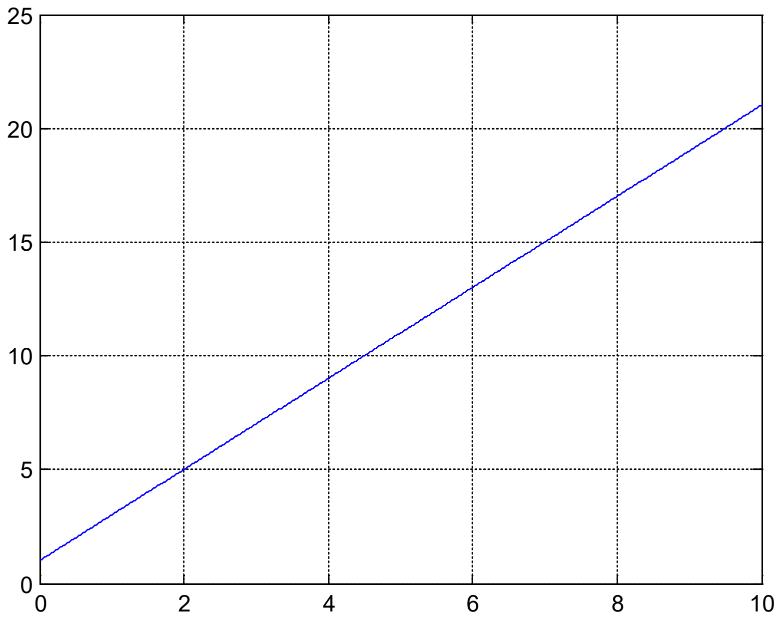
Introduction to Probability

Why we should study the probability?

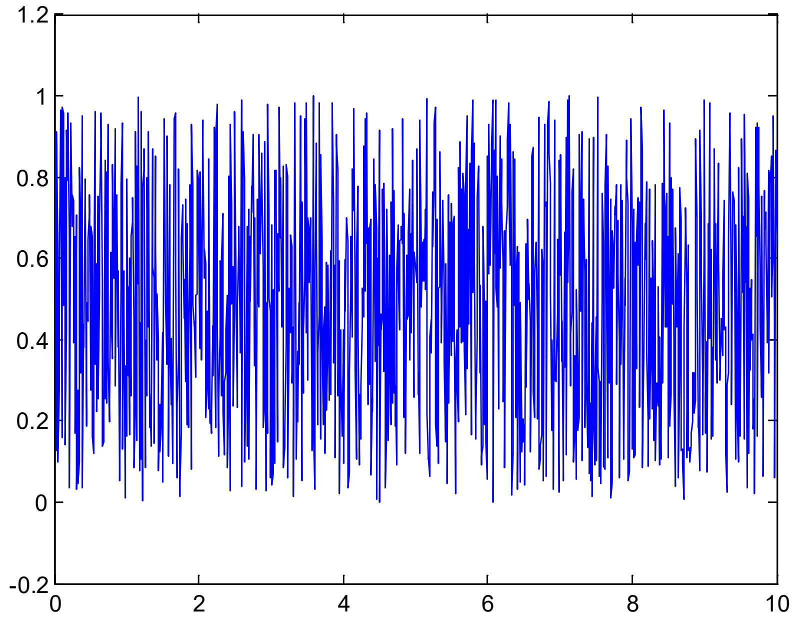
Our Interest and Goals

- Study tools to characterize uncertainty
  - Probability theory, random events, random variables
- Study related tools to characterize non-deterministic signals
  - Random processes, statistics
- Analyze systems with non-deterministic inputs and outputs.
  - Linear systems with random inputs
  - Communication channels with noise
  - Communication networks with uncertain delays

Random versus Deterministic



Y=2X+1



Random Noise

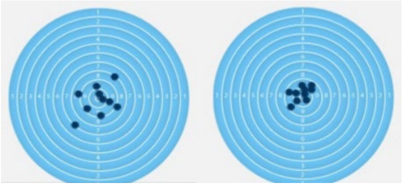
Randomness around us

- At bus stop
  - Waiting time for bus
  - The number of people to take a bus

- At a bank
  - Waiting time for service
  - Service time
  - The number of customers to wait for

