid A_j)}$$

This is just a shorthand version of the equation of Problem 10.25.

One practical application of Bayes' theorem is in calculating the probability that a particular hypothesis accounts for a particular observation. This calculation requires us to formulate a series of hypotheses which is pairwise mutually exclusive and exhaustive—the series $A_1, A_2, \ldots, A_n$ mentioned in the theorem. The observation is B. Now if we wish to learn the probability that a particular member A_i of the series of hypotheses accounts for B—that is, $P(A_i \mid B)$—Bayes' theorem gives a way of finding the answer.

There is a catch, however. To make the calculation, we need to know not only the probability of the observation, given each of the hypotheses—that is, $P(B \mid A_j)$ for each j such that $1 \leq j \leq n$—but also the probability of each of the hypotheses, or $P(A_j)$ for each such j. These latter probabilities are called *prior probabilities*, or sometimes just *priors*.[2] In many cases they are difficult or impossible to determine. It can be shown, however, that as observations accumulate, the priors have less and less influence on the outcome of the calculation. This phenomenon, known as "swamping the priors," permits useful application of Bayes' theorem even in cases in which the priors are only very roughly known.

SOLVED PROBLEM

10.26 Three machines at a certain factory produce girdles of the same kind. Three-tenths of the girdles produced by machine 1 are defective, as are two-tenths of the girdles produced by machine 2 and one-tenth of the girdles produced by machine 3. Machine 1 produces four-tenths of the factory's output, machine 2 produces three-tenths, and machine 3 produces three-tenths. Alma has acquired an egregiously defective girdle from the factory. What is the probability that the girdle was produced by machine 1, given that it is defective?

Solution

Here the observation B is that Alma's girdle is defective. This can be accounted for by three pairwise mutually exclusive and exhaustive hypotheses:

A_1 = Alma's girdle was produced by machine 1.

A_2 = Alma's girdle was produced by machine 2.

A_3 = Alma's girdle was produced by machine 3.

We may take the probability of B given each of these hypotheses as just the respective proportion of defective girdles produced by each machine. Thus

$$P(B \mid A_1) = \tfrac{3}{10}$$
$$P(B \mid A_2) = \tfrac{2}{10}$$
$$P(B \mid A_3) = \tfrac{1}{10}$$

Moreover, we have definite values for the priors in this case. These are just the proportions of the total output of the factory produced by each machine:

$$P(A_1) = \tfrac{4}{10}$$
$$P(A_2) = \tfrac{3}{10}$$
$$P(A_3) = \tfrac{3}{10}$$

[2]The notion of prior probability should not be confused with that of an a priori or logical probability, which was mentioned earlier in this chapter. Prior probabilities are usually obtained from subjective estimates or (as they will be in Problem 10.26) from relative frequencies of occurrence. A priori probabilities are not based on actual or estimated frequencies of events, but on the information content of statements.

Now, by Bayes' theorem,

$$P(A_1 \mid B) = \frac{P(A_1) \cdot P(B \mid A_1)}{P(A_1) \cdot P(B \mid A_1) + P(A_2) \cdot P(B \mid A_2) + P(A_3) \cdot P(B \mid A_3)}$$
$$= \frac{\frac{4}{10} \cdot \frac{3}{10}}{(\frac{4}{10} \cdot \frac{3}{10}) + (\frac{3}{10} \cdot \frac{2}{10}) + (\frac{3}{10} \cdot \frac{1}{10})}$$
$$= \frac{12}{21}$$

This is the solution to the problem.

Our next three theorems (Problems 10.27, 10.28, and 10.29) establish the behavior of negations, disjunctions, and conjunctions as the first terms of conditional probabilities. They are conditional analogues of Problems 10.3, 10.7, and 10.19, respectively.

