Foundation of Data Science and Analytics

Probability Distribution

Arun K. Timalsina

Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Clarification on Probability

What does it mean :

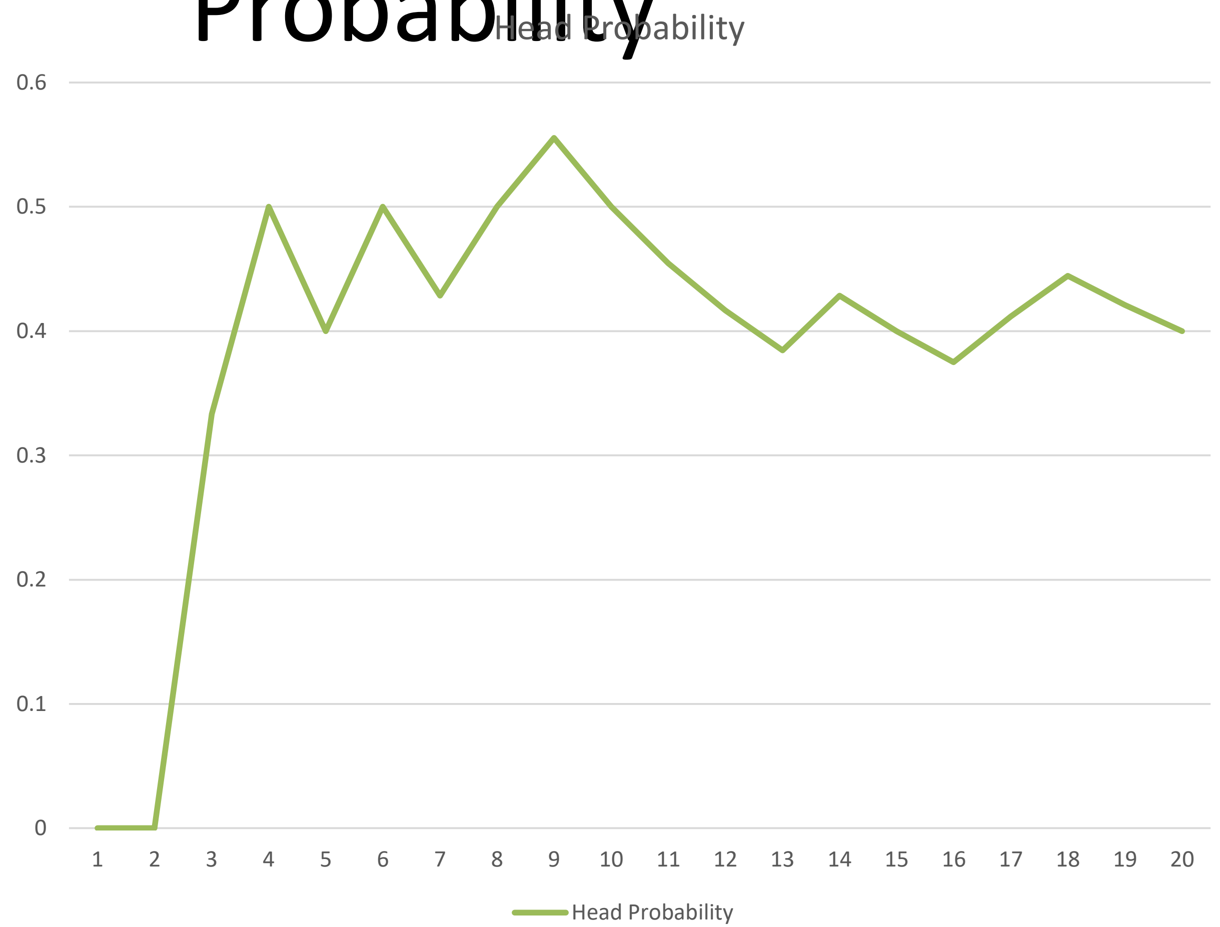
Probability of getting "***Head***" on coin toss experiment is ***0.5*** ?

Head / Tail Experiment

Event Count	Outcome	He
1	Head	
2	Tail	
3	Head	
4	Tail	
5	Tail	
6	Tail	
7	Tail	
8	Tail	
9	Head	
10	Tail	
11	Head	
12	Tail	
13	Head	
14	Tail	
15	Head	
16	Head	
17	Head	
18	Tail	
19	Tail	
20	Head	

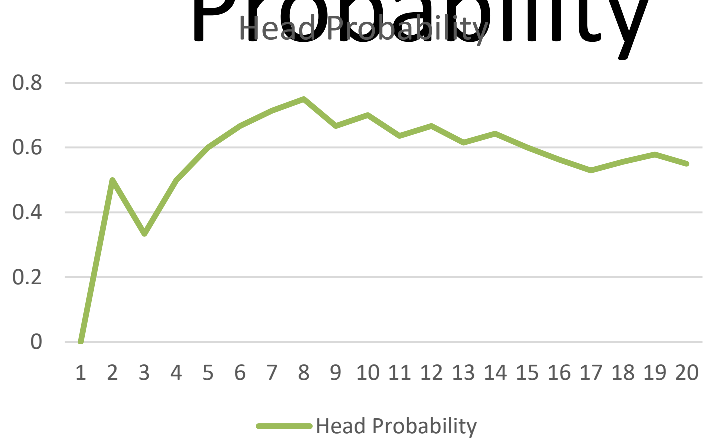
Probability

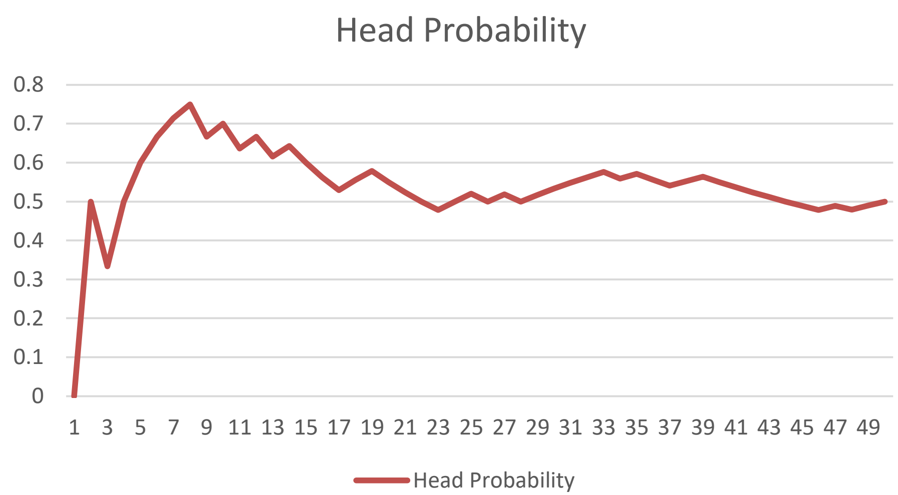
Event Count	Outcome	He
1	Head	
2	Tail	
3	Head	
4	Tail	
5	Tail	
6	Tail	
7	Tail	
8	Tail	
9	Head	
10	Tail	
11	Head	
12	Tail	
13	Head	
14	Tail	
15	Head	
16	Head	
17	Head	
18	Tail	
19	Tail	
20	Head	



Probability

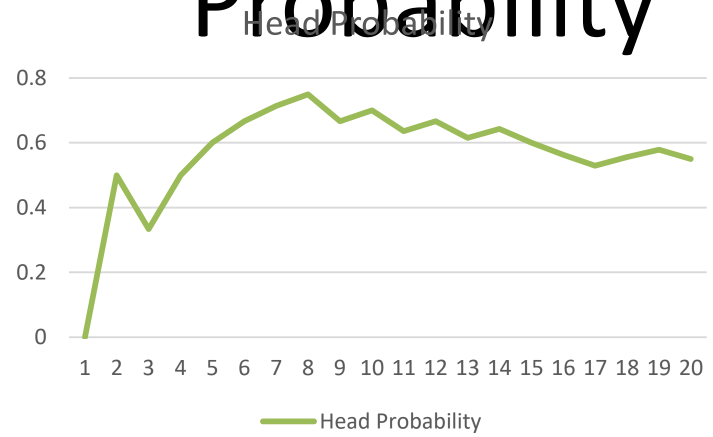
Event Count	Outcome	He
1	Head	
2	Tail	
3	Head	
4	Tail	
5	Tail	
6	Tail	
7	Tail	
8	Tail	
9	Head	
10	Tail	
11	Head	
12	Tail	
13	Head	
14	Tail	
15	Head	
16	Head	
17	Head	
18	Tail	
19	Tail	
20	Head	

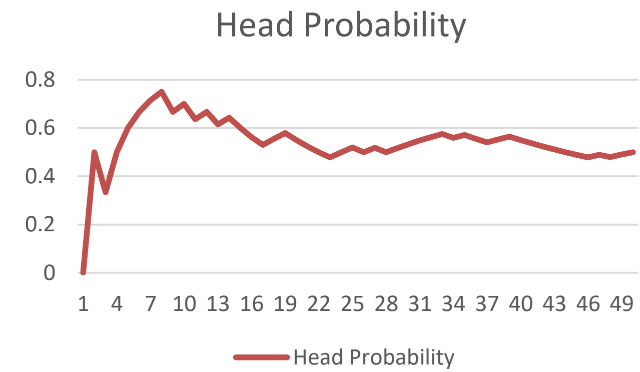


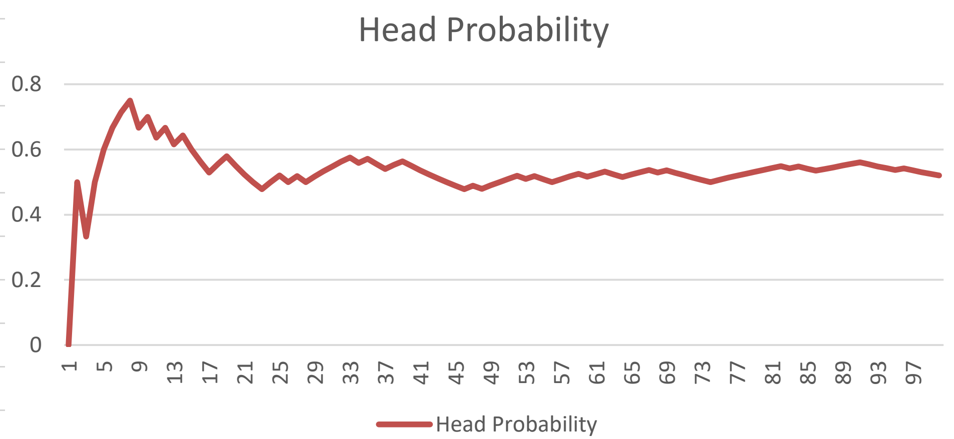


Probability

Event Count	Outcome	He
1	Head	
2	Tail	
3	Head	
4	Tail	
5	Tail	
6	Tail	
7	Tail	
8	Tail	
9	Head	
10	Tail	
11	Head	
12	Tail	
13	Head	
14	Tail	
15	Head	
16	Head	
17	Head	
18	Tail	
19	Tail	
20	Head	

