Probability and Statistics for Engineers

Chapter 3: Discrete Random Variables

Lecturer

Dr. Monjed H. Samuh

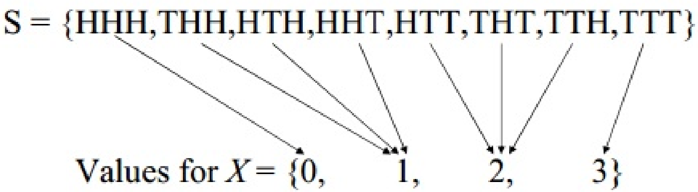
Applied Mathematics & Physics Department
Palestine Polytechnic University
(monjedsamuh@ppu.edu)

Term 192

Table of Contents

Discrete Random Variables

- A **random variable** X associates the outcomes of a random experiment to a number on the real line.
- **For example:**
  Toss a coin three times. Let the random variable X be the number of observed tails.



Example

The random variable is the number of nonconforming solder connections on a printed circuit board with 2000 connections.

Example

A batch of 500 machined parts contains 10 that do not conform to customer requirements. The random variable is the number of parts in a sample of five parts that do not conform to customer requirements.

Discrete Random Variables

Example

A batch of 500 machined parts contains 10 that do not conform to customer requirements. Parts are selected successively, without replacement, until a nonconforming part is obtained. The random variable is the number of parts selected.

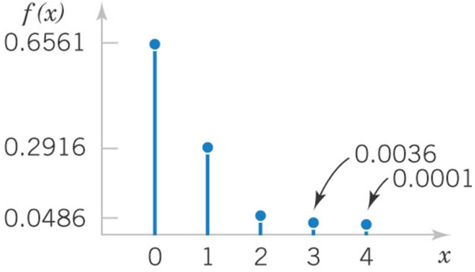
probability distribution and Probability Mass Functions

- The **probability distribution** of the random variable X is a description of the probabilities with the possible numerical values of X.

- A probability distribution of a discrete random variable can be:
  - A list of the possible values along with their probabilities.
  - A formula that is used to calculate the probability in response to an input of the random variable's value.

probability distribution and Probability Mass Functions

- A list of the possible values along with their probabilities.



x	0	1	2	3	4
$P(X = x) = p(x)$	0.6561	0.2916	0.0486	0.0036	0.0001

- We may write: $f(x)$ or $P(X = x)$ or $p(x)$ to denote *pmf*.

probability distribution and Probability Mass Functions

Definition (Probability Mass Functions)

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a **probability mass function (*pmf*)** $f(x)$ is a function such that

1. $f(x_i) \geq 0$.
2. $\sum_{i=1}^{n} f(x_i) = 1$.
3. $f(x_i) = P(X = x_i) = p(x_i)$.

Example

Toss a coin three times. Let the random variable X be the number of observed tails.

outcomes	$\{HHH\}$	$\{THH, HTH, HHT\}$	$\{HTT, THT, TTH\}$	$\{TTT\}$
x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Table: pmf for the number of tails observed

probability distribution and Probability Mass Functions

Example

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Table: pmf for the number of tails observed

1. $P(X = 2)$.
2. $P(X < 3)$.
3. $P(1 < X \leq 3)$.
4. $P(X = 1.5)$.
5. $P(X < 1.5)$.

Example

Let $f(x) = \frac{2x+c}{25}, \quad x = 0, 1, 2, 3, 4.$

1. Find the constant c.
2. $P(X = 4)$.
3. $P(X \leq 1)$.
4. $P(2 \leq X < 4)$.
5. $P(X > -10)$.

Example

An optical inspection system is to distinguish among different part types. The probability of a correct classification of any part is 0.98. Suppose that three parts are inspected and that the classifications are independent. Let the random variable X denote the number of parts that are correctly classified. Determine the probability mass function of X.

x	
$p(x)$	

Cumulative Distribution Functions

Definition (Cumulative Distribution Function)

The **cumulative distribution function (cdf)**, $F(x)$, of a random variable X is given by

$$F(x) = P(X \leq x).$$

For a discrete random variable X, $F(x)$ satisfies the following properties:

1. $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.
2. $0 \leq F(x) \leq 1$.
3. If $x \leq y$, then $F(x) \leq F(y)$.