SOLVED PROBLEM

10.27 Prove:

$$P(\sim A \mid B) = 1 - P(A \mid B)$$

Solution

By D1,

$$P(A \vee \sim A \mid B) = \frac{P((A \vee \sim A) \,\&\, B)}{P(B)}$$

But $(A \vee \sim A) \,\&\, B$ is truth-functionally equivalent to B. Hence, by Problem 10.6, $P((A \vee \sim A) \,\&\, B) = P(B)$, so that $P(A \vee \sim A \mid B) = P(B)/P(B) = 1$. Moreover, $(A \vee \sim A) \,\&\, B$ is also truth-functionally equivalent to $(A \,\&\, B) \vee (\sim A \,\&\, B)$, so that by Problem 10.6,

$$P(A \vee \sim A \mid B) = \frac{P((A \,\&\, B) \vee (\sim A \,\&\, B))}{P(B)} = 1$$

Now $A \,\&\, B$ and $\sim A \,\&\, B$ are mutually exclusive, so that by AX3,

$$\frac{P(A \,\&\, B) + P(\sim A \,\&\, B)}{P(B)} = \frac{P(A \,\&\, B)}{P(B)} + \frac{P(\sim A \,\&\, B)}{P(B)} = 1$$

So by D1, $P(A \mid B) + P(\sim A \mid B) = 1$; that is, $P(\sim A \mid B) = 1 - P(A \mid B)$.

Since as we noted earlier, the inductive probability of an argument is a kind of conditional probability (the probability of the conclusion given the conjunction of the premises), Problem 10.27 implies the important result that the probability of $\sim A$ given a set of premises is 1 minus the probability of A given those premises. (And, likewise, the probability of A given a set of premises is 1 minus the probability of $\sim A$ given those premises.) This was our justification for remarking in Section 2.3 that any argument whose inductive probability is less than .5 is weak, since the probability of the negation of its conclusion given the same premises is therefore greater than .5.

SOLVED PROBLEMS

10.28 Prove:

$$P(A \vee B \mid C) = P(A \mid C) + P(B \mid C) - P(A \,\&\, B \mid C)$$

10.32 Calculate $P(\sim A_1)$.

Solution

By Problem 10.3, we know that $P(\sim A_1) = 1 - P(A_1)$. We saw in Problem 10.31 that $P(A_1) = \frac{1}{6}$. Hence $P(\sim A_1) = 1 - \frac{1}{6} = \frac{5}{6}$.

10.33 Calculate $P(A_1 \,\&\, A_7)$.

Solution

We can do this directly by the classical definition, or indirectly via the probability calculus. Using the classical definition, we note that in only one of the 36 possible outcomes does $A_1 \,\&\, A_7$ occur. Hence $P(A_1 \,\&\, A_7) = \frac{1}{36}$.

To use the probability calculus, we must recognize that A_1 and A_7 are independent. That is (assuming the dice are tossed fairly, so that for example we don't start with both showing the same face and give them just a little flick, an arrangement which would tend to make them show the same face after the toss), we expect that $P(A_1 \mid A_7) = P(A_1)$ and $P(A_7 \mid A_1) = P(A_7)$. If this assumption is correct, we can apply Problem 10.19, so that $P(A_1 \,\&\, A_7) = P(A_1) \cdot P(A_7) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$. This agrees with our previous calculation.

As Problem 10.33 illustrates, we can calculate probabilities in several ways (both directly, from the classical definition, and indirectly, via the probability calculus). Checking one method of calculation by means of another helps eliminate mistakes. Problem 10.34 provides another illustration of this.

SOLVED PROBLEM

10.34 Find $P(A_1 \vee A_7)$.

Solution

To solve this problem via the calculus, we appeal to Problem 10.7:

$$P(A_1 \vee A_7) = P(A_1) + P(A_7) - P(A_1 \,\&\, A_7)$$

In Problem 10.31 we determined that $P(A_1) = P(A_7) = \frac{1}{6}$, and in Problem 10.33 we found that $P(A_1 \,\&\, A_7) = \frac{1}{36}$. Hence $P(A_1 \vee A_7) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$.

But we could just as well have proceeded directly, via the classical definition. There are 11 possible outcomes in which $A_1 \vee A_7$ occurs (these are just the outcomes listed in the leftmost column, along with those listed in the first horizontal row, of our table of the 36 outcomes). Hence, once again we get $P(A_1 \vee A_7) = \frac{11}{36}$.

