Fair Use Notice

The material used in this presentation i.e., pictures/graphs/text, etc. is solely intended for educational/teaching purpose, offered free of cost to the students for use under special circumstances of Online Education due to COVID-19 Lockdown situation and may include copyrighted material - the use of which may not have been specifically authorised by Copyright Owners. It's application constitutes Fair Use of any such copyrighted material as provided in globally accepted law of many countries. The contents of presentations are intended only for the attendees of the class being conducted by the presenter.

4. Review of Basic Probability and Statistics

Outline:

4.1. Random Variables and Their Properties

4.2. Simulation Output Data and Stochastic Processes

4.3. Estimation of Means and Variances

4.4. Confidence Interval for the Mean

4.1. Random Variables and Their Properties

A random variable is a function that assigns a real number greater than $-\infty$ and less than ∞ to each point in the sample space S.

The distribution function *F(x)* of random variable X is defined for each real number x as follows:

$F(x) = P(X \leq x)$ **for** $-\infty < x < \infty$

Where $P(X \leq x)$ **means the probability associated with the event** $\{X \leq x\}$**.**

4.1. Random Variables and Their Properties

A random variable X is said to be *discrete* if it can take on at most a countable number of values, say, $x_1, x_2, \ldots$. The probability that X is equal to x_i is given by

$$p(x_i) = P(X = x_i) \text{ for } i = 1, 2, \ldots$$

And

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

where $p(x)$ is the *probability mass function*. The *distribution function* $F(x)$ is

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i)$$

for all $-\infty < x < \infty$.

<u>4.1. Random Variables and Their Properties</u>

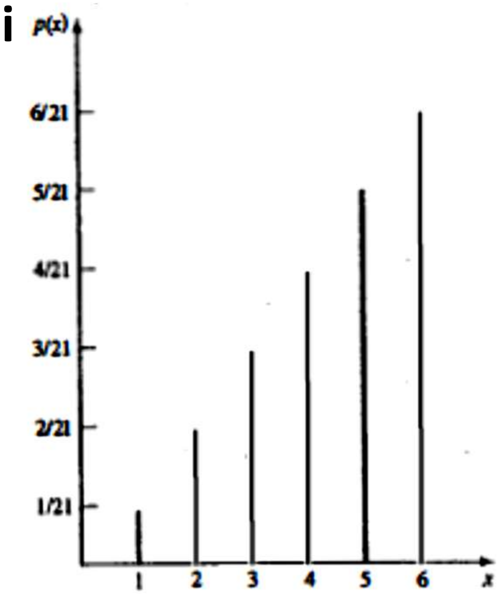
Example: Consider jobs arriving at a job shop.

- Let *X* be the number of jobs arriving each week at a job shop.
- R_x = possible values of *X* (range space of *X*) = $\{0,1,2,\ldots\}$
- $p(x_i)$ = probability the random variable is $x_i = P(X = x_i)$

Consider the experiment of tossing a single die. Define X as the number of spots on the up face of the die after a toss. Then $R_x = \{1, 2, 3, 4, 5, 6\}$. Assume the die is loaded so that the probability that a given face lands up is proportional to the number of spots showing. The discrete probability distributi experiment is given by

x_i	1	2	3	4	5	6
$p(x_i)$	1/21	2/21	3/21	4/21	5/21	6/21

The conditions stated earlier are satisfied-that is, $p(x_i) \geq 0$ for i= 1, 2, ... , 6 and $\sum_{i=1} p(x_i)$ = 1/21+ · ··+ 6/2 1 = 1.



5.1 Probability mass function for loaded-die example.

<u>Example 4.1:</u> Consider the demand-size random variable of Section 1.5 of Law (2006) that takes on the values 1, 2, 3, 4, with probabilities 1/6, 1/3, 1/3, 1/6. The probability mass function and the distribution function are given in Figure 4.1 and 4.2

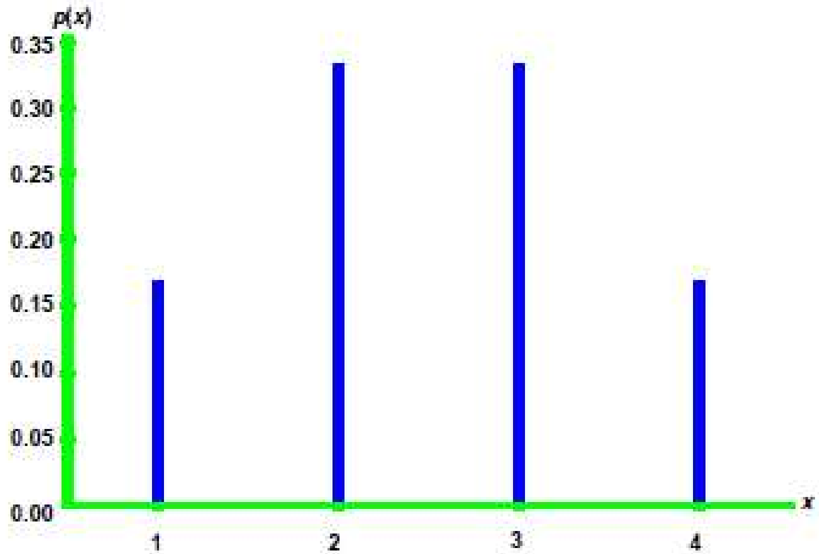


Figure 4.1. $p(x)$ for the demand-size random variable.