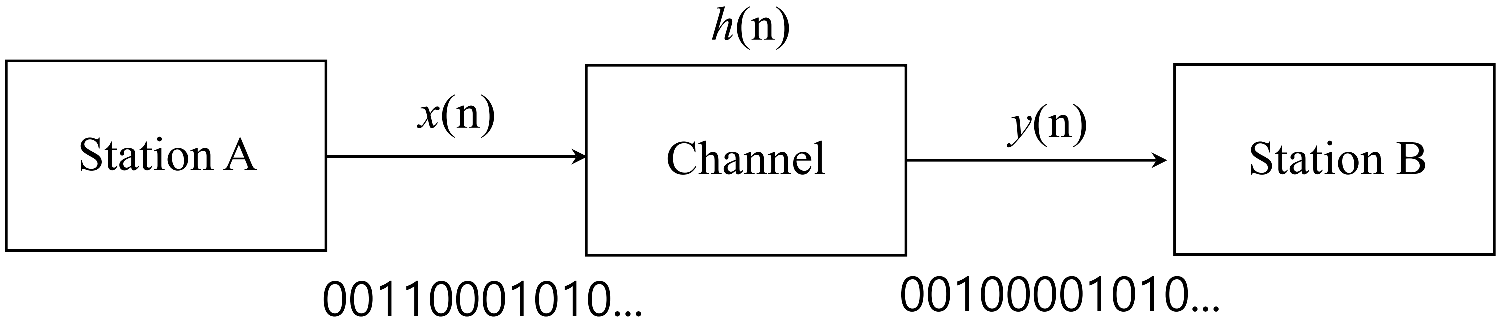
- Bet/Game
  - Toss a coin: Head or Tail
  - Roll a dice: 1, 2, 3, 4, 5, 6

- Shooting game



Typical Electrical Engineering Problems (1/3)

- A noisy binary communication channel
  - The channel can be twisted pair, coaxial cable, fiber optic cable, or wireless medium.
  - The channel introduces noise and thereby bit errors.



Typical Electrical Engineering Problems (2/3)

- Desired target signal is buried in noise.

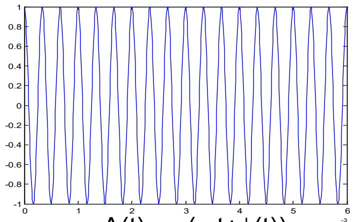
$$x(t) = A(t)\cos(\omega t + \varphi(t)) + n(t)$$

signal signal noise

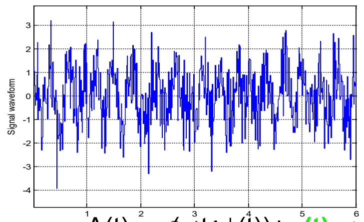
- Determine the presence or absence of the desired signal.

- Filter the signal out of noise.

- Demodulate the signal.



A(t)cos(ωt+ϕ(t))



A(t)cos(ωt+ϕ(t))+n(t)

