Chapter 5 Probability and Random Processes (I)

Sample Space, Events, and Probability (1/4)

- A random experiment is any experiment whose outcome cannot be predicted with certainty
- Flipping a coin, throwing a dice, and drawing a card from a deck of cards are examples of random experiments whose results (or *outcome*) of the experiment is uncertain
- In flipping of a coin, "head" and "tail" are the possible outcomes. In throwing a dice, 1,2,3,4,5, and 6 are the possible outcomes. The set of all possible outcomes is called the *sample space* and is denoted by Ω
- Outcomes are denoted by ω's, and ω lies in Ω, *i.e.*, $\omega \in \Omega$

Sample Space, Events, and Probability (2/4)

- A sample space is *discrete* if the number of its elements are finite or countably infinite, otherwise it is a non-discrete sample space
- If we randomly choose a number between 0 and 1, then the sample space corresponding to this random experiment is the set of all numbers between 0 and 1, which is infinite and uncountable. Such a sample space is non-discrete

Sample Space, Events, and Probability (3/4)

- An event is a collection of outcomes
- In throwing a dice, the event "the outcome is odd" consists of outcomes 1, 3, and 5; the event "the outcome is greater than 3" consists of outcomes 4, 5, and 6
- We define a *probability* (*measure*) P as a set function assigning nonnegative values to all event E such that the following conditions are satisfied:
  1. $0 \leqq P(E) \leqq 1$ for all events
  2. $P(\Omega)=1$
  3. For disjoint event $E_1, E_2, E_3, \ldots$ (*i.e.*, events for which $E_i \cap E_j = \Phi$ for all $i \neq j$, where Φ is the empty set), we have

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Sample Space, Events, and Probability (4/4)

- Some properties of probability:
  - $P(E^c)=1-P(E)$, where E^c denotes the complement of E
  - $P(\Phi)=0$
  - $P(E_1 \cup E_2)=P(E_1)+P(E_2)-P(E_1 \cap E_2)$
  - If $E_1 \subseteq E_2$ then $P(E_1) \leqq P(E_2)$

Conditional Probability (1/7)

- Let us assume that the two events, E_1 and E_2, have probabilities $P(E_1)$ and $P(E_2)$
- The conditional probability of the event E_1, given the event E_2, is defined by

$$P(E_1 \mid E_2) = \begin{cases} \frac{P(E_1 \cap E_2)}{P(E_2)}, & P(E_2) \neq 0 \\ 0, & otherwise \end{cases}$$

- If it happens that $P(E_1 \mid E_2) = P(E_1)$, then the knowledge of E_2 does not change the probability of the occurrence of E_1. In this case, the event E_1 and E_2 are said to be independent. For independent events, $P(E_1 \cap E_2) = P(E_1)P(E_2)$

Conditional Probability (2/7)

- **Example 5.1.1.** In throwing a fair dice, the probability of

$$A = \{\text{The outcomes are greater than 3}\}$$

is

$$P(A) = P(4) + P(5) + P(6) = 1/2.$$

The probability of

$$B = \{\text{The outcomes are even}\}$$

is

$$P(B) = P(2) + P(4) + P(6) = 1/2.$$

In this case,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(4) + P(6)}{\frac{1}{2}} = \frac{2}{3}$$

Conditional Probability (3/7)

- If the events $\{E_i\}_{i=1}^{n}$ are disjoint and their union makes the entire sample space, then they make a *partition* of the sample space Ω
- For any event A, we have the conditional probabilities $\{P(A|E_i)\}_{i=1}^{n}$, $P(A)$ can be obtained by applying the *total probability theorem* stated as

$$P(A) = \sum_{i=1}^{n} P(E_i)P(A|E_i)$$

- *Bayes's rule* gives the conditional probabilities $P(E_i|A)$ by the following relation

$$P(E_i|A) = \frac{P(E_i)P(A|E_i)}{\sum_{j=1}^{n} P(E_j)P(A|E_j)}$$

Conditional Probability (4/7)

- **Example 5.1.2.** In a certain city, 50% of the population drive to work, 30% take the subway, and 20% take the bus. The probability of being late for those who drive is 10%, for those who take the subway is 3%, and for those who take the bus is 5%.
  1. What is the probability that an individual in this city will be late for work?
  2. If an individual is late for work, what is the probability that he drove to work?
- Let D, S, and B denote the events of an individual driving, taking the subway, or taking the bus. Let L denote the event of being late

