Stony Brook University

Discrete Random Variables

Vibha Mane
The COSINE Lab
Department of Electrical and Computer Engineering

Topics

- Definition of Random Variables
- Characterization of Discrete Random Variables
  - Cumulative Distribution Function (CDF)
  - Probability Mass Function (PMF)
- Examples of Discrete Random Variables
  - Bernoulli
  - Binomial
- Additional Examples of Discrete Random Variables
  - Geometric
  - Poisson

I. Meaning of a Random Variable

- A **random variable (RV)** is a function whose domain is the sample space Ω.
- To every elementary outcome ω in Ω, we assign a real value X(ω).
- A **discrete random variable** takes values in a **countable set**, such as $0,1, 2, \cdots$
- For example, a coin toss has two outcomes $\Omega = \{h, t\}$. We assign the values $X(h) = 1$ and $X(t) = 0$.
- In this case, X is a discrete random variable which takes two values $\{0,1\}$. We call this set the **support** of X.

Definition of a Random Variable (Formal)

- A random variable is a **measurable function** $X: \Omega \rightarrow E$, from a set of possible outcomes Ω to a **measure space** E.
- The set of integers $\mathcal{Z}$, with counting measure, is an example of measure space
- The set of real numbers $\mathcal{R}$, with Lebesgue measure, is another example of measure space.
- A **discrete random variable *X*** takes values in a countable set, as in $\mathcal{Z}$.
- A **continuous random variables *X*** takes values in an uncountable set, as in $\mathcal{R}$.

Example - Three Coin Toss

Example

Toss three fair coins. There are $2^3 = 8$ outcomes given by

$\Omega = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$.

Define the random variable **X to be the number of heads**. Then X takes one of the values $\{0,1,2,3\}$, with probabilities given by

$P(X = 0) = P(\{ttt\}) = {}^{1}/_{8}$

$P(X = 1) = P(\{htt, tht, tth\}) = {}^{3}/_{8}$

$P(X = 2) = P(\{hht, hth, thh\}) = {}^{3}/_{8}$

$P(X = 3) = P(\{hhh\}) = {}^{1}/_{8}$

II. Characterization of Discrete Random Variables

- The discrete random variables are characterized by
  - Cumulative Distribution Function (CDF)
  - Probability Mass Function (PMF)

Cumulative Distribution Function

- The **cumulative distribution function (CDF)** of a random variable X is the function

  $F_X(x): \mathcal{R} \rightarrow [0,1]$, given by

  $F_X(x) = P(X \leq x)$, defined for both discrete and continuous random variables.

- Note that the uppercase X is used for the random variable and the lower case x is used for a specific value.

- For example, $F_X(3) = P(X \leq 3)$ means probability that the random variable X is less than or equal to 3.

Properties of CDF

The CDF of a random variable X has the following properties:

- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to \infty} F_X(x) = 1$
- If $x_1 < x_2$, then $F_X(x_1) \leq F_X(x_2)$, that is, $F_X(x)$ is a non-decreasing function of x.
- $F_X(x^+) = F_X(x)$, that is, $F_X(x)$ is right continuous.
- Note that the above properties are necessary and sufficient conditions for $F_X(x)$ to be a CDF.

Additional Properties of CDF

The CDF has the following additional properties:

- $P(X > x) = 1 - P(X \le x) = 1 - F_X(x)$
- $P(x_1 \le X \le x_2) = F_X(x_2) - F_X(x_1)$
- $P(X = x) = F_X(x) - F_X(x^-)$

Probability Mass Function

- The **probability mass function (PMF)** of a discrete random variable X is the function

  $f_X(x) \colon \mathcal{R} \to [0,1]$, given by

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

- For discrete random variable which takes values $\{x_1, x_2, x_3 \cdots\}$, we have

  $\sum_{i=1}^{\infty} f_X(x_i) = 1$, that is, **the PMF over its support sums up to 1**.

- **Notation**
  - We use uppercase $F_X(x)$ for CDF
  - We use lowercase $f_X(x)$ for PMF
  - We also use $p_X(x)$ for PMF

Example - Dice Roll (1 of 3)

Example

Roll a **fair** dice. Define the random variable X to be the number which shows on the dice. Determine the CDF and PMF of X.