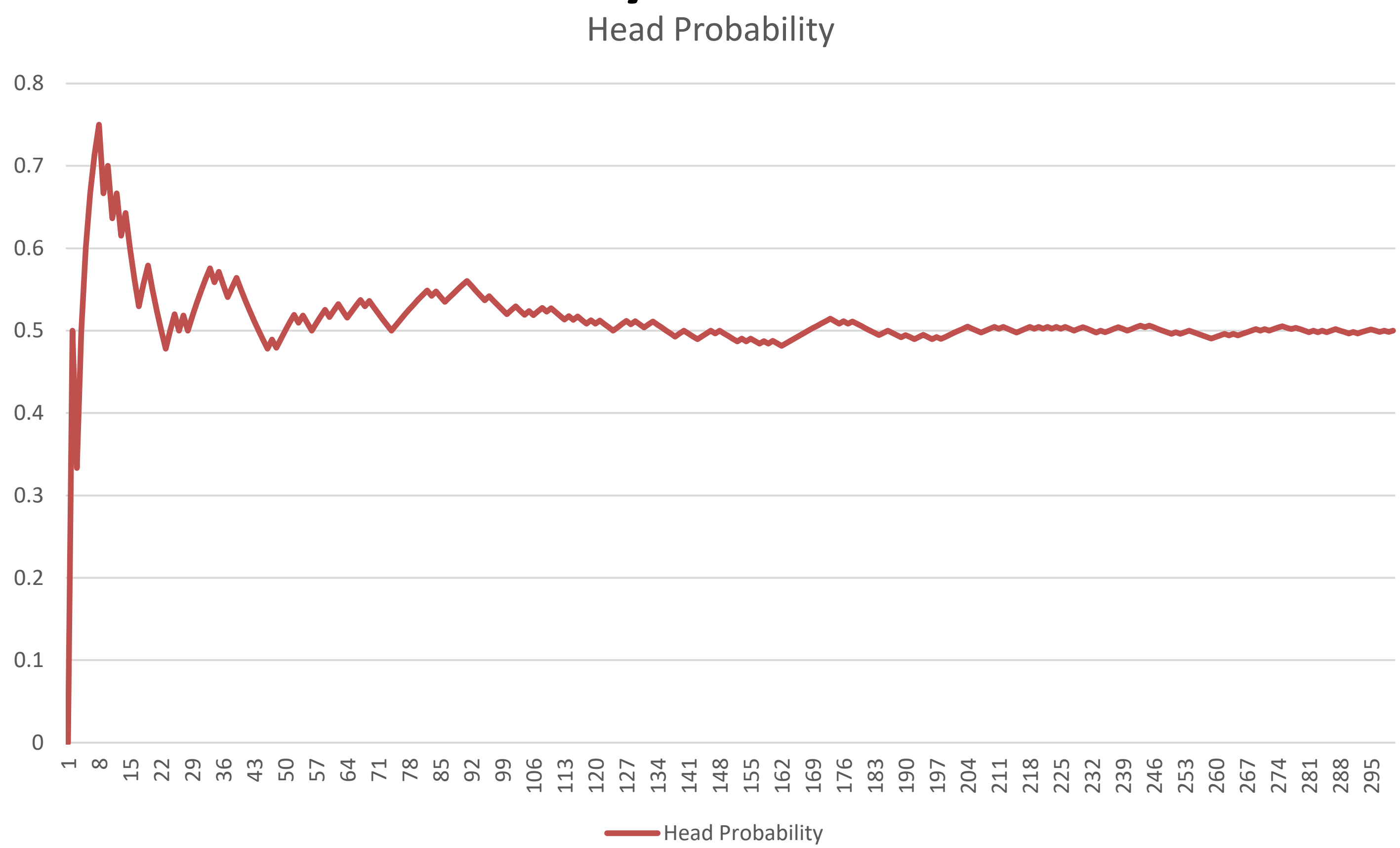






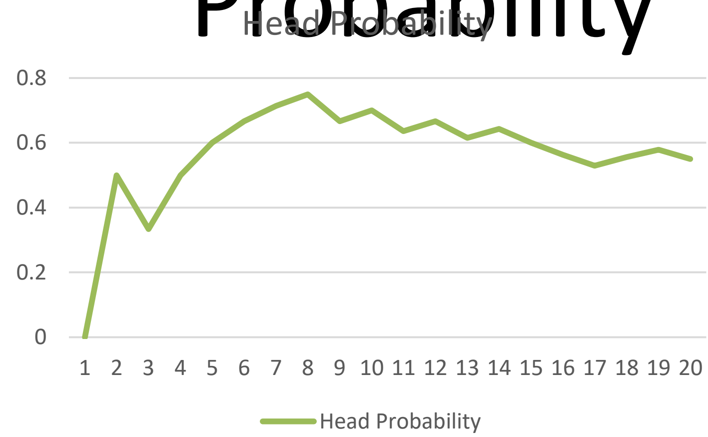
Probability

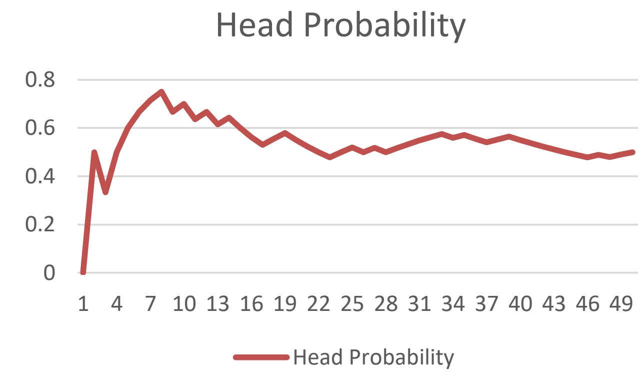
Event Count	Outcome	He
1	Head	
2	Tail	
3	Head	
4	Tail	
5	Tail	
6	Tail	
7	Tail	
8	Tail	
9	Head	
10	Tail	
11	Head	
12	Tail	
13	Head	
14	Tail	
15	Head	
16	Head	
17	Head	
18	Tail	
19	Tail	
20	Head	

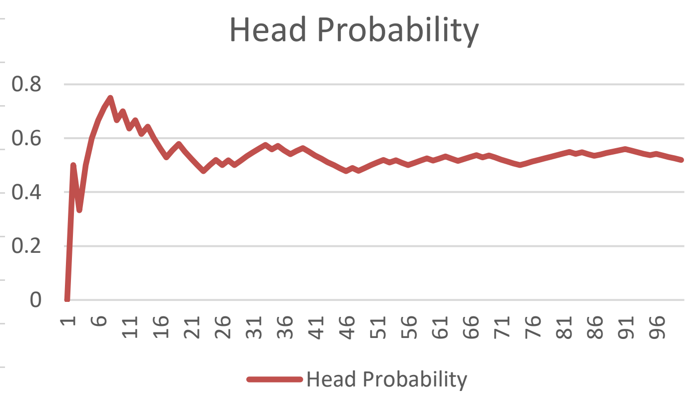


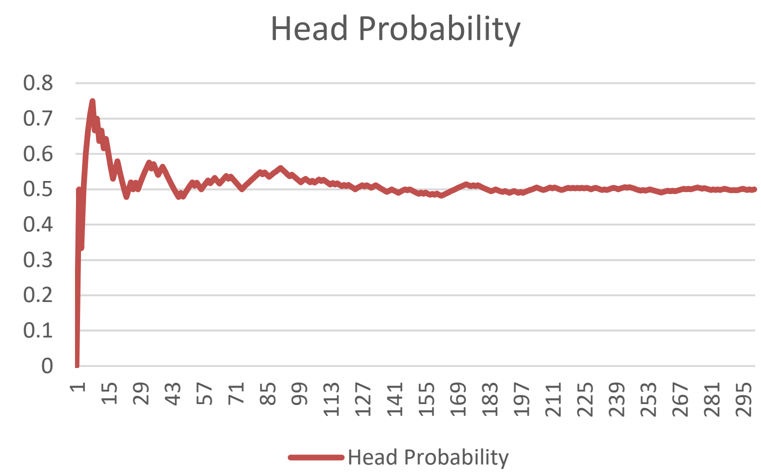
Probability

Event Count	Outcome	He
1	Head	
2	Tail	
3	Head	
4	Tail	
5	Tail	
6	Tail	
7	Tail	
8	Tail	
9	Head	
10	Tail	
11	Head	
12	Tail	
13	Head	
14	Tail	
15	Head	
16	Head	
17	Head	
18	Tail	
19	Tail	
20	Head	









Random Experiments

The goal is to understand, quantify and model the variation affecting a physical system's behavior. The model is used to analyze and predict the physical system's behavior as system inputs affect system outputs. The predictions are verified through experimentation with the physical system.

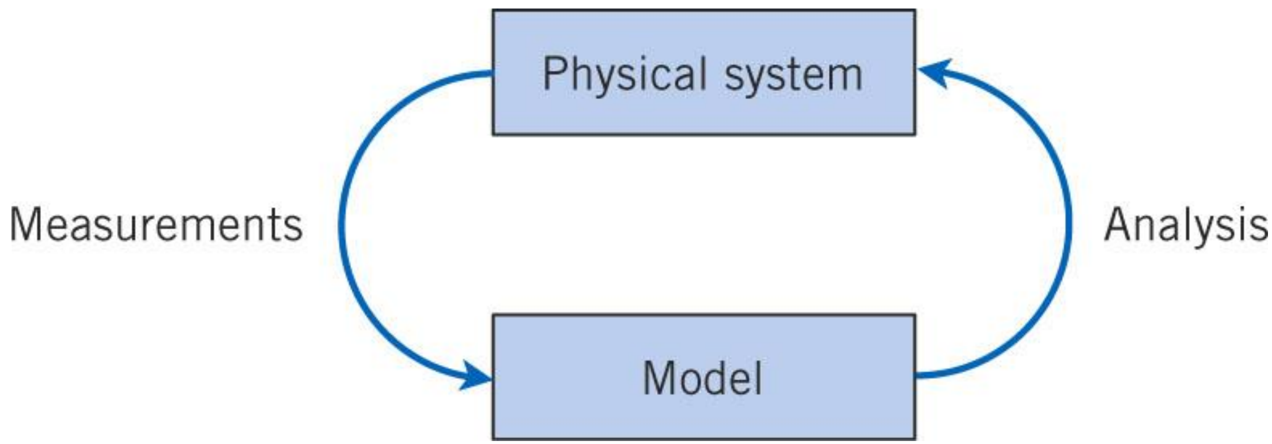


Figure 2-1 Continuous iteration between model and physical system.

Noise Produces Output Variation

Random values of the noise variables cannot be controlled and cause the random variation in the output variables. Holding the controlled inputs constant does not keep the output values constant.

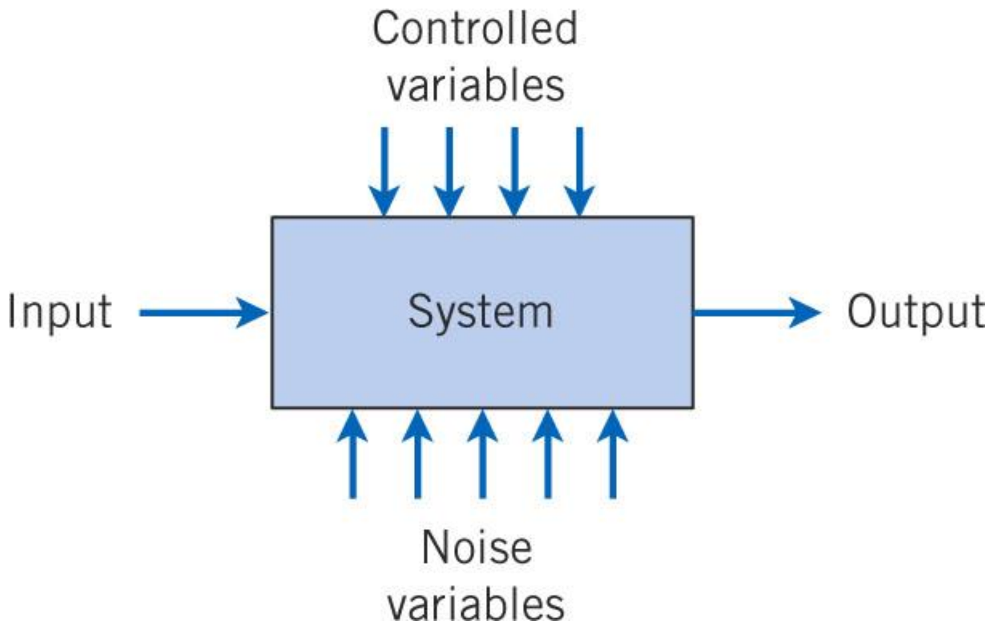


Figure 2-2 Noise variables affect the transformation of inputs to outputs.

Random Experiment

- An experiment is an operation or procedure, carried out under controlled conditions, executed to discover an unknown result or to illustrate a known law.
- An experiment that can result in different outcomes, even if repeated in the same manner every time, is called a random experiment.

Randomness Affects Natural Law

Ohm's Law current is a linear function of voltage. However, current will vary due to noise variables, even under constant voltage.