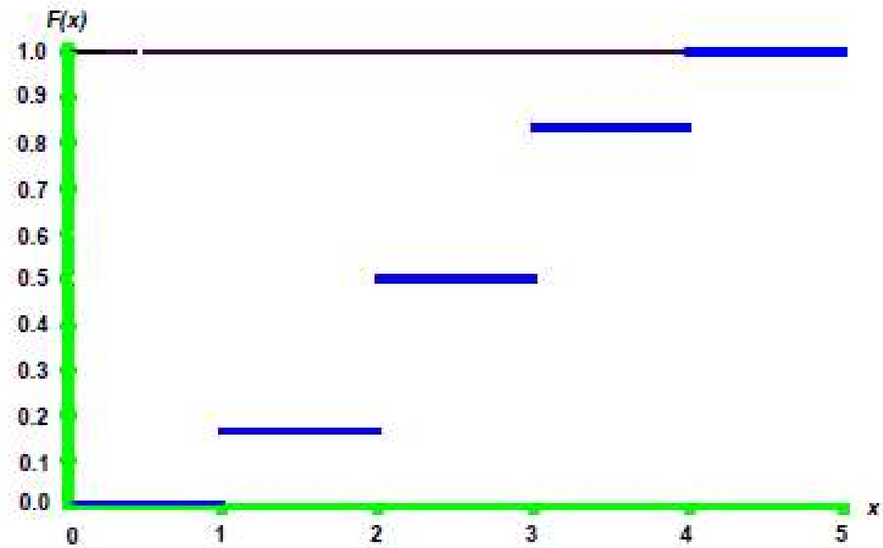


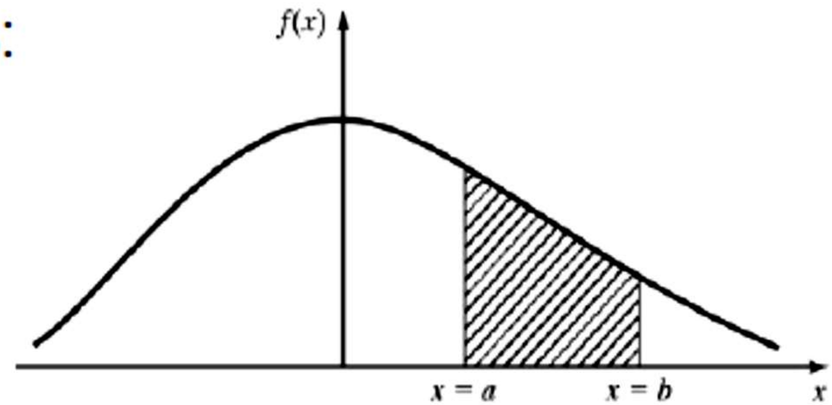
Figure 4.2. $F(x)$ for the demand-size random variable.

- X is a continuous random variable if its range space R_x is an interval or a collection of intervals.
- The probability that X lies in the interval *[a,b]* is given by:

$$P(a \le X \le b) = \int_a^b f(x)dx$$

- *f(x)*, denoted as the pdf of *X*, satisfies:

1. $f(x) \ge 0$, for all x in R_X
2. $\int_{R_X} f(x)dx = 1$
3. $f(x) = 0$, if x is not in R_X

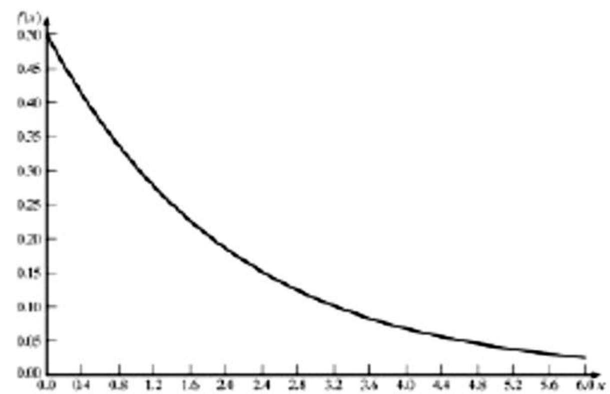


- Properties

  1. $P(X = x_0) = 0$, because $\int_{x_0}^{x_0} f(x)dx = 0$
  2. $P(a \le X \le b) = P(a \prec X \le b) = P(a \le X \prec b) = P(a \prec X \prec b)$

- Example: Life of an inspection device is given by *X*, a continuous random variable with pdf:

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2}, & \text{x} \ge 0 \\ 0, & \text{otherwise} \end{cases}$$



  - *X* has an exponential distribution with mean 2 years
  - Probability that the device's life is between 2 and 3 years is:

$$P(2 \le x \le 3) = \frac{1}{2}\int_2^3 e^{-x/2} dx = 0.14$$

Cumulative Distribution Function

- Cumulative Distribution Function (cdf) is denoted by *F(x)*, where *F(x)* = *P(X <= x)*
  - If *X* is discrete, then $$F(x) = \sum_{\substack{\text{all} \\ x_i \le x}} p(x_i)$$
  - If *X* is continuous, then $$F(x) = \int_{-\infty}^{x} f(t)dt$$
- Properties

  1. F is nondecreasing function. If $a \prec b$, then $F(a) \le F(b)$
  2. $\lim_{x \to \infty} F(x) = 1$
  3. $\lim_{x \to -\infty} F(x) = 0$
- All probability question about *X* can be answered in terms of the cdf, e.g.:

$$P(a \prec X \le b) = F(b) - F(a), \text{ for all } a \prec b$$

The cdf of die tossing example is given in figure

x	$(-\infty, 1)$	$[1, 2)$	$[2, 3)$	$[3, 4)$	$[4, 5)$	$[5, 6)$	$[6, \infty)$
$F(x)$	0	1/21	3/21	6/21	10/21	15/21	21/21

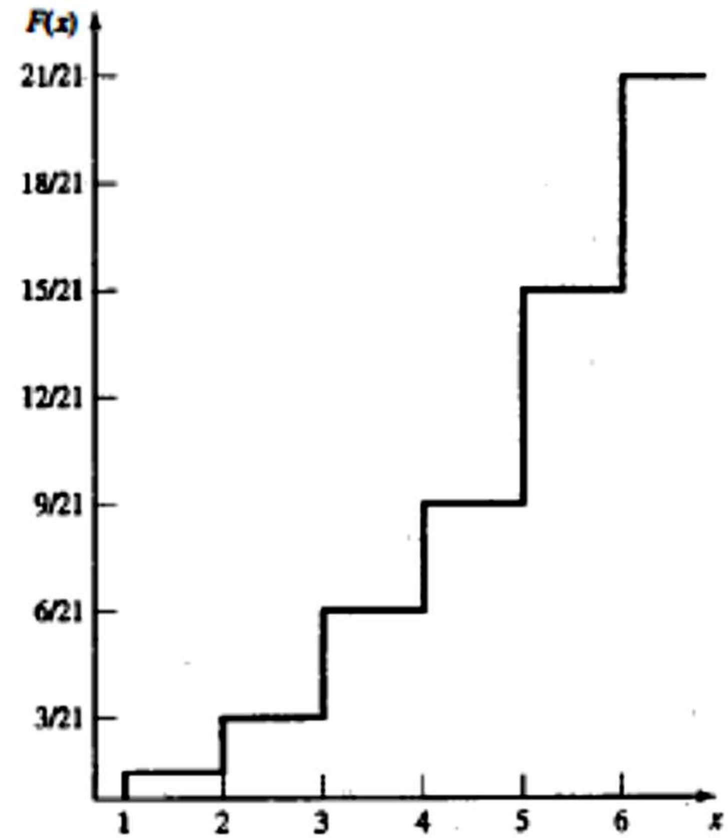


Figure 5.4 cdf for loaded-die example.