Our next set of problems involves the probabilities of dealing certain combinations of cards at random from a jokerless deck of 52 cards. (To say that the deal is random is to say that each card remaining in the deck is equally likely to be dealt. This provides the equally likely outcomes necessary for the classical interpretation.)

If only one card is dealt, there are 52 possible outcomes, one for each card in the deck. If two cards are dealt, the number of possible outcomes is $52 \cdot 51 = 2652$. (There are 52 possibilities for the first card and 51 for the second, since after the first card is dealt, 51 remain in the deck.) In general, the number of equally likely outcomes for a deal of n cards is $52 \cdot 51 \cdots (52 - (n - 1))$.

We adopt the following abbreviations:

J = Jack	K = King	H = Heart	D = Diamond
Q = Queen	A = Ace	S = Spade	C = Club

We use numerical subscripts to indicate the order of the cards dealt. Thus 'A_1' means "The first card dealt is an ace," 'D_3' means "The third card dealt is a diamond," and so on.

SOLVED PROBLEM

10.35 One card is dealt. What is the probability that it is either a queen or a heart?

Solution

We seek $P(Q_1 \vee H_1)$, which by Problem 10.7 is $P(Q_1) + P(H_1) - P(Q_1 \,\&\, H_1)$. Out of the 52 cards, 4 are queens, 12 are hearts, and 1 is the queen of hearts. Thus, by the classical definition of probability, we have $\frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$.

When two or more cards are dealt, we may wish to know the probability that one is of a certain type, given that certain other cards have been previously dealt. Such conditional probabilities can be computed by D1, but they are most efficiently obtained directly from consideration of the composition of the deck.

SOLVED PROBLEM

10.36 Two cards are dealt. What is the probability that the second is an ace, given that the first is?

Solution

We wish to find $P(A_2 \mid A_1)$. Given A_1, 51 cards remain in the deck, of which three are aces. Hence $P(A_2 \mid A_1) = \frac{3}{51} = \frac{1}{17}$. The same figure can be obtained with a little more work by appeal to D1. By D1 we have $P(A_2 \mid A_1) = P(A_2 \,\&\, A_1)/P(A_1)$. Now clearly $P(A_1) = \frac{4}{52} = \frac{1}{13}$. But to find $P(A_2 \,\&\, A_1)$, we must consider all the 2652 outcomes possible in a deal of two cards. Of these, only 12 give $A_2 \,\&\, A_1$. We list them below. ('*AH*' means "ace of hearts," '*AS*' means "ace of spades," and so on.)

First card	*Second card*	*First card*	*Second card*
AH	*AS*	*AD*	*AH*
AH	*AD*	*AD*	*AS*
AH	*AC*	*AD*	*AC*
AS	*AH*	*AC*	*AH*
AS	*AD*	*AC*	*AS*
AS	*AC*	*AC*	*AD*

Thus

$$\frac{P(A_1 \,\&\, A_2)}{P(A_1)} = \frac{\frac{12}{2652}}{\frac{1}{13}} = \frac{1}{17}$$

Clearly, however, the first method of calculation is simpler.

Using conditional probabilities obtained by the first method of calculation explained in Problem 10.36, we can easily obtain a variety of other probabilities by applying the theorems of the probability calculus.

SOLVED PROBLEMS

10.37 Two cards are dealt. What is the probability that they are both aces?

Solution

The desired probability is $P(A_1 \,\&\, A_2)$, which, by Problem 10.19, is $P(A_1) \cdot P(A_2 \mid A_1)$. As we saw in Problem 10.36, $P(A_1) = \frac{1}{13}$; and as we also saw in that problem, $P(A_2 \mid A_1) = \frac{1}{17}$. Hence

$P(A_1 \,\&\, A_2) = \frac{1}{13} \cdot \frac{1}{17} = \frac{1}{221}$. This is the same figure obtained above for $P(A_2 \,\&\, A_1)$, since $\frac{12}{2652} = \frac{1}{221}$.