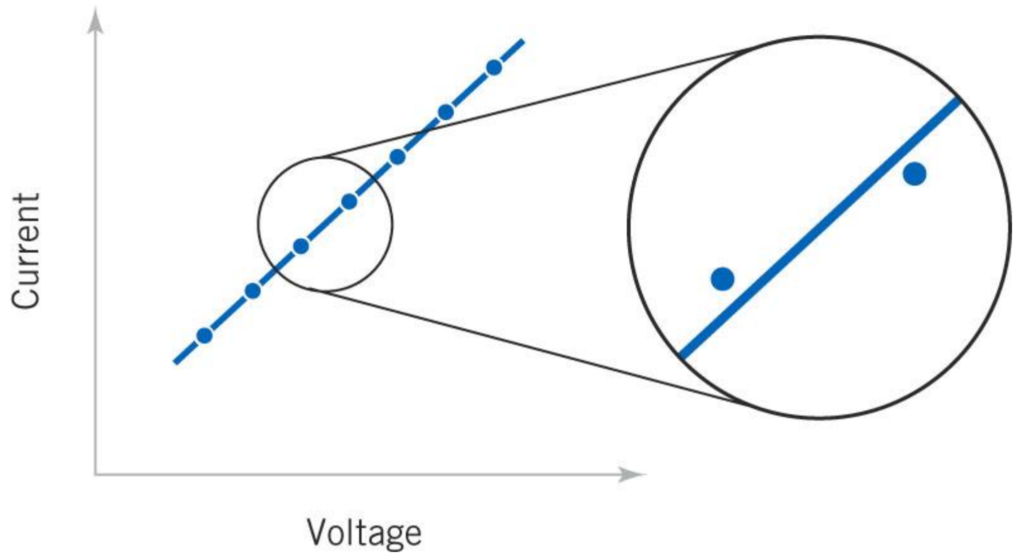


Figure 2-3 A closer examination of the system identifies deviations from the model.

Randomness Can Disrupt a System

- Telephone systems must have sufficient capacity (lines) to handle a random number of callers at a random point in time whose calls are of a random duration.
- If calls arrive exactly every 5 minutes and last for exactly 5 minutes, only 1 line is needed – a deterministic system.
- Practically, times between calls are random and the call durations are random. Calls can come into conflict as shown in following slide.
- Conclusion: Telephone system design must include provision for input variation.

Deterministic & Random Call Behavior

Calls arrive every 5 minutes. In top system, call durations are all of 5 minutes exactly. In bottom system, calls are of random duration, averaging 5 minutes, which can cause blocked calls, a "busy" signal.

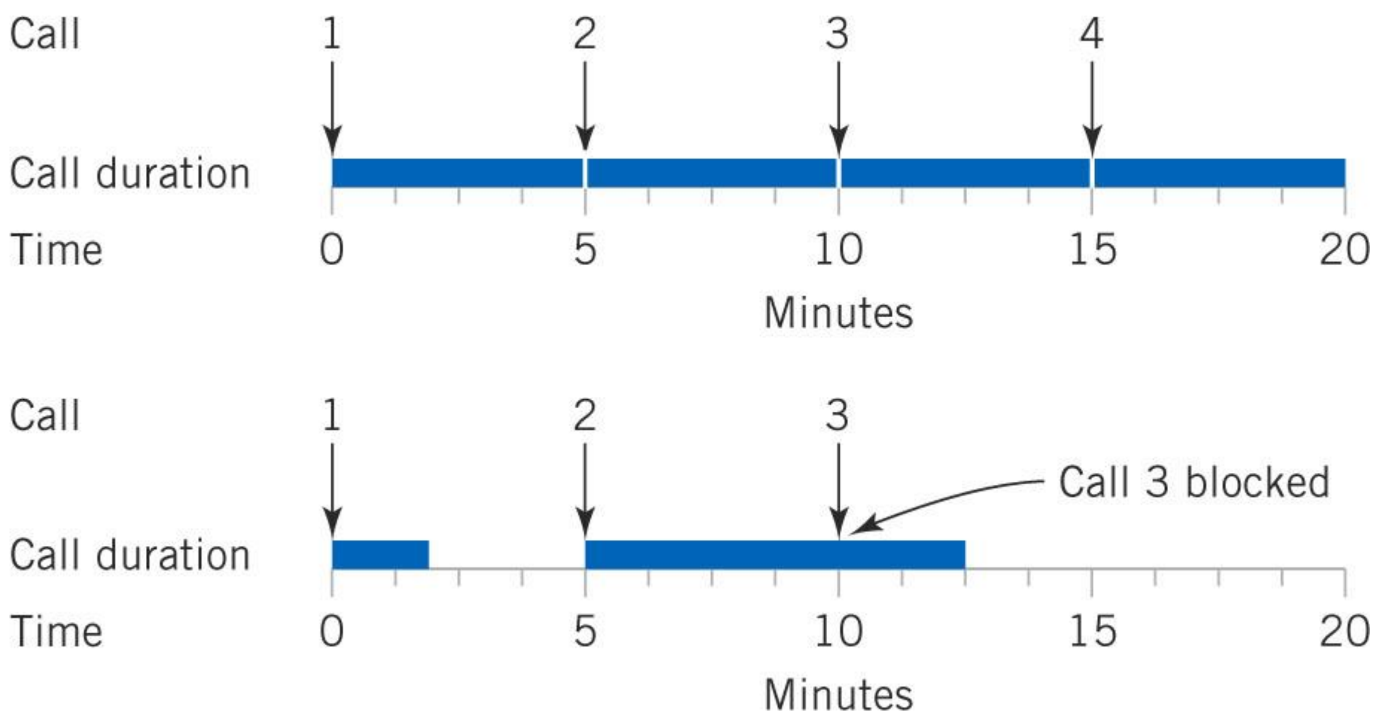


Figure 2-4 Variation causes disruption in the system.

Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Sample Spaces

- Random experiments have unique outcomes.
- The set of all possible outcome of a random experiment is called the sample space, *S*.
- *S* is discrete if it consists of a finite or countable infinite set of outcomes.
- S is continuous if it contains an interval (either a finite or infinite width) of real numbers.

Example 2-1: Defining Sample Spaces

- Randomly select and measure the thickness of a part. $S = R^{+} = \{x \mid x > 0\}$, the positive real line. Negative or zero thickness is not possible. S is continuous.
- It is known that the thickness is between 10 and 11 mm. $S = \{x \mid 10 < x < 11\}$, continuous.
- It is known that the thickness has only three values. $S = \{low, medium, high\}$, discrete.
- Does the part thickness meet specifications? $S = \{yes, no\}$, discrete.

Example 2-2: Defining Sample Spaces, n=2

- Two parts are randomly selected & measured.
  $S = R^+ * R^+$, S is continuous.
- Do the 2 parts conform to specifications?
  $S = \{yy, yn, ny, nn\}$, S is discrete.
- Number of conforming parts?
  $S = \{1, 1, 2\}$, S is discrete.
- Parts are randomly selected until a non-conforming part is found.
  $S = \{\text{n, yn, yyn, yyyn, …}\}$,
  S is countably infinite.

Sample Space Is Defined By A Tree Diagram

Example 2-3: Messages are classified as on-time or late. 3 messages are classified. There are $2^3 = 8$ outcomes in the sample space.

$$S = \{ooo, ool, olo, oll, loo, lol, llo, lll\}$$

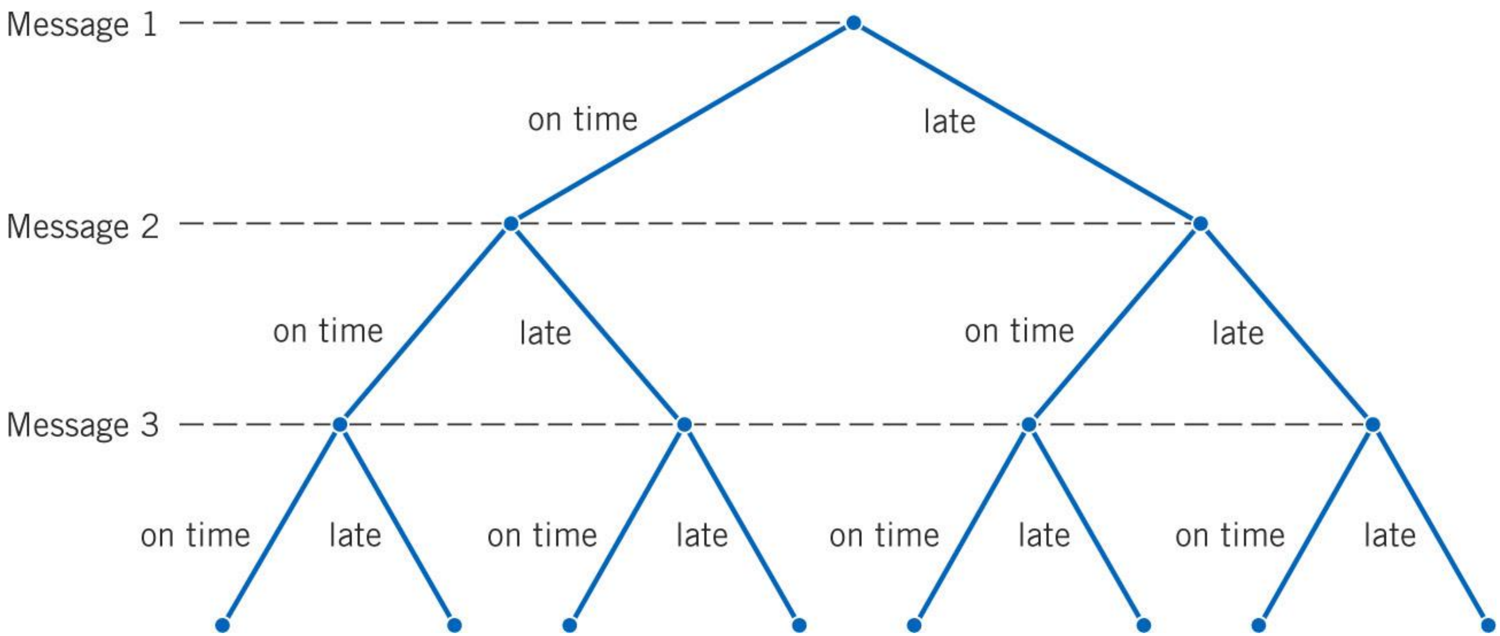


Figure 2-5 Tree diagram for three messages.

Tree Diagrams Can Fit The Situation

Example 2-4: New cars can be equipped with selected options as follows:

1. Manual or automatic transmission
2. With or without air conditioning
3. Three choices of stereo sound systems
4. Four exterior color choices

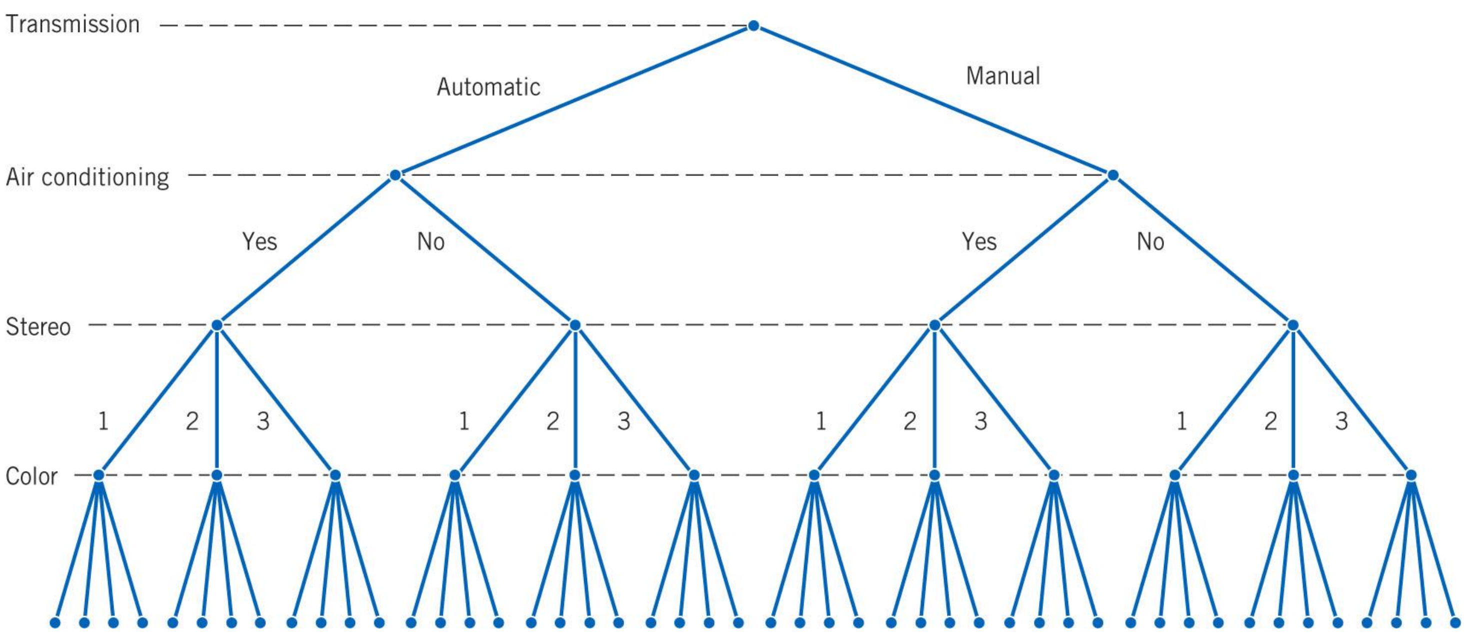


Figure 2-6 Tree diagram for different configurations of vehicles. Note that *S* has 2*2*3*4 = 48 outcomes.

