Topic 2

Probability Distribution

Section 1 – Discrete Random Variable

Section 1.1 Probability Distribution

Definitions:

A random variable is a variable whose value is determined by the outcome of a random experiment.

A random variable that assumes countable values is called a discrete random variable.

A random variable that can assume any value contained in one or more intervals is called a continuous random variable.

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Examples (of discrete random variables):

- The number of cars sold at a dealership during a given month.
- The number of people coming to a theater on a certain day.
- The number of complaints received at the office of an airline on a given day.
- The number of customers visiting a bank during any given hour.

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Examples (of continuous random variables):

- The height of a person.
- The time taken to complete an examination.
- The amount of milk in a gallon (note that we do not expect a gallon to contain exactly one gallon of milk but either slightly more or slightly less than a gallon).

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Example:

Suppose the following gives the frequency and relative frequency distributions of the vehicles owned by all 2000 families living in a small town.

Number of Vehicles Owned	Frequency
0	30
1	470
2	850
3	490
4	160

Frequency distribution of the owned vehicles

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Suppose one family is randomly selected from this population. The act of randomly selecting a family is called a random or chance experiment.

Let X denote the number of vehicles owned by the selected family. Then X can assume any of the five possible values (0, 1, 2, 3, and 4) listed in the first column of the Table.

The value assumed by X depends on which family is selected. Thus, this value depends on the outcome of a random experiment.

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Definitions:

The probability function or probability distribution $P(X)$ of a discrete random variable can be represented by a formula, a table or a graph, which provides the probabilities $P(x)$ or $P(X = x)$ or $f(x)$ corresponding to each and every value of *x*.

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Three Characteristics of a Probability Distribution

The probability distribution of a discrete random variable possesses the following three characteristics:

$0 \leq P(X) \leq 1$ for each value of x

$\sum P(X) = 1$

$P(x) = P(X = x)$

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Example (revisit):

Suppose the following gives the frequency and relative frequency distributions of the vehicles owned by all 2000 families living in a small town.

Number of Vehicles Owned	Frequency
0	30
1	470
2	850
3	490
4	160

Frequency distribution of the owned vehicles

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Let X be the number of vehicles owned by a randomly selected family.
To construct the probability distribution of X, we first compute the relative frequencies:

Number of Vehicles Owned	Frequency	Relative Frequency
0	30	0.015
1	470	0.235
2	850	0.425
3	490	0.245
4	160	0.080

Relative frequency distribution

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

When the relative frequencies are known for the population, they give the actual (theoretical) probabilities of outcomes. Using the relative frequencies, we can write the probability distribution of the discrete random variable X:

Number of Vehicles Owned (x)	Probability $P(x) = P(X = x)$
0	0.015
1	0.235
2	0.425
3	0.245
4	0.080

Probability distribution

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

From the probability distribution, the probability for any value X can be read.

For example, the probability that a randomly selected family from this town owns two vehicles is 0.425. The probability is written as

$$P(X = 2) = 0.425$$

The probability that the selected family owns more than two vehicles is given by the sum of the probabilities of three and four vehicles, respectively. This probability is written as

$$P(X > 2) = P(X = 3) + P(X = 4) = 0.325$$

Section 1 – Discrete Random Variable

Section 1.1 Probability Distribution

Definitions:

Let X be a discrete random variable with probability distribution $P(X)$.

The **mean** or **expected value** is

$$\boxed{\mu = E(X) = \sum x \cdot P(x)}$$

Let X be a discrete random variable with probability distribution $P(X)$ and mean μ.

The **variance** of X is

$$\sigma^2 = Var(X) = \sum (x-\mu)^2 \cdot P(x)$$

OR $$\sigma^2 = Var(X) = \left(\sum x^2 \cdot P(x)\right) - \mu^2$$

Section 1 – Discrete Random Variable

Section 1.1 Probability Distribution

Example:

The following table lists the probability distribution of the number of breakdowns per week for a machine based on past data.

Breakdowns per week	0	1	2	3
Probability	0.15	0.20	0.35	0.30

Find the probability that the number of breakdowns for this machine during a given week is (a) exactly two (b) zero to two (c) more than one (d) at most one.

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Solution:

Let X denote the number of breakdowns for this machine during a given week.

(a) the probability of exactly two breakdowns

$$P(X = 2) = 0.35$$

(b) the probability of zero to two breakdowns

$$P(0 \leq X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$= 0.15 + 0.20 + 0.35 = 0.70$$

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

(c) the probability of more than one breakdown

$$P(X > 1) = P(X = 2) + P(X = 3)$$

$$= 0.35 + 0.30 = 0.65$$

(d) the probability of at most one breakdown

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= 0.15 + 0.20 = 0.35$$

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

Further question:
Find the mean number of breakdowns per week for this machine.

Recall: $\mu = E(X) = \sum x \cdot P(x)$

x	$P(x)$	$x \cdot P(x)$
0	0.15	0.00
1	0.20	0.20
2	0.35	0.70
3	0.30	0.90

$$\sum x \cdot P(x) = 1.80$$

The mean is $\mu = 1.80$

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

On average, this machine is expected to break down 1.80 times per week over a period of time.

In other words, if this machine is used for many weeks, then for certain weeks we will observe no breakdowns; for some other weeks we will observe 1 breakdown per week; and for still other weeks we will observe 2 or 3 breakdowns per week.

The mean number of breakdowns is expected to be 1.80 per week for the entire period.

Section 1 – Discrete Random Variable

Section 1.1 Probability Distribution

Example:

The following table gives the probability distribution of a random variable *X*.

x	0	1	2	3
$P(X)$	0.2	0.1	0.3	0.4

Find the variance of the random variable *X*.

