Computer Security

Sets,
Permutations and Combinations,
Probability

Why?

- Cryptography uses a lot of math, some of which you may not have studied, so we need to give you an introduction to the basics, so you can make sense of what comes later.

- Sources:
  - *Applied Finite Mathematics*, by Soo Tang Tan, published by Prindle, Weber & Schmidt

Sets

- A *set* is a collection of objects or items
  - For example:
    - The set of days in the year 1999
    - The set of members of the Nashua Symphony Orchestra
    - The set of Jeep Cherokees produced in the year 2002
    - The set of positive integers
    - The set of real numbers
    - etc...
- The objects of a set are called the *elements* or *members* of the set
- Elements of a set are often denoted by lowercase letters
- Sets are often denoted by uppercase letters

Set Notation

- The conventional notation for a set uses braces, {}
  - A set may be represented by listing all the elements of the set within braces:

    A = {a, b, c}

    B = {a, b, c, ..., z}
  - Alternatively, a set may be represented by *set-builder notation*, which uses a rule to determine the elements of the set:

    A = {x | x is a letter of the English alphabet}

    B = {x | x is an integer > 4 and < 10}

    C = {x | x + 3 = 0; x an integer}
  - If a is an element of set A ("a belongs to A"), we write $a \in A$
  - If a is *not* an element of set A, we write $a \notin A$

Set Notation

- If every element of a set A is also an element of set B, then we say that A is a *subset* of B, and use the notation: $A \subseteq B$
- Two sets are *equal* (***A = B***) if and only if (iff):
  1) $A \subseteq B$

  and

  2) $B \subseteq A$
- If two sets, P and Q, are *not equal*, we say $P \neq Q$
- For example, what are, and are not, subsets in the following sets?

  A = {a, e, i, o, u}
  B = {a, i, o, e, u}
  C = {a, e, i, o}
  D = {a, e, i, o, x}

Set Notation

- If A and B are sets such that $A \subseteq B$ but $A \neq B$ then we say that A is a *proper subset* of B, written $A \subset B$
- If X and Y are sets such that X is *not a subset* of Y, we say $X \not\subset Y$
- A set that has no elements is called the empty set or null set, and is designated by the symbol $\varnothing$
  - For example:

    {x | x is a person with five eyes}
    {x | x is an animal with wings and fins and fingers}
    {x | x + 2 = 0; x is a positive integer}
- The empty set is a *subset of every set*

Set Union

- The *union* of sets A and B, written $A \cup B$, is the set of all elements that belong to either A or B or both:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B \text{ or both}\}$$

- For example:
  - If A = {a, b, c} and B = {a, c, d}, then the union of A and B is:

{a, b, c, d}

Set Intersection

- The *intersection* of sets A and B, written $A \cap B$, is the set of all elements that belong to either A or B or both:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

- For example:
  - If A = {a, b, c} and B = {a, c, d}, then the intersection of A and B is:

{a, c}

Disjoint Sets

- The sets A and B are said to be *disjoint* if they have no elements in common.
  - That is, $A \cap B = \varnothing$
- For example:
  - If F is the set of all female students enrolled in this course, and M is the set of all male students enrolled in this course, then F and M are disjoint.

The Number of Elements in a Finite Set

- Often called a *counting problem*
- Constitutes the field of study known as *combinatorics*
- One way is to simply count the number of elements in the set:
  - If $A = \{1, 2, 3, ..., 15\}$, $B = \{a, b, c\}$, $C = \{8\}$ then $n(A) = 15$, $n(B) = 3$, and $n(C) = 1$
- The empty set has zero elements (surprise!)
- If A and B are disjoint sets, then $n(A \cup B) = n(A) + n(B)$
- If A and B are not disjoint sets, then $n(A \cup B) = n(A) + n(B) - n(A \cap B)$

