Discrete Random Variables

Rosario Barone

Tor Vergata University of Rome

Statistical tools for decision making

Undergraduate Degree in Global Governance

A.Y. 2024/2025

Random Variables

Definition

A *random variable* is a function that assigns a real number to each outcome of a random experiment.

- Random variables are used to model and analyze uncertain or random phenomena in various fields.
- They can be discrete or continuous, depending on the nature of the outcomes.

Example

Consider the random variable X representing the number obtained when rolling a fair six-sided die. X can take on values 1, 2, 3, 4, 5, or 6.

- Random variables can be further categorized as:
  - Discrete Random Variables
  - Continuous Random Variables

Ramdom variables

Experiment	Number X	Possible Values of X
Roll two fair dice	Sum of the number of dots on the top faces	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Flip a fair coin repeatedly	Number of tosses until the coin lands heads	1, 2, 3,4, ...
Measure the voltage at an electrical outlet	Voltage measured	$118 \le x \le 122$
Operate a light bulb until it burns out	Time until the bulb burns out	$0 \le x < \infty$

Random variables

Definition

Formally, given Ω and a probability P on Ω, a r.v. X is a function $X(\omega)$ defined on Ω and taking values in $\mathbb{R}$

$$X : \Omega \to \mathbb{R}$$

and $\forall B \subseteq \mathbb{R}$

$$P(X \in B) = P(\omega \in \Omega : X(\omega) \in B)$$

Random variables will be generally indicated with the letters $X, Y, Z \ldots$

- A r.v. that may assume only a finite number or an infinite sequence of values is said to be discrete;
- A r.v. that may assume any value in some interval on the real number line is said to be continuous.

Ramdom variables

- A random variable representing the number of new cases of COVID-19 on one day would be discrete
- A random variable representing the weight of a person in kilograms would be continuous.
- Very often we work directly with random variables without knowing (or caring toknow) the underlying probability P on the space Ω
- In fact we will specify (model) directly the probabilities of the outcomes of the r.v.

Discrete Random Variables

Definition

A *discrete random variable* is a variable that can take on a countable number of distinct values. These values are often associated with the outcomes of a random experiment, and each value has an associated probability.

- Discrete random variables are used to model and analyze phenomena with distinct, separate outcomes.
- Number of heads in a series of coin flips or the count of customers arriving at a store in an hour.
- Discrete random variables are characterized by their Probability Mass Functions (PMFs), which provide a complete description of the variable's distribution.

Example

Suppose we toss an unbiased coin 2 times in succession. What is the probability of obtaining x heads ($x = 0, 1, 2$)?

Let X be the discrete r.v. describing the result of such experiments. The probability function is

	x
(T,T)	0
(T,H), (H,T)	1
(H,H)	2

Probability Mass Function (PMF)

Definition

For a discrete random variable X, the Probability Mass Function (PMF) denoted as $p_X(x)$ is a function that describes the probability of X taking on a specific value, x:

$$p_X(x) = P(X = x) = P(\omega \in \Omega : X(\omega) = x)$$

- The PMF characterizes the distribution of a discrete random variable.
- It assigns probabilities to each possible value of the random variable.

Properties

- $0 \leq p_X(x) \leq 1$ for all x.
- $\sum_{\text{all } x} p_X(x) = 1$ (the sum over all possible values of X is equal to 1).

Example

Suppose we toss an unbiased coin 2 times in succession. What is the probability of obtaining x heads ($x = 0, 1, 2$)?
Let X be the discrete r.v. describing the result of such experiments. The probability function is

	x	$p_X(x)$
(T,T)	0	1/4
(T,H), (H,T)	1	1/2
(H,H)	2	1/4

Cumulative Distribution Function (CDF)

Definition

The Cumulative Distribution Function (CDF) of a random variable X, denoted as $F_X(x)$, is a function that gives the probability that X takes on a value less than or equal to x for all possible values of x:

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

- CDF provides a comprehensive view of the probability distribution of a random variable.
- It is a non-decreasing, right-continuous function.