Cumulative Distribution Function

- Example: An inspection device has cdf:

$$F(x) = \frac{1}{2}\int_0^x e^{-t/2}\,dt = 1 - e^{-x/2}$$

  - The probability that the device lasts for less than 2 years:

$$P(0 \le X \le 2) = F(2) - F(0) = F(2) = 1 - e^{-1} = 0.632$$

  - The probability that it lasts between 2 and 3 years:

$$P(2 \le X \le 3) = F(3) - F(2) = (1 - e^{-(3/2)}) - (1 - e^{-1}) = 0.145$$

Expectation /Mean

- The expected value of X is denoted by $E(X)$
  - If X is discrete $$E(x) = \sum_{\text{all } i} x_i p(x_i)$$
  - If X is continuous $$E(x) = \int_{-\infty}^{\infty} x f(x) dx$$
  - a.k.a the mean, m, or the 1st moment of X
  - A measure of the central tendency
- The variance of X is denoted by $V(X)$ or $var(X)$ or σ^2
  - Definition: $V(X) = E[(X - E[X]^2]$
  - Also, $V(X) = E(X^2) - [E(x)]^2$
  - A measure of the spread or variation of the possible values of X around the mean
    - The standard deviation of X is denoted by σ
  - Definition: square root of $V(X)$
  - Expressed in the same units as the mean

Expectation

- Example: The mean of life of the previous inspection device is:

$$E(X) = \frac{1}{2}\int_0^\infty xe^{-x/2}dx = -xe^{-x/2}\Big|_0^\infty + \int_0^\infty e^{-x/2}dx = 2$$

- To compute variance of *X*, we first compute $E(X^2)$:

$$E(X^2) = \frac{1}{2}\int^\infty x^2e^{-x/2}dx = -x^2e^{-x/2}\Big|^\infty + \int^\infty e^{-x/2}dx = 8$$

- Hence, the variance and standard deviation of the device's life are:

$$V(X) = 8 - 2^2 = 4$$

$$\sigma = \sqrt{V(X)} = 2$$

The mean and variance of the die tossing experiment is given as follows

$$E(X) = 1\left(\frac{1}{21}\right) + 2\left(\frac{2}{21}\right) + \cdots + 6\left(\frac{6}{21}\right) = \frac{91}{21} = 4.33$$

$$E(X^2) = 1^2\left(\frac{1}{21}\right) + 2^2\left(\frac{2}{21}\right) + \cdots + 6^2\left(\frac{6}{21}\right) = 21$$

$$V(X) = 21 - \left(\frac{91}{21}\right)^2 = 21 - 18.78 = 2.22$$

$$\sigma = \sqrt{V(X)} = 1.49$$

Uniform random variable

A uniform random variable on the interval [0, 1] has the following probability density function:

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Furthermore, if $0 \leq x \leq 1$, then

$$F(x) = \int_0^x f(y)\,dy = \int_0^x 1\,dy = x$$

For the uniform random variable , the mean is given by

$$\mu = \int_0^1 x f(x)\,dx = \int_0^1 x\,dx = \frac{1}{2}$$

The variance is given by

$$E(X^2) = \int_0^1 x^2 f(x)\,dx = \int_0^1 x^2\,dx = \frac{1}{3}$$

$$\mathrm{Var}(X) = E(X^2) - \mu^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

Uniform random variable

For the demand-size random variable in Example 4.1, the mean is given by

$$\mu = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{3}\right) + 3\left(\frac{1}{3}\right) + 4\left(\frac{1}{6}\right) = \frac{5}{2}$$

The variance is computed as follows:

$$E(X^2) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{3}\right) + 3^2\left(\frac{1}{3}\right) + 4^2\left(\frac{1}{6}\right) = \frac{43}{6}$$

$$\mathrm{Var}(X) = E(X^2) - \mu^2 = \frac{43}{6} - \left(\frac{5}{2}\right)^2 = \frac{11}{12}$$

Joint Probability Mass Function

If X and Y are discrete random variables, then let

$$p(x, y) = P(X = x, Y = y) \qquad \text{for all } x, y$$

where $p(x, y)$ is called the *joint probability mass function* of X and Y. In this case, X and Y are *independent* if

$$p(x, y) = p_X(x) p_Y(y) \qquad \text{for all } x, y$$

where

$$p_X(x) = \sum_{\text{all } y} p(x, y)$$

$$p_Y(y) = \sum_{\text{all } x} p(x, y)$$

are the (marginal) probability mass functions of X and Y.

Joint Probability Mass Function

Suppose that X and Y are jointly discrete random variables with

$$p(x, y) = \begin{cases} \dfrac{xy}{27} & \text{for } x = 1, 2 \text{ and } y = 2, 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

$$p_X(x) = \sum_{y=2}^{4} \frac{xy}{27} = \frac{x}{3} \qquad \text{for } x = 1, 2$$

$$p_Y(y) = \sum_{x=1}^{2} \frac{xy}{27} = \frac{y}{9} \qquad \text{for } y = 2, 3, 4$$

Since $p(x, y) = xy/27 = p_X(x)p_Y(y)$ for all x, y, the random variables X and Y are independent.

Joint Probability density Function

The random variables X and Y are *jointly continuous* if there exists a nonnegative function $f(x, y)$, called the *joint probability density function* of X and Y, such that for all sets of real numbers A and B,

$$P(X \in A, Y \in B) = \int_B \int_A f(x, y)\, dx\, dy$$

In this case, X and Y are *independent* if

$$f(x, y) = f_X(x) f_Y(y) \qquad \text{for all } x, y$$

where

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

are the (marginal) probability density functions of X and Y, respectively.

Joint Probability density Function

EXAMPLE 4.11. Suppose that X and Y are jointly continuous random variables with

$$f(x, y) = \begin{cases} 24xy & \text{for } x \geq 0,\ y \geq 0, \text{ and } x + y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$f_X(x) = \int_0^{1-x} 24xy\, dy = 12xy^2 \Big|_0^{1-x} = 12x(1-x)^2 \qquad \text{for } 0 \leq x \leq 1$$

$$f_Y(y) = \int_0^{1-y} 24xy\, dx = 12yx^2 \Big|_0^{1-y} = 12y(1-y)^2 \qquad \text{for } 0 \leq y \leq 1$$

Since

$$f\left(\frac{1}{2}, \frac{1}{2}\right) = 6 \neq \left(\frac{3}{2}\right)^2 = f_X\left(\frac{1}{2}\right) f_Y\left(\frac{1}{2}\right)$$

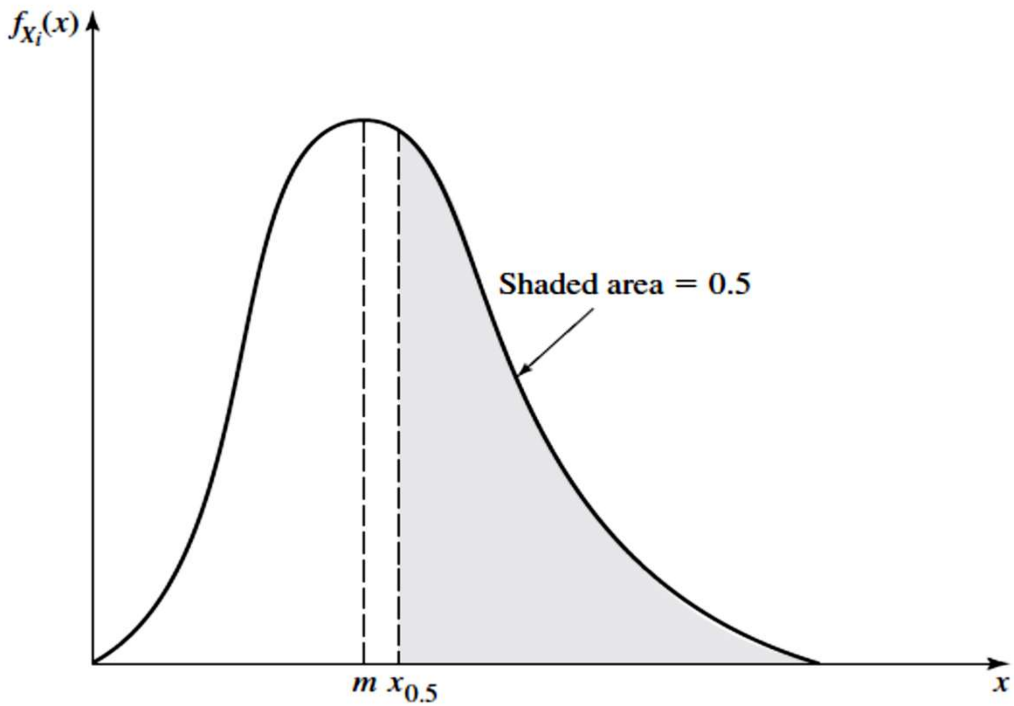
X and Y are not independent.