The Number of Elements in a Finite Set

- For example:
  - Of 100 coffee drinkers, 70 take sugar, 60 take cream, and 50 take both sugar and cream. How many take sugar *or* cream with their coffee?
  - U is the set of 100 coffee drinkers
  - A is the set of coffee drinkers who take sugar: $A = \{x \in U \mid x \text{ takes sugar with coffee}\}$
  - B is the set of coffee drinkers who take cream: $A = \{x \in U \mid x \text{ takes cream with coffee}\}$
  - $n(A) = 70$, $n(B) = 60$, $n(A \cap B) = 50$
  - Since the sets A and B are not disjoint:

$$\begin{aligned} n(A \cup B) &= n(A) + n(B) - n(A \cap B) \\ &= 70 + 60 - 50 \\ &= 80 \end{aligned}$$

The Multiplication Principle

- The *Multiplication Principle* states that, if there are m ways of selecting choice C_1, and n ways of selecting choice C_2, then there are $m \cdot n$ ways of selecting choice C_1, followed by the choice C_2.

- For example:
  - Joe's ice cream parlor offers 10 different flavors of ice cream, and 5 different toppings. How many combinations does Joe offer, which consist of one ice cream flavor, and one topping?
    - There are 10 ways of choosing an ice cream flavor
    - There are 5 ways of choosing a topping
    - So there are $10 \cdot 5$, or 50 possible combinations.

Permutations

- Given a set of (distinct) elements, a *permutation* of the set is an arrangement of those elements in a definite order.
- Consider the set A = {a, b, c}: How many permutations of the elements of this set are there?

Permutations

- Each permutation consists of a sequence of the elements a, b, c
  - Choose one of the elements of the set -- there are 3 choices
  - Now, select a second element of the set; we have already selected one element, so now we have 2 choices left
  - Finally, select the third element of the set; we have already selected two elements, so only one element is left to choose from
  - Using the multiplication principle, there are therefore $3 \cdot 2 \cdot 1 = 6$ permutations of the set.

Permutations

- For example:
  - How many ways can a soccer team of 11 players arrange themselves in a line for a group picture?

Permutations

- The general formula for the number of ways of permuting a set of *n* distinct objects is:

  $n \cdot (n - 1) \cdot (n - 2) \ldots 3 \cdot 2 \cdot 1$

  or *n* factorial, denoted by *n!*
- *n!*, with *n* a positive integer, is defined as follows:

$$n! = \begin{cases} 1 & \text{if } n = 0 \\ n \cdot (n-1)! & \text{if } n > 0 \end{cases}$$

- For example:
  - 1! = 1
  - 2! = 2
  - 3! = 6
  - 4! = 24
  - 5! = 120

    ...
  - 10! = 3,628,800

Try a Java applet that calculates factorials

Permutations

- We are often interested in determining the number of ways of permuting *n* distinct elements, taken *r* elements at a time:
  - Such a permutation may be viewed as being obtained by filling each of *r* places with one element from the set:
    - There are *n* ways of filling the first place, *(n - 1)* ways of filling the second, and so on, until the last, of which there are *(n - r + 1)* choices:

      $n \cdot (n - 1) \cdot (n - 2) \ldots (n - r + 1)$

      or:

      $$\frac{n \cdot (n-1) \cdot (n-2) \ldots (n-r+1) \cdot (n-r) \cdot (n-r-1) \ldots 3 \cdot 2 \cdot 1}{(n-r)(n-r-1) \ldots 3 \cdot 2 \cdot 1} = \frac{n!}{(n-r)!}$$

    - $P(n, r) = \frac{n!}{(n-r)!}$ Try a Java applet that calculates permutations

Permutations

- For example:
  - How many ways can a chairman, vice-chairman, and secretary be chosen from a committee of 8 members?

$$P(n, r) = \frac{n!}{(n-r)!}$$

Permutations with Non-Distinct Elements

- Sometimes we need to consider permutations on *sets with non-distinct elements*.
  - Imagine we have a set of *n* elements, with n_1 of those alike and of one kind, n_2 of them alike and of another kind, ... , and n_r of them alike and of yet another kind.
    - Denote the number of permutations of these *n* non-distinct elements, taken *n* at a time by *x*
    - If each of the subsets of n_m elements is distinct, then they each may be permuted $n_m!$ ways
    - Using the multiplication principle, there are $x \cdot n_1! \cdot n_2! \cdot \ldots \cdot n_r!$ permutations of these elements.
    - But the number of permutations of a set of *n* distinct objects taken *n* at a time is equal to *n!*, so $x \cdot (n_1! \cdot n_2! \cdot \ldots \cdot n_r!) = n!$
    - Therefore:

$$x = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_r!}$$

Permutations with Non-Distinct Elements

- Example:
  - How many permutations can be formed from all the letters in the word NASHUA ? (Note that not all the letters are distinct.)