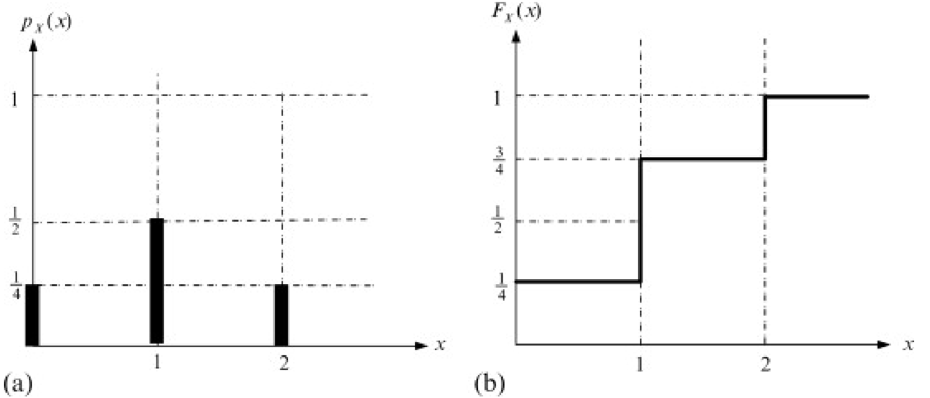
Properties

- $0 \leq F_X(x) \leq 1$ for all x.
- $F_X(-\infty) = 0$
- $F_X(\infty) = 1$
- $x < x' \Rightarrow F(x) \leq F(x')$

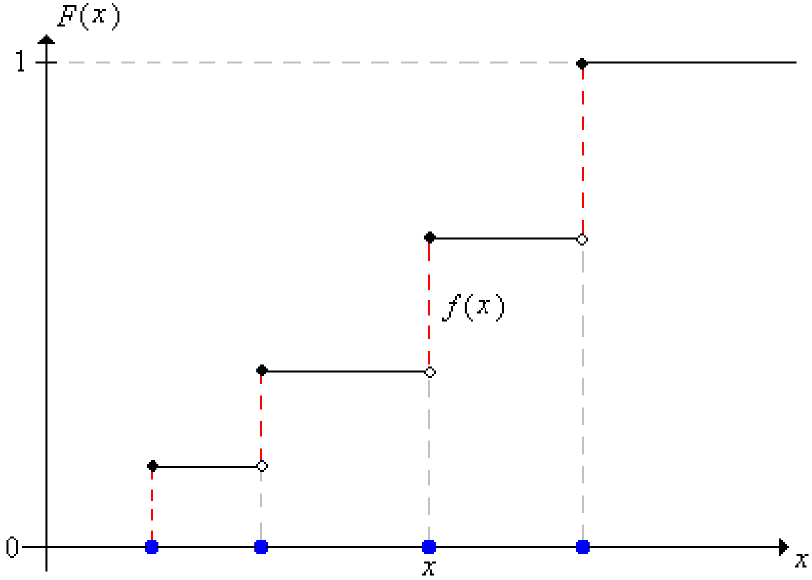
Example

Suppose we toss an unbiased coin 2 times in succession. What is the probability of obtaining x heads ($x = 0, 1, 2$)?
Let X be the discrete r.v. describing the result of such experiments. The probability function is

	x	$p_X(x)$	$F_X(x)$
(T,T)	0	1/4	1/4
(T,H), (H,T)	1	1/2	3/4
(H,H)	2	1/4	1



Probability distribution function and cumulative distribution function for the previous example



Cumulative distribution function of a discrete random variable

CDF and probability of intervals

By the distribution function and the rules of probability we can obtain other probabilities on the r.v. X

- $$P(X > a) = 1 - P(X \leq a) = 1 - F_X(a)$$

- Given two point a and b, with $a < b$, we can get $P(a < X \leq b)$ as $F_X(b) - F_X(a)$.

  Since for $a < b$, $(-\infty, b] = (-\infty, a] \cup (a, b]$

  $$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

  and

  $$P(a < X \leq b) = F_X(b) - F_X(a)$$

Moments of a Probability Distribution

In probability and statistics, moments are essential statistical measures that provide insights into the shape and characteristics of a probability distribution. Moments summarize the distribution's central tendencies and variability.

- Moments are often used to quantify the location, spread, skewness, and kurtosis of a probability distribution.