Section 1 – Discrete Random Variable

Section 1.1 Probability Distribution

Solution:

By definition $\sigma^2 = Var(X) = \sum (x-\mu)^2 \cdot P(x)$

x	$P(x)$	$x \cdot P(x)$	$(x-\mu)^2 \cdot P(x)$
0	0.2	0.0	0.722
1	0.1	0.1	0.081
2	0.3	0.6	0.003
3	0.4	1.2	0.484
		$\sum x \cdot P(x) = 1.9$	$\sum (x-\mu)^2 \cdot P(x) = 1.29$

The variance is $\sigma^2 = 1.29$

Section 1 – Discrete Random Variable
Section 1.1 Probability Distribution

By alternative formula

$$\sigma^2 = Var(X) = \left(\sum x^2 \cdot P(x)\right) - \mu^2$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0.2	0.0	0.0
1	0.1	0.1	0.1
2	0.3	0.6	1.2
3	0.4	1.2	3.6
		$\sum x \cdot P(x) = 1.9$	$\sum x^2 \cdot P(x) = 4.9$

The variance is $\sigma^2 = 4.9 - 1.9^2 = 1.29$

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

The Binomial Probability Distribution is one of the most widely used discrete probability distributions.

To apply the binomial probability distribution, the random variable X must be a discrete random variable and each repetition of the experiment must result in one of two possible outcomes.

An experiment that satisfies the four conditions (on next slide) is called a binomial experiment. Each repetition of a binomial experiment is called a trial or a Bernoulli trial (after Jacob Bernoulli).

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

- ✓ There are n identical trials. In other words, the given experiment is repeated n times. All these repetitions are performed under identical conditions.
- ✓ Each trial has two and only two outcomes. These outcomes are usually called a success and a failure.
- ✓ The probability of success is denoted by p and that of failure by $(1-p)$. The probabilities p and $(1-p)$ remain constant for each trial.
- ✓ The trials are independent. In other words, the outcome of one trial does not affect the outcome of another trial.

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

Example:

Five percent of all DVD players manufactured by a large electronics company are defective. Three DVD players are randomly selected from the production line of this company. The selected DVD players are inspected to determine if each of them is defective or good.

Is this experiment a binomial experiment?

Section 1 – Discrete Random Variable

Section 1.2 Binomial Probability Distribution

Solution:

- ✓ This example consists of three identical trials. A trial represents the selection of a DVD player.
- ✓ Each trial has two outcomes: A DVD player is defective or a DVD player is good. Let a defective VCR be called a success and a good DVD player be called a failure.
- ✓ Five percent of all DVD players are defective. So, the probability p that a DVD player is defective is 0.05. As a result, the probability $(1-p)$ that a DVD player is good is 0.95. These two probabilities add up to 1.

Section 1 – Discrete Random Variable

Section 1.2 Binomial Probability Distribution

- ✓ Each trial (DVD player) is independent. In other words, if one DVD player is defective it does not affect the outcome of another DVD player being defective or good. This is so because the size of the population is very large as compared to the sample size.

Since all four conditions of a binomial experiment are satisfied, this is an example of a binomial experiment.

Section 1 – Discrete Random Variable

Section 1.2 Binomial Probability Distribution

Definition: The random variable X that represents the number of successes in n trials for a binomial experiment is called a **binomial random variable**.

The probability distribution of X in such experiments is called the **binomial probability distribution** or simply **binomial distribution**.

Section 1 – Discrete Random Variable

Section 1.2 Binomial Probability Distribution

Binomial Formula

For a binomial experiment, the probability of exactly x successes in n trials is given by the binomial formula

$$P(x) = P(X = x) = \binom{n}{x} p^{x}(1-p)^{n-x} \qquad x = 0,1,\ldots,n$$

where n = total number of trials

p = probability of success

x = number of successes in n trials

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

$X \sim Bin(n, p)$ denotes that X follows a binomial distribution with parameters n and p.

The mean and variance of the binomial distribution are

$$\boxed{\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = Var(X) = np(1-p)}$$

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

Example:

The answer to a true/false question is either correct or incorrect. Assume that (*i*) an examination consists of four true/false questions and (*ii*) a student has no knowledge of the subject matter.

The chance (probability) that the student will guess the correct answer to the first question is 0.50. Likewise, the probability of guessing each of the remaining questions correctly is 0.50.

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

What is the probability of:

(a) Getting exactly none of four correct?
(b) Getting exactly one of four correct?
(c) Getting exactly two of four correct?

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

Solution:

Note that the binomial conditions are met:

✓There is a fixed number of trials (i.e. $n = 4$)

✓There are only two possible outcomes
(i.e. correct/incorrect)

✓There is a constant probability of success (i.e. $p = 0.5$)

✓The trials are independent

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

(a) The probability of getting exactly none of four correct

$$P(X=0)=\binom{4}{0}(0.5)^0(1-0.5)^{4-0}=0.0625$$

	$Q1$	$Q2$	$Q3$	$Q4$	Probability
Case 1	$\times$	$\times$	$\times$	$\times$	$0.5\times0.5\times0.5\times0.5$

$$\binom{4}{0}=1$$

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

(b) The probability of getting exactly one of four correct

$$P(X=1)=\binom{4}{1}(0.5)^1(1-0.5)^{4-1}=0.25$$

	$Q1$	$Q2$	$Q3$	$Q4$	Probability
Case 1	$\checkmark$	$\times$	$\times$	$\times$	$0.5\times0.5\times0.5\times0.5$
Case 2	$\times$	$\checkmark$	$\times$	$\times$	$0.5\times0.5\times0.5\times0.5$
Case 3	$\times$	$\times$	$\checkmark$	$\times$	$0.5\times0.5\times0.5\times0.5$
Case 4	$\times$	$\times$	$\times$	$\checkmark$	$0.5\times0.5\times0.5\times0.5$

$$\binom{4}{1}=4$$

Section 1 – Discrete Random Variable
Section 1.2 Binomial Probability Distribution