Solution

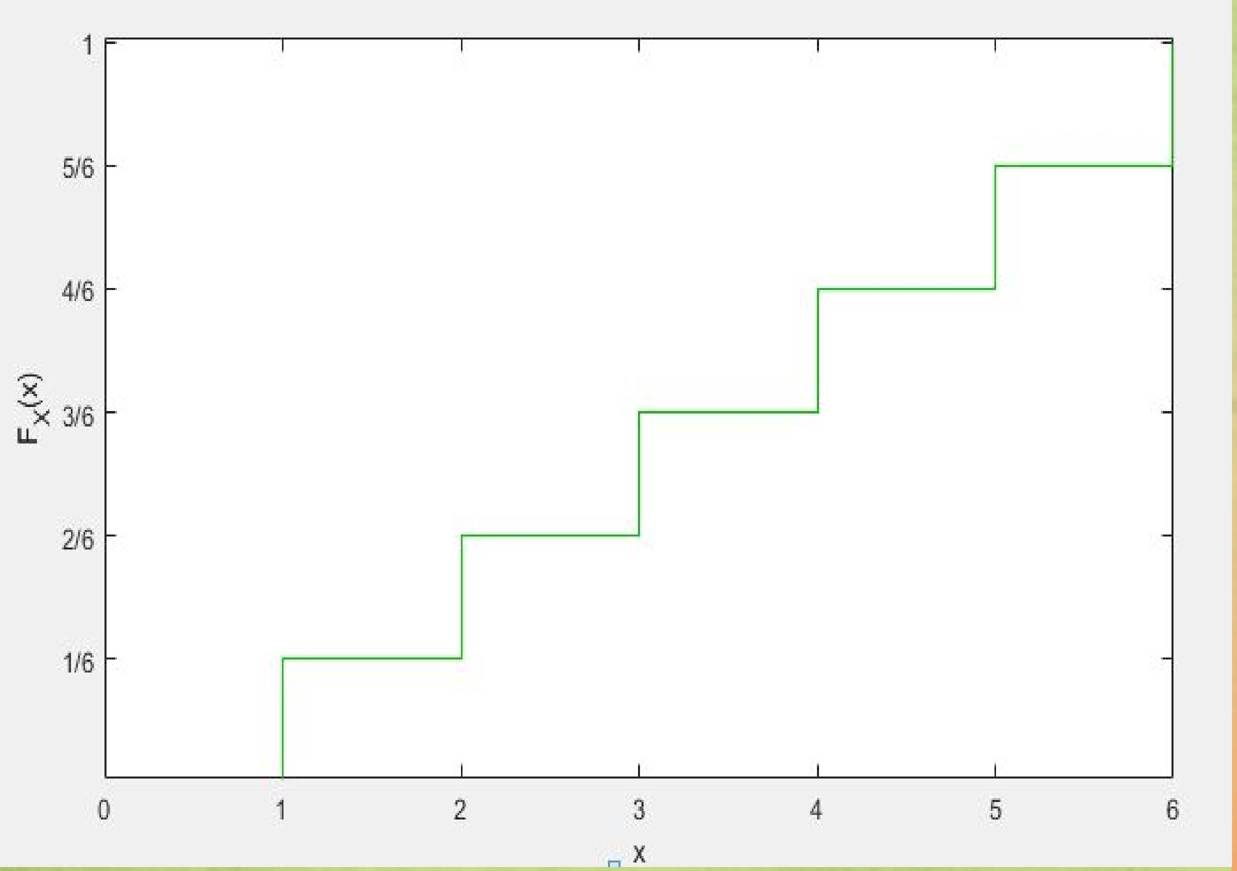
X takes a value in the set $\{1,2,3,4,5,6\}$, each with probability ${}^1/_6$.

The CDF is a **staircase function** with **"jumps"** at 1,2,3,4,5 and 6.

$F_X(1) = P(X \leq 1) = P(1) = {}^1/_6$

Also, $F_X(2^-) = {}^1/_6$, that is, $F_X(.)$ remains the same in the interval $[1,2)$

Example - Dice Roll (2 of 3)



