ST2334
AY22/23 Sem 1
github.com/jasonqiu212

01. Basic Concepts of Probability

Event Operations

- **Mutually Exclusive** - $A \cap B = \emptyset$
- **Contained** - $A \subset B$
- **Equivalence** - $A \subset B$ and $A \supset B \to A = B$
- **Distributive** - $A \cap (B \cup C) = (A \cup B) \cup (A \cup C)$
- **DeMorgan's** - $(A \cup B)' = A' \cap B'$
- $A = (A \cap B) \cup (A \cap B')$

Counting Methods

- **Multiplcation Principle** - Given r experiments performed **sequentially** and each has $n_1, n_2, ..., n_r$ outcomes. After r experiments, there are $n_1 n_2 ... n_r$ outcomes.
- **Addition Principle** - Given experiment can be done in k different ways and each has $n_1, n_2, ..., n_r$ ways. There are $n_1 + n_2 + ... + n_k$ total ways.
- **Permutation** - ${}_nP_r = \frac{n!}{(n-r)!}$
- **Combination** - $\binom{n}{r} = \frac{n!}{(n-r)!r!}$

Probability

Axioms of Probability

1. For any event A, $0 \leq P(A) \leq 1$
2. P(S) = 1
3. If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

- $P(A') = 1 - P(A)$
- $P(A) = P(A \cap B) + P(A \cap B')$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- If $A \subset B$, then $P(A) < P(B)$

Finite Sample Space with Equally Likely Outcomes

Given sample space $S = \{a_1, ..., a_k\}$ and all outcomes are **equally likely**, i.e. $P(a_1) = ... = P(a_k)$:

$$\text{For any event } A \subset S, P(A) = \frac{\text{No. of sample points in A}}{\text{No. of sample points in S}}$$

Conditional Probability

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

Independence

- $A \perp B \leftrightarrow P(A \cap B) = P(A)P(B)$
- $A \perp B \leftrightarrow P(A|B) = P(A)$

Law of Total Probability

- **Partition** - If $A_1, ..., A_n$ are mutually exclusive events and $\bigcup_{i=1}^n A_i = S$, then $A_1, ..., A_n$ are partitions
- If $A_1, ..., A_n$ are partitions of S, then for any event B:

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

Bayes' Theorem

Let $A_1, ..., A_n$ be partitions of S. For any event B:

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^n P(B|A_k)P(A_i)}$$

02. Random Variables

- Motivation: Assign value to outcome of experiment
- **Random Variable** - Let S be sample space. Function X which maps $\mathbb{R}$ to every $s \in S$

Probability Distribution

- Probability assigned to each possible X
- Given RV X with range of R_x:
  - **Discrete** - Numbers in R_x are finite or countable
  - **Continuous** - R_x is interval

Discrete Probability Distribution

- **Probability Function** - Given $R_x = \{x_1, ...\}$. For each x_i, there's some probability that $X = x_i$:

$$f(x) = P(X = x)$$

- $p.f.$ must satisfy:
  1. $f(x_i) = P(X = x_i)$ for $x_i \in R_x$
  2. $f(x_i) = 0$ for $x_i \notin R_x$
  3. $\sum_{i=1}^{\infty} f(x_i) = 1$
  4. $\forall B \subseteq \mathbb{R}, P(X \in B) = \sum_{x_i \in B \cap R_x} f(x_i)$
- **Probability Distribution** - Collection of pairs $(x_i, f(x_i))$

Continuous Probability Distribution

- **Probability Function** - Given R_x is interval. Quantifies probability that X is in some range.
- $p.f.$ must satisfy:
  1. $f(x) \geq 0$
  2. $f(x) = 0$ for $x \notin R_x$
  3. $\int_{R_x} f(x)dx = 1$
  4. $\forall a, b$ s.t. $a \leq b, P(a \leq X \leq b) = \int_a^b f(x)dx$
- Note: $P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$

Cumulative Distributive Function

Given RV X, which can be discrete or continuous:

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

- $F(x)$ is non-decreasing and $0 \leq F(x) \leq 1$
- **For discrete RV**: Step function

$$F(x) = \sum_{t \in R_x; t \leq x} f(t)$$

  - $P(a \leq X \leq b) = F(b) - \lim_{x \to a^-} F(x)$
  - $0 \leq f(x) \leq 1$
- **For continuous RV**:

$$F(x) = \int_{-\infty}^x f(t)dt$$

$$f(x) = \frac{d(F(x))}{dx}$$

  - $P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a)$
  - $0 \leq f(x)$ e.g. $f(x) = 3x^2$ is a valid $p.f.$ since $\int_{R_x} f(x)dx = 1$

Expectation of Random Variable

- **Mean of discrete RV**:

$$\mu = E(X) = \sum_{x \in R_x} x_i f(x_i) = \sum_{i=1}^{\infty} P(X \geq i)$$

  - Let g be some function. $E(g(x)) = \sum_{x \in R_x} g(x)f(x)$
- **Mean of continuous RV**:

$$\mu = E(X) = \int_{x \in R_x} xf(x)dx$$

  - Let g be some function. $E(g(x)) = \int_{x \in R_x} g(x)f(x)dx$
- $E(aX + b) = aE(X) + b$
- Linearity of expectation: $E(X + Y) = E(X) + E(Y)$

Variance of Random Variable

$$\sigma_X^2 = V(X) = E((X - \mu_X)^2)$$

- **Variance of discrete RV**:

$$V(X) = \sum_{x \in R_x} (x - \mu_X)^2 f(x)$$

- **Variance of continuous RV**:

$$V(X) = \int_{x \in R_x} (x - \mu_X)^2 f(x)dx$$

- $V(X) = 0$ when X is a constant
- $V(aX + b) = a^2 V(X)$
- $V(X) = E(X^2) - (E(X))^2$
- **Standard Deviation** - $\sigma_X = \sqrt{V(X)}$

03. Joint Distributions

04. Special Probability Distributions

Discrete Uniform Distribution

- If X has values $x_1, x_2, ..., x_k$ with **equal probability**
- p.f.: $f_X(x) = \frac{1}{k}$ where $x = x_1, ..., x_k$ and 0 otherwise
- Expectation: $\mu_X = E(X) = \sum_{i=1}^k x_i f_X(x_i) = \frac{1}{k} \sum_{i=1}^k x_i$
- Variance: $\sigma_X^2 = V(X) = E(X^2) - (E(X))^2 = \frac{1}{k} \sum_{i=1}^k x_i^2 - \mu_X^2$

Bernoulli

- **Bernoulli Trial** - Random experiment with 2 possible outcomes (success and failure)

Bernoulli Random Variable

- Number of successes in Bernoulli trial (Either 1 or 0)
- Let $0 \leq p \leq 1$ be the probability of success in Bernoulli trial

$$f_X(x) = P(X = x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \\ 0 & otherwise \end{cases}$$

- $f_X(x) = p^x(1-p)^{1-x}$ for $x = 0$ or 1
- Notation: $X \sim Ber(p)$ and $q = 1 - p$
- $\mu_X = E(X) = p$ and $\sigma_X^2 = V(X) = p(1-p)$

Bernoulli Process

- Sequence of repeatedly performed **independent and identical** Ber. trials
- Generates sequence of **independet and identically distributed (i.i.d.)** Ber. RVs: $X_1, X_2, ...$

Binomial Distribution

- **Binomial RV** - Counts the number of successes in n trials in a Ber. process
- Given n trials with each trial having probability p of success:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- Notation: $X \sim B(n, p)$
- $E(X) = np$ and $V(X) = np(1-p)$

Negative Binomial Distribution

- Let X = Number of i.i.d. Bernoulli(p) trials until kth success occurs

$$P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

- Notation: $X \sim NB(k, p)$
- $E(X) = \frac{1}{p}$ and $V(X) = \frac{(1-p)}{p^2}$

Geometric Distribution

- Let X = Number of i.i.d. Bernoulli(p) trials until 1st success occurs

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

- Notation: $X \sim G(p)$
- $E(X) = \frac{1}{p}$ and $V(X) = \frac{1-p}{p^2}$

Poisson Distribution

- **Poisson RV** - Denotes number of events happening in **fixed period of time**

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

- Notation: $X \sim Poisson(\lambda)$ where $\lambda > 0$ is expected number of occurences during some period
- $E(X) = \lambda$ and $V(X) = \lambda$
- **Poisson Process** - Continuous time process, where we count number of correucesn within some internval of time