Cumulative Distribution Functions

For example:

x	0	1	2	3
$p(x) = f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$F(x) = P(X \leq x)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	$\frac{8}{8}$

Table: pmf and cdf for the number of tails observed

The *pmf* is

$$f(x) = \begin{cases} \frac{1}{8} & \text{if } x = 0, 3 \\ \frac{3}{8} & \text{if } x = 1, 2 \end{cases},$$

The *cdf* is given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{4}{8} & \text{if } 1 \leq x < 2 \\ \frac{7}{8} & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases},$$

Cumulative Distribution Functions

Example

The thickness of wood paneling (in inches) that a customer orders is a random variable with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & \text{if } x < \frac{1}{8} \\ 0.2 & \text{if } \frac{1}{8} \leq x < \frac{1}{4} \\ 0.9 & \text{if } \frac{1}{4} \leq x < \frac{3}{8} \\ 1 & \text{if } \frac{3}{8} \leq x \end{cases},$$

1. What are the values of X.
2. Obtain the probability mass function of X.

Cumulative Distribution Functions

Example

In the previous example find the following:

1. $P(X \leq \frac{1}{18})$.
2. $P(X \leq \frac{1}{8})$.
3. $P(X \leq \frac{1}{4})$.
4. $P(X > \frac{1}{4})$.
5. $P(X = \frac{1}{4})$.
6. $P(X \leq \frac{1}{2})$.

Mean and Variance of a Discrete Random Variable

Definition (Mean of a Discrete Random Variable)

The **mean** or **expected value** or **expectation** of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_{\forall x} x f(x).$$

Definition (Variance of a Discrete Random Variable)

The **variance** of the discrete random variable X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_{\forall x} (x-\mu)^2 f(x).$$

- It can be shown that $\sigma^2 = E(X^2) - \mu^2$, where $E(X^2) = \sum_{\forall x} x^2 f(x)$.
- The **standard deviation** of X is σ.
- $E\left(h(X)\right) = \sum_{\forall x} h(x) f(x)$.

Mean and Variance of a Discrete Random Variable

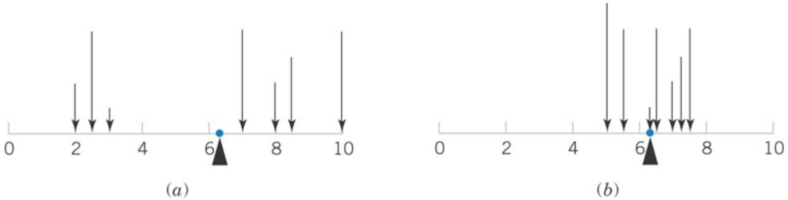


Figure: Parts (a) and (b) illustrate **equal means**, but Part (a) illustrates a **larger variance**.

Mean and Variance of a Discrete Random Variable

Example

x	0	1	2	3	Total
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1
$xf(x)$					
$x^2 f(x)$					

Table: pmf for the number of tails observed

- Mean: $\mu =$ ....
- Variance: $\sigma^2 =$ ....
- Standard Deviation: $\sigma =$ ....

Mean and Variance of a Discrete Random Variable

Example

The range of the random variable X is $\{0, 1, 2, 3, a\}$ where a is unknown. If each value is equally likely and the mean of X is 6, determine a.

Discrete Uniform Distribution

Definition (Discrete Uniform Distribution)

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x = x_1, x_2, \ldots, x_n \\ 0 & \text{otherwise} \end{cases},$$

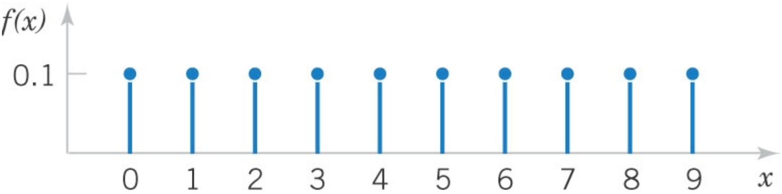


Figure: Example of a Discrete Uniform Distribution

Discrete Uniform Distribution

Definition (Mean and Variance)

Suppose X is a discrete uniform random variable on the consecutive integers $a, a+1, a+2, \ldots, b$, for $a \leq b$.

- The mean (or the expected value) of X is

$$\mu = E(X) = \frac{a+b}{2}.$$

- The variance of X is

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12}.$$

Discrete Uniform Distribution

Example

When a die is rolled, each element of the sample space $S = \{1, 2, 3, 4, 5, 6\}$ occurs with probability $\frac{1}{6}$. Therefore, we have a discrete uniform distribution, with

$$f(x) = \begin{cases} \frac{1}{6} & \text{if } x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases},$$

The mean of X is

The variance of X is

Example

The lengths of plate glass parts are measured to the nearest tenth of a millimeter. The lengths are uniformly distributed, with values at every tenth of a millimeter starting at 590.0 and continuing through 590.9. Determine the mean and variance of the lengths.