Tree Diagrams Help Count Outcomes

Example 2-5: The interior car color can depend on the exterior color as shown in the tree diagrams below. There are 12 possibilities without considering color combinations.

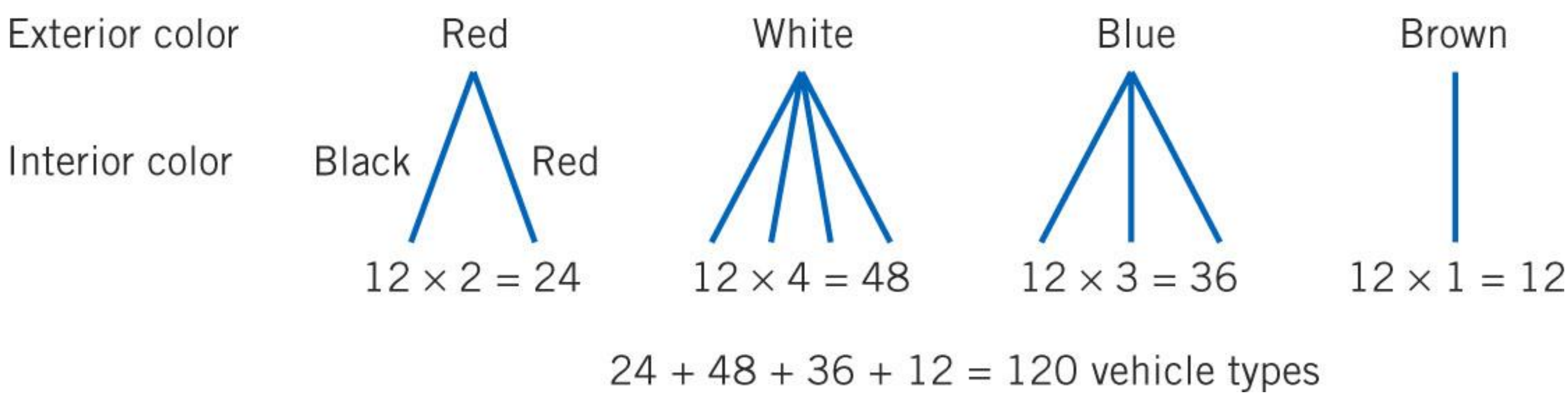


Figure 2-7 Tree diagram for different vehicle configurations with interior colors.

Events Are Sets of Outcomes

- An event (E) is a subset of the sample space of a random experiment, i.e., one or more outcomes of the sample space.
- Event combinations are:
  - Union of two events is the event consisting of all outcomes that are contained in <u>either</u> of two events, $E_1 \cup E_2$. Called E_1 <u>or</u> E_2.
  - Intersection of two events is the event consisting of all outcomes that contained in <u>both</u> of two events, $E_1 \cap E_2$. Called E_1 <u>and</u> E_2.
  - Complement of an event is the set of outcomes that are <u>not</u> contained in the event, E' or <u>not</u> E.

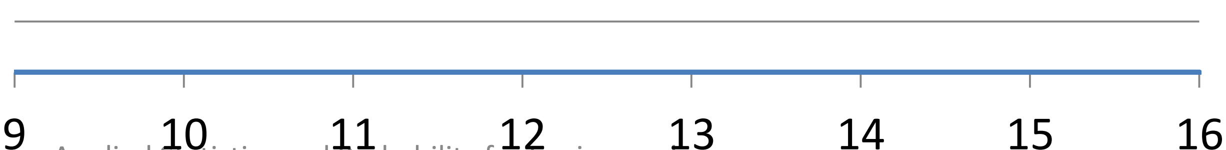
Example 2-6, Discrete Event Algebra

- Recall the sample space from Example 2-2, $S = \{yy, yn, ny, nn\}$ concerning conformance to specifications.
  - Let E_1 denote the event that at least one part <u>does</u> conform to specifications, $E_1 = \{yy, yn, ny\}$
  - Let E_2 denote the event that no part conforms to specifications, $E_2 = \{nn\}$
  - Let $E_3 = \emptyset$, the null or empty set.
  - Let $E_4 = S$, the universal set.
  - Let $E_5 = \{yn, ny, nn\}$, at least one part <u>does not</u> conform.
  - Then $E_1 \cup E_5 = S$
  - Then $E_1 \cap E_5 = \{yn, ny\}$
  - Then $E_1' = \{nn\}$

Example 2-7, Continuous Event Algebra

Measurements of the thickness of a part are modeled with the sample space: $S = R^+$.

- Let $E_1 = \{x | 10 \le x < 12\}$, show on the real line below.
- Let $E_2 = \{x | 11 < x < 15\}$
- Then $E_1 \cup E_2 = \{x | 10 \le x < 15\}$
- Then $E_1 \cap E_2 = \{x | 11 < x < 12\}$
- Then $E_1' = \{x | x < 10 \text{ or } x \ge 12\}$
- Then $E_1' \cap E_2 = \{x | 12 \ge x < 15\}$



Example 2-8, Hospital Emergency Visits

- This table summarizes the ER visits at 4 hospitals. People may leave without being seen by a physician(LWBS). The remaining people are seen, and may or may not be admitted.

	Hospital				
	1	2	3	4	Total
Total	5,292	6,991	5,640	4,329	22,252
LWBS	195	270	246	242	953
Admitted	1,277	1,558	666	984	4,485
Not admitted	3,820	5,163	4,728	3,103	16,814

Answers	
A $\cap$ B =	195
A' =	16,960
A $\cup$ B =	6,050

- Let A be the event of a visit to Hospital 1.
- Let B be the event that the visit is LWBS.
- Find number of outcomes in:
  - A ∩ B
  - A’
  - A ∪ B

Venn Diagrams Show Event Relations

Events A & B contain their respective outcomes. The shaded regions indicate the event relation of each diagram.

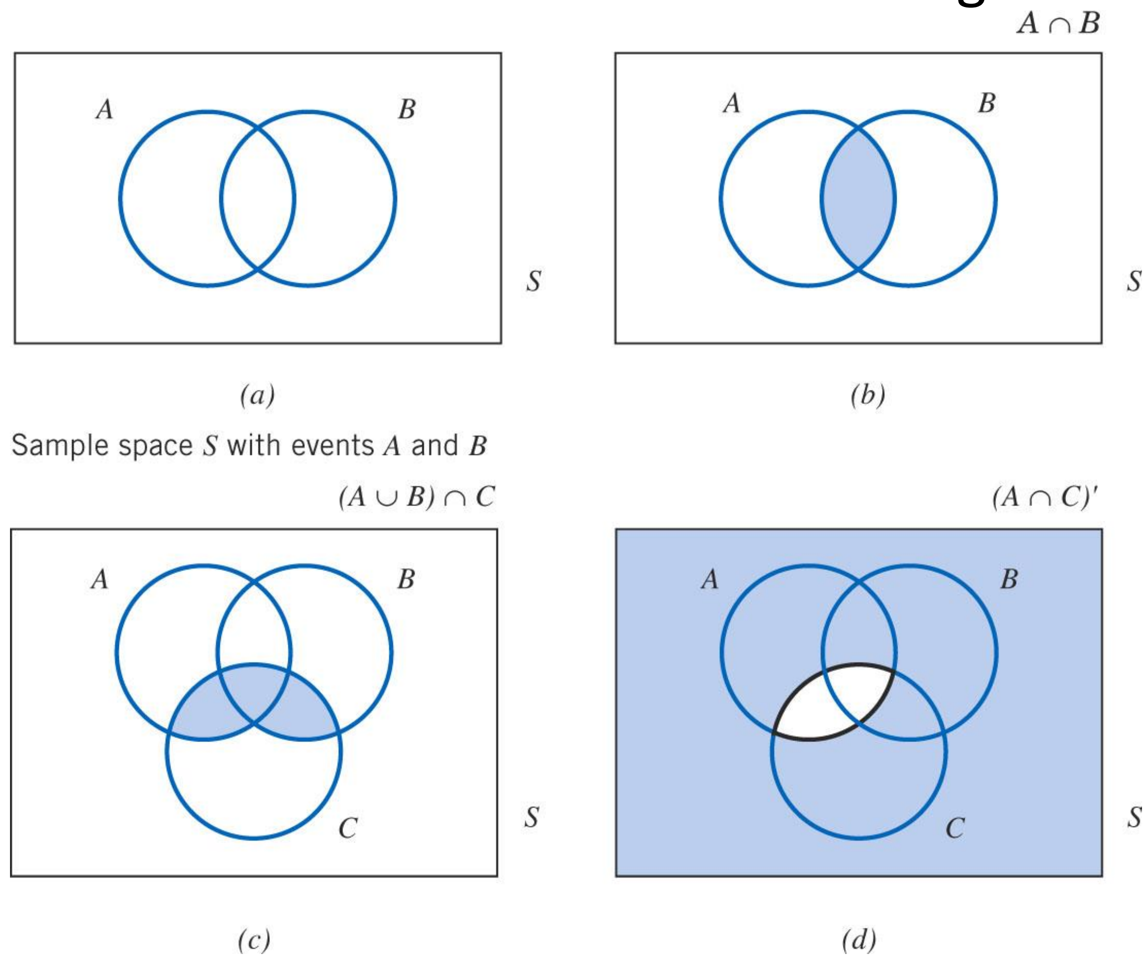


Figure 2-8 Venn diagrams

Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Venn Diagram of Mutually Exclusive Events

- Events A & B are mutually exclusive because they share no common outcomes.
- The occurrence of one event precludes the occurrence of the other.
- Symbolically, $A \cap B = \varnothing$

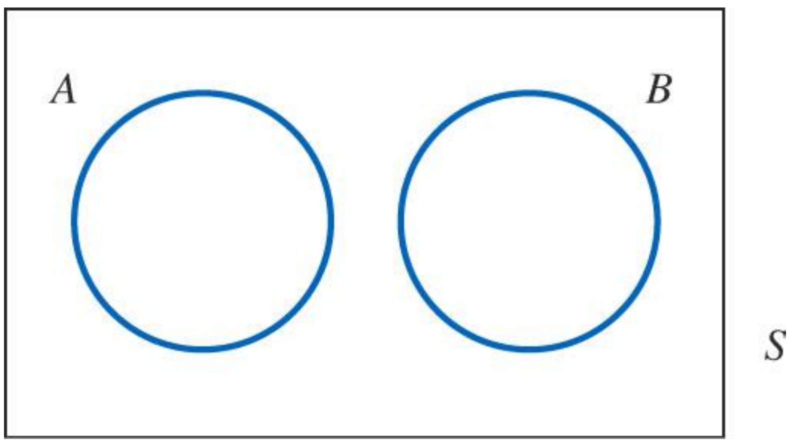


Figure 2-9 Mutually exclusive events

Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley &

Event Relation Laws

- Transitive law (event order is unimportant):
  - $A \cap B = B \cap A$ and $A \cup B = B \cup A$
- Distributive law (like in algebra):
  - $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
  - $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
- DeMorgan's laws:
  - $(A \cup B)' = A' \cap B'$ The complement of the union is the intersection of the complements.
  - $(A \cap B)' = A' \cup B'$ The complement of the intersection is the union of the complements.

Counting Techniques

- These are three special rules, or counting techniques, used to determine the number of outcomes in the events and the sample space.
- They are the:
  1. Multiplication rule
  2. Permutation rule
  3. Combination rule
- Each has its special purpose that must be applied properly – the right tool for the right job.