Conditional Probability (5/7)

- **Example 5.1.2. (Cont'd)**

  From the assumption, we have $P(D)=0.5$, $P(S)=0.3$, and $P(B)=0.2$. We also have

$$P(L\,|\,D)=0.1;$$

$$P(L\,|\,S)=0.03;$$

$$P(L\,|\,B)=0.05.$$

- From the total probability theorem,

$$P(L)=P(D)P(L\,|\,D)+P(S)P(L\,|\,S)+P(B)P(L\,|\,B)=0.069$$

- Applying Bayes's rule, we have

$$P(D\,|\,L)=\frac{P(D)P(L\,|\,D)}{P(D)P(L\,|\,D)+P(S)P(L\,|\,S)+P(B)P(L\,|\,B)}$$

$$\approx 0.725$$

Conditional Probability (6/7)

- **Example 5.1.3.** In a binary communication system, the input bits transmitted over the channel are either 0 or 1 with probabilities 0.3 and 0.7, respectively. When a bit is transmitted over the channel, it can be either received correctly or incorrectly (due to channel noise). Let us assume that if a 0 is transmitted, the probability of it being received in error (i.e., being received as 1) is 0.01, and if a 1 is transmitted, the probability of it being received in error (i.e., being received as 0) is 0.1.
  1. What is the probability that the output of this channel is 1?
  2. Assuming we have observed a 1 at the output of this channel, what is the probability that the input to the channel was a 1?

Conditional Probability (7/7)

- **Example 5.1.3. (Cont'd)**

  Let X denote the input and Y denote the output. From the problem assumptions, we have

$$P(X=0)=0.3; \quad P(X=1)=0.7;$$

$$P(Y=0 \mid X=0)=0.99; \ P(Y=1 \mid X=0)=0.01;$$

$$P(Y=0 \mid X=1)=0.1; \ P(Y=1 \mid X=1)=0.9$$

- $P(Y=1)=P(Y=1,X=0)+P(Y=1,X=1)$

$$=P(X=0)P(Y=1 \mid X=0)+P(X=1)P(Y=1 \mid X=1)$$

$$=0.633$$

- From Bayes's rule, we have

$$P(X=1 \mid Y=1)=\frac{P(X=1)P(Y=1 \mid X=1)}{P(X=0)P(Y=1 \mid X=0)+P(X=1)P(Y=1 \mid X=1)}=0.995$$

Random Variables (1/7)

- A *random variable* is a mapping from the sample space Ω to the set of real numbers
- In other words, a random variable is an assignment of real numbers to the outcome of a random experiment

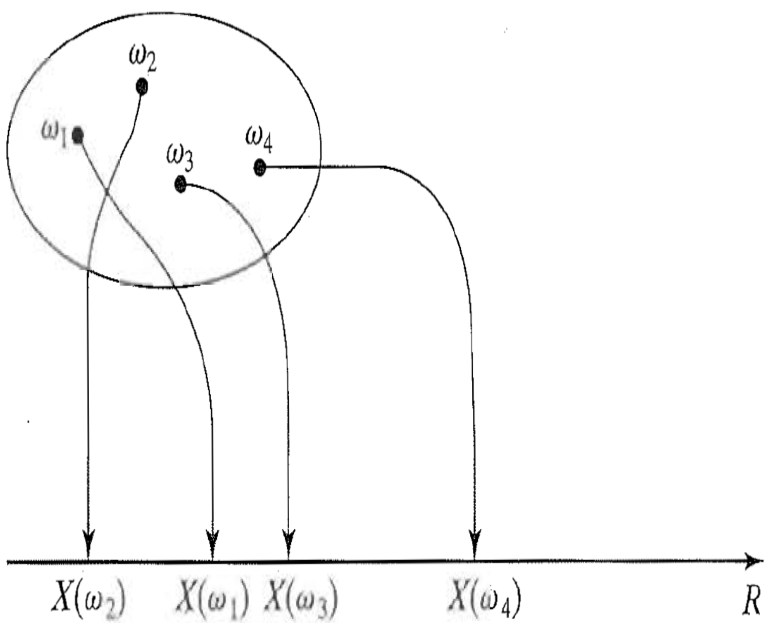

Figure 5.1 A random variable as a mapping from Ω to $\mathbb{R}$.