(c) The probability of getting exactly two of four correct

$$P(X = 2) = \binom{4}{2}(0.5)^2(1-0.5)^{4-2} = 0.375$$

	$Q1$	$Q2$	$Q3$	$Q4$	Probability
Case 1	✓	✓	×	×	$0.5 \times 0.5 \times 0.5 \times 0.5$
Case 2	✓	×	✓	×	$0.5 \times 0.5 \times 0.5 \times 0.5$
Case 3	✓	×	×	✓	$0.5 \times 0.5 \times 0.5 \times 0.5$
Case 4	×	✓	✓	×	$0.5 \times 0.5 \times 0.5 \times 0.5$
Case 5	×	✓	×	✓	$0.5 \times 0.5 \times 0.5 \times 0.5$
Case 6	×	×	✓	✓	$0.5 \times 0.5 \times 0.5 \times 0.5$

$$\binom{4}{2} = 6$$

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

The **Poisson probability distribution** is applied to experiments with random and independent occurrences.

The occurrences are random in the sense that they do not follow any pattern and, hence, they are unpredictable.

Independence of occurrences means that one occurrence (or nonoccurrence) of an event does not influence the successive occurrences or nonoccurrence of that event.

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

The occurrences are always considered with respect to an interval. The interval may be a time interval, a space interval, or a volume interval.

The actual number of occurrences within an interval is random and independent. If the average number of occurrences for a given interval is known, then by using the Poisson probability distribution we can compute the probability of a certain number of occurrences X in that interval.

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

Conditions to apply Poisson probability distribution

The following three conditions must be satisfied:

✓The experiment consists of counting the number, X, of times a particular event occurs during a given unit of time, or a given area or volume (or weight, or distance, or any other unit of measurement)

✓The probability that an event occurs in a given unit of time, area, or volume is the same for all the units.

✓The number of events that occur in one unit of time, area, or volume is independent of the number that occur in other units.

Section 1 – Discrete Random Variable

Section 1.3 Poisson Probability Distribution

In the Poisson probability distribution terminology, the average number or occurrences in an interval is denoted by μ .

The actual number of occurrences in that interval is denoted by X. Then, using the Poisson probability distribution, we find the probability of X occurrences during an interval given the mean occurrences are μ during that interval.

Section 1 – Discrete Random Variable

Section 1.3 Poisson Probability Distribution

Poisson Probability Distribution Formula

According to the Poisson probability distribution, the probability of x occurrences in an interval is

$$\boxed{P(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, \ldots}$$

where λ is the mean number of occurrences in that interval.

The mean and variance of the Poisson distribution are

$$\boxed{\mu = E(X) = \lambda \qquad \text{and} \qquad \sigma^2 = Var(X) = \lambda}$$

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

Example:

A washing machine in a laundry breaks down an average of three times per month.

Using the Poisson probability distribution formula, find the probability that during the next month this machine will have
(a) exactly two breakdowns;
(b) at most one breakdown.

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

Solution:

Let λ be the mean number of breakdowns per month and X be the actual number of breakdowns observed during the next month for this machine. Then, $\lambda = 3$.

$$\Rightarrow \ P(X = x) = \frac{3^x e^{-3}}{x!}$$

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

(a) The probability that exactly two breakdowns will be observed during the next month

$$P(X = 2) = \frac{3^2 e^{-3}}{2!} = 0.224$$

(b) The probability that at most one breakdown will be observed during the next month

$$P(X \leq 1) = \frac{3^0 e^{-3}}{0!} + \frac{3^1 e^{-3}}{1!} = 0.1992$$

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

Poisson Approximation to the Binomial Distribution

The computation of probabilities for a binomial distribution could be tedious if the number of trials n is large.

In the situation, the Poisson distribution can be used to approximate the binomial probabilities when the number of trials n is large and at the same time the probability p is small (generally such that $\mu = np \leq 7$).

Section 1 – Discrete Random Variable

Section 1.3 Poisson Probability Distribution

Proposition

Let X be the number of success resulting from n independent trials, each with probability of success p. The distribution of the number of success X is binomial, with mean np.

If the number of trials n is large and np is of only moderate size (i.e. $np \le 7$), this distribution can be approximated by the Poisson distribution with $\mu = np$.

Section 1 – Discrete Random Variable

Section 1.3 Poisson Probability Distribution

The probability of the approximating distribution is then:

$$\boxed{P(x) = P(X = x) = \frac{(np)^x e^{-np}}{x!} \qquad x = 0,1,2,\ldots}$$

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

Example:

An analyst predicted that 3.5% of small corporations would file for bankruptcy in the coming year.

For a random sample of 100 small corporations, estimate the probability that at least three will file for bankruptcy in the next year, assuming that the analyst's prediction is correct.

Section 1 – Discrete Random Variable
Section 1.3 Poisson Probability Distribution

Solution:

The distribution of X, the number of small corporations that will file for bankruptcy, is binomial with $n = 100$ and $p = 0.035$, so that the mean of the distribution is $np = 3.5$.

Using the Poisson distribution to approximate the probability of at least three bankruptcies:

$$P(X \geq 3) = 1 - P(X \leq 2)$$

$$= 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

$$= 1 - \frac{(3.5)^0 e^{-3.5}}{0!} - \frac{(3.5)^1 e^{-3.5}}{1!} - \frac{(3.5)^2 e^{-3.5}}{2!} = 0.679$$

Section 1.4 Hypergeometric Probability Distribution

One of the applications of the binomial probability distribution is its use as the probability distribution for the number X of favorable (or unfavorable) responses in a public opinion or market survey.

Ideally, it is applicable when sampling is with replacement, that is, each item drawn from the population is observed and returned to the population before the next item is drawn.