10.38 Five cards are dealt. What is the probability of a flush, i.e., five cards of a single suit?

Solution

We wish to find $P((H_1 \,\&\, H_2 \,\&\, H_3 \,\&\, H_4 \,\&\, H_5) \vee (S_1 \,\&\, S_2 \,\&\, S_3 \,\&\, S_4 \,\&\, S_5) \vee (D_1 \,\&\, D_2 \,\&\, D_3 \,\&\, D_4 \,\&\, D_5) \vee (C_1 \,\&\, C_2 \,\&\, C_3 \,\&\, C_4 \,\&\, C_5))$. Clearly these four disjuncts are mutually exclusive, and each is mutually exclusive from any disjunction of the others. Thus by Problem 10.7 the desired probability is simply the sum of the probabilities of these four disjuncts. We shall calculate $P(H_1 \,\&\, H_2 \,\&\, H_3 \,\&\, H_4 \,\&\, H_5)$. Clearly, the probabilities of the other disjuncts will be the same. Hence, to obtain our answer we simply multiply the result of our calculation by 4.

Now by four applications of Problem 10.19, $P(H_1 \,\&\, H_2 \,\&\, H_3 \,\&\, H_4 \,\&\, H_5) = P(H_1) \cdot P(H_2 \mid H_1) \cdot P(H_3 \mid H_1 \,\&\, H_2) \cdot P(H_4 \mid H_1 \,\&\, H_2 \,\&\, H_3) \cdot P(H_5 \mid H_1 \,\&\, H_2 \,\&\, H_3 \,\&\, H_4)$. Using the first method of Problem 10.36, we can see that this is $\frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \cdot \frac{10}{49} \cdot \frac{9}{48} = 0.0004952$. Multiplying by 4 gives the probability of a flush: 0.0019808.

10.39 In a game of blackjack, you have been dealt a king and an eight from a deck of 52 cards. If you ask for a third card, what is the probability that you will receive an ace, a two, or a three and thus not go over 21?

Solution

Let 'E' stand for an eight, 'W' for a two, and 'R' for a three. The probability we seek is either $P(A_3 \vee W_3 \vee R_3 \mid E_1 \,\&\, K_2)$ or $P(A_3 \vee W_3 \vee R_3 \mid K_1 \,\&\, E_2)$. Either of these will do; they are equal, and the order in which the king and the eight are dealt is not relevant to our problem. Hence we will calculate only the first.

We note that A_3 and $W_3 \vee R_3$ are mutually exclusive, as are W_3 and R_3. Hence, by two applications of Supplementary Problem I(12), below, $P(A_3 \vee W_3 \vee R_3 \mid K_1 \,\&\, E_2) = P(A_3 \mid K_1 \,\&\, E_2) + P(W_3 \mid K_1 \,\&\, E_2) + P(R_3 \mid K_1 \,\&\, E_2)$. Given $K_1 \,\&\, E_2$, the deck contains 50 cards, four of which are aces, four of which are twos, and four of which are threes. Hence the desired probability is $\frac{4}{50} + \frac{4}{50} + \frac{4}{50} = \frac{12}{50} = \frac{6}{25}$.

Supplementary Problems

I Prove the following theorems, using the axioms and theorems established above. (Keep in mind that all theorems containing expressions of the form '$P(A \mid B)$' hold only when $P(B) > 0$.)

(1) If $P(A) = 0$, then $P(A \vee B) = P(B)$.

(2) $P(A \vee B) = P(A) + P({\sim}A \,\&\, B)$.

(3) If A and B are mutually exclusive, then $P(A \,\&\, B) = 0$.

(4) If ${\sim}A$ and ${\sim}B$ are mutually exclusive, then $P(A \,\&\, B) = P(A) + P(B) - 1$.

(5) If $P(A) = P(B) = 1$, then $P(A \,\&\, B) = 1$.

(6) If A is a truth-functional consequence of B, then $P(A \vee B) = P(A)$.

(7) If A is a truth-functional consequence of B, then $P(A) \leqslant P(B)$.