$$x = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_r!}$$

Combinations

- In many cases, we are interested in determining the number of ways of arranging *r* objects from a set of *n* objects without regard to the order in which the objects are arranged. This is called a *combination*.
- To derive a formula for determining the number of combinations *C(n, r)* of *n* objects taken *r* at a time, note that each combination can be permuted in *r!* ways, so:
  - $r! \cdot C(n, r)$ gives the number of permutations of *n* objects, taken *r* at a time, so:

    $r! \cdot C(n, r) = P(n, r)$

    and therefore:

    $$C(n, r) = \frac{P(n,r)}{r!} = \frac{n!}{r!(n-r)!}$$

Try a Java applet that calculates combinations

Combinations

- For example:
  - How many 5-card poker hands may be dealt from a standard deck of 52 cards? (The order in which a poker hand is dealt is irrelevant.)

$$C(n, r) = \frac{n!}{r!(n-r)!}$$

Binomial Coefficients

- This result:

  $$C(n, r) = \frac{n!}{r!(n-r)!}$$

  is an important one, because this function is known as a *binomial coefficient*, related to an important property, the *binomial theorem*.

  Binomial coefficients are denoted as follows:

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

  and is often known as "*n* choose *r*" -- the number of ways to choose an *r*-element subset from an *n*-element set.

Probability: Experiments and Outcomes

- An *experiment* is an activity with observable results
- These results are referred to as *outcomes* of the experiment
  - Examples:
    - Tossing a coin and observing whether it falls "heads" or "tails"
    - Throwing a die and observing which of the numbers 1, 2, 3, 4, 5, or 6 is uppermost
    - Testing a semiconductor chip from a batch of 1000, and observing whether it is operational or defective

Probability: Sample Points and Sample Spaces

- An outcome of an experiment is called a *sample point* of the experiment
- The set of all possible sample points of an experiment is called the *sample space* of the experiment.
- A subset of the sample space of an experiment is called an *event* of the experiment.
- A sample space which has a finite number of possible outcomes (sample points) is called a *finite sample space*.

Probability: Sample Points and Sample Spaces

- For example:
  - The experiment of tossing a coin:
    - The two outcomes are "heads" or "tails"
    - The sample space is S = {H, T}
    - The events of the experiment (the subsets of S) are:
      ∅, {H}, {T}, S
  - The experiment of tossing a coin 3 times:
    - The sample points are:
      S = {HHH, HHT, ... ? ... }
    - The event that exactly two heads appear:
      E_1 = { ... ? ... }
    - The event that at least one head appears:
      E_2 = { ... ? ... }

Probability: Definition

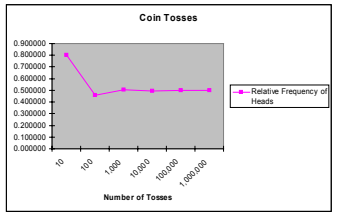
- Consider, again, the experiment of tossing a single coin:
  - There are 2 possible outcomes: H or T
  - Assuming that the coin is *unbiased*, there is *1 chance in 2* that we will obtain a head, and *1 chance in 2* that we will obtain a tail, so:

$$P(H) = \frac{1}{2},\ P(T) = \frac{1}{2}$$

Probability: Definition

- Another approach is based on continued experimentation, and does not depend on the assumption that the outcomes are equally likely.
  - Imagine tossing the coin many times, and obtaining the results (real results from running a coin toss simulator program:

Number of tosses (n)	Number of heads (m)	Relative Frequency (m/n)
10	8	0.800000
100	46	0.460000
1000	503	0.503000
10000	4948	0.494800
100000	49897	0.498970
1000000	500275	0.500275



  - Note that, when the number of trials is small, the relative frequencies vary more from the expected value (0.5) than when the number of trials becomes large.