Section 1.4 Hypergeometric Probability Distribution

Practically speaking, the sampling in surveys is rarely conducted with replacement. Nevertheless, the binomial probability distribution is still appropriate if the number N of elements in the population is large and the sample size n is small relative to N.

When sampling is without replacement, and the number of elements N in the population is small (or when the sample size n is large relative to N), the number of "successes" in a random sample of n items has a hypergeometric probability distribution.

Section 1 – Discrete Random Variable

Section 1.4 Hypergeometric Probability Distribution

Characteristics that define a Hypergeometric Random Variable:

The experiment consists of randomly drawing n elements without replacement from a set of N elements, r of which are "successes" and $N-r$ of which are "failures".

The hypergeometric random variable X is the number of "successes" in the draw of n elements.

Section 1 – Discrete Random Variable

Section 1.4 Hypergeometric Probability Distribution

Hypergeometric Probability Distribution Formula

For a Hypergeometric Probability Distribution, the probability of exactly x "successes" in n trials is given by the Hypergeometric Probability distribution formula

$$P(x) = P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

Section 1 – Discrete Random Variable
Section 1.4 Hypergeometric Probability Distribution

where N = number of elements in the population
r = number of successes in the population
n = sample size
x = number of successes in the sample

Assumptions:

•The sample of n elements is randomly selected from the N elements of the population.

•The value of X is restricted so that all factorials are nonnegative.

Section 1 – Discrete Random Variable
Section 1.4 Hypergeometric Probability Distribution

The mean and variance of the Hypergeometric distribution are

$$\mu = E(X) = \frac{nr}{N}$$

and

$$\sigma^2 = Var(X) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

The expression $\left(\frac{N-n}{N-1}\right)$ is known as the finite population correction factor.

Section 1 – Discrete Random Variable
Section 1.4 Hypergeometric Probability Distribution

Example:

Of six lecturers, four have been with the college five or more years. If three lecturers are chosen randomly from the group of six, find the probability that exactly two will have five or more years seniority.

Section 1 – Discrete Random Variable
Section 1.4 Hypergeometric Probability Distribution

Solution:

For this problem, the number of elements in the population is $N = 6$, the number of "successes" is $r = 4$ and the sample size is $n = 3$.

The required probability is

$$P(X = 2) = \frac{\binom{4}{2}\binom{6-4}{3-2}}{\binom{6}{3}} = 0.6$$

Section 1 – Discrete Random Variable
Section 1.5 Negative Binomial Distribution

Suppose that independent trials, each having probability p, where $0 < p < 1$, of being a success are performed until a total of r successes is accumulated.

If we let X be the number of trials required, then X follows the Negative Binomial Distribution.

Section 1 – Discrete Random Variable
Section 1.5 Negative Binomial Distribution

According to the Negative Binomial distribution, the probability of performing k independent trials until a total of r successes is accumulated is:

$$\boxed{P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad k = r, r+1, \dots}$$

where p is the probability of each trial being a success.

Section 1 – Discrete Random Variable

Section 1.5 Negative Binomial Distribution

$X \sim NegBin(r, p)$ denotes that X follows a negative binomial distribution with parameters r and p.

The mean and variance of the Negative Binomial distribution with parameters r, p are

$$\boxed{\mu = E(X) = \frac{r}{p} \qquad \text{and} \qquad \sigma^2 = Var(X) = \frac{r(1-p)}{p^2}}$$

Section 1 – Discrete Random Variable

Section 1.5 Negative Binomial Distribution

Example:

If independent trials, each resulting in a success with probability p, are performed, what is the probability of r successes occurring before m failures?

Solution:

Note that r successes will occur before m failures if and only if the r^{th} success occurs no later than the $(r+m-1)^{\text{th}}$ trial.

Hence, the desired probability is $\displaystyle\sum_{k=r}^{r+m-1} \binom{k-1}{r-1} p^r (1-p)^{k-r}$

Section 1 – Discrete Random Variable
Section 1.5 Negative Binomial Distribution

Example:

A company takes out an insurance policy to cover accidents that occur at its manufacturing plant. The probability that one or more accidents will occur during any given month is 0.6. The number of accidents that occur in any given month is independent of the number of accidents that occur in all other months.

Calculate the probability that there will be at least four months in which no accidents occur before the fourth month in which at least one accident occurs.

Section 1 – Discrete Random Variable
Section 1.5 Negative Binomial Distribution

Solution:

If a month with one or more accidents is regarded as success and X be the number of failures before the fourth success, then X follows a negative binomial distribution of the requested probability is

$$P(X \geq 4) = 1 - P(X \leq 3)$$

$$= 1 - \sum_{k=0}^{3} \binom{3+k}{k} (0.6)^4 (1-0.6)^k = 0.2898$$

Section 2 – Continuous Random Variable

Section 2.1 Probability Distribution

Definitions:

A continuous random variable is defined as a random variable whose values are not countable. A continuous random variable can assume any value over an interval or intervals.

Because the number of values contained in any interval is infinite, the possible number of values that a continuous random variable can assume is also infinite.

Section 2 – Continuous Random Variable

Section 2.1 Probability Distribution

Since there are infinitely many possible values for a continuous variable, we cannot possible list those values in a table, as we did with probability distributions for discrete random variables.

The distribution of a continuous *X* must be given not by a list or table of probabilities but by a continuous curve. The curve is called frequency curve, or a probability density curve.

The area under the curve between two limits on the horizontal scale is the probability that the random variable will take on a value lying between those two limits.

Section 2 – Continuous Random Variable

Section 2.1 Probability Distribution

Definition:

If X is a continuous random variable and $y = f(x)$ is a function such that $f(x) \geq 0$ for all X, then $y = f(x)$ is called the **probability density function (pdf)** of the continuous random variable X if for any numbers a and b, the area under the graph of $y = f(x)$ from $X = a$ to $X = b$ equals the probability that X lies between a and b.