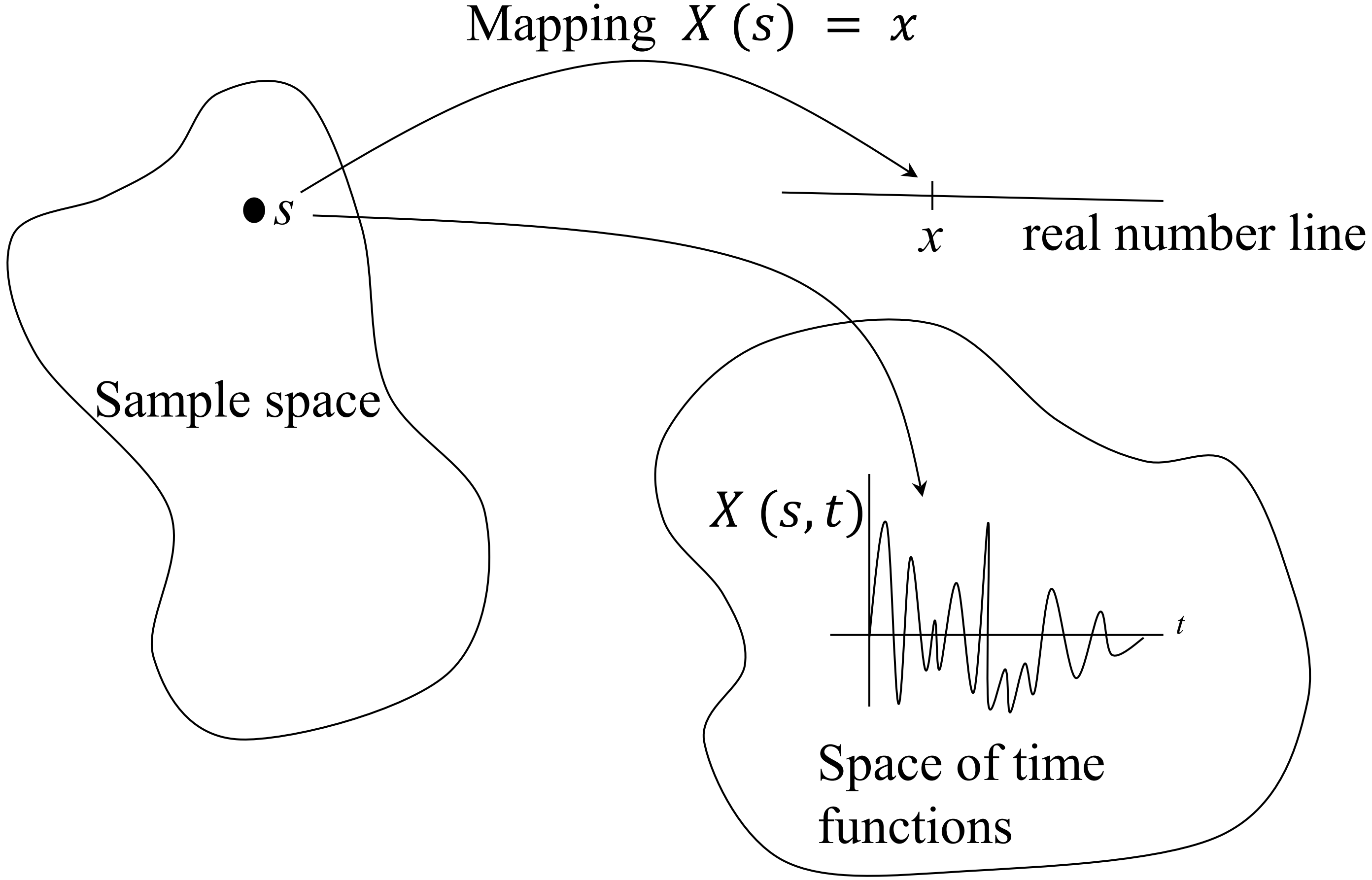
Typical Electrical Engineering Problems (3/3)

- In large computer networks, there are limited resources (e.g., bandwidth, routers, switches, printers and other devices) that need to be shared by the users.
  - User jobs/packets are queued and assigned service based on predefined criteria.
  - Demand is uncertain and service time is also uncertain.
  - Delay from the time the service is requested to the time it is completed is uncertain.

- Similar considerations exist for telephone networks, multiuser computer networks, and other communication networks.