Random Variables (2/7)

- **Example 5.1.4.** In throwing a dice, if the player wins the amount that the dice shows if the result is even and loses that amount if it is odd, then the random variable is

$$X(\omega) = \begin{cases} \omega, & \omega = 2,4,6 \\ -\omega, & \omega = 1,3,5 \end{cases}$$

- A random variable is discrete if the range of its values is either finite or countably infinite. The range is usually denoted by $\{x_i\}$
- A continuous random variable is one in which the range of values is a continuum

Random Variables (3/7)

- The cumulative distribution function (CDF) of a random variable X is defined as

$$F_X(x) = P\{\omega \in \Omega : X(\omega) \leq x\},$$

which can be simply written as

$$F_X(x) = P(X \leq x)$$

- The CDF has the following properties:
  - $0 \leqq F_X(x) \leqq 1$
  - $F_X(x)$ is nondecreasing
  - $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$
  - $F_X(x)$ is continuous from the right, *i.e.*, $\lim_{\varepsilon \to 0^+} F_X(x+\varepsilon) = F_X(x)$
  - $P(a < X \leqq b) = F_X(b) - F_X(a)$
  - $P(X = a) = F_X(a) - F_X(a^-)$

Random Variables (4/7)

- For *discrete* random variables, $F_X(x)$ is a staircase function. A random variable is *continuous* if $F_X(x)$ is a continuous function

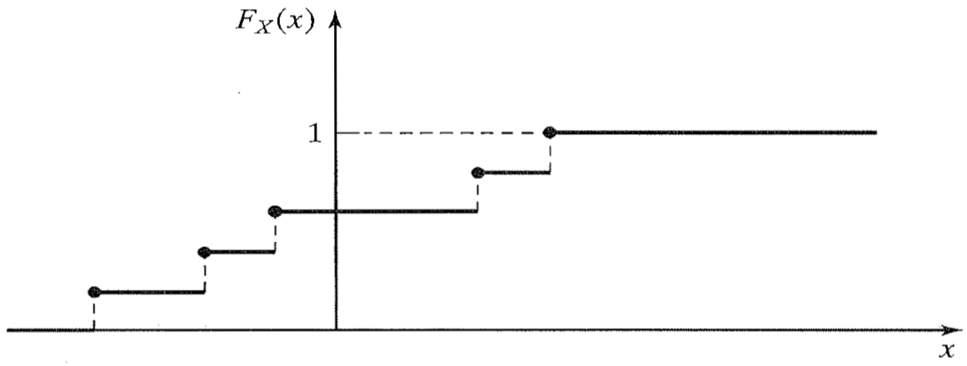

Figure 5.2 The CDF for a discrete random variable.

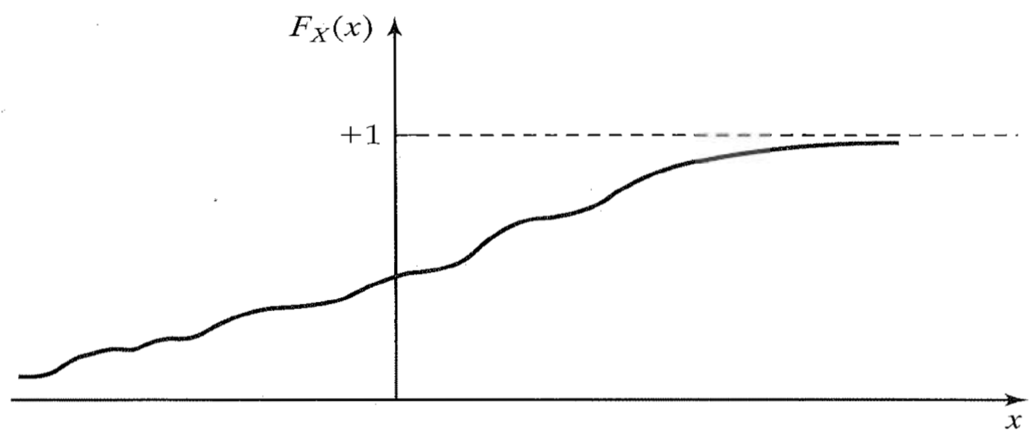

Figure 5.3 CDF for a continuous random variable.