Try a Java applet that simulates *n* coin tosses

Probability: Definition

- Suppose that in *n* trials, an event *E* occurs *m* times
  - We call the ratio *m/n* the *relative frequency* of the event *E* after *n* repetitions.
  - If this relative frequency approaches some value *P(E)* as *n* becomes larger, then *P(E)* is called the *empirical* probability* of *E*
  - The probability *P(E)* of an event *E* occurring is a measure of the proportion of the time that the event *E* will occur in the long run
  - Note that this measure of probability is effective even when the coin used in the experiment is biased.
- Since the *probability of an event* is a fraction, or is obtained by taking the "limit" of fractions which lie between 0 and 1 inclusive, it is itself a number lying between 0 and 1.
  - The larger the probability of an event, the more likely the event will occur

***em·pir·i·cal**
1 : originating in or based on observation or experience <*empirical* data>
2 : relying on experience or observation alone often without due regard for system and theory
3 : capable of being verified or disproved by observation or experiment <*empirical* laws>
4 : of or relating to empiricism

Probability: Basic Events

- Suppose we decide to determine the probabilities associated with certain events of an experiment.
  - We could compute *P(E)* directly for each event *E*. However, in practice the number of events may be quite large, making the approach impractical.
  - If the sample space of the experiment is finite (we will assume from now on that all sample spaces are finite), then we can do the following:
    - Let *S* be a finite sample space with *n* outcomes $s_1, s_2, s_3, \ldots, s_n$

      $$S = \{s_1, s_2, s_3, \ldots, s_n\}$$
    - Then the *n* singleton sets:

      $$\{s_1\}, \{s_2\}, \{s_3\}, \ldots, \{s_n\}$$

      are called the *elementary* or *basic events* of the experiment.
    - The basic events of the experiment are
      - *exhaustive* -- one of them must occur in a trial
      - *mutually exclusive* -- only one can occur in a trial

Probability Functions

- Properties of a *probability function P*:

  1) $0 \le P(s_i) \le 1 \qquad (i = 1,2,\ldots,n)$

  2) $\sum_{i=1}^{n} P(s_i) = P(s_1) + P(s_2) + \cdots + P(s_n) = 1$

  3) $P(\{s_i\} \cup \{s_j\}) = P(s_i) + P(s_j) \quad i \ne j \quad (i = 1,2,\ldots,n;\ j = 1,2,\ldots n)$

  4) $P(\{s_1\} \cup \{s_2\} \cup \cdots \cup \{s_m\}) = P(s_1) + P(s_2) + \cdots + P(s_m)$

  1) follows from the fact that $0 \le m \le n$, when we count *m* occurrences of an event s_i from *n* trials, and the relative frequency of the event is *m/n*

  2) follows from the fact that one of the s_i is certain to occur

  3) follows from the fact that s_i and s_j are mutually exclusive, and so obey the Multiplication Principle

  4) follows from 3), generalized to *m* mutually exclusive events

Probability Distributions

- If we perform *n* trials of an experiment, and determine the following:

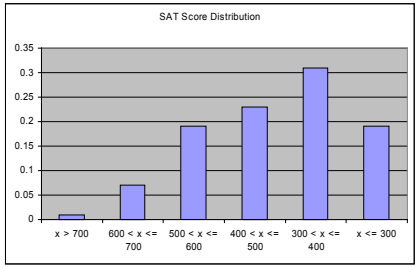
Basic Event	Probability
$\{s_1\}$	$P(s_1)$
$\{s_2\}$	$P(s_2)$
$\{s_3\}$	$P(s_3)$
.	.
.	.
.	.
$\{s_n\}$	$P(s_n)$

this is called a *probability distribution* for the experiment.