Random Variables and Random Processes

- We will study them details later

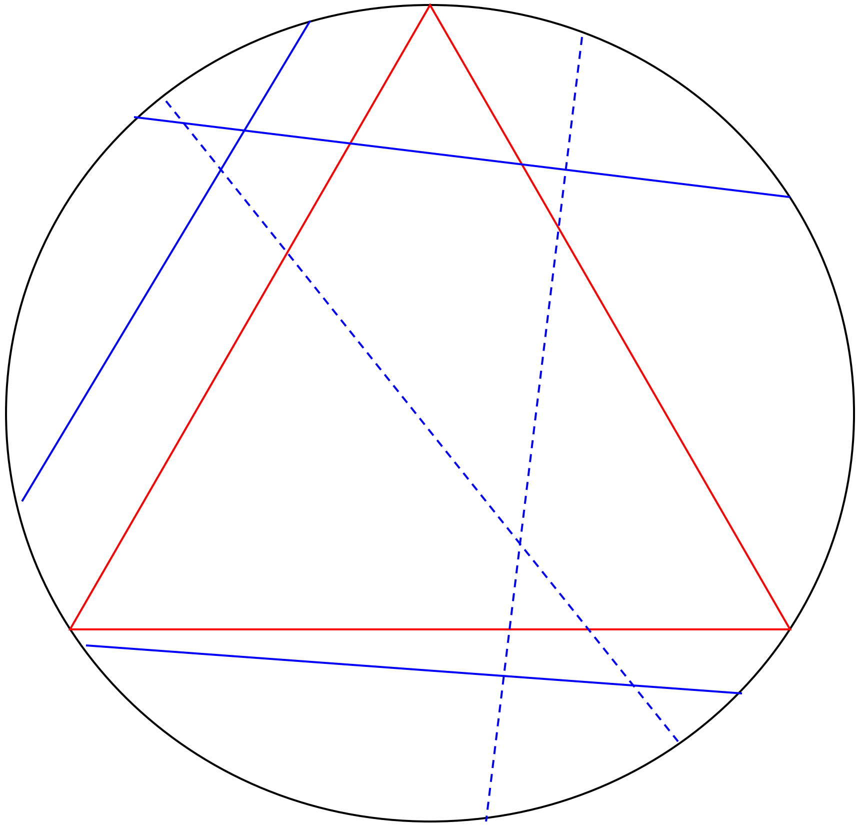


Why is the understanding of randomness important?

Bertrand Paradox

•Consider an equilateral triangle inscribed in a circle. Suppose a chord of the circle is chosen at random. What is the probability that the chord is longer than a side of the triangle?

1. A third
2. A half
3. A quarter



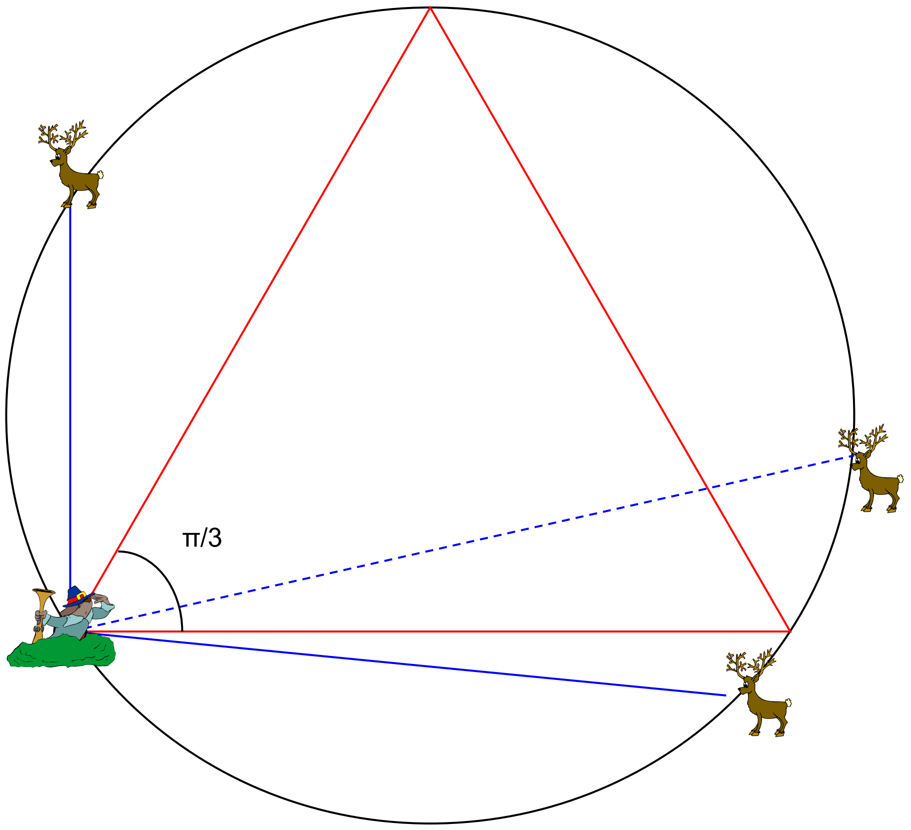
Solution 1: Bertrand Paradox

•Choose a point on the circumference and rotate the triangle so that the point is at one vertex.

•Choose another point on the circle and draw the chord joining it to the first point.

•For points on the arc between the endpoints of the side opposite the first point, the chord is longer than a side of the triangle.

•The length of the arc is one third of the circumference of the circle, therefore the probability a random chord is longer than a side of the inscribed triangle is one third.