Note: The probability that a continuous random variable assumes a single value is zero. That is,

$$P(X = c) = 0 \qquad P(X < c) = P(X \leq c)$$

Discrete Random Variable

Section 2 – ~~Continuous Random Variable~~

Section 2.1 Probability Distribution

Example:

An investment firm offers its customers municipal bonds that mature after varying numbers of years. Given that the cumulative distribution of T, the number of years to maturity for a randomly selected bond is

$$F(t) = \begin{cases} 0 & t < 1 \\ 0.25 & 1 \leq t < 3 \\ 0.5 & 3 \leq t < 5 \\ 0.75 & 5 \leq t < 7 \\ 1 & t \geq 7 \end{cases}$$

Discrete Random Variable

Section 2 – ~~Continuous Random Variable~~

Section 2.1 Probability Distribution

Calculate:

(a) $P(T=5)$ (b) $P(T>3)$ (c) $P(1.4<T<6)$

Solution:

(a) $P(T=5)=0.25$

(b) $P(T>3)=1-P(T\leq 3)=1-0.5=0.5$

(c) $P(1.4<T<6)=P(T\leq 6)-P(T\leq 1.4)=0.75-0.25=0.5$

Section 2 – Continuous Random Variable

Section 2.1 Probability Distribution

Example:

Assume the value of the continuous random variable X is between $x=1$ and 3.

It has a density function given by $f(x)=0.5$

Calculate the following probabilities:

(a) $P(2<X<2.5)$

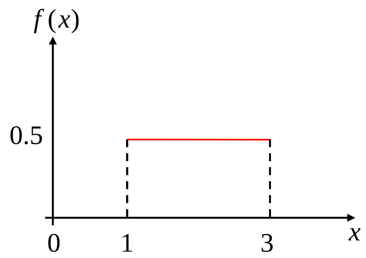
(b) $P(X\leq 16)$

Section 2 – Continuous Random Variable
Section 2.1 Probability Distribution

Solution:

(a) $P(2 < X < 2.5) = 0.25$

(b) $P(X \le 16) = 1$



Section 2 – Continuous Random Variable
Section 2.2 Normal Distribution

The **normal distribution** is the most important and most widely used of all the probability distributions. A large number of phenomena in the real world are normally distributed either exactly or approximately.

A random variable is said to have a **normal probability distribution** if and only if the probability density function of X is

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}}$$

Section 2 – Continuous Random Variable

Section 2.2 Normal Distribution

The random variable X that has a normal distribution is called a normal random variable.

Examples of continuous random variables that have all been observed to have a (approximate) normal distribution:

- the heights and weights of people
- scores on an examination
- weights of packages (e.g., cereal boxes, boxes of cookies)
- amount of milk in a gallon
- life of an item (such as a light bulb or a television set)
- the time taken to complete a certain job

Section 2 – Continuous Random Variable

Section 2.2 Normal Distribution

Characteristics of Normal Probability Distribution

A normal probability distribution, when plotted, gives a bell-shaped curve such that

- ✓ The total area under the curve is 1.0.
- ✓ The curve is symmetric about the mean.
- ✓ The two tails of the curve extend indefinitely.

Section 2 – Continuous Random Variable
Section 2.2 Normal Distribution

If X is a normal distributed random variable with parameters μ and σ, then

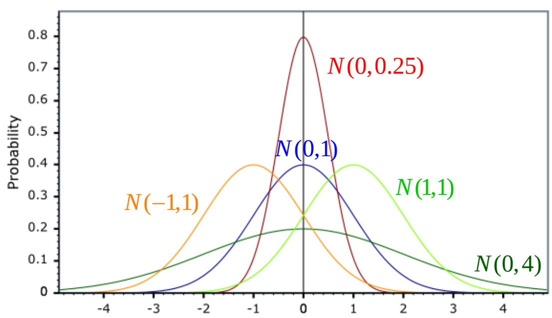
$$\boxed{E(X) = \mu \qquad \text{and} \qquad Var(X) = \sigma^2}$$

$X \sim N(\mu, \sigma^2)$ denotes that X follows a normal distribution with mean μ and σ.

The mean μ and the standard deviation σ are the parameters of the normal distribution. Each different set of values of μ and σ gives a different normal distribution.

Section 2 – Continuous Random Variable
Section 2.2 Normal Distribution

The value of μ determines the center of a normal distribution on the horizontal axis and the value of σ gives the spread of the normal distribution curve.



Section 2 – Continuous Random Variable
Section 2.2 Normal Distribution

The probability that a normal random variable X takes on a value in a specified interval (x_1, x_2) is found by determining the area enclosed by the x-axis, the lines $X = x_1$, $X = x_2$ and the bell shape curve.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Definition: The normal distribution with $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution.**

Notation: $Z \sim N(0, 1)$

The probability or area under the standard normal curve can be found from the standardized normal table.

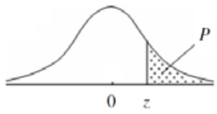
Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Table of the Standardized Normal Distribution

The table gives the probability

$$P = \Pr(Z > z)$$

where $Z \sim N(0,1)$.



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233

Examples:

$$P(Z > 1.95) = 0.0256$$

$$P(Z < 1.32) = 0.9066$$

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Standardizing a Normal Distribution

Standard normal distribution table can be used to find areas under the standard normal curve. However, in real-world applications, a (continuous) random variable may have a normal distribution with values of the mean and standard deviation different from 0 and 1, respectively.

The first step is to convert the given normal distribution to the standard normal distribution. This procedure is called standardizing a normal distribution.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Converting an *X* value to a *Z* value

For a normal random variable X, a particular value of X can be converted to a Z value by using the formula

$$\boxed{Z = \frac{X - \mu}{\sigma}}$$

where μ and σ are the mean and standard deviation of the normal distribution of X.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Example:

Let X be a continuous random variable that is normally distributed with a mean of 25 and a standard deviation of 4.