Counting – Multiplication Rule

- Multiplication rule:
  - Let an operation consist of k steps and
    - n_1 ways of completing step 1,
    - n_2 ways of completing step 2, … and
    - n_k ways of completing step k.
  - Then, the total number of ways or outcomes are:
    - $n_1 * n_2 * \ldots * n_k$

Example 2-9: Multiplication Rule

- In the design for a gear housing, we can choose to use among:
  - 4 different fasteners,
  - 3 different bolt lengths and
  - 2 different bolt locations.
- How many designs are possible?
- Answer: 4 * 3 * 2 = 24

Counting – Permutation Rule

- A permutation is a unique sequence of distinct items.
- If $S = \{a, b, c\}$, then there are 6 permutations
  - Namely: abc, acb, bac, bca, cab, cba (order matters)
  - The # of ways 3 people can be arranged.
- # of permutations for a set of n items is n!
- n! (factorial function) = n*(n-1)*(n-2)*…*2*1
- 7! = 7*6*5*4*3*2*1 = 5,040 = FACT(7) in Excel
- By definition: 0! = 1

Counting – Sub-set Permutations

- To sequence only r items from a set of n items:

$$P_r^n = n(n-1)(n-2)...(n-r+1) = \frac{n!}{(n-r)!}$$

$$P_3^7 = \frac{7!}{(7-3)!} = \frac{7!}{4!} = \frac{7*6*5*4!}{4!} = 7*6*5 = 210$$

In Excel: permut(7,3) = 210

Example 2-10: Circuit Board Designs

- A printed circuit board has eight different locations in which a component can be placed. If four different components are to be placed on the board , how many designs are possible?
- Answer: order is important, so use the permutation formula with n = 8, r = 4.

$$P_4^8 = \frac{8!}{(8-4)!} = \frac{8*7*6*5*4!}{4!} = 8*7*6*5 = 1,680$$

Counting - Similar Item Permutations

- Used for counting the sequences when not all the items are different.
- The number of permutations of:
  - $n = n_1 + n_2 + \ldots + n_r$ items of which
    - n_1 are identical,
    - n_2 are identical, … , and
    - n_r are identical.
- Is calculated as:

$$\frac{n!}{n_1!n_2! \ldots n_r!}$$

Example 2-11: Machine Shop Schedule

- In a machining operation, a piece of sheet metal needs two identical-diameter holes drilled and two identical-size notched cut. The drilling operation is denoted as d and the notching as n.
  - How many sequences are there?

$$\frac{4!}{2!2!} = \frac{4*3*2!}{2*1*2!} = 6$$

  - What is the set of sequences?

{ddnn, dndn, dnnd, nddn, ndnd, nndd}

Example 2-12: Bar Codes

- A part is labeled with 4 thick lines, 3 medium lines, and two thin lines. Each sequence is a different label.
  – How many unique labels can be created?

$$\frac{9!}{4!3!2!} = \frac{9*8*7*6*5*4!}{2*1*3*2*1*4!} = 1,260$$

  – In Excel:

1,260	=FACT(9) / (FACT(4)*FACT(3)*FACT(2))

Counting – Combination Rule

- A combination is a selection of r items from a set of n where order does not matter.
- If *S* = {*a, b, c*}, n =3, then there is 1 combination.
  - If r =3, there is 1 combination, namely: abc
  - If r=2, there are 3 combinations, namely ab, ac, bc
- # of permutations ≥ # of combinations

$$C_r^n = \frac{n!}{r!\left(n = r\right)!} \qquad (2\text{-}4)$$

Example 2-13: Applying the Combination Rule

- A circuit board has eight locations in which a component can be placed. If 5 identical components are to be placed on a board, how many different designs are possible?
- The order of the components is not important, so the combination rule is appropriate.

$$C_5^8 = \frac{8!}{5!(8-5)!} = \frac{8*7*6*5!}{3*2*1*5!} = 56$$

- Excel:

56	= COMBIN(8,5)

Example 2-14: Sampling w/o Replacement-1

- A bin of 50 parts contains 3 defectives & 47 good parts. A sample of 6 parts is selected from the 50 without replacement.
- How many different samples of size 6 are there that contain exactly 2 defective parts?

$$C_2^3 = \frac{3!}{2!1!} = 3 \text{ different ways}$$

- In Excel:

3	=COMBIN(3,2)

Example 2-14: Sampling w/o Replacement-2

- Now, how many ways are there of selecting 4 parts from the 47 acceptable parts?

$$C_4^{47} = \frac{47!}{4!43!} = \frac{47*46*45*44*43!}{4*3*2*1*43!} = 178{,}365 \text{ different ways}$$

- In Excel:

178,365	=COMBIN(47,4)

Example 2-14: Sampling w/o Replacement-3

- Now, how many ways are there to obtain:
  - 2 from the 3 defectives, and
  - 4 from the 47 non-defectives?

$$C_2^3 C_4^{47} = 3*178{,}365 = 535{,}095 \text{ different ways}$$

  - In Excel:

535,095	=COMBIN(3,2)*COMBIN(47,4)

Example 2-14: Sampling w/o Replacement-4

- Furthermore, how many ways are there to obtain 6 parts (0-6 defectives) from the set of 50?

$$C_6^{50} = \frac{50!}{6!*44!} = 15,890,700$$

- So the ratio of obtaining 2 defectives out 6 to any number (0-6) defectives out of 6 is:

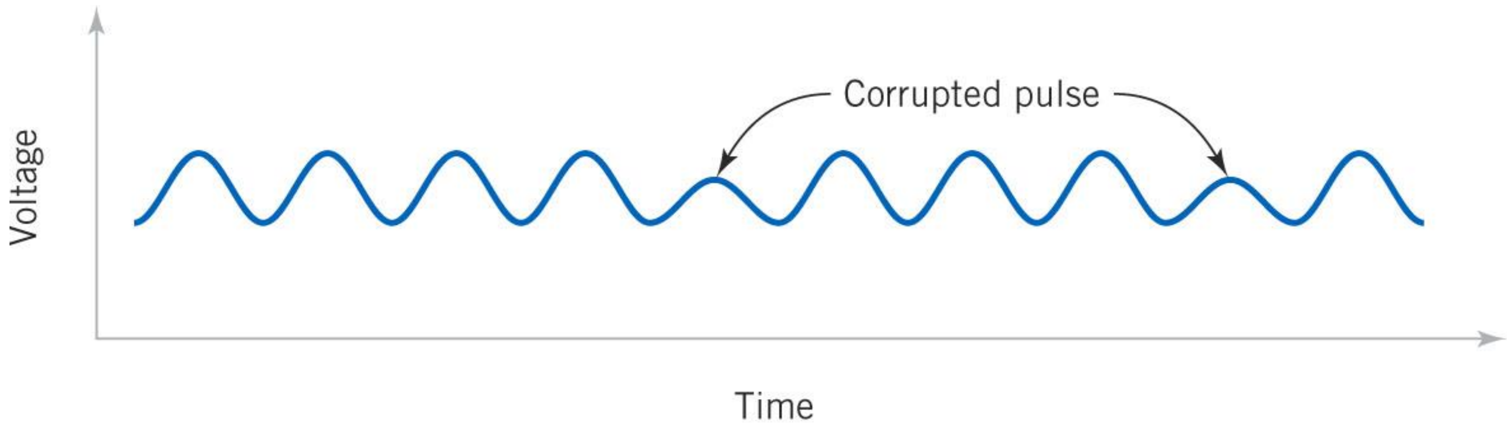
$$\frac{C_2^3 C_4^{47}}{C_6^{50}} = \frac{3*178,365}{15,890,700} = 0.034$$

What Is Probability?

- Probability is the likelihood or chance that a particular outcome or event from a random experiment will occur.
- Here, only finite sample spaces ideas apply.
- Probability is a number in the [0,1] interval.
- May be expressed as a:
  - proportion (0.15)
  - percent (15%)
  - fraction (3/20)
- A probability of:
  - 1 means certainty
  - 0 means impossibility

Types of Probability

- Subjective probability is a "degree of belief."
  - "There is a 50% chance that I'll study tonight."
- Relative frequency probability is based how often an event occurs over a very large sample space.



Relative frequency of corrupted pulse = $\frac{2}{10}$

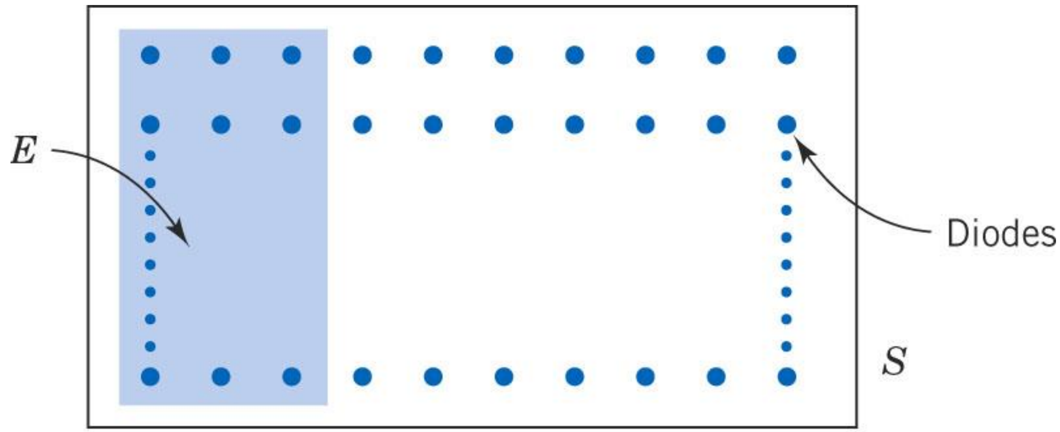
Figure 2-10 Relative frequency of corrupted pulses over a communications channel

Probability Based on Equally-Likely Outcomes

- Whenever a sample space consists of N possible outcomes that are equally likely, the probability of each outcome is $1/N$.
- Example: In a batch of 100 diodes, 1 is colored red. A diode is randomly selected from the batch. Random means each diode has an equal chance of being selected. The probability of choosing the red diode is 1/100 or 0.01, because each outcome in the sample space is equally likely.

Example 2-15: Laser Diodes

- Assume that 30% of the laser diodes in a batch of 100 meet a customer requirements.
- A diode is selected randomly. Each diode has an equal chance of being selected. The probability of selecting an acceptable diode is 0.30.



$$P(E) = 30(0.01) = 0.30$$

Figure 2-11 Probability of the event *E* is the sum of the probabilities of the outcomes in *E*.

Probability of an Event

- For a discrete sample space, the *probability of an event E*, denoted by *P(E),* equals the sum of the probabilities of the outcomes in *E*.
- The discrete sample space may be:
  - A finite set of outcomes
  - A countably infinite set of outcomes.
- Further explanation is necessary to describe probability with respect to continuous sample spaces.

Example 2-16: Probabilities of Events

- A random experiment has a sample space $\{w,x,y,z\}$. These outcomes are <u>not</u> equally-likely; their probabilities are: 0.1, 0.3, 0.5, 0.1.
- Event A $=\{w,x\}$, event B $= \{x,y,z\}$, event C $= \{z\}$
  - $P(A) = 0.1 + 0.3 = 0.4$
  - $P(B) = 0.3 + 0.5 + 0.1 = 0.9$
  - $P(C) = 0.1$
  - $P(A') = 0.6$ and $P(B') = 0.1$ and $P(C') = 0.9$
  - Since event $A \cap B = \{x\}$, then $P(A \cap B) = 0.3$
  - Since event $A \cup B = \{w,x,y,z\}$, then $P(A \cup B) = 1.0$
  - Since event $A \cap C = \{\text{null}\}$, then $P(A \cap C) = 0.0$

Example 2-17: Contamination Particles

- An inspection of a large number of semiconductor wafers revealed the data for this table. A wafer is selected randomly.

Number of Contamination Particles	Proportion of Wafers
0	0.40
1	0.20
2	0.15
3	0.10
4	0.05
5 or more	0.10
Total	1.00

- Let E be the event of selecting a 0 particle wafer. P(E) = 0.40
- Let E be the event of selecting a wafer with 3 or more particles. P(E) = 0.10+0.05+0.10 = 0.25

Example 2-18: Sampling w/o Replacement

- A batch of parts contains 6 parts {*a*,*b*,*c*,*d*,*e*,*f*}. Two are selected at random. Suppose part f is defective. What is the probability that part f appears in the sample?
- How many possible samples can be drawn?
  - Excel:

15	=COMBIN(6,2)

$$C_2^6 = \frac{6!}{2!4!} = 15$$

- How many samples contain part *f*?
  - 5 by enumeration: {*af*,*bf*,*cf*,*df*,*ef*}
- P(defective part) = 5/15 = 1/3.

Axioms of Probability

- Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties:
  1. $P(S) = 1$
  2. $0 \leq P(E) \leq 1$
  3. For each two events E_1 and E_2 with $E_1 \cap E_2 = \varnothing$,
     $$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$
- These imply that:
  - $P(\varnothing) = 0$ and $P(E') = 1 - P(E)$
  - If E_1 is contained in E_2, then $P(E_1) \leq P(E_2)$.