Solution 2: Bertrand Paradox

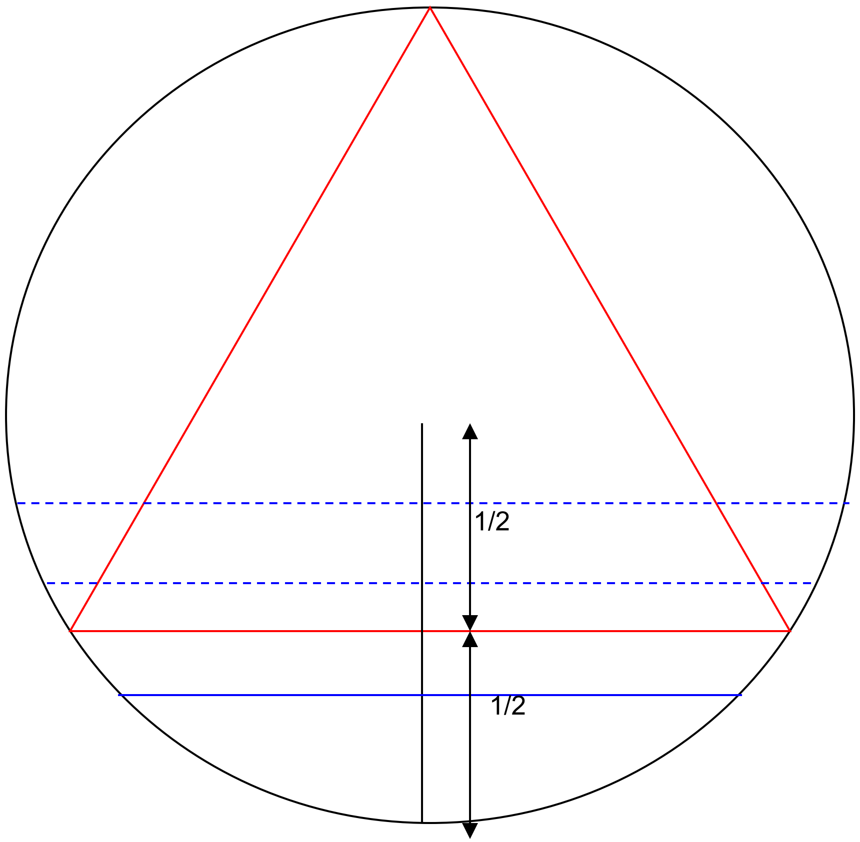
- Choose a radius of the circle and rotate the triangle so a side is perpendicular to the radius.

- Choose a point on the radius and construct the chord whose midpoint is the chosen point.

- The chord is longer than a side of the triangle if the chosen point is nearer the center of the circle than the point where the side of the triangle intersects the radius.

- Since the side of the triangle bisects the radius, it is equally probable that the chosen point is nearer or farther.

- Therefore the probability a random chord is longer than a side of the inscribed triangle is one half.

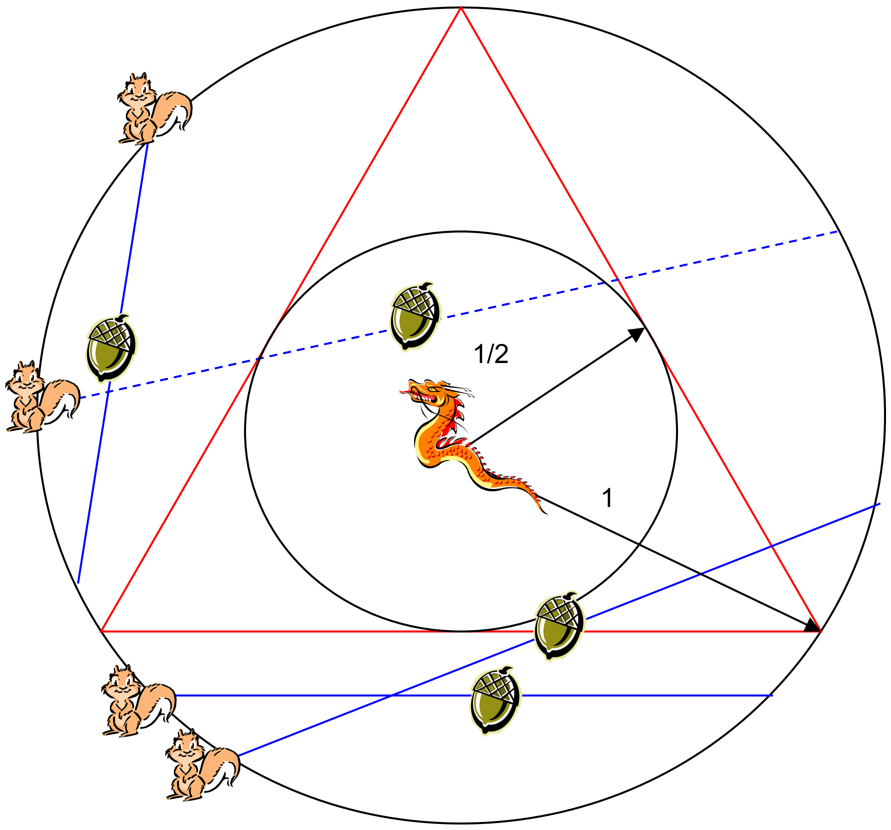


Solution 3: Bertrand Paradox

- Choose a point anywhere within the circle and construct a chord with the chosen point as its midpoint.