Random Variables (5/7)

- A random variable is *mixed* if it is neither discrete nor continuous

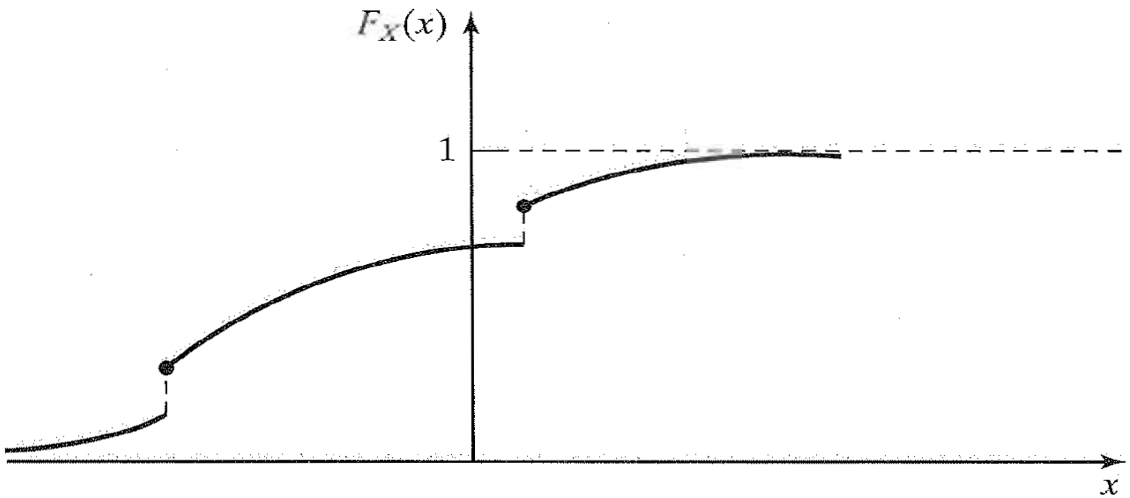

Figure 5.4 CDF for a mixed random variable.

Random Variables (6/7)

- The *probability density function*, or PDF, of a continuous random variable X is defined as the derivative of its CDF. It is denoted by $f_X(x)$, *i.e.*,

$$f_X(x) = \frac{d}{dx} F_X(x)$$

- Properties of PDF are as follows:
  - $f_X(x) \geqq 0$
  - $\int_{-\infty}^{\infty} f_X(x)dx = 1$
  - $\int_a^b f_X(x)dx = P(a < X \leq b)$
  - In general, $P(X \in A) = \int_A f_X(x)dx$
  - $F_X(x) = \int_{-\infty}^{x} f_X(u)du$

Random Variables (7/7)

- For discrete random variables, we generally define the probability mass function, or PMF, which is defined as $\{p_i\}$, where $p_i = P(X = x_i)$
- For all i, we have $p_i \geqq 0$ and $\sum_i p_i = 1$

Important Random Variables (1/12)

- **Bernoulli random variable.** This is a discrete random variable taking two values, 1 and 0, with probabilities p and 1-p
- A Bernoulli random variable can be employed to model the channel errors
- **Binomial random variable.** This is a discrete random variable giving the number of 1's in a sequence of n-independent Bernoulli trials
- The PMF is given by

$$P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & 0 \le x \le n \\ 0, & otherwise \end{cases}$$

Important Random Variables (2/12)