Mean , Median and mode

The *median* $x_{0.5}$ of the random variable *Xi*, which is an alternative measure of central tendency, is defined to be the smallest value of *x* such that $F_{X_i}(x) \geq 0.5$ If *Xi* is a continuous random variable, then $F_{X_i}(x_{0.5}) = 0.5$

Mode is an alternative measure of central tendency, is defined to be the smallest value of x such That

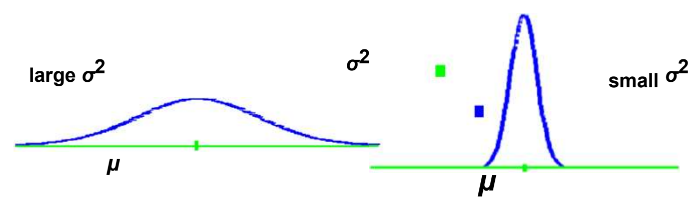
$$fx_i(x)[px_i(x)]$$



Consider a discrete random variable *X* that takes on each of the values, 1, 2, 3, 4, and 5 with probability 0.2. Clearly, the mean and median of *X* are 3. Consider now a random variable *Y* that takes on each of the values 1, 2, 3, 4, and 100 with probability 0.2. The mean and median of *Y* are 22 and 3, respectively. Note that the median is insensitive to this change in the distribution.

Properties:

1. $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$
2. $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ if X, Y are independent



Density functions for continuous random variables with large and small variances.

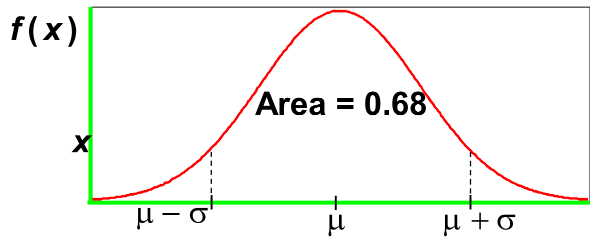


Figure 4.7. Density function for a N(μ, σ^2) distribution.

The *covariance* between the random variables *X* and *Y*, denoted by Cov(*X*, *Y*), is defined by

$$\mathrm{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$$
$$= E(XY) - E(X)E(Y)$$

The covariance is a measure of the dependence between *X* and *Y*. Note that Cov(*X*, *X*) = Var(*X*).

Definitions:

Cov(X, Y)	X and Y are
= 0	*uncorrelated*
> 0	*positively correlated*
< 0	*negatively correlated*

Independent random variables are also uncorrelated.

Note that, in general, we have

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)$$

If X and Y are independent, then

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$$

The *correlation* between the random variables X and Y, which is a measure of <u>linear</u> dependence , is denoted by Cor(X, Y) and defined

$$\mathrm{Cor}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

It can be shown that

$$-1 \leq \mathrm{Cor}(X,Y) \leq 1$$

Suppose that $Y = aX + b$, where a and b are constants. Then

$$\mathrm{Cor}(X,Y)=\begin{cases} 1 & \text{if } a>0 \\ -1 & \text{if } a<0 \end{cases}$$

Flip a fair coin 3 times. Let X be the number of heads in the first 2 flips and let Y be the number of heads on the last 2 flips (so there is overlap on the middle flip). Compute Cov(X;Y).

With 3 tosses there are only 8 out comes {HHH,HHT,...}, so we can create the joint probability table directly.

$X \backslash Y$	0	1	2	$p(x_i)$
0	1/8	1/8	0	1/4
1	1/8	2/8	1/8	1/2
2	0	1/8	1/8	1/4
$p(y_j)$	1/4	1/2	1/4	1

	X	Y
HHH	2	2
HHT	2	1
HTH	1	1
THH	1	2
TTH	0	1
THT	1	1
TTH	0	1
TTT	0	0

From the marginals we compute E(X)=1=E(Y)

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

$$E(XY) = 1 \cdot \frac{2}{8} + 2\frac{1}{8} + 2\frac{1}{8} + 4\frac{1}{8} = \frac{5}{4}.$$

$$\mathrm{Cov}(X, Y) = \frac{5}{4} - 1 = \frac{1}{4}$$

Example 2. (Zero covariance does not imply independence.) Let X be a random variable that takes values $-2,-1,0,1,2$; each with probability $1/5$. Let $Y = X^2$. Show that $\text{Cov}(X,Y) = 0$ but X and Y are not independent.

<u>**answer:**</u> We make a joint probability table:

$Y \backslash X$	-2	-1	0	1	2	$p(y_j)$
0	0	0	1/5	0	0	1/5
1	0	1/5	0	1/5	0	2/5
4	1/5	0	0	0	1/5	2/5
$p(x_i)$	1/5	1/5	1/5	1/5	1/5	1

X	Y
-2	4
-1	1
0	0
1	1
2	4

Using the marginals we compute means $E(X) = 0$ and $E(Y) = 2$.

Next we show that X and Y are not independent. To do this all we have to do is find one place where the product rule fails, i.e. where $p(x_i, y_j) \neq p(x_i)p(x_j)$:

$$P(X = -2, Y = 0) = 0 \quad \text{but} \quad P(X = -2) \cdot P(Y = 0) = 1/25.$$

E(XY) = $-2\left(\frac{1}{5}\right) - 1\left(\frac{1}{5}\right) + 1\left(\frac{1}{5}\right) + 2\left(\frac{1}{5}\right) = 0$

Cov(X,Y) =0

For the jointly continuous random variables X and Y, covariance is computed as

$$E(XY) = \int_0^1 \int_0^{1-x} xy\, f(x, y)\, dy\, dx$$

$$= \int_0^1 x^2 \left(\int_0^{1-x} 24y^2\, dy \right) dx$$

$$= \int_0^1 8x^2(1-x)^3\, dx$$

$$= \frac{2}{15}$$

$$E(X) = \int_0^1 x f_X(x)\, dx = \int_0^1 12x^2(1-x)^2\, dx = \frac{2}{5}$$

$$E(Y) = \int_0^1 y f_Y(y)\, dy = \int_0^1 12y^2(1-y)^2\, dy = \frac{2}{5}$$

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

$$= \frac{2}{15} - \left(\frac{2}{5}\right)\left(\frac{2}{5}\right)$$

$$= -\frac{2}{75}$$

$$\mathrm{Cor}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} = \frac{-\frac{2}{75}}{\frac{1}{25}} = -\frac{2}{3}$$

4.2. Simulation Output Data and Stochastic processes

A *stochastic process* is a collection of "similar" random variables ordered over time all defined relative to the same experiment. If the collection is X_1, X_2, ... , then we have a *discrete-time stochastic process.*

If the collection is $\{X(t), t \geq 0\}$, then we have a *continuous-time* stochastic process.