- The chord is longer than a side of the inscribed triangle if the chosen point falls within a concentric circle of radius 1/2.

- The area of the smaller circle is one fourth the area of the larger circle, therefore the probability a random chord is longer than a side of the inscribed triangle is one fourth.



Probability Space

- A probability space $(S, F(S), P)$ is a triple made up of three elements
  1. S: Sample space
  2. $F(S)$: A collection of sets S, event space
  3. The probabilities of $P(E)$ for each $E \in F(S)$
     $P(.)$: $F(S) \rightarrow [0, 1]$

Sample Space $(S, F(S), P)$

- Experimental outcomes are unpredictable

- The set of all possible outcomes is known

- The sample space S of an experiment is defined as the set of all possible outcomes of the experiment.
  - The flipping of a coin: $S = \{H, T\}$
  - The rolling of a die: $S = \{1,2,3,4,5,6\}$
  - The flipping of two coins: $S = \{(H,H),(H,T),(T,H),(T,T)\}$
  - The lifetime of a car: $S = [0, \infty)$

- An Event is any subset E of the sample space S $(E \subset S \; or \; E \in F(S))$
  - $E = \{H\}, E = \{1\}, E = \{1,3,5\}\; E = \{(H,H),(T,T)\}$

Basic Set Theory Definitions

- The set containing all possible elements of interest is called the universe, universal set or space S.
- The set containing no elements is called the empty set or null set.
- For any two events E and $F \subset S$, we define the event
  - $E \cup F$: the event E∪F will occur if either E or F (the union of E and F)
  - $E \cap F$ (or EF): the event EF consists of all outcomes which are both in E and in F
    - $E = \{1,3,5\}, F = \{1,2,3\}, E \cup F = ?\ E \cap F = ?$
  - If $EF = 0$, then E and F are said to be mutually exclusive.
  - For more than tow events, we can define unions and intersections such as $\bigcup_{n=1}^{\infty} E_n$ and $\bigcap_{n=1}^{\infty} E_n$, where E_ns are events
- E^c : the complement of E, E^c will occur if and only if E does not occur
  - $S = \{1,2,3,4,5,6\}$, $E = \{1,3,5\}$, $E^c = \{2,4,6\}$, $S^c = 0$ or $\{\}$
- Two sets E and F are equal if they contain exactly the same elements.

Event Space $(S, F(S), P)$

- Intuitively, is a collection of events which we are interested in computing the probability of.

- Mathematically, is a family of subset of sample space S closed under certain set operations as below:
  1. If $E \in F(S)$, then $E^c \in F(S)$
  2. If $E_1, E_2 \in F(S)$, then $E_1 \cup E_2 \in F(S)$
  3. If $E_1, E_2, E_3 \ldots \in F(S)$, then $\cup_{i=\{1,..,\infty\}} E_i \in F(S)$
  - A family of sets satisfying these three properties is called a σ-field

Probabilities Defined on Events $(S, F(S), P)$

- For each event E of a sample space S, we assume that a number $P(E)$ is defined and satisfies the following three conditions:
  1. $0 \le P(E) \le 1$
  2. $P(S) = 1$
  3. For any sequence of events $E_1, E_2, \ldots$ that are mutually exclusive, that is , events for which $E_n E_m = 0$ when $n \neq m$, then
     $p(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n)$

- We refer to P(E) as the probability of the event E.
  - (e.g) the coin flipping $P(\{H\}) = P(\{T\}) = ½$
  - The die rolling:
    $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$
    $P(\{1,3,5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = \frac{1}{2}$

Properties of Probability

- $P(S) = 1$
  - $P(E) + P(E^c) = P(S)$
  - $P(E) = 1 - P(E^c)$
  - $P(\emptyset) = 0$
  - If $EF = \emptyset$, then $P(EF) = 0$

- $P(E \cup F) = P(E) + P(F) - P(EF)$
  - If $E \subset F$, then $P(E) \leq P(F)$