Moments of a Probability Distribution

Types of Moments

1. **First Moment: Mean (Expectation)**
   - The first moment is the expected value of the random variable and represents the center of the distribution.
2. **Second Moment: Variance**
   - The second moment quantifies the spread or dispersion of the distribution.
3. **Third Moment: Skewness**
   - The third moment measures the asymmetry or skew of the distribution.
4. **Fourth Moment: Kurtosis**
   - The fourth moment describes the tails or thickness of the distribution.

- Moments of higher order also exist, but the first four moments are the most commonly used.

Expected Value of a Discrete Random Variable

Definition

The *expected value* of a discrete random variable X, denoted as $E(X)$ or μ_X, is a measure of the center or average of its distribution. It is defined as:

$$E(X) = \sum_{x} x \cdot p_X(x)$$

- The expected value represents a weighted sum of all possible values of the random variable, where the weights are the probabilities of those values.
- It provides a single number that summarizes the central tendency of the random variable.

Expected Value of a Discrete Random Variable

Properties

- $E(X)$ is a constant, not a random variable.
- It is a linear operator, meaning that the expected value of a sum of random variables is the sum of their individual expected values.

- The expected value is a fundamental concept in probability and statistics, often used to make predictions and decisions.

Variance of a Discrete Random Variable

Definition

The *variance* of a discrete random variable X, denoted as $\text{Var}(X)$, is a measure of the spread or dispersion of its probability distribution. It is defined as:

$$\text{Var}(X) = E\left[(X - \mu_X)^2\right] = \sum_x (x - \mu_X)^2 \cdot p_X(x)$$

- The variance quantifies how much individual values of the random variable deviate from its expected value.
- It's a key measure of uncertainty in the distribution.

Variance of a Discrete Random Variable

Interpretation

- If $\text{Var}(X) = 0$, it implies that all values of X are the same, and there is no variability.
- A larger variance indicates greater spread or dispersion in the values of X.

- Variance is a fundamental concept in statistics and plays a crucial role in understanding and characterizing random variables.

Skewness of a Discrete Random Variable

Skewness

- **Definition:** Skewness is a measure of the asymmetry in the probability distribution of a random variable.
- It quantifies whether the distribution is skewed to the left (negative skew) or to the right (positive skew) of the mean.
- **Interpretation:**
  - Negative skewness indicates a longer tail on the left.
  - Positive skewness indicates a longer tail on the right.
  - A skewness of 0 implies a symmetric distribution.

Kurtosis of a Discrete Random Variable

Kurtosis

- **Definition:** Kurtosis measures the heaviness of the tails and the peakedness of a probability distribution.
- It quantifies whether the distribution is more or less peaked than a normal distribution.
- **Interpretation:**
  - Positive kurtosis (excess kurtosis > 0) lighter tails and a more peaked distribution than the normal distribution.
  - Negative kurtosis (excess kurtosis < 0) indicates heavier tails and a flatter distribution than the normal distribution.
  - A kurtosis of 3 is subtracted in the formula to make the kurtosis of a standard normal distribution equal to 0.

Families of Discrete Probability Distributions

Discrete probability distributions are categorized into different families that provide a framework for modeling and understanding random phenomena. The common discrete probability distributions are:

- Bernoulli Distribution
- Binomial Distribution
- Geometric Distribution
- Poisson Distribution

The choice of distribution depends on the specific problem and underlying assumptions.

Bernoulli Probability Distribution

Characteristics

$X \sim$ Bernoulli(p) with $p \in (0, 1)$.

- A Bernoulli random variable, typically denoted as X, takes on two values: 1 (success) or 0 (failure).
- The probability of success is denoted as p, and the probability of failure is $q = 1 - p$.
- The probability mass function (PMF) of a Bernoulli distribution:

$$p_X(x) = \begin{cases} p, & \text{if } x = 1 \\ q, & \text{if } x = 0 \end{cases}$$

Applications

- Bernoulli trials are commonly used to model success-failure experiments, such as coin flips, yes-no questions, or pass-fail tests.
- It serves as the basis for more complex distributions, like the binomial and geometric distributions.