Example 4.3:

For the single-server queuing system of Chapter 1, assume the following:

- **The (arrival times) A_i's are independent and identically distributed (IID)**
- **The (Service times) S_i's are IID**
- **The A_i's and S_i's are independent**

Relative to the experiment of generating the A_i's and P_i's, one can define the discrete-time stochastic process of delays in queue D_1, D_2, ... as follows:

$$D_1 = 0$$

$$D_{i+1} = \max\{D_i + S_i - A_{i+1}, 0\} \text{ for } i = 1, 2, \ldots$$

Thus, the simulation maps the input random variables into the output process of interest.

Problem 4.2: Are D_i and D_{i+1} independent, positively correlated, or negatively correlated?

Other examples of stochastic processes:

- **N_1, N_2, ... , where N_i = number of parts produced in the ith hour for a manufacturing system.**
- **T_1, T_2, ... , where T_i = time in system of the ith part for a manufacturing system.**
- **$\{Q(t), t \geq 0\}$, where $Q(t)$ = number of customers in queue at time t.**
- **C_1, C_2, ... , where C_i = total cost in the ith month for an inventory system**
- **E_1, E_2, ... , where E_i = end-to-end delay of ith message to reach its destination in a communications network**

Example 4.4: **Consider the delay-in- queue process D_1, D_2, ... for the $M/M/1$ queue with utilization factor ρ.**

Then the correlation function ρ_j between D_i and D_{i+j} is given in Figure 4.8.

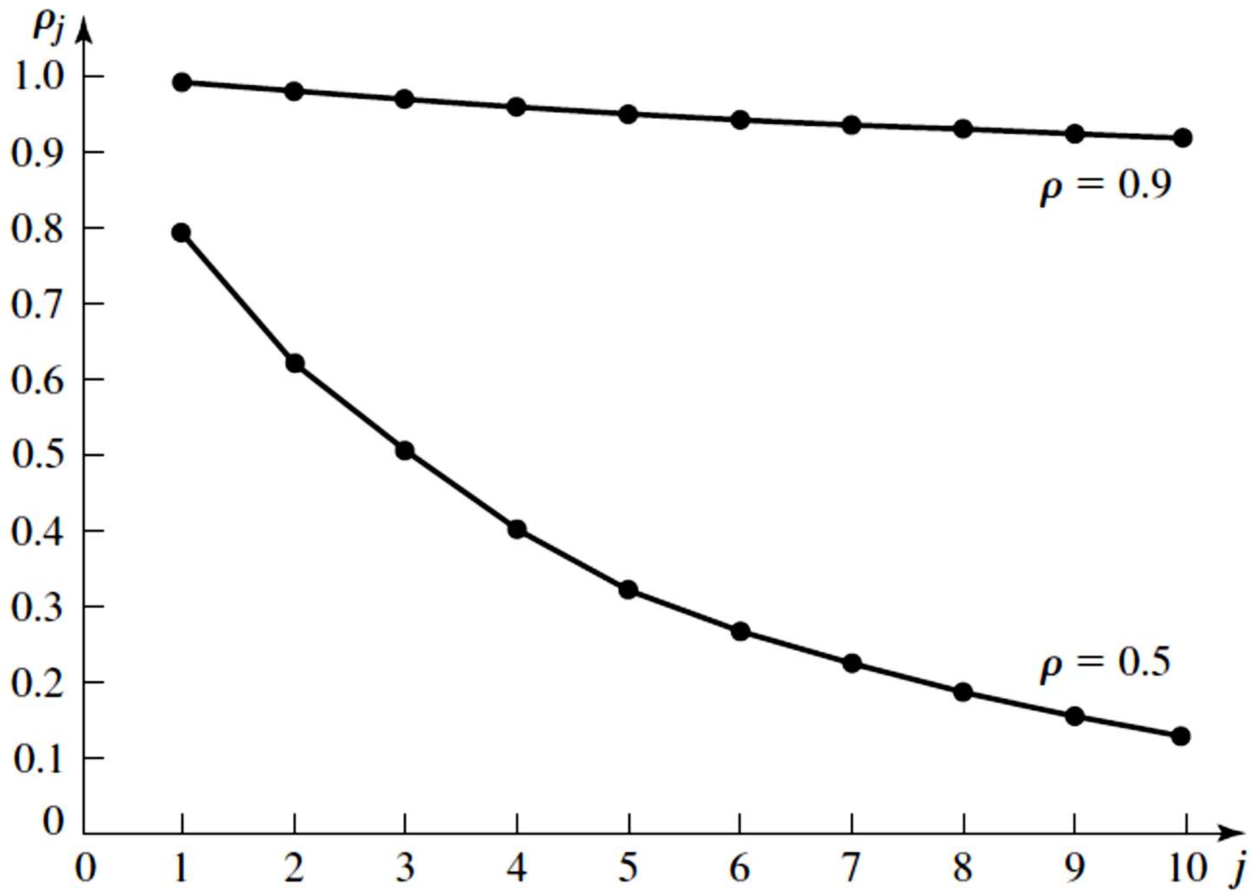


FIGURE 4.10
Correlation function ρ_j of the process $D_1, D_2, \ldots$ for the $M/M/1$ queue.

4.3. Estimation of Means and Variances

Let X_1, X_2, ..., X_n be IID random variables with population mean and variance μ and σ^2, respectively.

Population parameter	Sample estimate	
μ	Sample mean $\overline{X}(n) = \dfrac{\sum_{i=1}^{n} X_i}{n}$	(1)
σ^2	Sample variance $S^2(n) = \dfrac{\sum_{i=1}^{n} [X_i - \overline{X}(n)]^2}{n-1}$	(3)
$\mathrm{Var}[\overline{X}(n)] = \dfrac{\sigma^2}{n}$ (4)	$\widehat{\mathrm{Var}}[\overline{X}(n)] = \dfrac{S^2(n)}{n}$	(5)

Note that $\overline{X}(n)$ is an *unbiased estimator* of μ, i.e.,
$E[\overline{X}(n)] = E(X) = \mu$. (2)

Problem 4.3: Show that $X(n)$ is an unbiased estimator of μ .
The difficulty with using $X(n)$ as an estimator of μ without any additional information is that we have no way of assessing how close $X(n)$ is to μ .
Because $X(n)$ is a random variable with variance $Var[X(n)]$, on one experiment it may be close to μ while on another it may differ from μ by a large amount (see Figure 4.9).