- $$\begin{aligned} P(E \cup F \cup G) &= P(E) + P(F) - P(EF) + P(G) - P(EG \cup FG) \\ &= P(E) + P(F) - P(EF) + P(G) - P(EG) - P(FG) + P(EFG) \\ &= P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG) \end{aligned}$$

- $P(E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_n) = \sum_i P(E_i) - \sum_{i < j} P(E_i E_j) + \sum_{i < j < k} P(E_i E_j E_k) + \cdots + (-1)^{n+1} P(E_1 E_2 \ldots E_n)$

The Principle of Total Probability

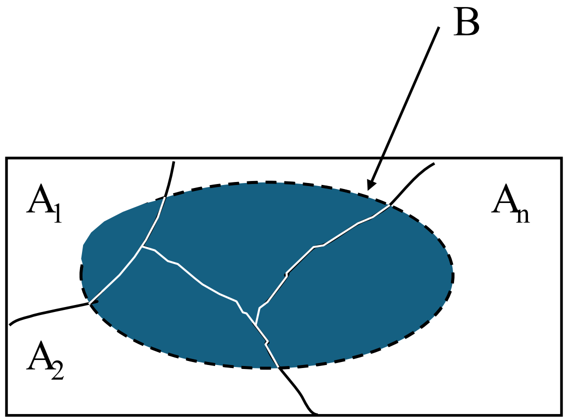
- Let $A_1, A_2, \ldots, A_n$ be a set of mutually exclusive and collectively exhaustive events:

$$A_k A_j = 0 \qquad k \neq j$$

$\bigcup_{j=1}^{n} A_j = S$ then $\sum_{j=1}^{n} \Pr[A_j] = 1$

- Now let B be <u>any</u> event in S. Then,

$\Pr[B] = \Pr[BA_1] + \Pr[BA_2] + \ldots + \Pr[BA_n]$



Conditional Probability

- If we let E and F denote the events, then the probability $P(E|F)$ is the conditional probability that E occurs given that F has occurred
  - Because we know that F has occurred, it follows that F becomes our new sample space and hence the event EF occurs will equal the probability of EF relative to the probability of F. That is
    $P(E|F) = P(EF)/P(F)$ (1.5)

- Ex) A family has two children. What is the conditional probability that boys given that at least one of them is a boy? Assume that the sample space S is given by
  $S = \{(b, b), (b, g), (g, b), (g, g)\}$, and all outcomes are equally likely
  - $P(B|A) = P(BA)/P(A) = P(\{(b, b)\})/P(\{(b, b), (b, g), (g, b)\}) = \frac{\left(\frac{1}{4}\right)}{\left(\frac{3}{4}\right)} = \frac{1}{3}$

Independent Events

- Two events E and F are said to independent if $P(EF) = P(E)P(F)$
- By (1.5) this implies that E and F are independent if $P(E|F) = P(E)$
  - $P(E|F) = P(EF)/P(F) = P(E)P(F)/P(F) = P(E)$
- Two event E and F that are not independent are said to be dependent
- The definition of independence can be extended to more than two events. The events $E_1, E_2, \ldots, E_n$ are said to be independent if for every subset $E_1, E_2, \ldots, E_r$ $r \leq n$ of these events

$$P(E_1E_2 \ldots E_r) = P(E_1)P(E_2)P(E_r)$$

- Ex) Let $E = \{1,2\}, F = \{1,3\}, G = \{1,4\}, P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\})$
  - $P(EF) = P(E)P(F) = ¼$
  - $P(EG) = P(E)P(G) = ¼$
  - $P(FG) = P(F)P(G) = ¼$ (Mutually independent)
  - However, $\frac{1}{4} = P(EFG) \neq P(E)P(F)P(G) = \frac{1}{8}$. Hence, the events EFG are not jointly independent.