(8) $0 \leqslant P(A \mid B) \leqslant 1$.

(9) $P(A \mid B) = 0$ if and only if $P(B \mid A) = 0$.

(10) If A is tautologous, $P(A \mid B) = 1$.

(11) If A is contradictory, $P(A \mid B) = 0$.

(12) If A and B are mutually exclusive, $P(A \vee B \mid C) = P(A \mid C) + P(B \mid C)$.
(13) If A is a truth-functional consequence of B, then $P(A \mid B) = 1$.
(14) If B is a truth-functional consequence of C, then $P(A \,\&\, B \mid C) = P(A \mid C)$.
(15) If A is a truth-functional consequence of $B \,\&\, C$, then $P(A \,\&\, B \mid C) = P(B \mid C)$.
(16) $P(A \mid B) = P(A \,\&\, C \mid B) + P(A \,\&\, {\sim}C \mid B)$.
(17) If $P(A) > 0$ and $P(B) > 0$ and A and B are mutually exclusive, then A and B are not independent.
(18) $P(A \mid {\sim}B) = \dfrac{P(A) - P(A \,\&\, B)}{1 - P(B)}$
(19) $P(A \mid B \vee C) = \dfrac{P(A \,\&\, B) + P(A \,\&\, C) + P(A \,\&\, B \,\&\, C)}{P(B) + P(C) - P(B \,\&\, C)}$
(20) $P(A \mid B \,\&\, C) = \dfrac{P(A \,\&\, B \mid C)}{P(B \mid C)}$

II Calculate and compare the values for the following pairs of expressions:

(1) $P(A \mid A)$, $P(A \rightarrow A)$.
(2) $P({\sim}A \mid A)$, $P(A \rightarrow {\sim}A)$.
(3) $P(A \mid B \,\&\, {\sim}B)$, $P((B \,\&\, {\sim}B) \rightarrow A)$.
(4) $P(A_1 \mid A_2)$, $P(A_2 \rightarrow A_1)$, where A_1 is the event of rolling a one and A_2 is the event of rolling a two on a single toss of a die (use the classical definition of probability).

III (1) Using the data given in Problem 10.26, calculate the probability that Alma's girdle was produced by machine 2, given that it is defective, and the probability that it was produced by machine 3, given that it is defective.

(2) Consider the following game. There are two boxes containing money: box 1 contains three \$1 bills and three fifties; box 2 contains thirty \$1 bills and one fifty. A blindfolded player is allowed to make a random draw of one bill from one box, but the choice of box is determined by the toss of a fair coin; if the result is heads, the person must draw from box 1; if it is tails, the draw is made from box 2. The coin is tossed and a player makes a draw. Use Bayes' theorem to calculate the probability that the draw was made from box 1, given that the player draws a \$50 bill.

(3) Let Q be the hypothesis that some version of the quark theory of subatomic physics is true. Let N be the hypothesis that no proton decay is observed over a period of a year in a certain quantity of proton-rich liquid. Suppose that $P(N \mid Q) = .001$ and $P(N \mid {\sim}Q) = .99$. Suppose further that we estimate the prior probabilities of Q and ${\sim}Q$ as $P(Q) = .7$, $P({\sim}Q) = .3$. Using Bayes' theorem, calculate $P(Q \mid N)$. Suppose $P(Q) = .3$, $P({\sim}Q) = .7$. How strongly does this change in the priors affect our confidence in Q, given N?

IV Two fair dice are tossed once. Using the probability calculus and/or the classical definition of probability, calculate the following. (The abbreviations used are those given just before Problem 10.31.)