Probability Distributions

- Example: The observed probability distribution of test scores of students in a school district:

Score	Probability
x > 700	0.01
600 < x <= 700	0.07
500 < x <= 600	0.19
400 < x <= 500	0.23
300 < x <= 400	0.31
x <= 300	0.19



Probability Distributions

- Example:
  - Given the probability distribution of student test scores:

Score	Probability
x > 700	0.01
600 < x <= 700	0.07
500 < x <= 600	0.19
400 < x <= 500	0.23
300 < x <= 400	0.31
x <= 300	0.19

what is the probability that a randomly selected student will obtain a test score of:

- more than 400?
- less than or equal to 500?
- greater than 400, but less than or equal to 600?

Probability Distributions

- The following property of the probability function, P, applies when the two events s_i and s_j are mutually exclusive:

  3) $P(\{s_i\} \cup \{s_j\}) = P(s_i) + P(s_j) \quad i \neq j \quad (i = 1,2,\ldots,n; j = 1,2,\ldots n)$

- If the two events are *not* mutually exclusive, the above formula is generalized to:

  5) $P(\{s_i\} \cup \{s_j\}) = P(s_i) + P(s_j) - P(\{s_i\} \cap \{s_j\})$

  $i \neq j \quad (i = 1,2,\ldots,n; j = 1,2,\ldots n)$

  - For example:
    - If a card is drawn from a well-shuffled pack of 52 playing cards, what is the probability that it is an ace or a spade? (Note that these events are not mutually exclusive)

Probability Distributions

- For sample spaces whose outcomes are *equally likely* (a.k.a. *uniform sample spaces*), the assignment of probabilities to the basic events is simple. If we have n sample points in the sample space S = $\{s_1, s_2, s_3, \ldots, s_n\}$ then we assign equal values to the probabilities:

$$P(s_1) = P(s_2) = \cdots = P(s_n) = \frac{1}{n}$$

Probability: Counting Techniques

- If S is a *uniform sample space*, and E is any event for that space, then:

$$P(E) = \frac{\text{number of favorable outcomes in } E}{\text{number of possible outcomes in } S} = \frac{n(E)}{n(S)}$$

  - Example:
    - An unbiased coin is tossed 6 times.
      a) What is the probability that the coin will land "heads" all 6 times?
      b) What is the probability that the coin will land "heads" on the first and the last toss?

Probability: Counting Techniques

- An unbiased coin is tossed 6 times.
  a) What is the probability that the coin will land "heads" all 6 times?
  - Each outcome of the experiment may be represented as a sequence of heads and tails
  - Using the generalized multiplication principle:
    - the number of outcomes of one toss of the coin is 2
    - the number of outcomes of two tosses of the coin is $2 \cdot 2 = 2^2 = 4$
    - the number of outcomes of three tosses of the coin is $2 \cdot 2 \cdot 2 = 2^3 = 8$
    - the number of outcomes of n tosses of the coin is 2^n
  - So the number of outcomes of 6 tosses is $2^6 = 64$
  - The event, E, that the coin will land "heads" all six times can only occur in one way, so the probability of this event is:

$$P(E) = \frac{n(E)}{n(S)} = \frac{1}{64} = 0.015625$$

Probability: Counting Techniques

- An unbiased coin is tossed 6 times.
  b) What is the probability that the coin will land "heads" on the first and the last toss?
  - The event, F, that the coin lands "heads" on the first and last tosses can occur in a number of different ways:
    - "Heads" on the first toss (1 possibility)
    - 4 tosses, where the result is irrelevant (2 possibilities for each toss)
    - "Heads" on the last toss (1 possibility)

    or: $1 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 1 = 2^4 = 16$
  - So the probability of F occurring is:

$$P(F) = \frac{n(F)}{n(S)} = \frac{16}{64} = \frac{1}{4} = 0.25$$

Conditional Probability

- The probability of an event is often affected by the occurrence of other events. In general, given two events A and B of an experiment, one may compute the probability of the event B, given that the event A has already occurred.
- This probability, denoted by $P(B \mid A)$, is called the *conditional probability of B, given A*.
  - Once event A has occurred, the outcomes favorable to the event B are those in the set $A \cap B$
  - So the conditional probability of B, given A is:

$$P(B \mid A) = \frac{\text{number of elements in } A \cap B}{\text{number of elements in } A} = \frac{n(A \cap B)}{n(A)} \qquad [n(A) \neq 0]$$