The usual way to access the precision of $X(n)$ as an estimator of μ is to construct a confidence interval for μ.

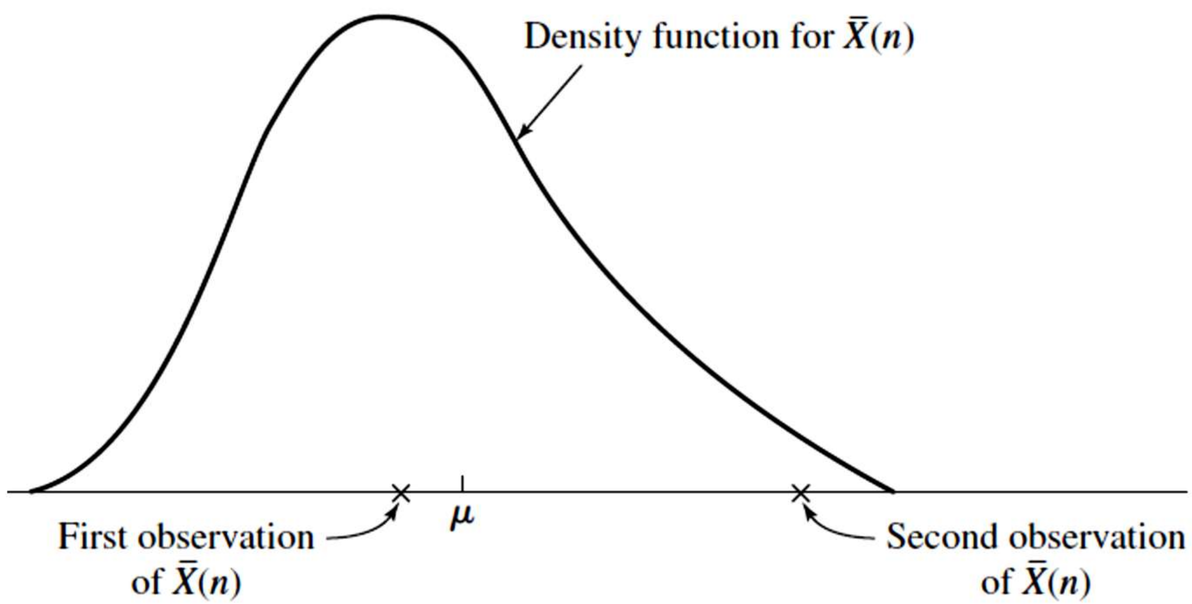


Figure 4.9. Two observations of the random variable $X(n)$.

<u>Example 4.5:</u> Consider the bank with 5 tellers on p. 486-487 of Law. The following are the average delays in queue resulting from 10 independent replications of the simulation model:

1.53, 1.66, 1.24, …, 2.60

Since these observations are IID, they can be plugged into (1) through (5).

However, the delays in queue from <u>one</u> particular replication are not independent.

4.4. Confidence Interval for the Mean

Let X_1, X_2, ..., X_n be IID random variables with mean μ. Then an (approximate) 100(1 - α) percent (0 < α < 1) *confidence interval* for μ Is

$$\overline{X}(n) \pm t_{n-1,1-\alpha/2}\sqrt{S^2(n)/n}$$

where $t_{n-1,\,1-\alpha/2}$ is the upper 1 - α/2 critical point for a *t* distribution with *n* - 1 df (see Figure 4.10).

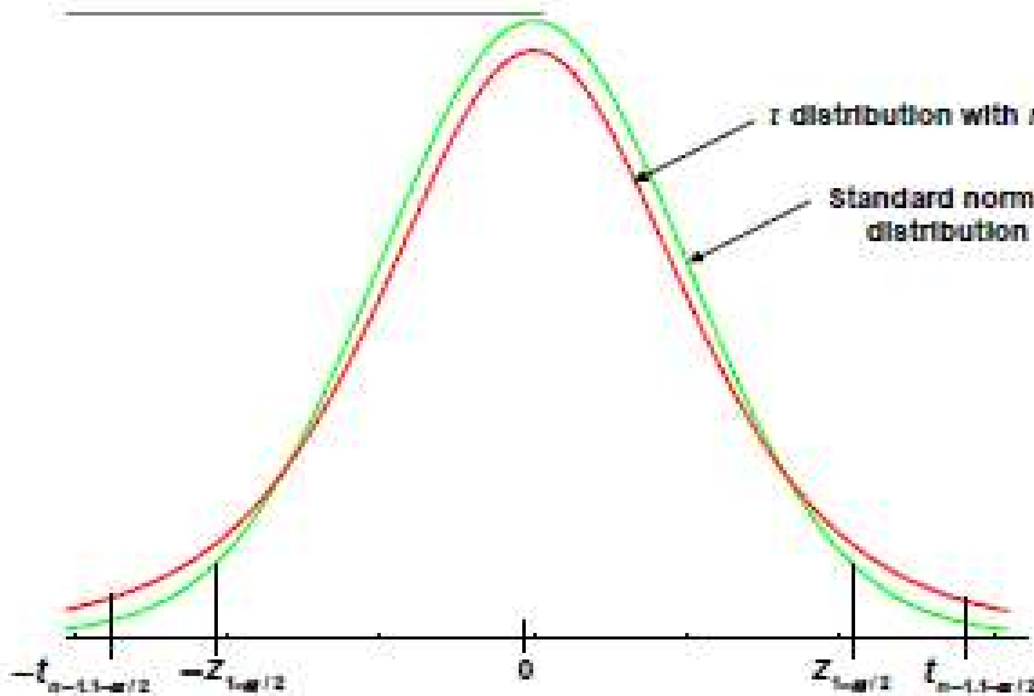


Figure 4.10. Standard normal distribution and t distribution with n - 1 df.

Interpretation of a confidence interval:

If one constructs a very large number of independent 100(1 - α) percent confidence intervals for μ each based on *n* observations, where *n* is <u>sufficiently large</u>, then the proportion of these confidence intervals that contain μ should be 1 - α (regardless of the distribution of X)

Alternatively, if *X* is N(μ,σ^2), then the coverage probability will be 1- α regardless of the value of *n*. If *X* is <u>not</u> N(μ,σ^2), then there will be a degradation in coverage for "small" *n*. The greater the skewness of the distribution of *X*, the greater the degradation (see pp. 256-257).

We used *t* $t_{n-1,\,1-\alpha/2}$ rather than $z_{1-\alpha/2}$ in (6) to help lessen the effect of skewness in the distribution of *X* and of "small" *n*.