(1) $P(A_1 \mid A_7)$
(2) $P(A_7 \mid A_1)$
(3) $P(A_1 \mid A_2)$
(4) $P(A_1 \,\&\, A_2)$
(5) $P(A_1 \mid A_1 \vee A_2 \vee A_3)$
(6) $P(A_1 \vee A_2 \mid A_7 \vee A_8)$

(7) $P(A_3 \vee \sim A_3)$
(8) $P(A_1 \& \sim A_1)$
(9) $P(\sim A_1 \& \sim A_2)$
(10) $P(A_2 \mid A_1 \& \sim A_1)$

V Five cards are dealt at random from a deck of 52. Calculate the probabilities of the following:

(1) The first four cards dealt are aces.
(2) The fifth card dealt is an ace, given that none of the first four was.
(3) The fifth card dealt is a heart, given that the first four were also hearts.
(4) The fifth card dealt is not a heart, given that the first four were hearts.
(5) At least one of the first two cards dealt is an ace.
(6) The ten, jack, queen, king, and ace of hearts are dealt, in that order.
(7) Neither of the first two cards dealt is a king.
(8) The first card dealt is an ace, and the second one is not an ace.
(9) None of the five cards is a club.
(10) The fifth card deal is an ace or a king, given that four have been dealt, two of which are aces and two of which are kings.

VI A fair coin is tossed three times. The outcomes of the tosses are independent. Find the probabilities of the following:

(1) All three tosses are heads.
(2) None of the tosses is heads.
(3) The first toss is heads and the second and third are tails.
(4) One of the tosses is heads; the other two are tails.
(5) At least one of the tosses is heads.
(6) The first toss is heads, given that all three tosses are heads.
(7) The second toss is heads, given that the first is heads.
(8) The first toss is heads, given that the second is heads.
(9) The first toss is heads, given that exactly one of the three is heads.
(10) Exactly two heads occur, given that the first toss is heads.

Answers to Selected Supplementary Problems

I (5) Suppose $P(A) = P(B) = 1$. Then, by Problem 10.11, $P(A \& B) = P(B) = 1$.

(10) Suppose A is tautologous. Then, by AX2, $P(A) = 1$; and by Problem 10.11, $P(A \& B) = P(B)$. But by D1,

$$P(A \mid B) = \frac{P(A \& B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

(15) Suppose A is a truth-functional consequence of $B \& C$. Then clearly $(A \& B) \& C$ is a truth-functional consequence of $B \& C$. Moreover, $B \& C$ is a truth-functional consequence of

$(A \,\&\, B) \,\&\, C$, so that $(A \,\&\, B) \,\&\, C$ and $B \,\&\, C$ are truth-functionally equivalent. Thus, by Problem 10.6, $P((A \,\&\, B) \,\&\, C) = P(B \,\&\, C)$. Now, by D1,

$$P(A \,\&\, B \mid C) = \frac{P((A \,\&\, B) \,\&\, C)}{P(C)} = \frac{P(B \,\&\, C)}{P(C)}$$

which, again by D1, is just $P(B \mid C)$.

(20) By D1 and Problem 10.6,

$$P(A \mid B \,\&\, C) = \frac{P(A \,\&\, (B \,\&\, C))}{P(B \,\&\, C)} = \frac{P((A \,\&\, B) \,\&\, C)/P(C)}{P(B \,\&\, C)/P(C)} = \frac{P(A \,\&\, B \mid C)}{P(B \mid C)}$$

III (3) Q and $\sim Q$ are mutually exclusive and exhaustive. Therefore, by Bayes' theorem,

$$P(Q \mid N) = \frac{P(Q) \cdot P(N \mid Q)}{P(Q) \cdot P(N \mid Q) + P(\sim Q) \cdot P(N \mid \sim Q)}$$

$$= \frac{.7 \cdot .001}{(.7 \cdot .001) + (.3 \cdot .99)} = .00235$$

For $P(Q) = .3$, $P(\sim Q) = .7$, the answer is .00043. The change in the prior probabilities is drastic, but the change in $P(Q \mid N)$ is rather small; Q is still very unlikely, given the observations N. Within a wide range of values for the priors, Q remains very unlikely, given N.