Conditional Probability

- Starting from the result in the previous slide, and dividing top and bottom by $n(S)$:

$$P(B \mid A) = \frac{n(A \cap B)}{n(A)} \qquad [n(A) \neq 0]$$

$$= \frac{\frac{n(A \cap B)}{n(S)}}{\frac{n(A)}{n(S)}} = \frac{P(A \cap B)}{P(A)} \qquad [P(A) \neq 0]$$

Conditional Probability

- Sometimes, the probability of an event B occurring, given that A has occurred, written $P(B \mid A)$, is known, and we wish to find the probability of both A **and** B occurring
  - Using: $P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad [P(A) \neq 0]$

    (from the previous slide) and multiplying both sides by $P(A)$ gives:

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

    which is known as the *Product Rule*.

Conditional Probability

- Example:
  - Two cards are drawn (without replacement) from a well-shuffled deck of 52 playing cards

    a) What is the probability that the first card drawn is an ace?

    b) What is the probability that the second card drawn is an ace, given that the first card drawn *was not* an ace?

    c) What is the probability that the second card drawn is an ace, given that the first card drawn *was* an ace?

Conditional Probability

- Two cards are drawn (without replacement) from a well-shuffled deck of 52 playing cards

  a) What is the probability that the first card drawn is an ace?

  - There are 52 equally likely possible outcomes
  - 4 of those possible outcomes are aces
  - So the probability that the first card drawn will be an ace is:

$$P(E) = \frac{4}{52} = 0.0769 \text{ (approx)}$$

Conditional Probability

- Two cards are drawn (without replacement) from a well-shuffled deck of 52 playing cards
  b) What is the probability that the second card drawn is an ace, given that the first card drawn *was not* an ace?
  - After the first card has been drawn, there are 51 cards left in the deck -- we have a *reduced sample space*
  - If the card drawn first *was not* an ace, the the 51 card sample space still contains 4 possible favorable outcomes (an ace)
  - So the probability of the second card being an ace is:

$$P(E) = \frac{4}{51} = 0.07843 \text{ (approx)}$$

Conditional Probability

- Two cards are drawn (without replacement) from a well-shuffled deck of 52 playing cards
  c) What is the probability that the second card drawn is an ace, given that the first card drawn *was* an ace?
  - If the first card drawn was an ace, then there are 3 aces left in the deck of 51 cards
  - So the probability of the second card being an ace is:

$$P(E) = \frac{3}{51} = 0.0588 \text{ (approx)}$$

Conditional Probability

- Example:
  - HQ Corporation manufactures PCs at three locations and then ships them to the main distribution plant.
  - Plants A, B, and C manufacture 50%, 30%, and 20%, respectively, of the PCs. The QA department has determined that 1% of the PCs from plant A are defective, while 2% of the PCs from plants B and C are defective.
  - If a PC is selected at random, what is the probability that the PC will be found to be defective?
    - Let *A, B,* and *C* denote events that the PC chosen was manufactured in plant A, B and C
    - Let D denote the event that a PC is defective. Taking the product of probabilities, and adding them:

$$\begin{aligned} P(D) &= (0.5)(0.01) + (0.3)(0.02) + (0.2)(0.02) \\ &= 0.005 + 0.006 + 0.004 \\ &= 0.015 \end{aligned}$$

Probability: Independent Events

- In the previous example of drawing aces from a deck, event *B* was *dependent* on event *A*.
- There are occasions when we have *independent events*:
  - Examples:
    - Tossing a coin twice
      - The outcome of the second toss is in no way affected by the result of the first toss.
    - Buying a lottery ticket on two successive days
      - Despite folklore, the likelihood of winning with the second lottery ticket is *not* dependent on whether you won with the first lottery ticket.

Probability: Independent Events

- Two events, *A* and *B* are independent if:

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

or:

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

so, by the product rule:

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

Probability: Independent Events

- Example:
  - A survey found that, of 2000 women, 680 were heavy smokers and 50 had emphysema*. Of those who had emphysema, 42 were also heavy smokers.
  - Are the events "being a heavy smoker" and "having emphysema" independent events?