Find the area

(a) between $X = 25$ and $X = 32$

(b) between $X = 18$ and $X = 34$

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Solution:

For the given information, $\mu = 25$ $\sigma = 4$

(a) The area between $X = 25$ and $X = 32$

$$X = 25 \Rightarrow Z = \frac{25-25}{4} = 0$$

$$X = 32 \Rightarrow Z = \frac{32-25}{4} = 1.75$$

The area between $X = 25$ and $X = 32$ under the given normal distribution curve is equivalent to the area between $Z = 0$ and $Z = 1.75$ under the standard normal curve.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

The required area can be written as probability

$$P(25 < X < 32) = P(0 < Z < 1.75) = 0.4599$$

(b) The area between $X = 18$ and $X = 34$

$$P(18 < X < 34) = P\left(\frac{18-25}{4} < Z < \frac{34-25}{4}\right)$$

$$= P\left(-1.75 < Z < 2.25\right) = 0.9477$$

Section 2 – Continuous Random Variable

Section 2.3 Standard Normal Distribution

Determining the *Z* values when an area under the Normal Distribution Curve is Known

Example:
Find the value of Z such that the area under the standard normal curve in the right tail is 0.05.

Solution:

$$P(Z > z) = 0.05$$
$$\Rightarrow \; z = 1.645$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
			⋮					⋮		
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233

Section 2 – Continuous Random Variable

Section 2.3 Standard Normal Distribution

Example:
Find the value of Z such that the area under the standard normal curve in the left tail is 0.005.

Solution:

$$P(Z < z) = 0.005 \;\Rightarrow\; z = -2.575$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
			⋮					⋮		
2.5	0.00621	0.00604	0.00587	0.00570	0.00554	0.00539	0.00523	0.00508	0.00494	0.00480
2.6	0.00466	0.00453	0.00440	0.00427	0.00415	0.00402	0.00391	0.00379	0.00368	0.00357
2.7	0.00347	0.00336	0.00326	0.00317	0.00307	0.00298	0.00289	0.00280	0.00272	0.00264
2.8	0.00256	0.00248	0.00240	0.00233	0.00226	0.00219	0.00212	0.00205	0.00199	0.00193
2.9	0.00187	0.00181	0.00175	0.00169	0.00164	0.00159	0.00154	0.00149	0.00144	0.00139

Section 2 – Continuous Random Variable

Section 2.3 Standard Normal Distribution

Example:

Find a point Z such that the area under the standard normal curve between 0 and Z is 0.4251 and the value of Z is positive.

Solution: $P(0 < Z < z) = 0.4251 \Rightarrow P(Z > z) = 0.0749$

$$\Rightarrow z = 1.44$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
			⋮					⋮		
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681

Section 2 – Continuous Random Variable

Section 2.3 Standard Normal Distribution

Finding an *X* value for a Normal Distribution

To find an X value when an area under a normal distribution curve is given, first we find the Z value corresponding to that X value from the normal distribution table.

Then, the X value is calculated using the formula

$$Z = \frac{X - \mu}{\sigma} \Rightarrow \boxed{X = \mu + Z\sigma}$$

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Example:

It is known that the life of a calculator manufactured by Intal Corporation has a normal distribution with a mean of 54 months and a standard deviation of 7.65 months.

What should the warranty period be to replace a malfunctioning calculator if the company does not want to replace more than 1% of all the calculators sold?

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Solution:

Let X be the life of a calculator $\Rightarrow \; X \sim N\left(54, 7.65^2\right)$

The calculators that would be replaced are the ones that start malfunctioning during the warranty period. The company's objective is to replace at most 1% of all the calculators sold. We are to find the value of X so that the area to the left of X under the normal curve is 1% or 0.01.

$$P(Z < z) = 0.01 \;\Rightarrow\; z = -2.33$$

$$\Rightarrow\; X = 54 + (7.65)(-2.33) = 36.1755$$

Thus, the warranty period should be 36.1755 months (which can be rounded down to 36 months) if the company is to replace no more than 1% of the calculators.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Normal Distribution as an Approximation to Binomial Distribution

Usually, the normal distribution is used as an approximation to the binomial distribution when both $np > 5$ and $n(1-p) > 5$.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Using the normal distribution as an approximation to the binomial involves the following three steps:

Step 1: Compute μ and σ for the binomial distribution

$$\boxed{\mu = np \qquad \sigma = \sqrt{np(1-p)}}$$

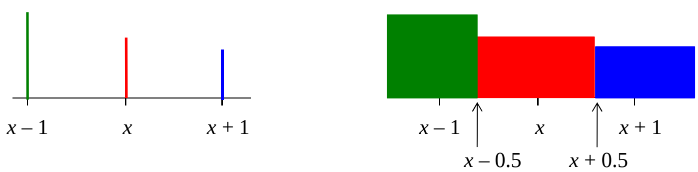
Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Step 2: Convert the discrete random variable to a continuous random variable

The normal distribution applies to a continuous random variable, whereas the binomial distribution applies to a discrete random variable. The second step in applying the normal approximation to the binomial distribution is to convert the discrete random variable to a continuous random variable by making the correction for continuity.

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

The addition of 0.5 and/or subtraction of 0.5 from the value(s) of x when the normal distribution is used as an approximation to the binomial distribution, where x is the number of successes in n trials, approximated by the interval $(x-0.5,\ x+0.5)$.