The next "jump" is at $x = 2$

$F_X(2) = P(X \leq 2) = P(\{1,2\}) = {}^2/_6$

Continuing this way, we have

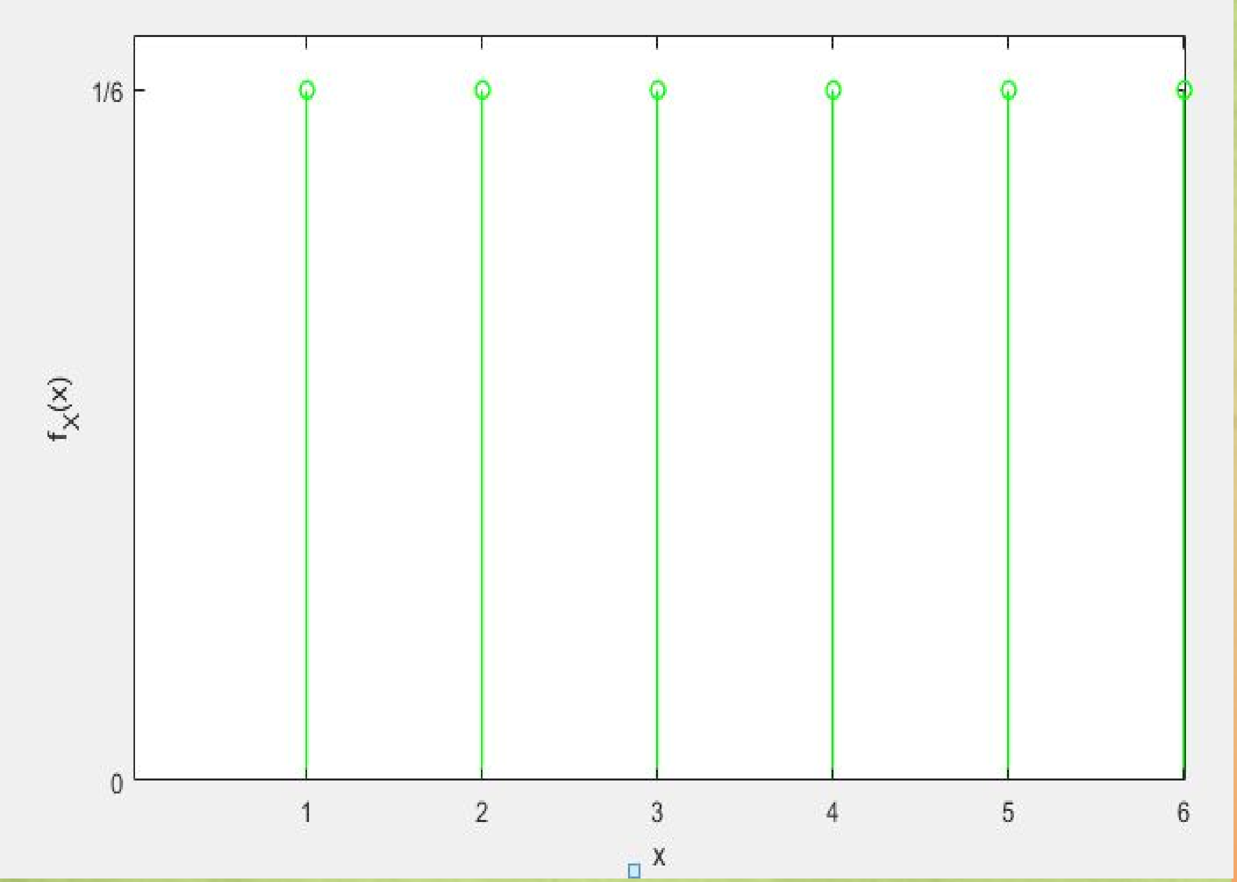
$F_X(5) = P(X \leq 5) = P(\{1,2,3,4,5\}) = {}^5/_6$

$F_X(6) = P(X \leq 6) = P(\{1,2,3,4,5,6\}) = 1$

As can be seen from the plot, the CDF satisfies the properties $F_X(-\infty) = 0$ and $F_X(\infty) = 1$.

You can verify the additional properties of the CDF.

Example - Dice Roll (3 of 3)



The PMF is given by

$f_X(x) = P(X = i) = {}^1/_6$, for i = 1,2,3,4,5,6

It is a **comb function**, as shown here.

Relationship between CDF and PMF

We can compute CDF from PMF and vice versa. The relationship is given by

$F_X(x) = \sum_{u=-\infty}^{x} f_X(u)$,

where u is a **dummy variable** of summation.

Example - Roll A Dice Twice

Example

A fair dice is rolled twice. Define the random variable X to be the **maximum of two rolls**. Determine the CDF and PMF of X.