***em·phy·se·ma**
: a condition characterized by air-filled expansions of body tissues; *specifically* : a condition of the lung marked by abnormal dilation of its air spaces and distension of its walls and frequently by impairment of heart action

Probability: Independent Events

- Let *A* denote "being a heavy smoker", and *B* denote "having emphysema"
- The probability that a women is a heavy smoker *and* has emphsyema is:

$$P(A \cap B) = \frac{42}{2000} = 0.021$$

- The probability that a woman is a heavy smoker is:

$$P(A) = \frac{680}{2000} = 0.34$$

- The probability that a woman has emphysema is:

$$P(B) = \frac{50}{2000} = 0.025$$

- Since $P(A) \cdot P(B) = 0.34 \cdot 0.025 = 0.0085 \neq P(A \cap B)$, we conclude that *A* and *B* are *not independent events*

Probability: Bayes' Theorem

- We have been calculating probabilities that give the likelihood that an event *will* occur -- *a priori probabilities*
- There are also probabilities that are calculated *after* the outcomes of experiments -- *a posteriori probabilities*
  - Example:
    - Three machines, A, B, and C produce similar electronic components.
    - Machine A produces 45% of the total components
    - Machine B produces 30%
    - Machine C produces 25%
    - In normal production:
      - 6% of machine A's components do not meet specifications
      - 4% of machine B's components do not meet specifications
      - 3% of machine B's components do not meet specifications
    - One component is selected at random from the total output, and is found to be defective.
    - What is the probability that the component selected was produced by machine A?

Probability: Bayes' Theorem

- We can answer this question by determining the *a posteriori probability* for the event that the component selected was produced by machine A
  - Let *A, B, C* denote the event that a component is produced by machine A, machine B, and machine C.
  - The events *A, B, C* are mutually exclusive events, and form a *partition* of the sample space, *S*. Their union is *S*.
  - The event D that a component is defective may be expressed as:
    $$D = (A \cap D) \cup (B \cap D) \cup (C \cap D)$$
  - The event that a component is defective and produced by machine A is given by $A \cap D$
  - Thus, the a posteriori probability that a defective component was selected by machine A is given by:

$$P(A \mid D) = \frac{n(A \cap D)}{n(D)}$$

Probability: Bayes' Theorem

- Dividing top and bottom by n(S), and noting that the events
  $$A \cap D, B \cap D, \text{and } C \cap D$$
  are mutually exclusive, we obtain:

$$P(A \mid D) = \frac{n(A \cap D)}{n(D)} = \frac{\dfrac{n(A \cap D)}{n(S)}}{\dfrac{n(D)}{n(S)}}$$

$$= \frac{P(A \cap D)}{P(D)} = \frac{P(A \cap D)}{P(A \cap D) + P(B \cap D) + P(C \cap D)}$$

Probability: Bayes' Theorem

- Now, by using the product rule, we can express:

$$P(A \cap D) = P(A) \times P(D \mid A)$$
$$P(B \cap D) = P(B) \times P(D \mid B)$$
$$P(C \cap D) = P(C) \times P(D \mid C)$$

so that the previous formula becomes:

$$P(A \mid D) = \frac{P(A \cap D)}{P(A \cap D) + P(B \cap D) + P(C \cap D)}$$

$$= \frac{P(A) \times P(D \mid A)}{P(A) \times P(D \mid A) + P(B) \times P(D \mid B) + P(C) \times P(D \mid C)}$$

which is a special case of the result known as *Bayes' Theorem*

Probability: Bayes' Theorem

- The *general form of Bayes' Theorem*:
  - Let $A_1, A_2, \ldots, A_n$ be a partition of a sample space S (i.e., $A_1, A_2, \ldots, A_n$ are mutually exclusive events, and $S = A_1 \cup A_2 \cup \ldots \cup A_n$), and let E be an event of the experiment such that $P(E) \neq 0$. Then the *a posteriori* probability $P(A_1 \mid E)$ $(1 \leq i \leq n)$ is given by:

$$P(A_i \mid E) = \frac{P(A_i) \times P(E \mid A_i)}{P(A_1) \times P(E \mid A_1) + P(A_2) \times P(E \mid A_2) + \cdots + P(A_n) \times P(E \mid A_n)}$$

