A. Probability Distribution

1. Concept of Random Variables

A **random variable** is defined as a real-valued function whose domain is the sample space of a random experiment. In simpler terms, it maps every possible outcome of an experiment to a unique real number.

- **Example:** Consider tossing a fair coin twice. The sample space is $S = HH, HT, TH, TT$. If we define a random variable X as the "number of heads," its values are:
  - $X(HH) = 2$
  - $X(HT) = 1$
  - $X(TH) = 1$
  - $X(TT) = 0$.
- **Multiple Variables:** It is possible to define more than one random variable on the same sample space. For instance, a variable Y could represent the "number of heads minus the number of tails," resulting in values like $Y(HH) = 2$ and $Y(HT) = 0$.

2. Discrete Random Variables

A random variable is **discrete** if it can take only a finite or countably infinite number of values. The CUET syllabus primarily focuses on these discrete sample spaces.

3. Probability Distribution of a Discrete R.V.

The probability distribution of a random variable X is a system of numbers that lists the possible values of X ($x_1, x_2, \ldots, x_n$) along with their associated probabilities ($p_1, p_2, \ldots, p_n$).

The conditions for a valid probability distribution are:

1. Each probability must be non-negative: $p_i \geq 0$ for all i.
2. The sum of all probabilities must be exactly 1: $\sum_{i=1}^{n} p_i = 1$

B. Mathematical Expectation (Expected Value)

The **expected value** (or **mean**) of a random variable is the arithmetic mean of its probability distribution. It represents the long-term average outcome if the experiment is repeated many times.

1. Formula for Expected Value

If X is a discrete random variable with possible values $x_1, x_2, \ldots, x_n$ and corresponding probabilities $p_1, p_2, \ldots, p_n$, the expectation $E(X)$, often denoted by μ, is: $E(X) = \mu = \sum_{i=1}^{n} x_i p_i$

2. Application

This concept allows us to quantify the "center" of the distribution. It is analogous to the weighted average used in frequency distributions, where the probabilities act as weights.

C. Variance and Standard Deviation

While the mean tells us where the distribution is centered, **variance** measures the "spread" or dispersion of the random variable's values around that mean.

1. Variance

The variance of X, denoted by $Var(X)$ or σ^2, is calculated using the following formula: $Var(X) = \sigma^2 = E(X^2) - [E(X)]^2$ Where:

- $E(X^2) = \sum x_i^2 p_i$
- $E(X) = \sum x_i p_i$

2. Standard Deviation (S.D.)

The **standard deviation** is the positive square root of the variance. It is expressed in the same units as the random variable itself, making it easier to interpret. $\sigma = \sqrt{Var(X)}$

D. Binomial Distribution

The **Binomial distribution** is one of the most important discrete probability distributions, originally discovered by Jacob Bernoulli. It describes the number of successes in a fixed number of independent trials.

1. Bernoulli Trials

A sequence of trials is called **Bernoulli trials** if they satisfy the following conditions:

1. The number of trials (n) is finite.
2. The trials are **independent**.
3. Each trial has exactly two outcomes: **Success** (**p**) or **Failure** (**q**).
4. The probability of success (p) remains constant for each trial.
   - Note: $p + q = 1$.

2. The Binomial Formula

The probability of obtaining exactly r successes in n Bernoulli trials is given by: $P(X = r) = \binom{n}{r} p^r q^{n-r}$
Where:

- $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

3. Parameters of Binomial Distribution

The distribution is entirely determined by the parameters n and p.

- **Mean:** $\mu = np$.
- **Variance:** $\sigma^2 = npq$.
- **Standard Deviation:** $\sigma = \sqrt{npq}$.

E. Poisson Distribution

(Note: The following information is not present in the provided sources and should be independently verified.)

The Poisson distribution is another discrete distribution used to model the number of times an event occurs in a fixed interval of time or space.

1. Conditions for Poisson Distribution

1. The events occur independently.
2. The probability of an event occurring in a very small interval is proportional to the length of the interval.
3. The probability of more than one occurrence in such a small interval is negligible.
4. It is often used as a limiting case of the Binomial distribution when n **is very large** and p **is very small**, such that the product $np = \lambda$ (a constant).

2. Mean and Variance

A unique property of the Poisson distribution is that its mean and variance are equal.

- **Mean:** $E(X) = \lambda$
- **Variance:** $Var(X) = \lambda$
- **Standard Deviation:** $\sigma = \sqrt{\lambda}$

F. Normal Distribution

(Note: The following information is not present in the provided sources and should be independently verified.)

Unlike the previous sections, the **Normal distribution** is a **continuous** probability distribution. It is often called the "Gaussian distribution" and is characterized by a bell-shaped curve.

1. Continuous Nature

In a continuous distribution, the random variable can take any value within a range (e.g., height, weight). The probability of the variable taking an exact single value is zero; instead, we calculate the probability of it falling within an interval, which corresponds to the **area under the curve**.

2. Standard Normal Variate (Z)

To compare different normal distributions, we transform the random variable X into a **Standard Normal Variate (Z)**. This process is called "standardization." $Z = \frac{X - \mu}{\sigma}$

- The Z-distribution has a **mean of 0** and a **standard deviation of 1**.

3. Area Relationship (The 68-95-99.7 Rule)

The area under the normal curve relates directly to the standard deviation:

- Approximately **68.27%** of the data falls within **1 S.D.** of the mean ($\mu \pm 1\sigma$).
- Approximately **95.45%** of the data falls within **2 S.D.** of the mean ($\mu \pm 2\sigma$).
- Approximately **99.73%** of the data falls within **3 S.D.** of the mean ($\mu \pm 3\sigma$).

Summary Table for CUET Preparation

Distribution	Type	Mean (μ)	Variance (σ^2)	Key Conditions
General Discrete	Discrete	$\sum x_i p_i$	$E(X^2) -$	$\sum p_i = 1$
Binomial	Discrete	np	npq	Independent trials, constant p
Poisson	Discrete	λ	λ	Large n, small p, rare events
Normal	Continuous	μ	σ^2	Bell curve, Z-score standardization

Key Historical Concepts for Context

- **Kolmogorov:** Formulated the axiomatic approach to probability used in modern mathematics.
- **Pascal and Fermat:** Laid the first foundations of probability science in the 17th century through correspondence regarding gambling problems.
- **Leibnitz:** Introduced the notation for integration ($\int$ and dy), which is the mathematical basis for calculating areas in continuous distributions like the Normal distribution.

These notes provide the depth required for the CUET Unit IV syllabus. Students should focus on mastering the **matrix-like calculation of expectation and variance** for discrete variables and the **independent trial logic** of the Binomial distribution.