Step 3: Compute the required probability using the normal distribution

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Example:

Assume that $X \sim Bin(100, 0.36)$.
Use the normal approximation of binomial distribution to find the following probability:

(a) $P(X > 51)$
(b) $P(X \geq 51)$
(c) $P(X < 45)$
(d) $P(X \leq 45)$
(e) $P(30 < X < 40)$
(f) $P(30 < X \leq 40)$
(g) $P(30 \leq X < 40)$
(h) $P(30 \leq X \leq 40)$

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Solution:

Since $np = 36 > 5$ and $n(1-p) = 64 > 5$, we can use the normal approximation of binomial distribution. That is,

$$X \sim N(100 \times 0.36, 100 \times 0.36 \times 0.64) \Rightarrow X \sim N(36, 23.04)$$

(a) $P(X > 51) = P(X \geq 51.5) = P\left(Z \geq \dfrac{51.5-36}{\sqrt{23.04}}\right)$

$= P(Z \geq 3.22) = 0.00064$

Section 2 – Continuous Random Variable

Section 2.3 Standard Normal Distribution

(b) $P(X \geq 51) = P(X \geq 50.5) = P\left(Z \geq \dfrac{50.5-36}{\sqrt{23.04}}\right)$

$= P(Z \geq 3.02) = 0.00126$

(c) $P(X < 45) = P(X \leq 44.5) = P\left(Z \geq \dfrac{44.5-36}{\sqrt{23.04}}\right)$

$= P(Z \leq 1.77) = 0.9616$

(d) $P(X \leq 45) = P(X \leq 45.5) = P(Z \leq 1.98) = 0.9761$

Section 2 – Continuous Random Variable

Section 2.3 Standard Normal Distribution

(e) $P(30 < X < 40) = P(30.5 \leq X \leq 39.5)$

$= P(-1.15 \leq Z \leq 0.73) = 0.6422$

(f) $P(30 < X \leq 40) = P(30.5 \leq X \leq 40.5)$

$= P(-1.15 \leq Z \leq 0.94) = 0.7013$

(g) $P(30 \leq X < 40) = P(29.5 \leq X \leq 39.5)$

$= P(-1.35 \leq Z \leq 0.73) = 0.6788$

(h) $P(30 \leq X \leq 40) = P(29.5 \leq X \leq 40.5)$

$= P(-1.35 \leq Z \leq 0.94) = 0.7379$

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Example:

According to an estimate, 50% of the people in America have at least one credit card. If a random sample of 30 persons is selected, what is the probability that 19 of them will have at least one credit card?

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Solution:

Let X be the number of persons in the sample who have at least one credit card.

$$\Rightarrow \; X \sim Bin(30, 0.5)$$

Required probability:

$$P(X = 19) = \binom{30}{19}(0.5)^{19}(1-0.5)^{11} = 0.0509$$

(Exact Solution)

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Now let us solve this problem using the normal distribution as an approximation to the binomial distribution. For this example,

$$np = 15 > 5 \qquad n(1-p) = 15 > 5$$

→ We can use the normal distribution as an approximation to solve the problem.

$$X \sim N(30\times 0.5,\, 30\times 0.5\times 0.5) \;\Rightarrow\; X \sim N(15,\, 7.5)$$

Section 2 – Continuous Random Variable
Section 2.3 Standard Normal Distribution

Required probability:

$$P(X = 19) = P(18.5 \le X \le 19.5)$$

$$= P\left(\frac{18.5-15}{\sqrt{7.5}} \le Z \le \frac{19.5-15}{\sqrt{7.5}}\right)$$

$$= P(1.28 \le Z \le 1.64) = 0.0498$$

Using the binomial formula, the exact probability is 0.0509. The error due to using the normal approximation is 0.0011. Thus, the exact probability is underestimated by 0.0011 if the normal approximation is used.

Section 2 – Continuous Random Variable

Section 2.4 Uniform Probability Distribution

A continuous uniform probability distribution is a simple distribution with a rectangular shape, and it is useful in a diverse number of applications.

For example:
The time that a commuter waits to board a MTR train from Central to Causeway Bay has a uniform distribution.

Section 2 – Continuous Random Variable

Section 2.4 Uniform Probability Distribution

Definition: A continuous random variable X is said to have a continuous uniform probability distribution on the interval (a, b) if and only if the probability density function of X is

$$\boxed{f(x) = \begin{cases} 1/(b-a) & \text{if } a \le x \le b \\ 0 & \text{elsewhere} \end{cases}}$$

Section 2 – Continuous Random Variable
Section 2.4 Uniform Probability Distribution

Remarks:

The values of a and b are parameters or constants, and their particular values depend on the probability problem.
The smallest value the random variable can take is a and the largest is b.

There is a different uniform distribution from the family of uniform distributions for each pair of values for a and b.

Section 2 – Continuous Random Variable
Section 2.4 Uniform Probability Distribution

The mean and the variance of a continuous uniform probability distribution X are

$$\boxed{\mu = E(X) = \frac{a+b}{2} \qquad \text{and} \qquad \sigma^2 = Var(X) = \frac{(b-a)^2}{12}}$$

Section 2 – Continuous Random Variable
Section 2.4 Uniform Probability Distribution

The probability that a continuous uniform random variable X takes on a value in a specified interval (x_1, x_2) is found by determining the corresponding area of a rectangle under the probability density function.

That is, $P(x_1 < X < x_2) = (width)\times(height) = \boxed{(x_2 - x_1)\times\frac{1}{b-a}}$ for $a \le x_1 < x_2 \le b$.

If $x_1 < a$ or $x_2 > b$, then the width of the rectangular area is reduced accordingly.