Suppose that the 10 observations 1.20, 1.50, 1.68, 1.89, 0.95, 1.49, 1.58, 1.55, 0.50, and 1.09 are from a normal distribution with unknown mean m and that our objective is to construct a 90 percent confidence interval for m. From these data we get

$$\overline{X}(10) = 1.34 \qquad \text{and} \qquad S^2(10) = 0.17$$

which results in the following confidence interval for μ:

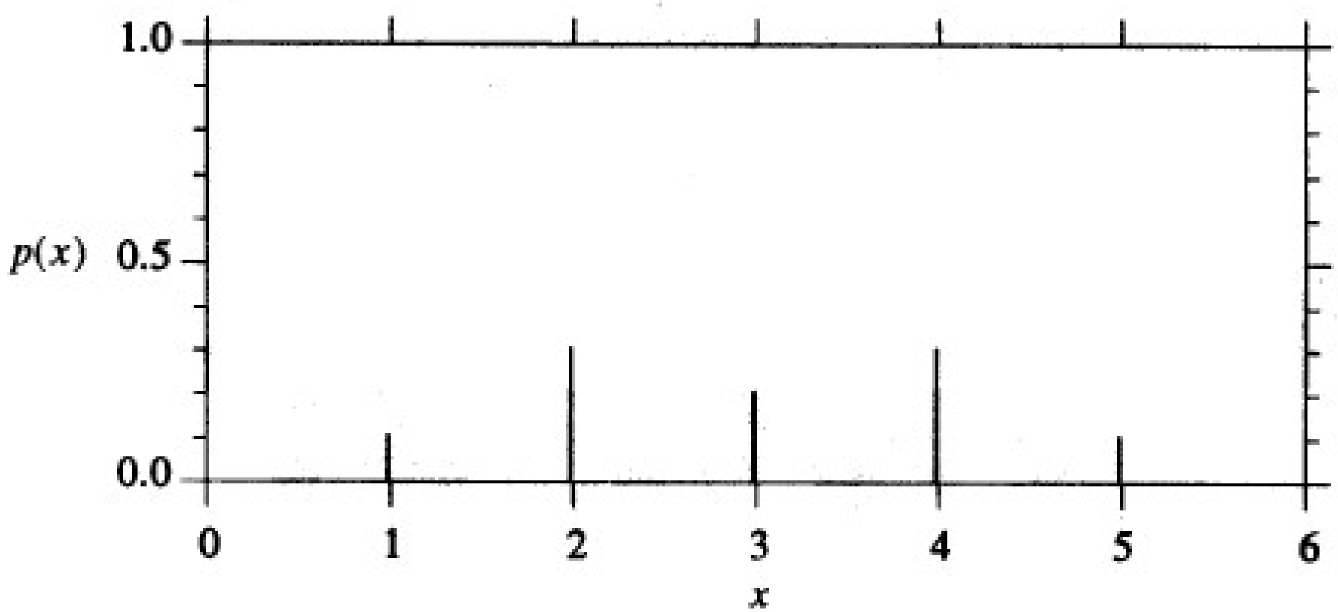
$$\overline{X}(10) \pm t_{9,0.95}\sqrt{\frac{S^2(10)}{10}} = 1.34 \pm 1.83\sqrt{\frac{0.17}{10}} = 1.34 \pm 0.24$$

We claim with 90 percent confidence that m is in the interval [1.10, 1.58]

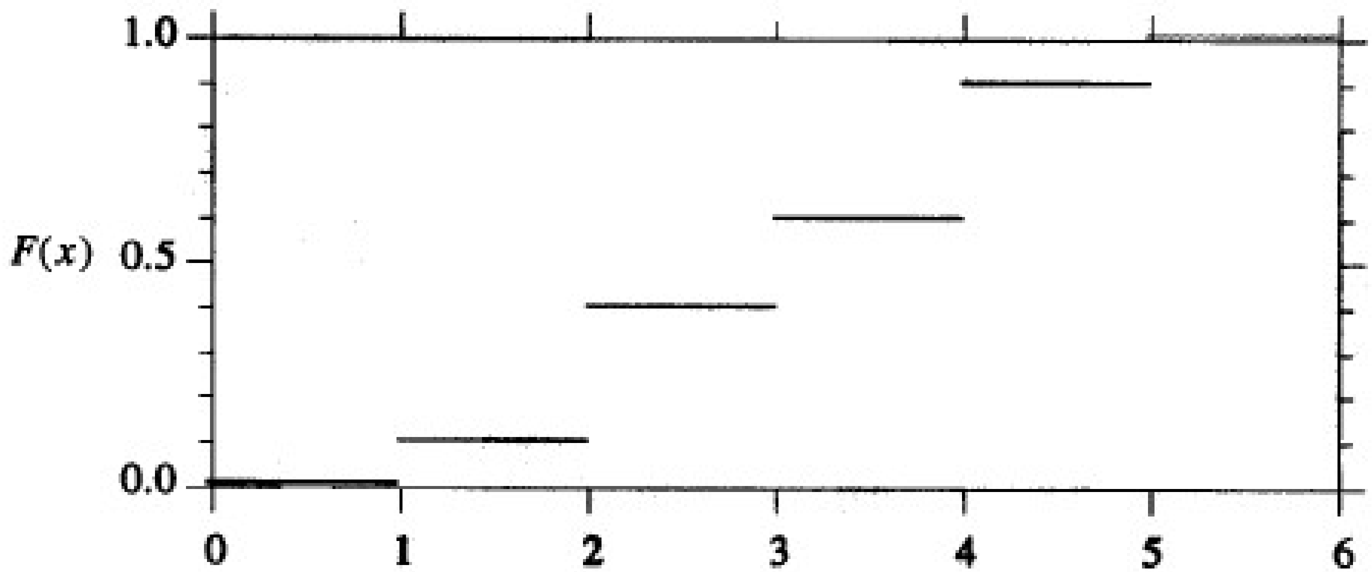
4.1. Suppose that X is a discrete random variable with probability mass function given by

$$p(1) = \tfrac{1}{10}, \qquad p(2) = \tfrac{3}{10}, \qquad p(3) = \tfrac{2}{10}, \qquad p(4) = \tfrac{3}{10}, \qquad \text{and} \qquad p(5) = \tfrac{1}{10}$$

(*a*) Plot $p(x)$.
(*b*) Compute and plot $F(x)$.
(*c*) Compute $P(1.4 \le X \le 4.2)$, $E(X)$, and $\mathrm{Var}(X)$.



$$F(x)=\begin{cases}0.0 & \text{if } x<1\\ 0.1 & \text{if } 1\le x<2\\ 0.4 & \text{if } 2\le x<3\\ 0.6 & \text{if } 3\le x<4\\ 0.9 & \text{if } 4\le x<5\\ 1.0 & \text{if } 5\le x\end{cases}$$



(c).

$P(1.4 \le X \le 4.2) = 3/10 + 2/10 + 3/10 = 4/5.$
$E(X) = 1(1/10) + 2(3/10) + 3(2/10) + 4(3/10) + 5(1/10) = 3.$
$E(X^2) = 1^2(1/10) + 2^2(3/10) + 3^2(2/10) + 4^2(3/10) + 5^2(1/10) = 52/5$, so
$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = 52/5 - 3^2 = 7/5.$

4.2. Suppose that X is a continuous random variable with probability density function given by

$$f(x) = x^2 + \tfrac{2}{3}x + \tfrac{1}{3} \qquad \text{for } 0 \le x \le c$$

(*a*) What must be the value of c?