IV (5) Note first that A_1 is truth-functionally equivalent to $A_1 \,\&\, (A_1 \vee A_2 \vee A_3)$. So, by D1 and Problem 10.6,

$$P(A_1 \mid A_1 \vee A_2 \vee A_3) = \frac{P(A_1 \,\&\, (A_1 \vee A_2 \vee A_3))}{P(A_1 \vee A_2 \vee A_3)} = \frac{P(A_1)}{P(A_1 \vee A_2 \vee A_3)}$$

Since $A_1 \vee A_2$ and A_3 are mutually exclusive, as are A_1 and A_2, we have by two applications of AX3:

$$P(A_1 \vee A_2 \vee A_3) = P(A_1) + P(A_2) + P(A_3)$$

But the probabilities of A_1, A_2, and A_3 are each $\frac{1}{6}$. Thus the desired value is:

$$\frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{6} + \frac{1}{6}} = \frac{1}{3}$$

(10) This probability is undefined, since $P(A_1 \,\&\, \sim A_1) = 0$.

V (5) Using 'A_1' for 'The first card is an ace' and 'A_2' for 'The second card is an ace', the probability we seek is $P(A_1 \vee A_2)$, which is, by Problem 10.6,

$$P((A_1 \,\&\, A_2) \vee (A_1 \,\&\, \sim A_2) \vee (\sim A_1 \,\&\, A_2))$$

Since the disjuncts here are mutually exclusive, two applications of AX3, give simply

$$P(A_1 \,\&\, A_2) + P(A_1 \,\&\, \sim A_2) + P(\sim A_1 \,\&\, A_2)$$

which by Problem 10.19 becomes:

$$P(A_1) \cdot P(A_2 \mid A_1) + P(A_1) \cdot P(\sim A_2 \mid A_1) + P(\sim A_1) \cdot P(A_2 \mid \sim A_1)$$

i.e.,

$$\tfrac{4}{52} \cdot \tfrac{3}{51} + \tfrac{4}{52} \cdot \tfrac{48}{51} + \tfrac{48}{52} \cdot \tfrac{4}{51} = \tfrac{12}{2652} + \tfrac{192}{2652} + \tfrac{192}{2652} = \tfrac{396}{2652} = \tfrac{33}{221}$$

(10) If two aces and two kings have been dealt, there are two aces and two kings among the 48 remaining cards. That is, there are four chances out of 48 to draw an ace or a king. The probability of drawing an ace or a king, given that two aces and two kings have been dealt, is thus $\frac{4}{48} = \frac{1}{12}$.

VI (5) Let the probability of getting a head on the nth toss be H_n. We wish to determine $P(H_1 \vee H_2 \vee H_3)$. By Problem 10.3, this is equal to $1 - P(\sim(H_1 \vee H_2 \vee H_3))$, which, since $\sim(H_1 \vee H_2 \vee H_3)$ is truth-functionally equivalent to $\sim H_1 \mathbin{\&} \sim H_2 \mathbin{\&} \sim H_3$, is just:

$$1 - P(\sim H_1 \mathbin{\&} \sim H_2 \mathbin{\&} \sim H_3)$$

Because the tosses are independent, two applications of Problem 10.22 to $P(\sim H_1 \mathbin{\&} \sim H_2 \mathbin{\&} \sim H_3)$ give:

$$P(\sim H_1) \cdot P(\sim H_2) \cdot P(\sim H_3) = \tfrac{1}{2} \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{8}$$

Thus $P(H_1 \vee H_2 \vee H_3) = 1 - \frac{1}{8} = \frac{7}{8}$.

(10) Using the notation of the answer to Supplementary Problem VI(5), the probability we wish to determine is $P((H_2 \mathbin{\&} \sim H_3) \vee (\sim H_2 \mathbin{\&} H_3) \mid H_1)$. Since the tosses are independent, this equals $P((H_2 \mathbin{\&} \sim H_3) \vee (\sim H_2 \mathbin{\&} H_3))$. The disjuncts here are mutually exclusive; so, by AX3, this becomes

$$P(H_2 \mathbin{\&} \sim H_3) + P(\sim H_2 \mathbin{\&} H_3)$$

and since the tosses are independent, this is, by Problem 10.22,

$$P(H_2) \cdot P(\sim H_3) + P(\sim H_2) \cdot P(H_3) = \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{2}$$