# Introduction to Probability Distributions

**Distribution –** The possible values a variable can take and how frequently the occur.
**Probability** is the likelihood of an outcome.

**Types of Probability Distributions:**

- Discrete – finite number of outcomes
- Continuous – infinitely many outcomes

## Discrete Distributions:

**Uniform Distribution (U)** – equiprobable; all outcomes are equally likely.
**Bernoulli Distribution (Bern)** – events with only 2 possible outcomes; true or false.
**Binomial Distribution (B)** – carrying out a similar experiment several times in a row.

- 2 outcomes per iteration
- Many iterations

**Poisson Distribution (Po)** – test out how unusual an event frequency is for a given interval.

## Continuous Distributions:

**Normal Distribution (N) –** often observed in nature
**Student's-T Distribution (t)** – a small sample approximation of a Normal Distribution; when limited data. It accommodates extreme values significantly better.
**Chi-Squared Distribution** - asymmetric and only consists of non-negative values.
**Exponential Distribution** – events that are rapidly changing early on.

**Logistic Distribution** – useful in forecast analysis, determining a cut-off point for a successful outcome.

**Transformation** – a way in which we alter every element of a distribution to get a new distribution. For normal distributions we can add, subtract, multiply, or divide every element of a distribution and get a new distribution with similar characteristics.

**Standardizing** – a special kind of transformation where we make the expected value; $E(X) = 0$, and the variance; $Var(X) = 1$.

# Probability & Statistics

## ① Binomial Distribution - $B(n,p)$

↳ a sequence of Bernoulli events

# Probability Function: $p(y) = \binom{n}{y} \times p^y \times (1-p)^{n-y}$

# Expected Values : $E(X) = x_0 \cdot p(x_0) + x_1 \cdot p(x_1) + \dots x_n \cdot p(x_n)$

$Y \sim B(n,p) \Rightarrow E(Y) = n \cdot p$

# Variance: $= E(y^2) - E(y)^2$

$= n \cdot p \cdot (1-p)$

## ② Bernoulli Distribution - $Bern(p)$

↳ 1 trial , 2 possible outcomes

# Expected Values : $E(X) = 1 \cdot p + 0 \cdot (1-p) = p$

# Variance : $\sigma^2 = p(1-p)$

## ③ Poisson Distribution - $Po(\lambda)$

↳ frequency with which an event occurs

# Probability Function: $P(Y) = \dfrac{\lambda^y \cdot e^{-\lambda}}{y!}$

- $e \rightarrow \approx 2.72$
- $e^{-\lambda} = \frac{1}{e^{\lambda}}$

# Expected Values : $E(y) = y_0 \cdot P(y_0) + y_1 \cdot p(y_1) + \dots = \lambda$

# Variance : $\sigma^2 = (y_0 - \mu)^2 + (y_1 - \mu)^2 + \dots = \lambda$

- $\mu = \sigma^2 = \lambda$

## ④ Normal Distribution - $N(\mu, \sigma^2)$

↳ appearing in nature

# Expected Value : $E(X) = \mu$

# Variance : $Var(X) = \sigma^2$

$= E(X^2) - E(X)^2$

/ /

## ⑤ Student's-T Distribution - $t(k)$

↳ small sample size approximation of a Normal Distribution

→ Certain Characteristics + Suffient data = Normal Distribution

→ Certain Characteristics + Insufficient data = Student's T Distribution

→ $k$ → degrees of freedom

if $k > 2$

↳ $E(Y) = \mu$

↳ $Var(Y) = \frac{S^2 \cdot k}{k-2}$

## ⑥ Chi-Squared Distribution - $\chi^2(k)$

↳ hypothesis testing & computing confidence intervals

# Expected Values : $E(X) = k$

# Variance : $Var(X) = 2k$

# Conditional Probability

Conditional probability is the probability of an event occurring given that another event has already occurred. It is denoted as **P(A | B)**, meaning the probability of event **A** happening given that event **B** has happened. It is calculated using the formula**: P(A | B) = P(A and B) / P(B)**

Independent events are events where the occurrence of one does not affect the probability of the other. For example, rolling a die and flipping a coin are independent because the result of one does not influence the other. **P(A | B) = P(A)**

# Baye's Theorem

Baye's Theorem describes how to update probabilities based on new evidence.

P(A|B) = P(B|A) * P(A) / P(B)

Where:

- P(A|B) - probability of event A given B has occurred (posterior probability).
- P(B|A) - probability of event B given A (likelihood).
- P(A) - prior probability of A.
- P(B) - probability of B.