Solution

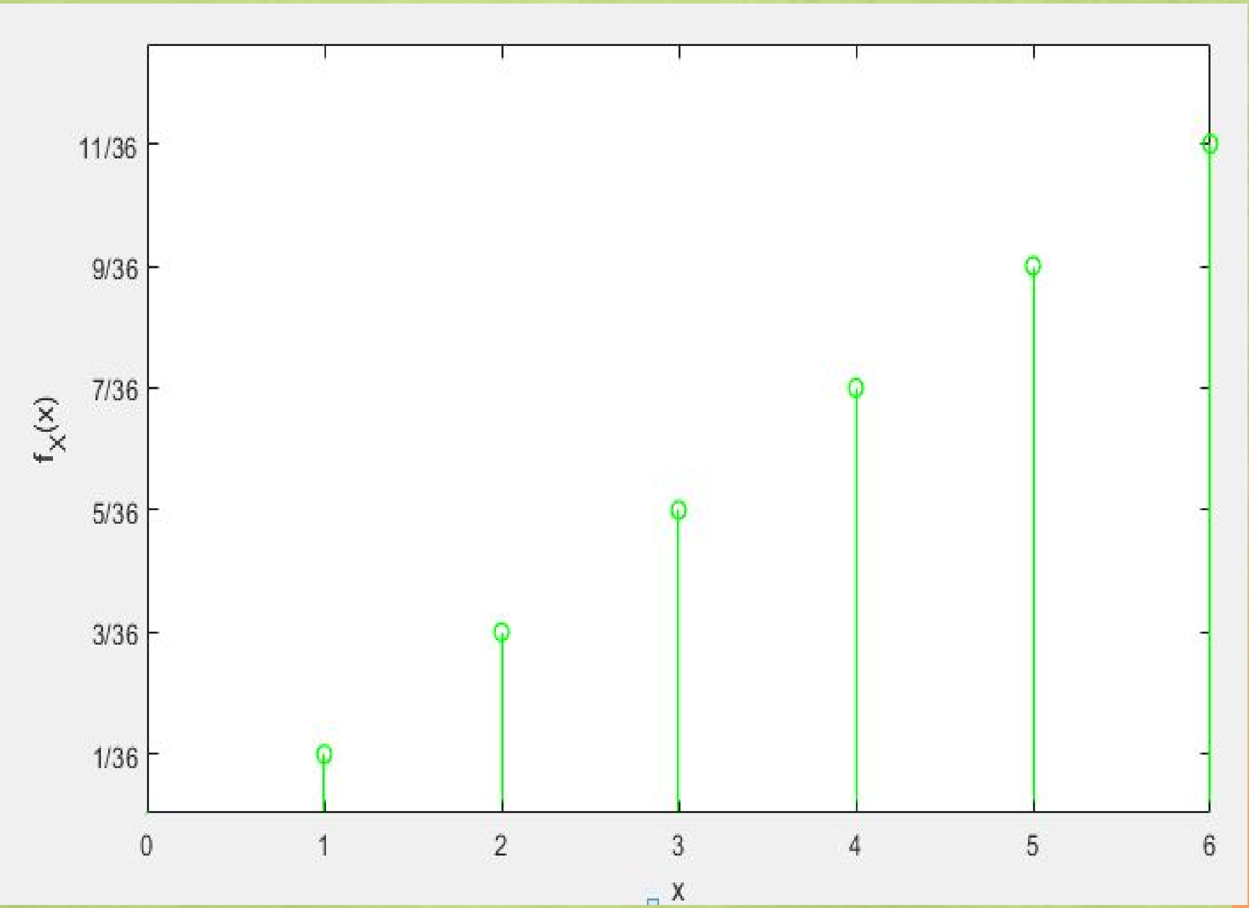
Recall that the sample space Ω for two rolls of a dice has $6 \times 6 = 36$ outcomes. The random variable X takes a value in the set $\{1,2,3,4,5,6\}$.

Next, we compute the PMF of X, that is, the probability for each value that X takes.

$f_X(1) = P(X = 1) = P((1,1)) = {}^1/_{36}$, as the max of the two rolls is 1 for outcome $(1,1)$.

$f_X(2) = P(X = 2) = P\big((1,2),(2,1),(2,2)\big) = {}^3/_{36}$, as the max of the two rolls is 2 for outcomes $(1,2),(2,1),(2,2)$.

Example - Roll a Dice Twice, cont.



Continuing in this manner, we obtain,

$f_X(3) = P(X = 3) = P\big((1,3),(2,3),(3,3),(3,2),(3,1)\big) = {}^{5}/_{36}$

$f_X(4) = P(X = 4) = {}^{7}/_{36}$

$f_X(5) = P(X = 5) = {}^{9}/_{36}$

$f_X(6) = P(X = 6) = {}^{11}/_{36}$

The PMF of X is shown in this plot.

Exercise: Determine the CDF of X using the relationship between CDF and PMF.

III. Examples of Discrete Random Variables

We will study properties of the following discrete random variables

- Bernoulli
- Binomial

Bernoulli Random Variable

- A Bernoulli random variable X takes two values 1 and 0, with probabilities p and $1 - p$, respectively.
- That is, $P(X = 0) = 1 - p$ and $P(X = 1) = p$. This is the **PMF of X**.
- Note that $\sum_x f_X(x) = (1 - p) + p = 1$.
- A Bernoulli random variable has a **single parameter p**.
- As an example, we toss a coin with two outcomes $\Omega = \{h, t\}$. Let $P(h) = p$. We assign the values $X(h) = 1$ and $X(t) = 0$.
- We designate the outcome 1 as **success** and the outcome 0 as **failure**.