Assuming this value of c, do the following:

(*b*) Plot $f(x)$.

(*c*) Compute and plot $F(x)$.

(*d*) Compute $P(\tfrac{1}{3} \le X \le \tfrac{2}{3})$, $E(X)$, and $\mathrm{Var}(X)$.

(a)

$1 = \int_0^c f(x)\,dx = \int_0^c \left(x^2 + \frac{2}{3}x + \frac{1}{3}\right)dx = \left(\frac{1}{3}x^3 + \frac{1}{3}x^2 + \frac{1}{3}x\right)\Big|_{x=0}^{x=c} = \frac{1}{3}c\left(c^2 + c + 1\right)$. The solution to this $c = 1$,

Problem 1. The probability mass function p_X of some discrete real-valued random variable X is given by the following table :

x	0	1	2	3	4
$p_X(x)$	0.35	0.15		0.1	0.2

(a) Give the missing value $p_X(2)$.

Answer. Since p_X is a probability mass function, we have that $\sum_{x=0}^{4} p_X(x) = 1$. Therefore

$$\begin{aligned} p_X(2) &= 1 - (p_X(0) + p_X(1) + p_X(3) + p_X(4)) \\ &= 1 - (0.35 + 0.15 + 0.1 + 0.2) \\ &= 0.2. \end{aligned}$$

(b) Draw the histogram of p_X.

Answer.

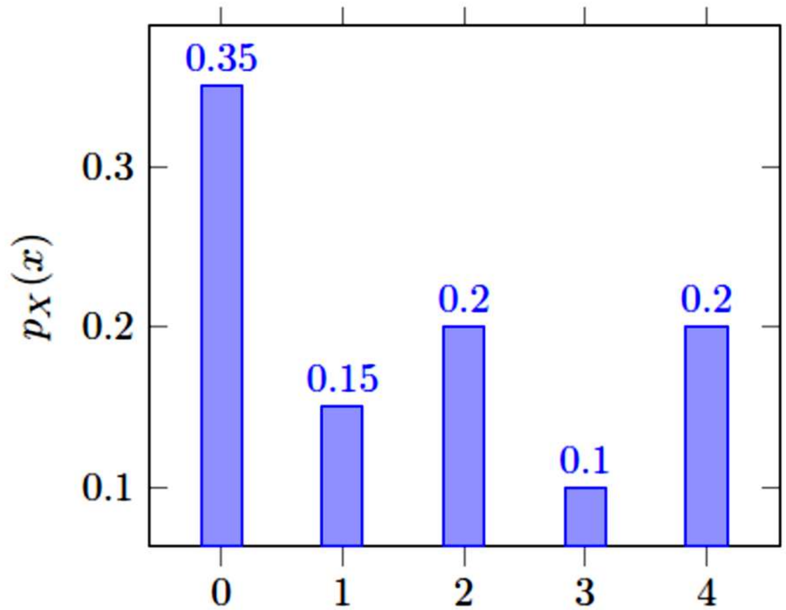


Figure 1: Histogram of p_X

(c) Give the cumulative distribution function F_X of X.

Answer. By definition, for any real number x, $F_X(x) = \mathbb{P}(X \leq x)$.

$$F_X(x) = \sum_{y \leq x} p_X(y)$$

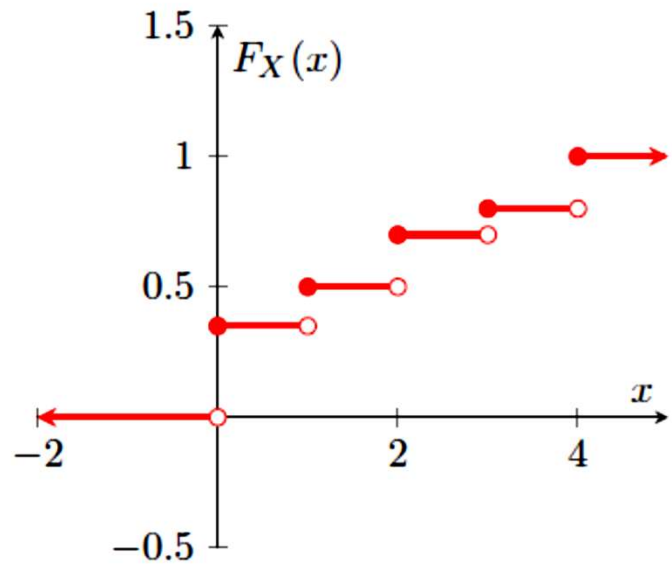
For instance, $F_X(2.2) = \sum_{y \leq 2.2} p_X(y) = p_X(0) + p_X(1) + p_X(2) = 0.7$

In our case, X is a discrete random variable, so its cumulative distribution function F_X is piecewise constant:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 0.35 & 0 \leq x < 1 \\ 0.5 & 1 \leq x < 2 \\ 0.7 & 2 \leq x < 3 \\ 0.8 & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

(d) Graph F_X.

Answer.



Problem 3. Calculate the probability mass function of X whose cumulative distribution function is given by :

$$F_X(b) = \begin{cases} 0 & b < 0 \\ 1/2 & 0 \leq b < 1 \\ 3/5 & 1 \leq b < 2 \\ 4/5 & 2 \leq b < 3 \\ 9/10 & 3 \leq b < 3.5 \\ 1 & b \geq 3.5 \end{cases}$$

Answer. By definition, for any real number x, $F_X(x) = \mathbb{P}(X \leq x)$. Here, we notice that the distribution function F_X is piecewise constant, which means that the underlying random variable X is discrete and the values that X takes on correspond to points y where F_X is discontinuous, that is $X(\Omega) = \{0, 1, 2, 3, 3.5\}$. The probability mass function p_X of X can be calculated in the following fashion

- For any $0 \leq x < 1$, $1/2 = F_X(x) = \sum_{y \leq x} p_X(x) = p_X(0)$. So $p_X(0) = 1/2$.
- For any $1 \leq x < 2$, $3/5 = F_X(x) = \sum_{y \leq x} p_X(x) = p_X(0) + p_X(1)$. So $p_X(1) = 3/5 - p_X(0) = 1/10$.
- For any $2 \leq x < 3$, $4/5 = F_X(x) = \sum_{y \leq x} p_X(x) = p_X(0) + p_X(1) + p_X(2)$. So $p_X(2) = 4/5 - (p_X(0) + p_X(1)) = 1/5$.
- For any $3 \leq x < 3.5$, $9/10 = F_X(x) = \sum_{y \leq x} p_X(x) = p_X(0) + p_X(1) + p_X(2) + p_X(3)$. So $p_X(3) = 9/10 - (p_X(0) + p_X(1) + p_X(2)) = 1/10$.
- Since p_X is a pmf, $p_X(0) + p_X(1) + p_X(2) + p_X(3) + p_X(3.5) = 1$. So $p_X(3.5) = 1 - (p_X(0) + p_X(1) + p_X(2) + p_X(3)) = 1/10$.