Discrete Uniform Distribution

Example

Suppose that X has a discrete uniform distribution on the integers 0 through 9. Determine the mean, variance, and standard deviation of the random variable $Y = 5X$.

Binomial Distribution

Definition (Bernoulli Trial)

A Bernoulli trial is a random experiment with exactly two possible outcomes, "success" and "failure", in which the probability of success is the same every time the experiment is conducted.

Definition (Binomial Experiment)

A random experiment consists of n Bernoulli trials such that:

1. The trials are independent.
2. Each trial results in only two possible outcomes, labeled as "success" and "failure".
3. The probability of a success in each trial, denoted as p, remains constant.

Binomial Distribution

Definition (Binomial Distribution)

The random variable X that equals the number of trials that result in a success has a binomial random variable with parameters n and $0 < p < 1$. The probability mass function of X is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Definition (Mean and Variance)

If X is a binomial random variable with parameters n and p, then

$$\mu = E(X) = np,$$

and

$$V(X) = \sigma^2 = np(1-p).$$

Binomial Distribution

Example

Toss a fair coin 3 times. Let X be the number of Heads observed.

Binomial Distribution

Example

The random variable X has a binomial distribution with $n = 10$ and $p = 0.01$. Determine the following probabilities.

1. $P(X = 5)$.
2. $P(X \leq 2)$.
3. $P(X \geq 9)$.
4. $P(3 \geq X < 5)$.

Binomial Distribution

Example

Determine the cumulative distribution function of a binomial random variable with $n = 3$ and $p = \frac{1}{2}$.

Binomial Distribution

Example

A multiple-choice test contains 25 questions, each with four answers. Assume a student just guesses on each question.

1. What is the probability that the student answers more than 20 questions correctly?

2. What is the probability the student answers less than five questions correctly?

Binomial Distribution

Example

An electronic product contains 40 integrated circuits. The probability that any integrated circuit is defective is 0.01, and the integrated circuits are independent. The product operates only if there are no defective integrated circuits. What is the probability that the product operates?

Binomial Distribution

Example

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Let X be the number of samples that contain the pollutant in the next 18 samples analyzed.

1. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

2. Compute the expected value and variance of X.

Binomial Distribution

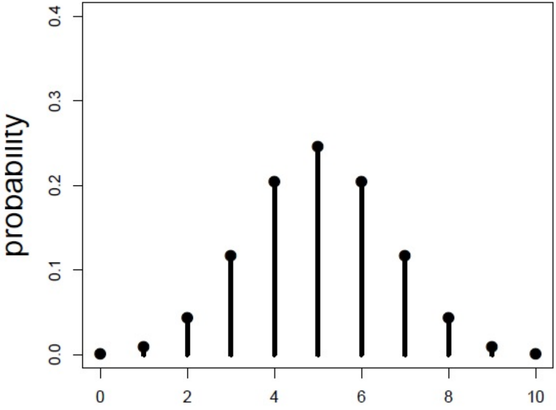


Figure: n=10, p=0.5: equal chance of success/failure

Binomial Distribution

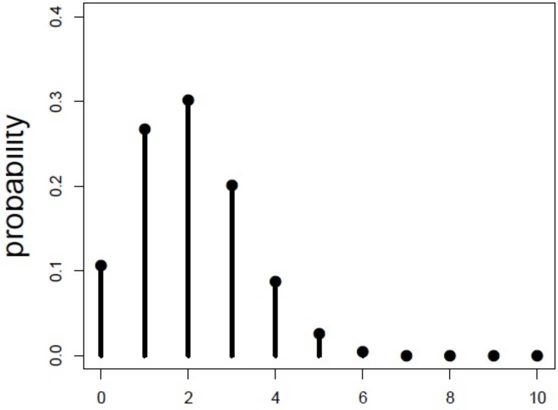


Figure: n=10, p=0.2: small chance of success

Binomial Distribution

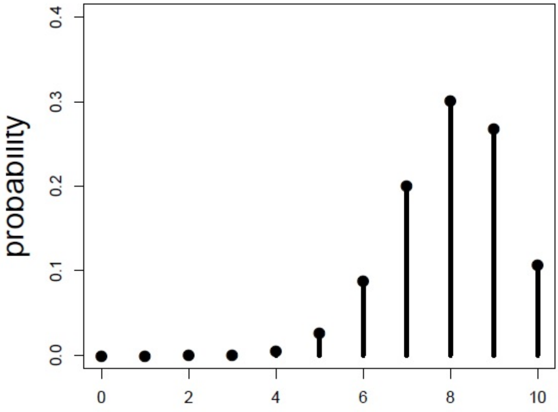


Figure: n=10, p=0.8: large chance of success

Geometric Distribution

Definition (Geometric Distribution)