How the Binomial Distribution Arises

- Conduct ***n* independent repetitions** of Bernoulli trial.
- Let ***p* denote the probability of success** in each trial. Here p takes a value between 0 and 1, and it is the same for each trial.
- Let ***X* denote the number of successes** in n trials. Therefore, X takes the values $0, 1, 2, \cdots n$.
- The random variable X is called **binomial random variable**.
- n and p are called the **parameters** of the binomial distribution.

PMF of Binomial Distribution

- Conduct n independent repetitions of Bernoulli trials.
- Probability of a specific sequence of x successes is given by $p^x(1-p)^{n-x}$.
- For example, if $n = 6$ and $x = 2$, a specific sequence is hhtttt, and its probability is given by $p^2(1-p)^4$.
- Another specific sequence is tthhtt, and its probability is also given by $p^2(1-p)^4$.
- Out of n trials, there are a total of $\binom{n}{x}$, that is, n choose x ways of obtaining x heads.
- Therefore, $P(X = x) = \binom{n}{x} p^x(1-p)^{n-x}$.

PMF of Binomial Distribution, cont.

- The **PMF of a binomial random variable** X is given by

$$f_X(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0,1,2 \cdots n.$$

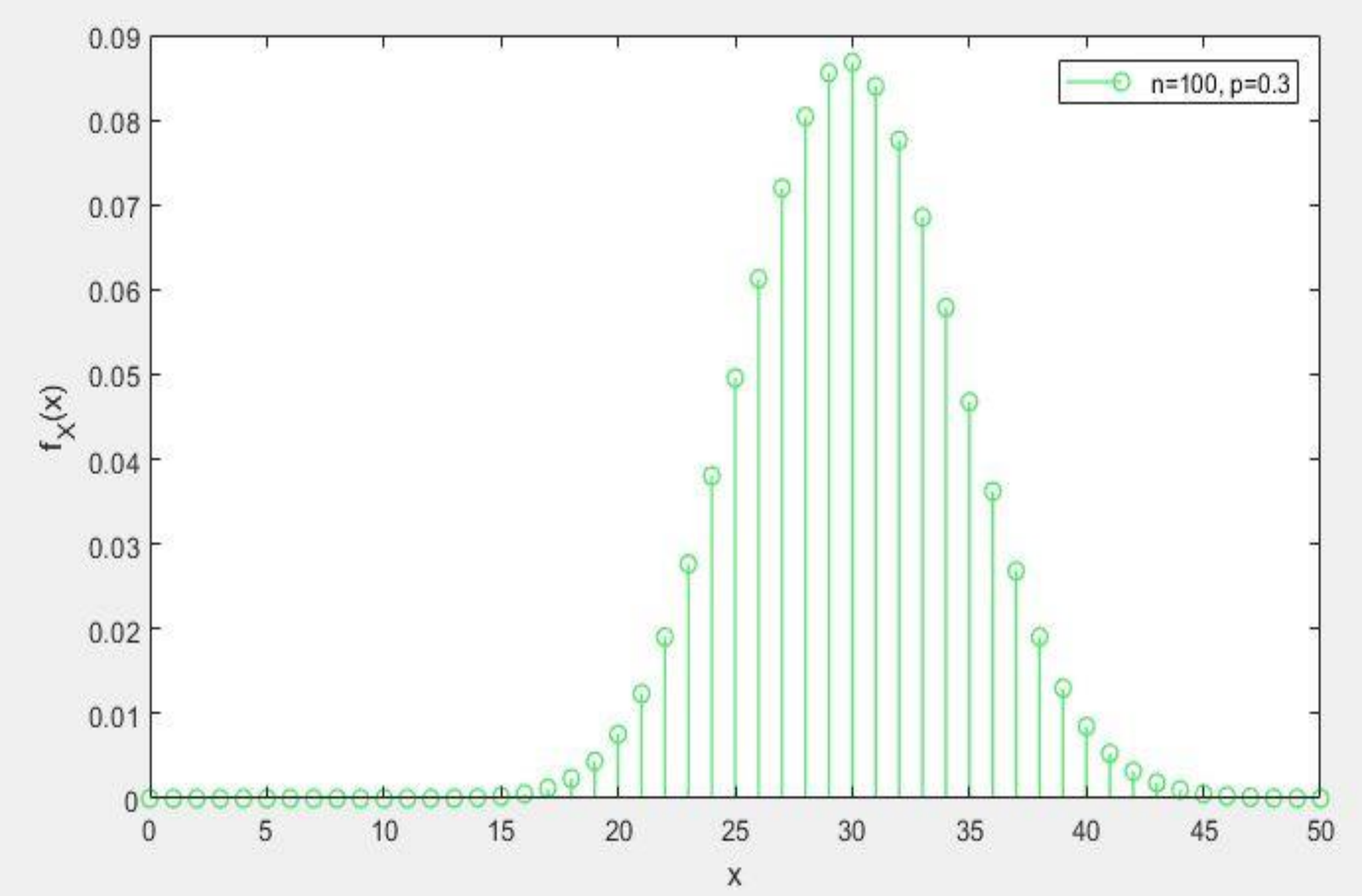
- We also have that

$$\sum_{x=0}^{n} f_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x},$$

$= [p + (1-p)]^n$, using Binomial Theorem (see Appendix A)