Addition Rules

- Joint events are generated by applying basic set operations to individual events, specifically:
  - Unions of events, $A \cup B$
  - Intersections of events, $A \cap B$
  - Complements of events, A'
- Probabilities of joint events can often be determined from the probabilities of the individual events that comprise it. And conversely.

Example 2-19: Semiconductor Wafers

A wafer is randomly selected from a batch as shown in the table.

- Let H be the event of high concentrations of contaminants. Then $P(H) = 358/940$.
- Let C be the event of the wafer being located at the center of a sputtering tool used in manufacture. Then $P(C) = 626/940$.
- $P(H \cap C) = 112/940$

Table 2-1	Location of Tool		
Contamination	Center	Edge	Total
Low	514	68	582
High	112	246	358
Total	626	314	940

- $P(H \cup C) = P(H) + P(C) - P(H \cap C) = (358+626-112)/940$ This is the addition rule.

- The probability of a union:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{2-5}$$

and, as rearranged:

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

- If events A and B are mutually exclusive:

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

therefore:

$$P(A \cup B) = P(A) + P(B) \tag{2-6}$$

Example 2-20: Contaminants & Location

Wafers in last example are now classified by degree of contamination per table of proportions.

- E_1 is the event that a wafer has 4 or more particles. $P(E_1) = 0.15$
- E_2 is the event that a wafer was on edge. $P(E_2) = 0.28$
- $P(E_1 \cap E_2) = 0.04$
- $P(E_1 \cup E_2)$
  $= 0.15 + 0.28 - 0.04$
  $= 0.39$

Number of	Table 2-2		
Contamination	Location of Tool		
Particles	Center	Edge	Total
0	0.30	0.10	0.40
1	0.15	0.05	0.20
2	0.10	0.05	0.15
3	0.06	0.04	0.10
4	0.04	0.01	0.05
5 or more	0.07	0.03	0.10
Totals	0.72	0.28	1.00

Addition Rule: 3 or More Events

$$P\left(A \cup B \cup C\right) = P\left(A\right) + P\left(B\right) + P\left(C\right)$$
$$-P\left(A \cap B\right) - P\left(A \cap C\right) - P\left(B \cap C\right)$$
$$+P\left(A \cap B \cap C\right) \qquad \text{(2-7)}$$

Note the alternating signs.

If a collection of events E_i is mutually exclusive,

thus for all pairs: $E_i \cap E_j = \varnothing$

Then: $P\left(E_1 \cup E_2 \cup \ \ldots \ \cup E_k\right) = \sum_{i=1}^{k} E_i$ (2-8)

Venn Diagram of Mutually Exclusive Events

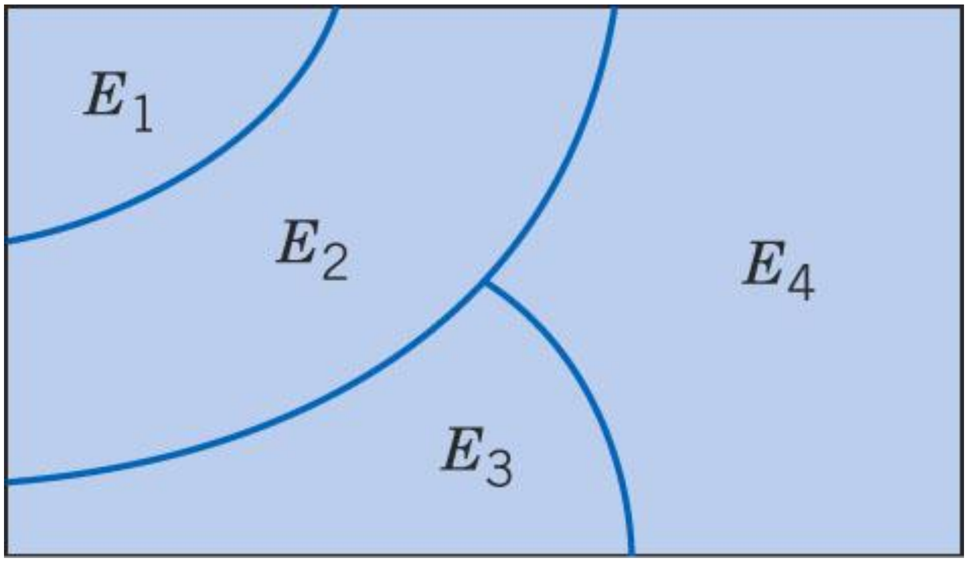


Figure 2-12 Venn diagram of four mutually exclusive events. Note that no outcomes are common to more than one event, i.e. all intersections are null.

Example 2-21: pH

- Let X denote the pH of a sample. Consider the event that $P(6.5 < X \leq 7.5) =$

$$P(6.5 < X \leq 7.0) + P(7.0 < X \leq 7.5) + P(7.5 < X \leq 7.8)$$

- The partition of an event into mutually exclusive subsets is widely used to allocate probabilities.

Conditional Probability

- Probabilities should be reevaluated as additional information becomes available.
- $P(B|A)$ is called the probability of event B occurring, given that event A has already occurred.
- A communications channel has an error rate of 1 per 1000 bits transmitted. Errors are rare, but do tend to occur in bursts. If a bit is in error, the probability that the next bit is also an error ought to be greater than 1/1000.

An Example of Conditional Probability

- In a thin film manufacturing process, the proportion of parts that are not acceptable is 2%. However the process is sensitive to contamination that can increase the rate of parts rejection.
- If we know that the plant is having filtration problems that increase film contamination, we would presume that the rejection rate has increased.

Another Example of Conditional Probability

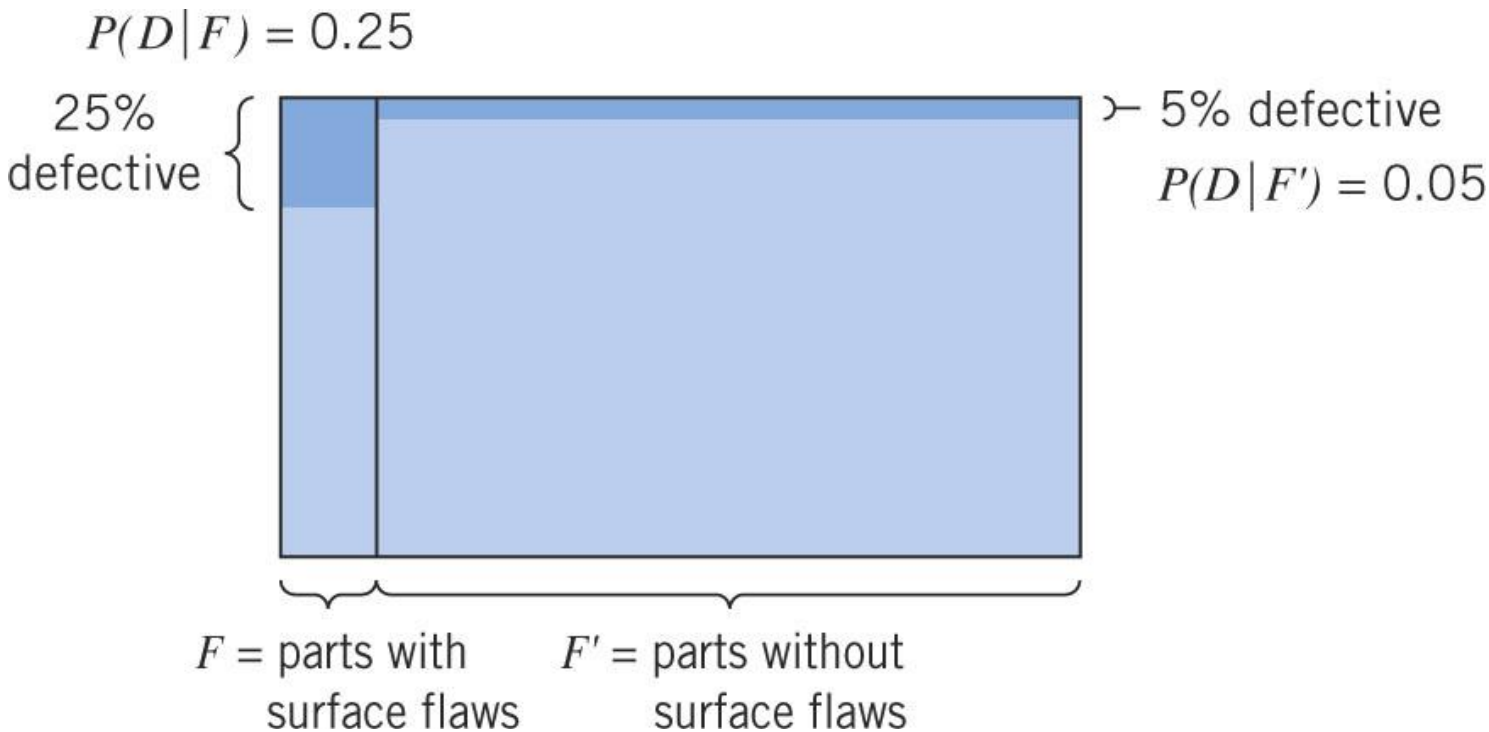


Figure 2-13 Conditional probability of rejection for parts with surface flaws and for parts without surface flaws. The probability of a defective part is not evenly distributed. Flawed parts are five times more likely to be defective than non-flawed parts, i.e., $P(D|F)$ / $P(D|F')$.

Example 2-22: A Sample From Prior Graphic

- Table 2-3 shows that 400 parts are classified by surface flaws and as functionally defective. Observe that:
  - $P(D|F) = 10/40 = 0.25$
  - $P(D|F') = 18/360 = 0.05$

Table 2-3 Parts Classified

	Surface Flaws		
Defective	Yes (F)	No (F')	Total
Yes (D)	10	18	28
No (D')	30	342	372
Total	40	360	400

Conditional Probability Rule

- The conditional probability of event B given event A, denoted as $P(B|A)$, is:

$$P(B|A) = P(A \cap B) \;/\; P(A) \qquad (2\text{-}9)$$

for $P(A) > 0$.

- From a relative frequency perspective of n equally likely outcomes:
  - $P(A)$ = (number of outcomes in A) / n
  - $P(A \cap B)$ = (number of outcomes in $A \cap B$) / n

Example 2-23: More Surface Flaws

Refer to Table 2-3 again. There are 4 probabilities conditioned on flaws.

Table 2-3 Parts Classified			
	Surface Flaws		
Defective	Yes (F)	No (F')	Total
Yes (D)	10	18	28
No (D')	30	342	372
Total	40	360	400

$$P(F) = 40/400 \text{ and } P(D) = 28/400$$

$$P(D \mid F) = P(D \cap F)/P(F) = \tfrac{10}{400} \Big/ \tfrac{40}{400} = \tfrac{10}{40}$$

$$P\left(D' \mid F\right) = P\left(D' \cap F\right)/P\left(F\right) = \tfrac{30}{400} \Big/ \tfrac{40}{400} = \tfrac{30}{40}$$

$$P\left(D \mid F'\right) = P\left(D \cap F'\right)/P\left(F'\right) = \tfrac{18}{400} \Big/ \tfrac{360}{400} = \tfrac{18}{360}$$

$$P\left(D' \mid F'\right) = P\left(D' \cap F'\right)/P\left(F'\right) = \tfrac{342}{400} \Big/ \tfrac{360}{400} = \tfrac{342}{360}$$

Example 2-23: Tree Diagram

Tree illustrates sampling two parts without replacement:

- At the 1st stage (flaw), every original part of the 400 is equally likely.
- At the 2nd stage (defect), the probability is conditional upon the part drawn in the prior stage.