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote **the number of trials until the first success**. Then X is a geometric random variable with parameter $0 < p < 1$ and

$$f(x) = p(1-p)^{x-1}, \quad x = 1, 2, \ldots.$$

Definition (Mean and Variance)

- The mean of a geometric random variable is given by

$$\mu = E(X) = \frac{1}{p}.$$

- The variance of a geometric random variable is given by

$$\sigma^2 = V(X) = \frac{1-p}{p^2}.$$

Geometric Distribution

Example

Suppose the random variable X has a geometric distribution with $p = 0.5$. Determine the following probabilities:

1. $P(X = 4)$.

2. $P(X > 2)$.

Geometric Distribution

Example

Suppose the random variable X has a geometric distribution with a mean of 2.5. Determine the following:

1. $P(X = 4)$.

2. $V(X)$.

Geometric Distribution

Example

The probability of a successful optical alignment in the assembly of an optical data storage product is 0.8. Assume the trials are independent.

1. What is the probability that the first successful alignment requires exactly four trials?

2. What is the probability that the first successful alignment requires at most four trials?

3. What is the probability that the first successful alignment requires at least four trials?

Hypergeometric Distribution

Definition (Hypergeometric Distribution)

A set of N objects contains K objects classified as successes $N-K$ objects classified as failures. A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$. Let the random variable X denote **the number of successes in the sample**. Then X is a hypergeometric random variable and

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \quad x = \max\{0, n+K-N\}, \ldots, \min\{K, n\}.$$

Definition (Mean and Variance)

- The mean of a hypergeometric random variable is given by

$$\mu = E(X) = np, \quad p = \frac{K}{N}.$$

- The variance of a hypergeometric random variable is given by

$$\sigma^2 = V(X) = np(1-p)\left(\frac{N-n}{N-1}\right), \quad p = \frac{K}{N}.$$

Hypergeometric Distribution

Example

Suppose the random variable X has a hypergeometric distribution with $N = 100$, $n = 4$ and $K = 20$. Determine the following:

1. $P(X = 2)$.
2. $P(X \leq 2)$.
3. $E(X)$.
4. σ^2.

Example

Printed circuit cards are placed in a functional test after being populated with semiconductor chips. A lot contains 140 cards, and 20 are selected without replacement for functional testing.

1. If 20 cards are defective, what is the probability that at least 1 defective card is in the sample?

2. If 5 cards are defective, what is the probability that at least one defective card appears in the sample?

Poisson Distribution

- In many applications, we are interested in counting the number of occurrences of an event in a certain time period or in a certain region in space.
- The Poisson random variable arises in situations where the events occur completely at random in time or space.
- **Examples:**
  - The number of errors a typist makes in a 5 minute period.
  - The number of telephone calls per hour received by an office.
  - The number of bomb hits in a given area.
  - The number crossover events along a section of paired chromosomes.

Poisson Distribution

Definition (Poisson Distribution)

The random variable X that equals the number of events in a Poisson process is a Poisson random variable with parameter $\lambda > 0$, and the probability mass function of X is

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots.$$

Definition (Mean and Variance)

- The mean of a Poisson random variable is given by

$$\mu = E(X) = \lambda.$$

- The variance of a Poisson random variable is given by

$$\sigma^2 = V(X) = \lambda.$$

Poisson Distribution

Example

Suppose the random variable X has a Poisson distribution with a mean of 4. Determine the following:

1. $P(X = 0)$.
2. $P(X \leq 2)$.
3. $E(X)$.
4. σ^2.

Poisson Distribution

Example

Let the random variable X be the number of customers who enter a bank in an hour, and suppose that $P(X = 0) = 0.05$. Determine the mean and variance of X.

Poisson Distribution

Example

In a 400-page manuscript, there are 200 randomly distributed misprints. If a page is selected at random,

1. find the probability that it has exactly 1 misprint.
2. find the probability that it has at least 2 misprints.
3. find the probability that it has at most 1 misprint.

Poisson Distribution

Example

The number of failures of a testing instrument from contamination particles on the product is a Poisson random variable with a mean of 0.02 failure per hour.

1. What is the probability that the instrument does not fail in an eight-hour shift?

2. What is the probability of at least one failure in a 24-hour day?