Mean and Variance of a Bernoulli Distribution

Mean (Expected Value):

The mean (μ) of a Bernoulli distribution is calculated as:

$$\mu = E(X) = \sum_{x} x \cdot p_X(x) = 1 \cdot p + 0 \cdot (1 - p) = p$$

Variance:

The variance (σ^2) of a Bernoulli distribution is calculated as:

$$\sigma^2 = \mathrm{Var}(X) = \sum_{x} (x - \mu_X)^2 \cdot p_X(x) =$$

$$\sigma^2 = (1 - p)^2 \cdot p + (0 - p)^2 \cdot (1 - p) =$$

$$\sigma^2 = p(1 - p)(1 - p + p) = p(1 - p).$$

Geometric Probability Distribution

The Geometric distribution is a probability distribution that models the number of Bernoulli trials required for the first success to occur. It is used to describe the probability of success on the x-th trial.

Characteristics

$X \sim \text{Geometric}(p)$ with $p \in (0, 1)$.

- A Geometric random variable, typically denoted as X, takes on non-negative integer values: $0, 1, 2, \ldots$.
- The probability of success on each trial is denoted as p.
- The probability mass function (PMF) of a Geometric distribution:

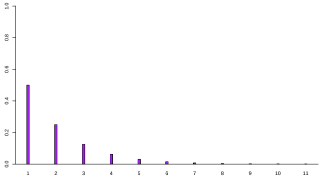
$$p_X(x) = (1 - p)^{x-1} \cdot p, \text{ for } x = 1, 2, 3, \ldots$$

Geometric Probability Distribution

Applications

- The Geometric distribution is commonly used to model situations where you're interested in the number of trials required for a success in a sequence of independent Bernoulli trials.
- Examples include the number of coin flips to get the first heads, the number of attempts to make the first sale, or the number of failures before a machine works.

Number of coin flips until the first head sshows up (assuming independent coin flips)



Probability distribution function for a Geometric r.v. with $p = 0.5$

Geometric Probability Distribution

Example

Suppose you are playing a game where you have a 1 in 5 chance of winning on each trial. You keep playing the game until you win. Calculate the probability that you will need exactly 4 trials to win.

$$
\begin{aligned}
p_X(4) = P(X = 4) &= (1 - 1/5)^{(4-1)}(1/5) \\
&= (4/5)^3(1/5) \\
&= (64/125)(1/5) \\
p_X(4) &= 64/625
\end{aligned}
$$

Geometric Probability Distribution

Example

Suppose you are playing a game where you have a 1 in 5 chance of winning on each trial. You keep playing the game until you win. Calculate the probability that you will need exactly 4 trials to win.

$$
\begin{aligned}
p_X(4) = P(X = 4) &= (1 - 1/5)^{(4-1)}(1/5) \\
&= (4/5)^3(1/5) \\
&= (64/125)(1/5) \\
p_X(4) &= 64/625
\end{aligned}
$$

Mean and Variance of a Geometric Distribution

Mean (Expected Value)

The mean (expected value) of X is calculated as:

$$\mu = \frac{1}{p}$$

Variance

The variance of X is calculated as:

$$\sigma^2 = \frac{1 - p}{p^2}$$

Binomial Probability Distribution

The Binomial distribution is a discrete probability distribution that models the number of successes (e.g., yes/no, heads/tails) in a fixed number of independent Bernoulli trials.

Characteristics

$X \sim \text{Binomial}(n, p)$ with $n > 0$ and $p \in (0, 1)$.

- A Binomial random variable, typically denoted as X, represents the number of successes in n independent Bernoulli trials.
- The probability of success in each trial is denoted as p.
- The probability mass function (PMF) of a Binomial distribution:

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, 2, \ldots, n$$

or equivalently

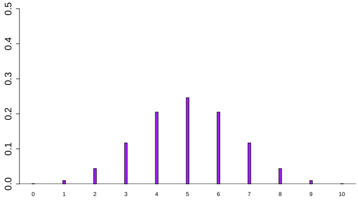
$$p_X(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, 2, \ldots, n$$

Binomial Probability Distribution

Applications

- The Binomial distribution is widely used to model real-world scenarios, such as the number of successful sales out of a fixed number of sales attempts, the number of heads in a fixed number of coin flips, or the number of defective items in a batch of products.
- It's a fundamental distribution in statistics and probability theory.