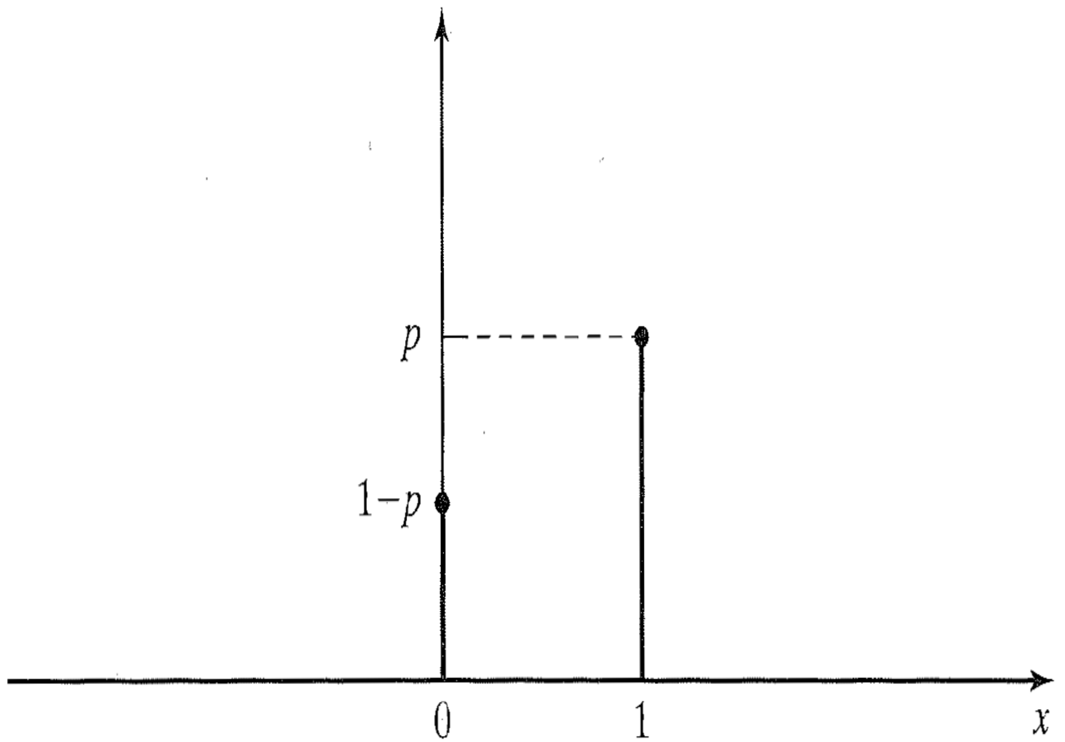

Figure 5.5 The PMF for the Bernoulli random variable.

Important Random Variables (3/12)