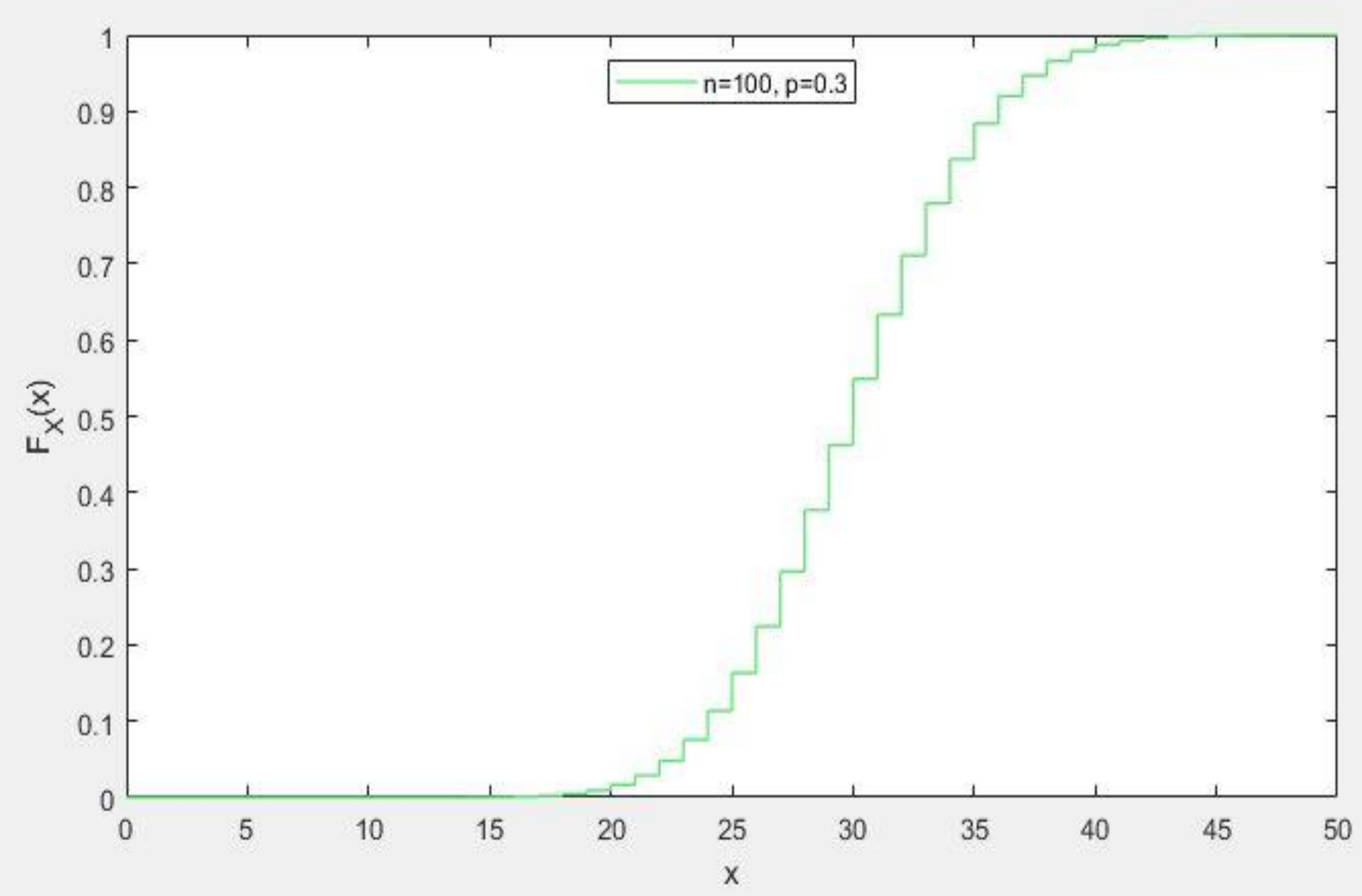
$= 1.$

Binomial Distribution Plots - PMF



This plot shows the PMF (probability mass function) of a binomial distribution. The parameters are $n = 100$, and $p = 0.3$

Binomial Distribution Plots - CDF



This plot shows the CDF (cumulative distribution function) of a binomial distribution. The parameters are $n = 100$, and $p = 0.3$

Example - Communication Channel

Example

In a binary communication channel, the probability of received zeros is 0.4 and the probability of received ones is 0.6. In a frame of 10 received bits, what is the probability that 4 of the bits will be zero.

