POPH90148: Probability & Distribution Theory - Notes

MODULE 1 - Introduction to Probability

Set Theory

$A \subset B$ (**subset**); A is contained in B, if every point in A is also in B
$\varnothing$ (**null set**); subset of every set
$A \cup B$ (**union**); the set of all points in A or B or both
$A \cap B$ (**intersection**); the set of all points in both A and B
A' (**complement**); the set of points that are in S (universal set) but not in A
$A \cap B = \varnothing$ (**mutually exclusive**); disjoint sets that have no points in common

Distributive laws given by:
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

De Morgan's Laws given by:
$(A \cap B)' = A' \cup B'$
$(A \cup B)' = A' \cap B'$

Probability Axioms

- An **experiment** is the process by which an observation is made.
- A **simple event** is an event that cannot be decomposed, corresponding to one and only one **sample point** (a distinct point).
- The **sample space** associated with an experiment is the set consisting of all possible sample points.
- A **discrete sample space** is one that contains either a finite or a countable number of distinct sample points.
- An **event** in a discrete sample space S is a collection of sample points - that is, any subset of the sample space.

Suppose S is a sample space associated with an experiment. To every event A in S (A is a subset of S), we assign a number, *P(A)*, called the **probability** of A, so the following axioms hold:

(A1) P(A) $\geq$ 0
(A2) P(S) = 1
(A3) For a sequence of pairwise mutually exclusive events in S:

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

Calculating Probability Using Combinatorics

$${}_n\mathbf{P}_r = \frac{n!}{(n-r)!}$$

$${}_n\mathbf{C}_r = \frac{n!}{r!(n-r)!}$$

- An ordered arrangement of r distinct objects is called a **permutation**; the number of ways of ordering n distinct objects taken r at a time.

The number of ways of partitioning n distinct objects into k distinct groups is:
$$N = \binom{n}{n_1\ n_2\ \ldots\ n_k} = \frac{n!}{n_1!n_2!\ldots n_k!}$$
The terms in the bracket are called **multinomial coefficients** because they occur in the expansion of the multinomial term.

The number of **combinations** of n objects taken r at a time is the number of unordered subsets, each of size r, that can be formed from the n objects.

The **binomial coefficient** is given by:
$$\binom{n}{r}$$
They are called binomial coefficients because they occur in the **binomial expansion**:
$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$$

Conditional Probability

The **conditional probability** of an event A, given that event B has occurred, is given by:
$$P(A\,|\,B) = \frac{P(A \cap B)}{P(B)}$$

- If the probability of event A is unaffected by occurrence or non-occurrence of event B then the two events are **independent**. Formally, this means:
- $P(A\,|\,B) = P(A)$
- $P(A \cap B) = P(A)P(B)$

- The **sensitivity** is the probability that someone with a condition will test positive.
- The **specificity** is the probability that someone without a condition will test negative.
- The **positive predictive value** is the probability that someone has a condition, given they tested positive.
- The **prevalence** is the probability of having a condition in a population.
- The **Law of Total Probability** states that if the sample space, S, has been split into mutually exclusive subsets then the probability of an event A in S can be calculated as:

$$P(A) = \sum_{i=1}^{k} P(A\,|\,B_i)P(B_i)$$
Thus, **Bayes Rule** is given by:
$$P(B_j\,|\,A) = \frac{P(A\,|\,B_j)P(B_j)}{\sum_{i=1}^{k} P(A\,|\,B_j)P(B_j)}$$

MODULE 2 - Discrete Random Variables

Random Variables

- A **random variable** is a real-valued function defined on the sample space; it is a variable (measured quantity) whose value is not known ahead of time. The RV assigns a number to each sample point in the sample space.
- A **probability distribution** assigns a probability to each possible value y of Y.
- A **discrete random variable** take on a finite or countable number of possible values.

For any discrete probability distribution, the following must be true:

1. $0 \leq p(y) \leq 1$ for all y
2. $\sum_{y} p(y) = 1$, where the summation is over all values of y with nonzero probability.

Bernoulli Distribution

Bernoulli distribution is a distribution for a discrete RV that:

- Can take on only two possible values, 0 and 1.
- $P(X = 1) = p$ and $P(X = 0) = 1 - p$ for some p in the range $0 \leq p \leq 1$.

The pmf is given by:

$P(X = x) = p^x(1-p)^{1-x}$ with mean $E(X) = p$ and $Var(X) = p(1 - p)$

Binomial Distribution

The **Binomial distribution** is a distribution for a discrete RV, denoted by $X \sim Bin(n,\ p)$, where, for $x = 0, 1, 2, \ldots, n$:

$P(X = x) = \binom{n}{x} p^x(1-p)^{n-x}$ with mean $E(X) = np$ and $Var(X) = np(1 - p)$

A **binomial experiment** possess the following properties:

1. The experiment consists of a fixed number, n, of identical trials.
2. Each trial results in one of two outcomes: success or failure.
3. The probability of success on a single trial is equal to some value p and remains the same from trial to trial.
4. The trials are independent.
5. The random variable of interest is X, the number of successes observed during the n trials.

Poisson Distribution

The **Poisson distribution** is a distribution for a discrete RV, denoted by $X \sim Pois(\lambda)$, where for any $x = 0, 1, 2, \ldots, n$:

$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$ with mean $E(X) = \lambda$ and $Var(X) = \lambda$

Note that the Poisson distribution often provides a good model for the probability distribution of the number of rare events that occur in space, time, volume or any other dimension.