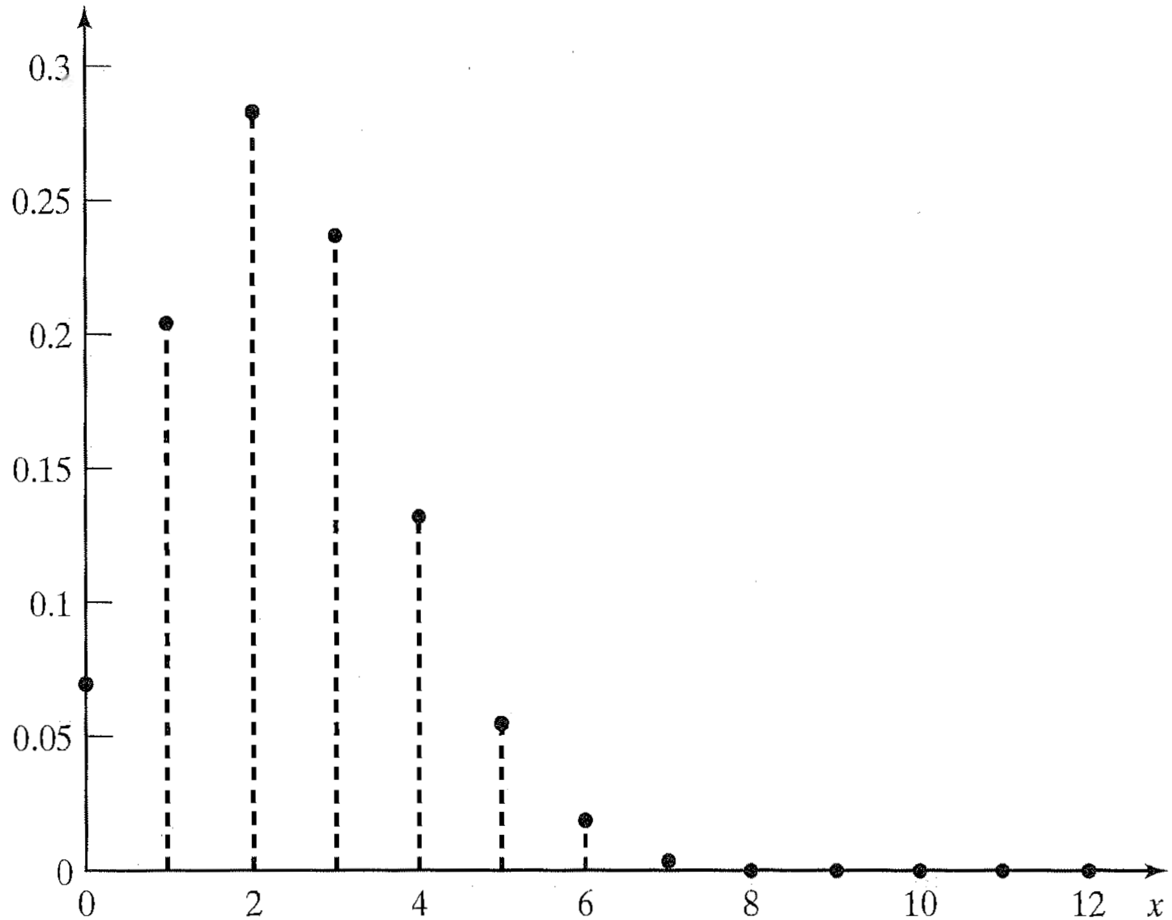

Figure 5.6 The PMF for the binomial random variable.

Important Random Variables (4/12)

- A binomial random variable can be used to model the total number of bits received in error when a sequence of n bits is transmitted over a channel with a bit-error probability of p
- **Example 5.1.5.** Assume 10,000 bits are transmitted over a channel in which the error probability is 10^{-3}. What is the probability that the total number of errors is less than 3?
- In this example, n=10,000, p=0.001, and we are looking for $P(X<3)$. We have

$$
\begin{aligned}
P(X<3) &= P(X=0)+P(X=1)+P(X=2) \\
&= 0.0028
\end{aligned}
$$

Important Random Variables (5/12)

- **Uniform random variable.** This is a continuous random variable taking values between *a* and *b* with equal probabilities for intervals of equal length
- The density function is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & otherwise \end{cases}$$

- When the phase of a sinusoid is random, it is usually modeled as a uniform random variable between 0 and 2π

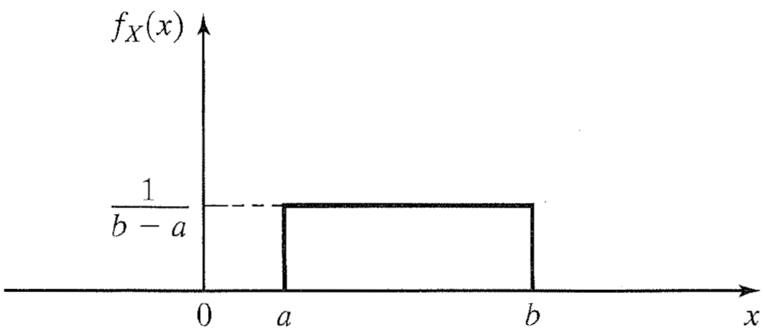

Figure 5.7 The PDF for the uniform random variable.

Important Random Variables (6/12)

- **Gaussian or normal variable.** The Gaussian, or normal, random variable is a continuous random variable
- The density function of a Gaussian random variable is described as

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

- There are two parameters involved in the definition of the Gaussian random variable. The parameter m is called the *mean* and can assume any finite value
- The parameter σ is called the standard deviation and can assume any finite and positive value. The square of the standard deviation σ^2 is call the variance

Important Random Variables (7/12)