Solution

Let X be the number of zeros. We can utilize binomial distribution with parameters $n = 10$ and $p = 0.4$. We want to compute the probability of receiving 4 bits, that is, $\mathrm{x} = 4$.

The required probability is given by

$$P(X = 4) = \binom{10}{4}(0.4)^4(0.6)^6 = 0.25$$

Example - Cancer Rate

Example

In a given nation, 1.5% of the population has cancer. We select at random 100 people from the population. What is the probability that **at least one person** has cancer?

Solution

Let X be the number of people with cancer, among the selected population. We can utilize binomial distribution with parameters $n = 100$ and $p = 0.015$.

Note that $X = 1$, means **exactly** one person has cancer, $X = 2$ means **exactly** two people have cancer, and so forth.

Therefore, the required probability is computed by adding up the probabilities for $X = 1,2,3 \cdots 100$.

Example - Cancer Rate, cont.

Therefore,

$$P(\text{at least one person has cancer}) = \sum_{x=1}^{100} \binom{100}{x} (0.015)^x (0.985)^{n-x}$$

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

$$= 1 - \left[\binom{100}{0} (0.015)^0 (0.985)^{100}\right]$$

$$= 0.779$$

IV. Additional Examples of Discrete Random Variables

We will study properties of the following discrete random variables

- Geometric
- Poisson

Geometric Random Variable

- Suppose that independent trials, each having a probability p of success are performed.
- Let ***X* denote the number of trials required for success**. Then

$$f_X(x) = P(X = x) = P(\text{failure in first } (x-1) \text{ trials, followed by success in } x\text{th trial})$$

$$= \underbrace{(1-p)(1-p)\cdots(1-p)}_{(x-1)\text{ times}}\, p$$

$$= (1-p)^{x-1}p.$$

Geometric Random Variable, cont.

- The **PMF of the geometric random variable** X is given by

$$f_X(x) = P(X = x) = (1-p)^{x-1}p.$$

- Since we require success for the first time in xth trial, hence it is necessary (and sufficient) that the first $(x-1)$ trials be failures and the xth trial be a success.
- The geometric random variable has a single **parameter $\boldsymbol{p}$**.
- Note that its **support** is given by $x = 1,2,3,\cdots$

Example - Dice Roll

Example

A fair dice is rolled. (a) What is the probability of scoring a six for the first time in second trial? (b) What is the probability of scoring a six for the first time in third trial?

Solution

(a) Let X be the **number of trials required** to score a six. Since it is a fair dice, the probability of scoring a six in one trial is $^1/_6$. Therefore, the probability of scoring a six for the first time in **second trial** is given by

$$\mathrm{P}(\mathrm{X}=2)=(1-p)p=\frac{5}{6}\times\frac{1}{6}=\frac{5}{36}.$$

(b) The probability of scoring a six for the first time in **third trial** is given by

$$\mathrm{P}(\mathrm{X}=3)=(1-p)^2p=\left(\frac{5}{6}\right)^2\times\frac{1}{6}=\frac{25}{216}.$$

Example - Balls and Urns (1 of 3)

Example

An Urn contains 5 white balls and 7 black balls. A ball is drawn repeatedly, with replacement, until a black ball is drawn.

(a) What is the probability that **exactly 5** draws are needed?

Solution

Let X be the **number of trials required** to draw a black ball. Therefore, X is a geometric random variable. The probability that we require 5 trials is given by

$$\mathrm{P}(\mathrm{X} = 5) = (1 - p)^4 p,$$

where probability of success in one trial is given by

$$p = \frac{7}{7+5} = \frac{7}{12}.$$

Example - Balls and Urns (2 of 3)

(b) What is the probability that **at least 5** draws are needed?

Solution

We want the probability that the number of draws are 5 or more. This is given by

$$\mathrm{P}(\mathrm{X} \geq 5) = 1 - P(X < 5)$$

$$= 1 - \sum_{x=1}^{4} (1-p)^{x-1} p$$

$$= (1-p)^4.$$

The above sum is done using the **geometric sum expression** (Appendix B), with ratio between successive terms being

$$(1-p) < 1.$$

Example - Balls and Urns (3 of 3)

$$S = \sum_{x=1}^{4} (1-p)^{x-1} p$$

Let $x' = x - 1$; therefore,

$$S = p \sum_{x'=0}^{3} (1-p)^{x'},$$

$$= p \left(\frac{1-(1-p)^4}{1-(1-p)} \right)$$

$$= p \left(\frac{1-(1-p)^4}{p} \right)$$

$$= (1 - (1-p)^4).$$

Poisson Random Variable

- The PMF of a Poisson random variable with **parameter λ** is given by

$$f_X(x) = P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!},$$

- Its **support** is given by $x = 0,1,2,\cdots$,
- The parameter λ satisfies the condition $\lambda > 0$.