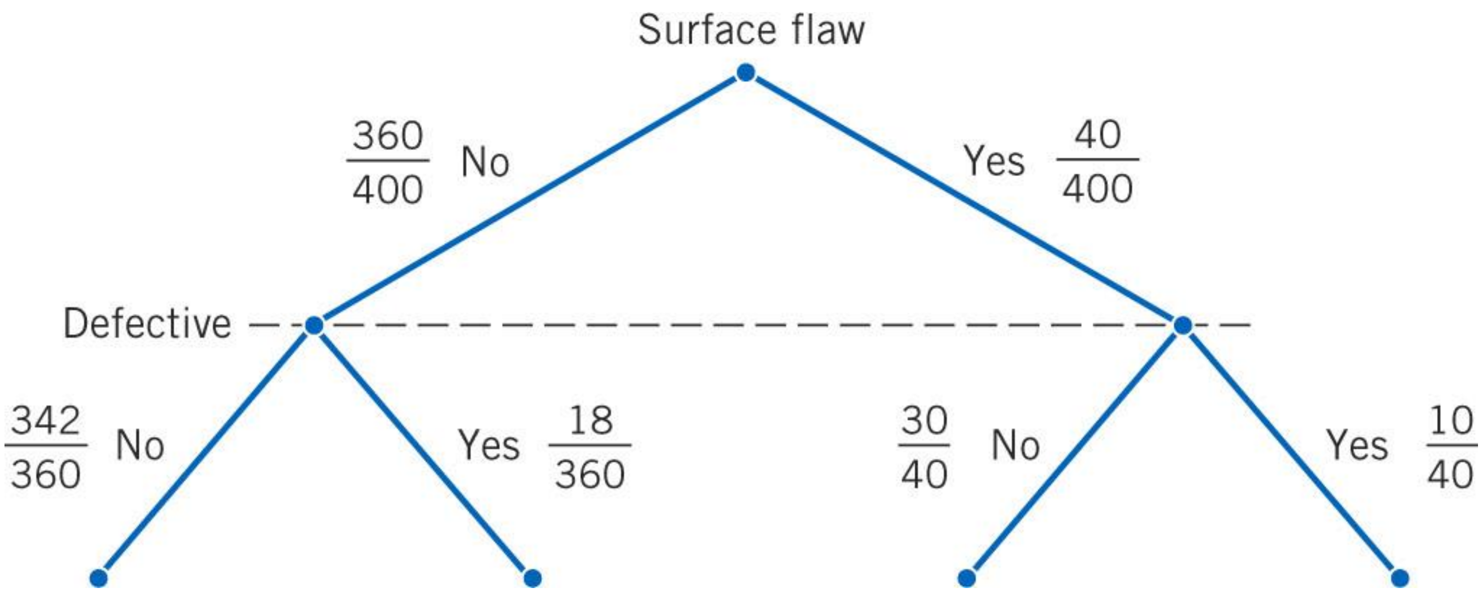


Figure 2-14 Tree diagram for parts classification

Random Samples & Conditional Probabilities

- Random means each item is equally likely to be chosen. If more than one item is sampled, random means that every sampling outcome is equally likely.
- 2 items are taken from $S = \{a,b,c\}$ without replacement.
- Ordered sample space: $S = \{ab,ac,bc,ba,bc,cb\}$
- Unordered sample space: $S = \{ab,ac,bc\}$
- This is done by enumeration – too hard ☹

Sampling Without Enumeration

- Use conditional probability to avoid enumeration. To illustrate: A batch of 50 parts contains 10 made by Tool 1 and 40 made by Tool 2. We take a sample of n=2.
- What is the probability that the 2^{nd} part came from Tool 2, given that the 1^{st} part came from Tool 1?
  - P(1^{st} part came from Tool 1) = 10/50
  - P(2^{nd} part came from Tool 2) = 40/49
  - P(Tool 1, then Tool 2 part sequence) = (10/50)*(40/49)
- To select randomly implies that, at each step of the sample, the items remaining in the batch are equally likely to be selected.

Example 2-24: Sampling Without Replacement

- A production lot of 850 parts contains 50 defectives. Two parts are selected at random.
- What is the probability that the 2nd is defective, given that the first part is defective?
- Let *A* denote the event that the 1st part selected is defective.
- Let *B* denote the event that the 2nd part selected is defective.
- Probability desired is $P(B|A) = 49/849$.

Example 2-25: Continuing Prior Example

- Now, 3 parts are sampled randomly.
- What is the probability that the first two are defective, while the third is not?

$$P\left(ddn\right)=\frac{50}{850}*\frac{49}{849}*\frac{800}{848}=0.0032$$

- In Excel:

0.0032	=(50*49*800)/(850*849*848)

Multiplication Rule

- The conditional probability definition of Equation 2-9 can be rewritten to generalize it as the multiplication rule.

- $P(A \cap B) = P(B|A)*P(A) = P(A|B)*P(B) \qquad (2\text{-}10)$

- The last expression is obtained by exchanging the roles of *A* and *B*.

Example 2-26: Machining Stages

- The probability that, a part made in the 1st stage of a machining operation passes inspection, is 0.90. The probability that, it passes inspection after the 2nd stage, is 0.95.
- What is the probability that the part meets specifications?
- Let A & B denote the events that the 1st & 2nd stages meet specs.
- $P(A \cap B) = P(B|A) * P(A) = 0.95 * 0.90 = 0.955$

Two Mutually Exclusive Subsets

- A & A' are mutually exclusive.
- $A \cap B$ and $A' \cap B$ are mutually exclusive
- $B = (A \cap B) \cup (A \cap B)$

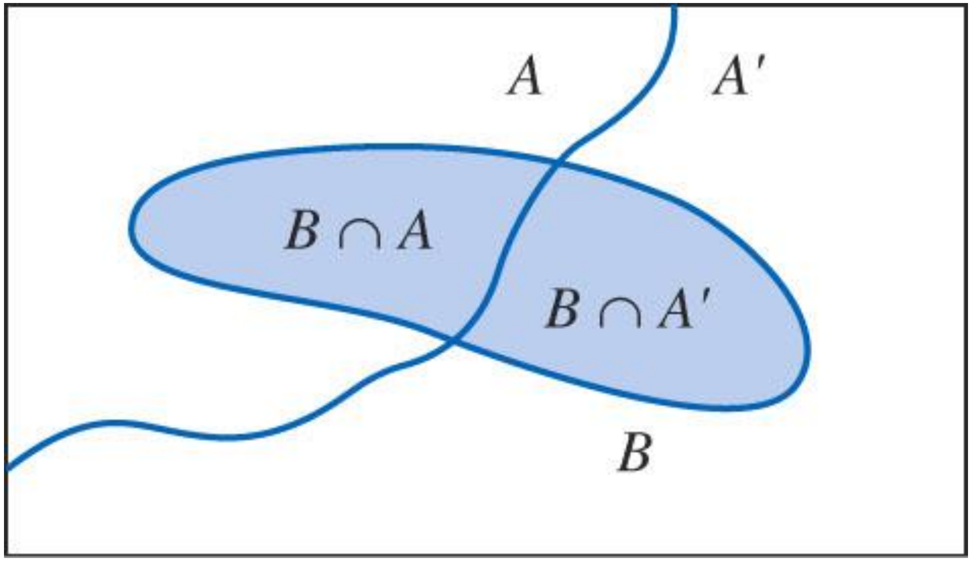


Figure 2-15 Partitioning an event into two mutually exclusive subsets.

Total Probability Rule

For any two events *A* and *B*:

$$
\begin{aligned}
P(B) &= P(B \cap A) + P(B \cap A') \\
&= P(B \mid A) * P(A) + P(B \mid A') * P(A') \qquad (2\text{-}11)
\end{aligned}
$$

Example 2-27: Semiconductor Contamination

- Information about product failure based on chip manufacturing process contamination.

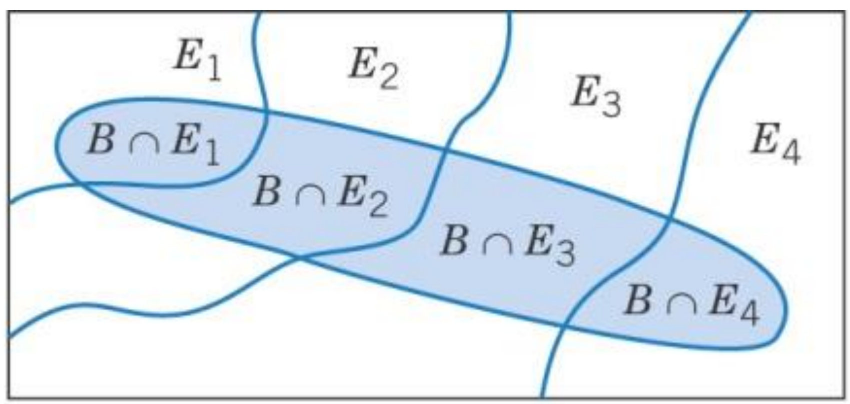
Probability of Failure	Level of Contabination	Probability of Level
0.100	High	0.2
0.005	Not High	0.8

  - F denotes the event that the product fails.
  - H denotes the event that the chip is exposed to high contamination during manufacture.
  - $P(F|H) = 0.100$ & $P(H) = 0.2$, so $P(F \cap H) = 0.02$
  - $P(F|H') = 0.005$ and $P(H') = 0.8$, so $P(F \cap H') = 0.004$
  - $P(F) = P(F \cap H) + P(F \cap H') = 0.020 + 0.004 = 0.024$

Total Probability Rule (multiple events)

- Assume E_1, E_2, … E_k are k mutually exclusive & exhaustive subsets. Then:

$$P(B) = P(B \cap E_1) + P(B \cap E_2) + \ \dots \ + P(B \cap E_k)$$
$$= P(B \mid E_1) * P(E_1) + P(B \mid E_2) * P(E_2) + \ \dots \ + P(B \mid E_k) * P(E_k) \quad \text{(2-11)}$$



$$B = (B \cap E_1) \cup (B \cap E_2) \cup (B \cap E_3) \cup (B \cap E_4)$$

Figure 2-16 Partitioning an event into several mutually exclusive subsets.

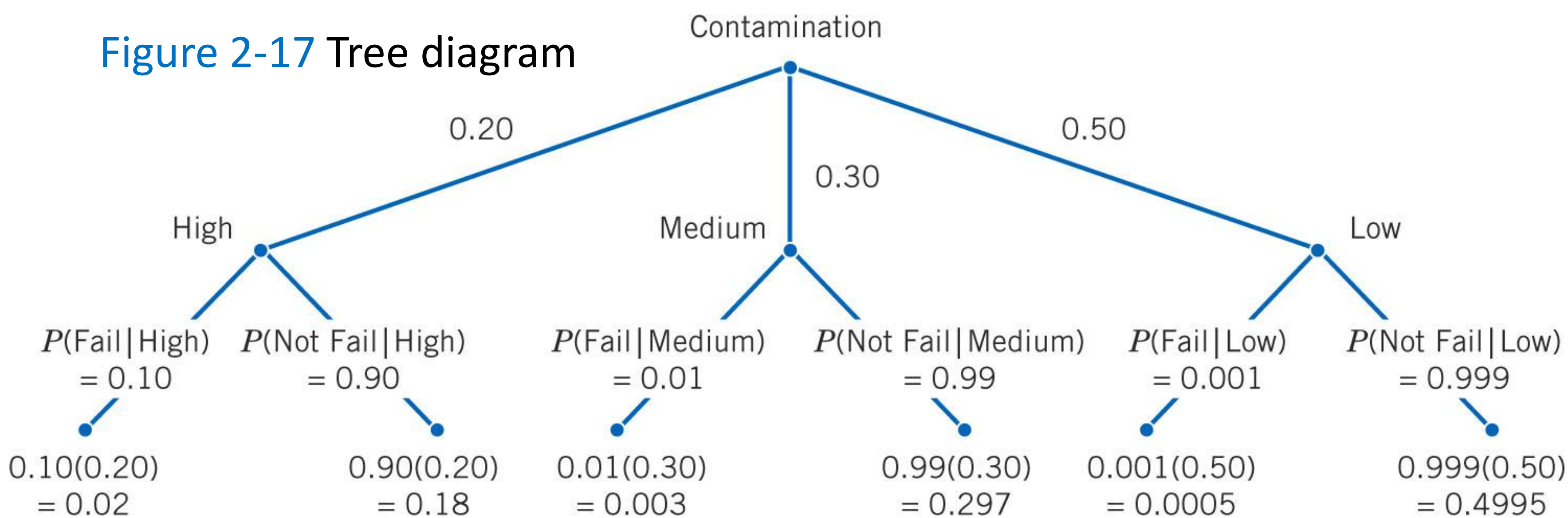
Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Example 2-28: Refined Contamination Data

Continuing the discussion of contamination during chip manufacture:

Probability of Failure	Level of Contabination	Probability of Level
0.100	High	0.2
0.010	Medium	0.3
0.001	Low	0.5

Figure 2-17 Tree diagram



$$P(\text{Fail}) = 0.02 + 0.003 + 0.0005 = 0.0235$$

Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Event Independence

- Two events are independent if any one of the following equivalent statements are true:
    1. $P(B|A) = P(A)$
    2. $P(A|B) = P(B)$
    3. $P(A \cap B) = P(A)*P(B)$
- This means that occurrence of one event has no impact on the occurrence of the other event.