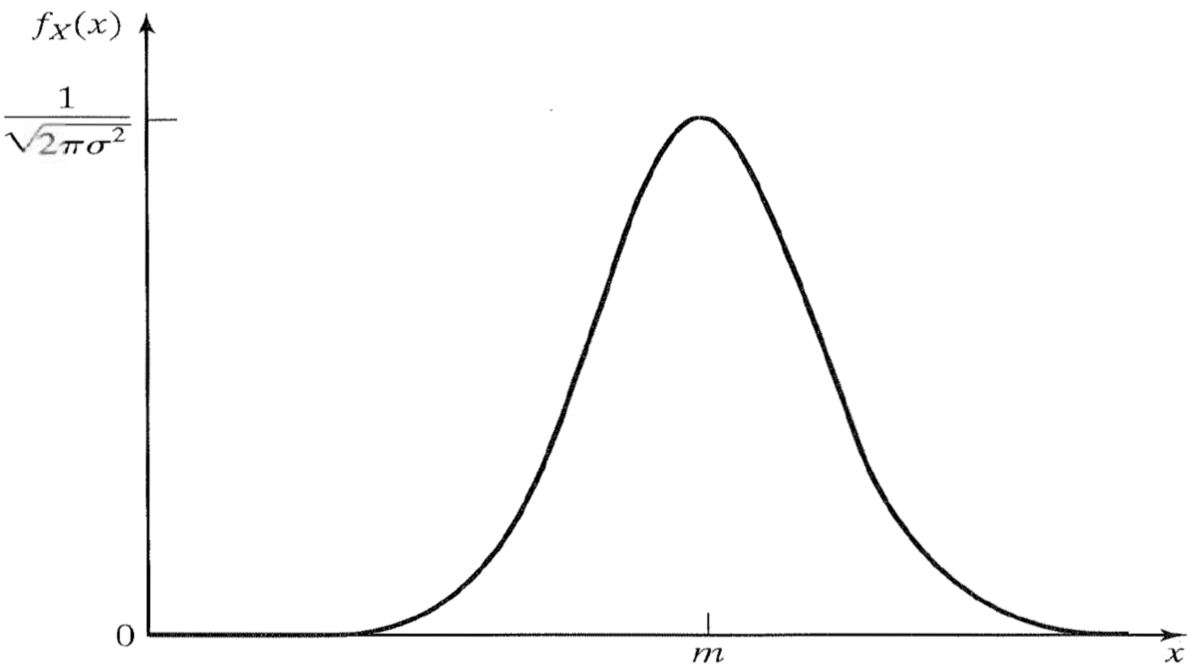

Figure 5.8 The PDF for the Gaussian random variable.

Important Random Variables (8/12)

- A Gaussian random variable with mean m and variance σ^2 is denoted by $\mathcal{N}(m,\sigma^2)$. The random variable $\mathcal{N}(0,1)$ is called *standard normal*
- The Gaussian random variable is the most important and frequently encountered random variable in communications
- Thermal noise is the major source of noise in communication systems and has a Gaussian distribution
- Assuming that X is a standard normal random variable, we define the function $Q(x)$ as $P(X>x)$. The Q-function is given by the relation

$$Q(x) = P(X > x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Important Random Variables (9/12)

- It is easy to see that $Q(x)$ satisfies the following relations:

$$Q(-x) = 1 - Q(x)$$

$$Q(0) = \frac{1}{2}$$

$$Q(\infty) = 0$$

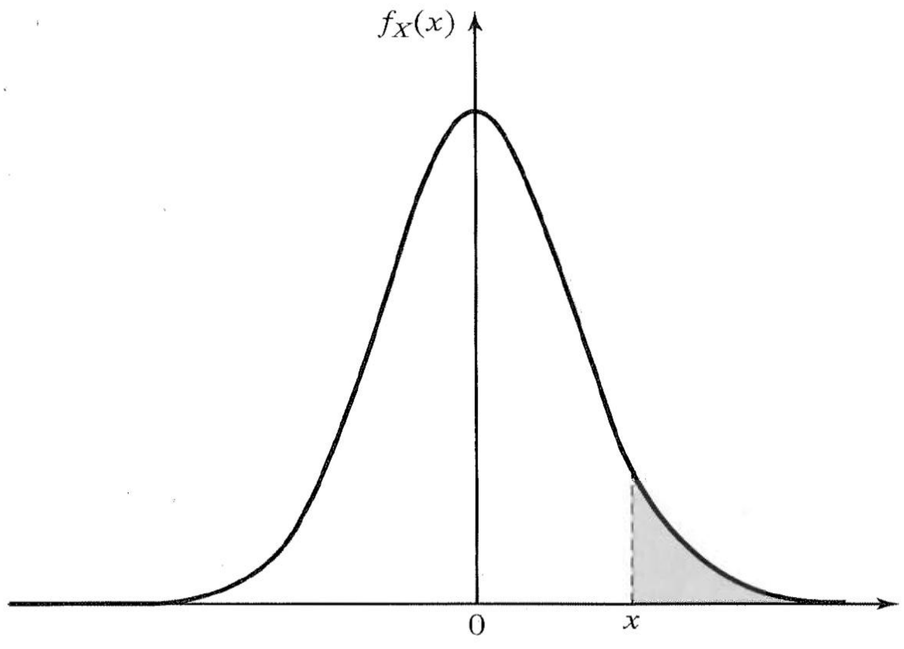

Figure 5.9 The Q-function as the area under the tail of a standard normal random variable.

Important Random Variables (10/12)

- Two important upper bounds on the Q-function are

$$Q(x) \le \frac{1}{2} e^{-\frac{x^2}{2}} \qquad \text{for all } x \geqq 0$$

and

$$Q(x) < \frac{1}{\sqrt{2\pi}\, x} e^{-\frac{x^2}{2}} \qquad \text{for all } x \geqq 0$$

- A frequently used lower bound is

$$Q(x) > \frac{1}{\sqrt{2\pi}\, x} \left(1 - \frac{1}{x^2}\right) e^{-\frac{x^2}{2}} \qquad \text{for all } x \geqq 0$$

- If X is a $\mathcal{N}(\mathrm{m}, \sigma^2)$ random variable, we have

$$P(X > x) = Q\left(\frac{x - m}{\sigma}\right)$$

Important Random Variables (11/12)

TABLE 5.1 TABLE OF THE Q FUNCTION

0	5.000000e–01	2.4	8.197534e–03	4.8	7.933274e–07
0.1	4.601722e–01	2.5	6.209665e–03	4.9	4.791830e–07
0.2	4.207403e–01	2.6	4.661189e–03	5.0	2.866516e–07
0.3	3.820886e–01	2.7	3.466973e–03	5.1	1.698268e–07
0.4	3.445783e–01	2.8	2.555131e–03	5.2	9.964437e–06
0.5	3.085375e–01	2.9	1.865812e–03	5.3	5.790128e–08
0.6	2.742531e–01	3.0	1.349898e–03	5.4	3.332043e–08
0.7	2.419637e–01	3.1	9.676035e–04	5.5	1.898956e–08
0.8	2.118554e–01	3.2	6.871378e–04	5.6	1.071760e–08
0.9	1.840601e–01	3.3	4.834242e–04	5.7	5.990378e–09
1.0	1.586553e–01	3.4	3.369291e–04	5.8	3.315742e–09
1.1	1.356661e–01	3.5	2.326291e–04	5.9	1.817507e–09
1.2	1.150697e–01	3.6	1.591086e–04	6.0	9.865876e–10
1.3	9.680049e–02	3.7	1.077997e–04	6.1	5.303426e–10
1.4	8.075666e–02	3.8	7.234806e–05	6.2	2.823161e–10
1.5	6.680720e–02	3.9	4.809633e–05	6.3	1.488226e–10
1.6	5.479929e–02	4.0	3.167124e–05	6.4	7.768843e–11
1.7	4.456546e–02	4.1	2.065752e–05	6.5	4.016001e–11
1.8	3.593032e–02	4.2	1.334576e–05	6.6	2.055790e–11
1.9	2.871656e–02	4.3	8.539898e–06	6.7	1.042099e–11
2.0	2.275013e–02	4.4	5.412542e–06	6.8	5.230951e–12
2.1	1.786442e–02	4.5	3.397673e–06	6.9	2.600125e–12
2.2	1.390345e–02	4.6	2.112456e–06	7.0	1.279813e–12
2.3	1.072411e–02	4.7	1.300809e–06		

Important Random Variables (12/12)

- **Example 5.1.6.** X is a Gaussian random variable with mean 1 and variance 4. Find the probability that X is between 5 and 7
- We have $m=1$ and $\sigma=2$. Thus,

$$P(5 < X < 7) = P(X > 5) - P(X \geq 7)$$
$$= Q(\frac{5-1}{2}) - Q(\frac{7-1}{2})$$
$$= Q(2) - Q(3)$$
$$\approx 0.0214$$