Number of heads in n independent coin flips



Probability distribution function for a Binomial r.v. with $n = 10$ and $p = 0.5$

Binomial Probability Distribution

Example

Suppose a factory produces light bulbs, and the probability of any single light bulb being defective is 0.1. The factory produces a batch of 50 light bulbs.Calculate the probability of finding exactly 5 defective light bulbs in the batch.

We have $n = 50$, $x = 5$, and $p = 0.1$:

$$p_X(5) = P(X = 5) = \binom{50}{5} \cdot (0.1)^5 \cdot (0.9)^{45}$$

$$p_X(5) = 0.185$$

Binomial Probability Distribution

Example

Suppose a factory produces light bulbs, and the probability of any single light bulb being defective is 0.1. The factory produces a batch of 50 light bulbs.Calculate the probability of finding exactly 5 defective light bulbs in the batch.
We have $n = 50$, $x = 5$, and $p = 0.1$:

$$p_X(5) = P(X = 5) = \binom{50}{5} \cdot (0.1)^5 \cdot (0.9)^{45}$$

$$p_X(5) = 0.185$$

Mean and Variance of a Binomial Distribution

Mean (Expected Value)

The mean (expected value) of X is calculated as:

$$\mu = np$$

Variance

The variance of X is calculated as:

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

Poisson Probability Distribution

The Poisson distribution is a discrete probability distribution that models the number of events occurring within a fixed interval of time or space, given a known average rate of occurrence.

Characteristics

$X \sim \text{Pois}(\lambda)$ with $\lambda > 0$.

- A Poisson random variable, typically denoted as X, represents the count of events in a fixed interval.
- The parameter λ represents the average rate of events per interval.
- The probability mass function (PMF) of a Poisson distribution:

$$p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \text{ for } x = 0, 1, 2, \ldots$$

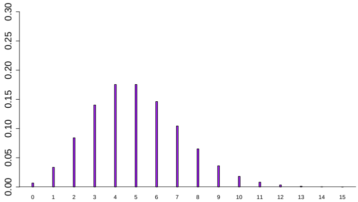
- The expected value is $E(X) = \lambda$, and the variance is $\text{Var}(X) = \lambda$.

Poisson Probability Distribution

Applications

- The Poisson distribution is used in a wide range of applications, including modeling the number of phone calls to a call center in an hour, the number of accidents at an intersection in a day, or the number of emails arriving in a mailbox in a minute.
- It is particularly useful in situations where events are rare but occur randomly over time.

Number of customers in one hour



Probability distribution function for a Poisson r.v. with $\lambda = 5$.

Exercise: Poisson Distribution and Probability

Example

Suppose that, on average, there are 3 customers arriving at a coffee shop per hour (rate $\rightarrow \lambda = 3$). We can model this with a Poisson distribution.

1. Calculate the probability of 5 customers arriving in an hour:

The probability mass function (PMF) of the Poisson distribution is given by:

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

where λ is the average rate of arrivals (in this case, $\lambda = 3$) and x is the number of arrivals.

$$P(X = 5) = \frac{e^{-3} \cdot 3^5}{5!}$$

The probability of 5 customers arriving in an hour is approximately 0.1008.

Exercise: Poisson Distribution and Probability

Example

2. Calculate the probability of at least 2 customers arriving in 15 minutes.

To calculate the probability of at least 2 customers arriving in 15 minutes, we need to consider the rate for 15 minutes, which is $\lambda/4$. We can use the complementary probability approach:

$$P(X \geq 2) = 1 - P(X < 2)$$

$$P(X < 2) = P(X = 0) + P(X = 1)$$

$$P(X \geq 2) = 1 - \left(\frac{e^{-3/4} \cdot (3/4)^0}{0!} + \frac{e^{-3/4} \cdot (3/4)^1}{1!} \right)$$

The probability of at least 2 customers arriving in 15 minutes is approximately 0.1734.

Mean and Variance of a Poisson Distribution

Mean (Expected Value)

The mean (expected value) of X is equal to the parameter λ:

$$\mu = \lambda$$

Variance

The variance of X is also equal to the parameter λ:

$$\sigma^2 = \lambda$$