Probability: Bayes' Theorem

- Example:
  - HQ Corporation manufactures PCs at three locations and then ships them to the main distribution plant.
  - Plants A, B, and C manufacture 50%, 30%, and 20%, respectively, of the PCs. The QA department has determined that 1% of the PCs from plant A are defective, while 2% of the PCs from plants B and C are defective.
  - If a PC is selected at random, and it is found to be defective, what is the probability that the PC was manufactured in plant C?
    - Let *A, B,* and *C* denote events that the PC chosen was manufactured in plant A, B and C
    - Let *D* denote the event that a PC is defective. Using the information, plus Bayes' Theorem, the a posteriori probability is given by:

$$P(C \mid D) = \frac{P(C)\cdot P(D \mid C)}{P(A)\cdot P(D \mid A) + P(B)\cdot P(D \mid B) + P(C)\cdot P(D \mid C)}$$
$$= \frac{(0.2)(0.02)}{(0.5)(0.01) + (0.3)(0.02) + (0.2)(0.02)}$$
$$= 0.27$$

Probability: Bayes' Theorem

- Example:
  - A study of the annual incomes of married couples, in which the husbands were the sole providers and of those in which the husbands and wives were both employed, found the following results:

Annual income ($)	% of Married Couples	% of Income Group with both spouses working
75,000 and over	4	65
50,000 to 74,999	10	73
35,000 to 49,999	21	68
25,000 to 34,999	24	63
15,000 to 24,999	30	43
Under $15,000	11	28

a) What is the probability that a couple selected at random has two incomes?

b) If a randomly chosen couple has two incomes, what is the probability that the annual income of this couple is over $75,000?

c) If a randomly chosen couple has two incomes, what is the probability that the annual income of this couple is greater than $24,999?

Probability: Bayes' Theorem

a) What is the probability that a couple selected at random has two incomes?

- Let *A, B, C, D, E,* and *F* denote the events listed in the table, starting from the top
- Let *T* denote the event that a couple is a two-income family.
- The probability that a couple selected at random has two incomes is given by:

$$P(T) = P(A)\cdot P(T \mid A) + P(B)\cdot P(T \mid B) + P(C)\cdot P(T \mid C)$$
$$+ P(D)\cdot P(T \mid D) + P(E)\cdot P(T \mid E) + P(F)\cdot P(T \mid F)$$
$$= (0.04)(0.65) + (0.10)(0.73) + (0.21)(0.68)$$
$$+ (0.24)(0.63) + (0.30)(0.43) + (0.11)(0.28)$$
$$= 0.5528$$

Probability: Bayes' Theorem

b) If a randomly chosen couple has two incomes, what is the probability that the annual income of this couple is over $75,000?

- Using the result from part a), and Bayes' Theorem, the probability that a randomly chosen couple has an annual income over $75,000, given that both spouses are working is:

$$P(A \mid T) = \frac{P(A)\cdot P(T \mid A)}{P(T)} = \frac{(0.04)(0.65)}{0.5528} = 0.047$$

Probability: Bayes' Theorem

c) If a randomly chosen couple has two incomes, what is the probability that the annual income of this couple is greater than $24,999?

– The probability that a randomly chosen couple has an annual income greater than $24,999, given that both spouses are working is:

$$P(A \mid T) + P(B \mid T) + P(C \mid T) + P(D \mid T)$$

$$= \frac{P(A) \cdot P(T \mid A) + P(B) \cdot P(T \mid B) + P(C) \cdot P(T \mid C) + P(D) \cdot P(T \mid D)}{P(T)}$$

$$= \frac{(0.04)(0.65) + (0.1)(0.73) + (0.21)(0.68) + (0.24)(0.63)}{0.5528}$$

$$= 0.711$$