Poisson Random Variable, cont.

We want to verify the PMF relation $\sum_x f_X(x) = 1$.

$$\sum_x f_X(x)$$

$$= \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}.$$

We have that $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}$, using the **Taylor series expansion for e^{λ}**.

Therefore, $\sum_x f_X(x) = e^{-\lambda} e^{\lambda} = 1$.

Examples which follow Poisson Distribution

- Number of phone calls arriving at a call center per hour.
- Number of customers arriving at a bank on a given day.
- Number of misprints on a page of a book.
- Number of meteors, greater than one meter, that strike the earth every year.
- Number of people in a community living to 100 years of age.

Examples which violate Poisson Distribution

- Number of students arriving at a university café, because the rate of arrival is not a constant; students arrive in bursts after class.
- Number of magnitude 5 earthquakes in California may not follow a Poisson distribution, if one large earthquake is followed by several smaller aftershocks.

Example - Typographical Errors

Example

Suppose that the average number of typographical errors on a single page of a book has a Poisson distribution with parameter $\lambda = 1/2$.

(a) Calculate the probability of **exactly 3** errors on the page.

Solution

Let X denote the number of errors on this page. Then we have,

$P(X = 3) = e^{-1/2} \frac{(1/2)^3}{3!} = 0.013,$

where we have used $\lambda = 1/2$ in the expression for Poisson PMF.

Example - Typographical Errors, cont.

(b) What is the probability that there is **at least 1** error on the page?

Solution

We want the probability that the number of errors is 1 or more. This is given by

$$P(X \geq 1) = 1 - P(X = 0)$$
$$= 1 - e^{-\frac{1}{2}} = 0.395$$

Probability of Events in a Time Interval

- For probability of events in a **time interval $\boldsymbol{t}$**, we write the PMF of the Poisson random variable as follows:

$$f_{X(t)}(x) = P(X(t) = x) = e^{-\lambda t}\frac{(\lambda t)^x}{x!}.$$

- In utilizing the above expression, we make sure that λ and t have appropriate units such that **$\boldsymbol{\lambda t}$ is dimensionless**.
- For example, if $t = 2$ hours, then $\lambda = 4/\text{hour}$ is fine.
- As another example, if $t = 1$ month, then $\lambda = 3/\text{month}$ is fine.

Example - Earthquakes

Example

Suppose that earthquakes occur in a region with an average value of $\lambda = 2/\text{month}$. Given that the number of earthquakes follow a Poisson distribution, find the probability of an earthquake in the next month.

Solution

Let $X(t)$ represent the number of earthquakes in time interval t. Then we have,

$\lambda = 2/\text{month}$

$t = 1 \text{ month}$

$\lambda t = 2$, dimensionless

Probability of one earthquake in one month is

$P(X(t) = 1) = e^{-2}\frac{(2)^1}{1!} = 0.27.$

Appendix A: Binomial Theorem

The binomial theorem is given by:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \, y^{n-k}$$

Appendix B: Geometric Series

- The sum of the first $(N+1)$ terms of a geometric series is given by

$$\sum_{j=0}^{N} r^j = \frac{1-r^{N+1}}{1-r}, \quad r \neq 1$$

- For $N \rightarrow \infty$, the above sum is given by

$$\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}, \quad |r| < 1$$

Appendix C: Binomial-Poisson Approximation

- The Poisson random variable can be used as an **approximation** to the binomial random variable, under certain conditions, as described below.
- The mean or **average value** of a binomial random variable with **parameters n and p is np**.
- The mean or **average value** of a Poisson random variable with **parameter λ is λ**.
- A binomial random variable X, with **large n and small p**, can be approximated by a Poisson random variable, with $\lambda = np$.

Appendix C: Binomial-Poisson Approximation, cont.

- That is, we perform n independent trials, each of which results in success with probability p.
- As $n \to \infty$ and $p \to 0$, such that np is moderate, then the number of successes is approximately a Poisson random variable, with parameter $\lambda = np$.

References

1. Charles Boncelet, Probability, Statistics and Random Signals, Oxford University Press, 2016.
2. Sheldon Ross, A First Course in Probability, Macmillan Publishing Company, 1988.
3. R. D. Yates, et al., Probability and Stochastic Processes, John Wiley, 2005.