Example 2-29: Sampling With Replacement

- A production lot of 850 parts contains 50 defectives. Two parts are selected at random, but the first is replaced before selecting the 2nd.
- Let A denote the event that the 1st part selected is defective. $P(A) = 50/850$
- Let B denote the event that the 2nd part selected is defective. $P(B) = 50/850$
- What is the probability that the 2nd is defective, given that the first part is defective? The same.
- Probability that both are defective is:

  $P(A)*P(B) = 50/850 *50/850 = 0.0035.$

Example 2-30: Flaw & Functions

The data shows whether the events are independent.

Table 2-3 Parts Classified				Table 2-4 Parts Classified (data chg'd)			
	Surface Flaws				Surface Flaws		
Defective	Yes (F)	No (F')	Total	Defective	Yes (F)	No (F')	Total
Yes (D)	10	18	28	Yes (D)	2	18	20
No (D')	30	342	372	No (D')	38	342	380
Total	40	360	400	Total	40	360	400
	P(D|F) =	10/40 =	0.25		P(D|F) =	2/40 =	0.05
	P(F) =	40/400 =	0.10		P(F) =	20/400 =	0.05
			not same				same
	Events D & F are **dependent**				Events D & F are **independent**		

Example 2.31: Conditioned vs. Unconditioned

- A production lot of 850 parts contains 50 defectives. Two parts are selected at random, without replacement.
- Let A denote the event that the 1st part selected is defective. $P(A) = 50/850$
- Let B denote the event that the 2nd part selected is defective. $P(B|A) = 49/849$
- Probability that the 2nd is defective is:

$$P(B) = P(B|A)*P(A) + P(B|A')*P(A')$$
$$P(B) = (49/849)*(50/850) + (50/849)*(800/850)$$
$$P(B) = (49*50+50*800) / (849*850)$$
$$P(B) = 50*(49+800) / (849*850)$$

$P(B) = 50/850$ is unconditional, same as $P(A)$

- Since $P(B|A) \neq P(A)$, then A and B are dependent.

Independence with Multiple Events

The events $E_1, E_2, \ldots, E_k$ are independent if and only if, for any subset of these events:

$$P(E_1 \cap E_2 \cap \ldots, \cap E_k) = P(E_1)^* \, P(E_2)^* \ldots ^* \, P(E_k) \quad (2\text{-}14)$$

Be aware that, if E_1 & E_2 are independent, E_2 & E_3 may or may not be independent.

Example 2-32: Series Circuit

This circuit operates only if there is a path of functional devices from left to right. The probability that each device functions is shown on the graph. Assume that the devices fail independently. What is the probability that the circuit operates?

Let L & R denote the events that the left and right devices operate. The probability that the circuit operates is:

$$P(L \cap R) = P(L) * P(R) = 0.8 * 0.9 = 0.72.$$

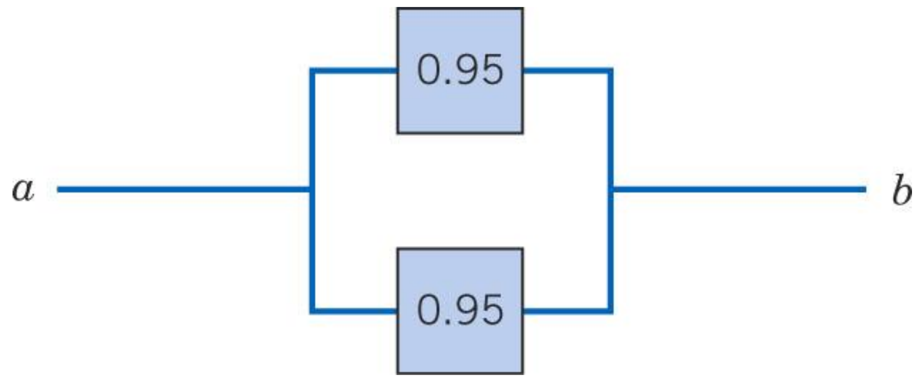
Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Example 2-33: Another Series Circuit

- The probability that a wafer contains a large particle of contamination is 0.01. The wafer events are independent.
- $P(E_i)$ denotes the event that the i^{th} wafer contain no particles and $P(E_i) = 0.99$.
- If 15 wafers are analyzed, what is the probability that no large particles are found?
- $P(E_1 \cap E_2 \cap \ldots \cap E_k) = P(E_1) * P(E_2) * \ldots * P(E_k)$
  $= (0.99)^{15} = 0.86.$

Example 2-34: Parallel Circuit

This circuit operates only if there is a path of functional devices from left to right. The probability that each device functions is shown. Each device fails independently.



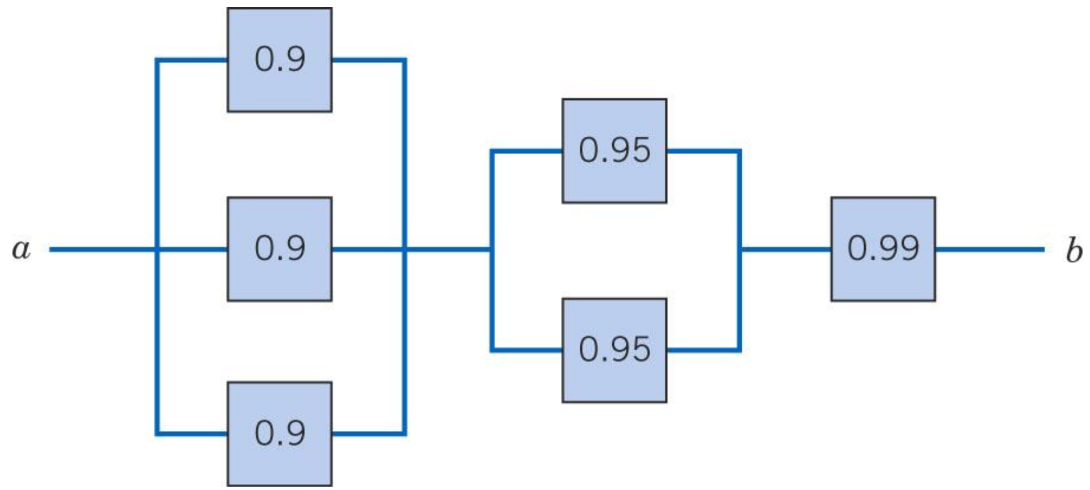
Let T & B denote the events that the top and bottom devices operate. The probability that the circuit operates is:

$P(T \cup B) = 1 - P(T' \cap B') = 1 - P(T')*P(B') = 1 – 0.05^2 = 1 – 0.0025 – 0.9975.$

(this is 1 minus the probability that they both don't fail)

Example 2-35: Advanced Circuit

This circuit operates only if there is a path of functional devices from left to right. The probability that each device functions is shown. Each device fails independently.



Partition the graph into 3 columns with L & M denoting the left & middle columns.

P(L) = 1- 0.1^3 , and P(M) = 1- 0.5^2, so the probability that the circuit operates is: $(1 – 0.1^3)(1-0.05^2)(0.99) = 0.9875$ (this is a series of parallel circuits). In Excel:

0.98752	= (1-0.01^3)*(1-0.05^2)*0.99

Bayes Theorem

- Thomas Bayes (1702-1761) was an English mathematician and Presbyterian minister.
- His idea is that we observe conditional probabilities through prior information.
- The short formal statement is:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \quad \text{for } P(B) > 0 \qquad (2\text{-}15)$$

- Note the reversal of the condition!

Example 2-36:

- From Example 2-27, find *P*(*F*) which is not given:

Probability of Failure	Level of Contabination	Probability of Level
0.100	High	0.2
0.005	Not High	0.8

$$P(H \mid F) = \frac{P(B \mid A) * P(A)}{P(F)} = \frac{0.10 * 0.20}{0.024} = 0.83$$

$$P(F) = P(F \mid H) * P(H) + P(F \mid H') * P(H')$$

$$= 0.1 * .2 + 0.005 * 0.8 = 0.024$$

Bayes Theorem with Total Probability

- If E_1, E_2, … E_k are k mutually exclusive and exhaustive events and B is any event,

$$P(E_1 \mid B) = \frac{P(B \mid E_1)P(E_1)}{P(B \mid E_1)P(E_1) + P(B \mid E_2)P(E_2) + \ldots + P(B \mid E_k)P(E_k)}$$

for $P(B) > 0$ (2-16)

- Note that the:
  - Total probability expression of the denominator
  - Numerator is always one term of the denominator.

Example 2-37: Medical Diagnostic-1

Because a new medical procedure has been shown to be effective in the early detection of a disease, a medical screening of the population is proposed. The probability that the test correctly identifies someone with the disease as positive is 0.99, and probability that the test correctly identifies someone without the disease as negative is 0.95. The incidence of the illness in the general population is 0.0001. You take the test and the result is positive. What is the probability that you have the illness?

Let D denote the event that you have the disease and let S denote the event that the test signals positive. Given info is:

- $P(S'|D') = 0.95$, so $P(S|D') = 0.05$, and $P(D) = 0.0001$,
- $P(S|D) = 0.99$. We desire $P(D|S)$.

Example 2-37: Medical Diagnostic-2

$$P(D \mid S) = \frac{P(S \mid D) * P(D)}{P(S \mid D) * P(D) + P(S \mid D') * P(D')}$$

$$= \frac{0.99 * 0.0001}{0.99 * 0.0001 + 0.05 * (1 - 0.0001)}$$

$$= 1/506 = 0.002$$

Excel:

0.00198	=(0.99*0.0001) / (0.99*0.0001+0.05*(1-0.0001))

Example 2-37: Medical Diagnostic-2

$$P(D|S) = \frac{P(S|D)*P(D)}{P(S|D)*P(D)+P(S|D')*P(D')}$$

$$= \frac{0.99*0.0001}{0.99*0.0001+0.05*(1-0.0001)}$$

$$= 1/506 = 0.002$$

Excel:	0.00198	=(0.99*0.0001) / (0.99*0.0001+0.05*(1-0.0001))

Before the test, your chance was 0.0001. After the positive result, your chance is 0.00198. So your risk of having the disease has increased 20 times = 0.00198/0.00010, but is still tiny.

Example 2-38: Bayesian Network-1

- Bayesian networks are used on Web sites of high-tech manufacturers to allow customers to quickly diagnose problems with products. A printer manufacturer obtained the following probabilities from its database. Printer failures are of 3 types: hardware $P(H) = 0.3$, software $P(S)=0.6$, and other $P(O)=0.1$. Also:
  - $P(F|H) = 0.9$, $P(F|S) = 0.2$, $P(F|O) = 0.5$.
- Find the max of $P(H|F)$, $P(S|F)$, $P(O|F)$ to direct the diagnostic effort.

Example 2-38: Bayesian Network-2

$$P(F) = P(F \mid H)P(H) + P(F \mid S)P(S) + P(F \mid O)P(O)$$
$$= 0.9(0.1) + 0.2(0.6) + 0.5(0.3) = 0.36$$

$$P(H \mid F) = \frac{P(F \mid H) * P(H)}{P(F)} = \frac{0.9 * 0.1}{0.36} = 0.250$$

$$P(S \mid F) = \frac{P(F \mid S) * P(S)}{P(F)} = \frac{0.2 * 0.6}{0.36} = 0.333$$

$$P(O \mid F) = \frac{P(F \mid O) * P(O)}{P(F)} = \frac{0.5 * 0.3}{0.36} = 0.417$$

Note that the conditionals based on Failure add to 1. Since the Other category is the most likely cause of the failure, diagnostic effort should be so initially directed.

Random Variables

- A variable that associates a number with the outcome of a random experiment is called a random variable.
- A random variable is a function that assigns a real number to each outcome in the sample space of a random experiment.
- Particular notation is used to distinguish the random variable (rv) from the real number. The rv is denoted by an uppercase letter, such as X. After the experiment is conducted, the measured value is denoted by a lowercase letter, such a x = 70. X and x are shown in italics, e.g., $P(X=x)$.

Continuous & Discrete Random Variables

- A discrete random variable is a rv with a finite (or countably infinite) range. They are usually integer counts, e.g., number of errors or number of bit errors per 100,000 transmitted (rate). The ends of the range of rv values may be finite ($0 \le x \le 5$) or infinite ($x \ge 0$).
- A continuous random variable is a rv with an interval (either finite or infinite) of real numbers for its range. Its precision depends on the measuring instrument.

Examples of Discrete & Continuous RVs

- Discrete rv's:
  - Number of scratches on a surface.
  - Proportion of defective parts among 100 tested.
  - Number of transmitted bits received in error.
  - Number of common stock shares traded per day.
- Continuous rv's:
  - Electrical current and voltage.
  - Physical measurements, e.g., length, weight, time, temperature, pressure.