Section 2 – Continuous Random Variable
Section 2.4 Uniform Probability Distribution

Example:

An industrial psychologist has determined that it takes a worker between 9 and 15 minutes to complete a task on an automobile assembly line.
If the continuous random variable–time to complete the task is uniformly distributed over the interval $9 \le X \le 15$, then determine for this distribution:

(a) $f(x), E(X), Var(X)$
(b) $P(X < 13)$
(c) $P(4 < X < 7)$
(d) $P(6 < X < 18)$

Section 2 – Continuous Random Variable
Section 2.4 Uniform Probability Distribution

Solution:

(a) $f(x), E(X), Var(X)$

$$f(x)=\begin{cases}\dfrac{1}{15-9}=\dfrac{1}{6} & \text{if } 9\le x\le 15\\ 0 & \text{elsewhere}\end{cases}$$

$$E(X)=\frac{a+b}{2}=\frac{9+15}{2}=12$$

$$Var(X)=\frac{(b-a)^2}{12}=\frac{(15-9)^2}{12}=3$$

Section 2 – Continuous Random Variable
Section 2.4 Uniform Probability Distribution

(b) $P(X<13)=\dfrac{13-9}{6}=\dfrac{2}{3}=0.6667$

(c) $P(4<X<7)=0$ because $f(x)=0$ everywhere except in the interval $9\le X\le 15$.

(d) $P(6<X<18)=P(9<X<15)=\dfrac{15-9}{6}=1$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

A continuous probability distribution that is often useful in describing the time it takes to complete a task is the exponential probability distribution.

The exponential random variable can be used to describe such things as:
- The time between arrivals at a car wash
- The time required to load a truck
- The distance between major defects in a highway

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Definition: A continuous random variable X is said to have an exponential distribution if and only if the probability density function of X is:

$$\boxed{f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad x \geq 0, \lambda > 0}$$

Denote $\lambda = \frac{1}{\mu}$.

The mean and the variance of the exponential distribution are

$$\boxed{E(X) = \mu \qquad \text{and} \qquad Var(X) = \mu^2}$$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Computing Probabilities for the Exponential Distribution

As with any continuous probability distribution, the area under the curve corresponding to some interval provides the probability that the random variable takes on a value in that interval. In order to compute exponential probabilities such as those described, we make use of the following formulas:

$$P(X \le c) = 1 - e^{-c\lambda}$$
$$P(X \ge c) = e^{-c\lambda}$$
$$P(c \le X \le d) = e^{-c\lambda} - e^{-d\lambda}$$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Example:

Suppose X has an exponential distribution with mean equal to 10. Determine the following:

(a) $P(X > 10)$
(b) $P(X < 20)$
(c) $P(15 < X < 30)$
(d) Find the value of x such that $P(X < x) = 0.95$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Solution:

Given $E(X) = \mu = 10 \Rightarrow \lambda = \frac{1}{\mu} = 0.1$

(a) $P(X > 10) = P(X \geq 10) = e^{-(10)(0.1)} = 0.3679$

(b) $P(X < 20) = P(X \leq 20) = 1 - e^{-(20)(0.1)} = 0.8647$

(c) $P(15 < X < 30) = P(15 \leq X \leq 30)$
$= e^{-(15)(0.1)} - e^{-(30)(0.1)} = 0.1733$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

(d) Find the value of x such that $P(X < x) = 0.95$

$$P(X < x) = 0.95$$

$$1 - e^{-(0.1)x} = 0.95$$

$$e^{-0.1x} = 0.05$$

$$-0.1x = \ln(0.05)$$

$$x = 29.957$$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Relationship Between the Poisson and Exponential Distributions

Recall that the Poisson distribution, as a discrete probability distribution, is often useful when dealing with the number of occurrences of an event over a specified interval of time or space.

The continuous exponential probability distribution is related to the discrete Poisson distribution in that, whilst the Poisson distribution provides an appropriate description of the number of occurrences per interval, the exponential distribution provides a description of the length of the interval between occurrences.

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

To provide an example that illustrates this relationship, suppose that the number of cars that arrive at a car wash during 1 hour is described by a Poisson probability distribution with a mean of 10 cars per hour.

Thus the Poisson probability function that provides the probability of *x* arrivals per hour is

$$f(x) = P(X = x) = \frac{e^{-\mu}\mu^{x}}{x!} = \frac{e^{-10}(10)^{x}}{x!}$$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Since the average number of arrivals is 10 cars per hour, the average time between cars arriving is

$$\frac{1 \text{ hour}}{10 \text{ cars}} = 0.1 \text{ hour / car}$$

Thus the corresponding exponential distribution that describes the time between the arrival of cars has a mean of $\mu = 1/\lambda = 0.1$ hour per car. Thus the exponential probability density function that give the time y hours between arrivals is

$$f(y) = 10e^{-10y}$$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Example:

The number of customers arriving at a teller's window at a bank follows the Poisson distribution with a mean rate of $\mu = 0.75$ customer per minute. If the time between arrivals is less than or equal to three minutes, then the teller can provide banking services without irritating customers with annoying waiting times.

(a) Find the mean and standard deviation of the time X between customer arrivals at the teller's window.

(b) Find the proportion of customers for whom the teller provides service without an annoying delay.

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

Solution:

(a) Find the mean and standard deviation of the time X between customer arrivals at the teller's window.

Since the number of arrivals in one minute has a Poisson distribution, the time between arrivals is exponentially distributed.

The mean and standard deviation of the exponential X is

$$E(X) = \frac{1}{\lambda} = \frac{1}{0.75 \text{ arrival/min}} = 1.33 \text{ min / arrival}$$

$$Var(X) = \frac{1}{\lambda^2} = (1.33)^2 \Rightarrow \sigma_X = \sqrt{Var(X)} = 1.33$$

Section 2 – Continuous Random Variable
Section 2.5 Exponential Distribution

(b) Find the proportion of customers for whom the teller provides service without an annoying delay.

The teller can provide services without an annoying delay if $X \leq 3$ minutes. Thus, the required probability is

$$P(X \leq 3) = 1 - e^{-3/1.33} = 0.895$$

Consequently, about 89.5% of the customers will not be irritated by an annoying delay.