Bayes' Formula

- Suppose that $F_1, F_2, \ldots, F_n$ are mutually exclusive events such that $S = \bigcup_{i=1}^{n} F_i$, then
  - $E = ES = E(\bigcup_{i=1}^{n} F_i) = \bigcup_{i=1}^{n} EF_i$
  - $P(E) = \sum_{i=1}^{n} P(EF_i) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i)$ from (1.5) (1.8)
  - $P(F_i|E) = \frac{P(E|F_i)}{P(E)}$ from (1.5)
    $= \frac{P(E|F_i)P(F_i)}{\sum_{i=1}^{n} P(EF_i)P(F_i)}$ from (1.8) (1.9)

- Equation (1.9) is known as Bayes' Formula

- Ex) $F \cup F^c = S$
  - $P(E) = P(ES) = P(E(F \cup F^c)) = P(EF \cup EF^c) = P(EF) + P(EF^c)$
    $= P(E|F)P(F) + P(E|F^c)\, P(F^c) = P(E|F)P(F) + P(E|F^c)\,(1 - P(F))$
  - $P(F|E) = P(E|F)P(F)/P(E) = P(E|F)P(F)/(P(E|F)P(F) + P(E|F^c)\, P(F^c))$

Example of Bayes' Formula

- Ex. 1.13] In answering a question on a multiple-choice test a student either knows the answer or guesses. Let p be the probability that she knows the answer and $1 - p$ the probability that she guesses. Assume that a student who guesses at the answer will be correct with probability $\frac{1}{m}$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?
  - Let C and K denote respectively the event that the student answers the question correctly and the event that she actually knows the answer. Now
  - $P(K|C) = P(KC)/P(C)$
    $= P(C|K)P(K)/(P(C|K)P(K) + P(C|K^c)P(K^c))$
    $= p/(p + (1/m)(1 - p))$
    $= mp/(1 + (m - 1)p)$
  - Thus, for example, if $m = 5, p = \frac{1}{2}$, then the probability that a student knew the answer to a question she correctly answered is $\frac{5}{6}$.

Example 1 of Probability

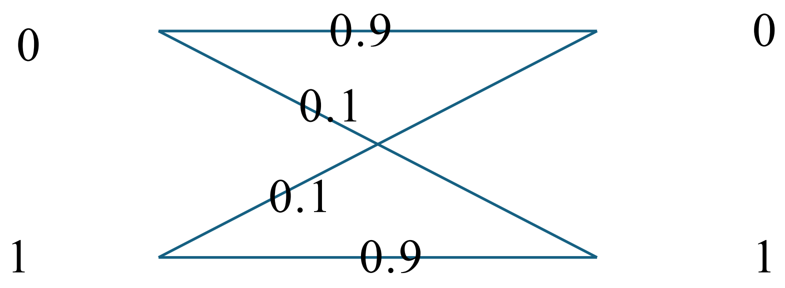
- Consider an experiment with a coin and a die. We define event A as an outcome with "head" of the coin and an even number of the die. Even B is defined as an outcome with a number greater than 3. Find the following answers
  - $P(A)$
  - $P(B)$
  - $P(AB)$
  - $P(A|B)$
  - $P(B|A)$

Answer: Example 1

- $P(A) = 3/12 = 1/4$
- $P(B) = 6/12 = 1/2$
- $AB\ =\ \{H4, H6\}, P(AB) = 2/12 = 1/6$
- $P(A|B) = P(AB)/P(B) = 1/3$
- $P(B|A) = P(AB)/P(B) = 2/3$

Example 2 of Probability

- Let A_1 (A_0) and B_1 (B_0) be the event that 1 (0) is sent and the event that 1(0) is received, respectively.
- Assumption
  - $P(A_0) = 0.8, P(A_1) = 1 - P(A_0) = 0.2,$
  - The probability of error, i.e., $p = P(B_1|A_0) = P(B_0|A_1)$, is 0.1



- Find
  - The error probability at the receiver
  - The probability that 1 is sent when the receiver decides 1.

Answer: Example 2

- Parameters
  - $P(A_0) = 0.8,\ \ P(A_1) = 1 - P(A_0) = 0.2,\ \ P(B_1|A_0) = P(B_0|A_1) = 0.1$
- The error probability at the receiver
  - $P_e = P(A_0B_1) + P(A_1B_0) = P(B_1|A_0)P(A_0) + P(B_0|A_1)P(A_1) = 0.1$
- The probability that 1 is sent when the receiver decides 1.
  - $P(A_m|B) = \frac{P(B|A_m)P(A_m)}{P(B)} = \frac{P(B|A_m)P(A_m)}{\sum_{n=1}^{N} P(B|A_n)P(A_n)}$
  - $P(A_1|B_1) = \frac{P(B_1|A_1)P(A_1)}{P(B_1)} = \frac{P(B_1|A_1)P(A_1)}{P(B_1|A_1)P(A_1)+P(B_1|A_0)P(A_0)} = \frac{0.9\times0.2}{0.9\times0.2+0.1\times0.8} = 0.69$