REVIEW OF STAT 265

A simple event is an event that cannot be decomposed. The sample space associated with an experiment is the set consisting of all possible sample points, denoted by $\mathcal{S}$. That is $\mathcal{S} = \{E_0, E_1, ...\}$

A discrete sample space is the one that contains either a finite or a countable number of distinct sample points.

An event in a discrete sample space is a collection of sample points (any subset of $\mathcal{S}$).

Probability : Suppose $\mathcal{S}$ is a sample space associated with an experiment. To every event A in $\mathcal{S}$ we assign a number (measure), $P(A)$, called the probability of A such that the following axioms hold:

i). $P(A) \geq 0$

ii). $P(\mathcal{S}) = 1$

iii) If $A_1, A_{2,}...$ form a sequence of pair-wise mutually exclusive events in $\mathcal{S}$ (that is , $A_i \cap A_j = \phi$ if $i \neq j = 1, 2, .., \infty.$), then

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

Results on permutation and combination:

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 is given by

$$P_r^n = \ _nP_r = \frac{n!}{(n-r)!},$$

where, $n! = n(n-1)(n-2), ...3.2.1$ (multiplication of n consecutive integers from 1 to n).

The number of ways of choosing r distinct objects from n distinct objects ($r \leq n$) is given by

$$C_r^n = \binom{n}{r} =_n C_r = \frac{n!}{r!(n-r)!}.$$

The number of ways of partitioning n distinct objects into k distinct groups containing $n_1.n_2, ...n_k$ distinct objects, respectively, where each object appears in exactly one group and $\sum_{i=1}^{k} n_i = n$, is given by

$$M = \frac{n!}{n_1!n_2!....n_k!}.$$

Conditional Probability and the Independence of events:

Consider two events A and B.

The conditional probability of an event A, given that an event B has occurred is equal to

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

provided $P(B) > 0$. The notation $P(A|B)$ is read 'probability of A given B'. Similarly, the probability of B given A is given by $P(B|A) = \frac{P(A \cap B)}{P(A)}$ provided $P(A) > 0$.

Two events A and B are said to independent if any one of the following holds: i). $P(A|B) = P(A)$ ii). $P(B|A) = P(B)$

iii)$P(A \cap B) = P(A)P(B)$

otherwise, the events are said to be dependent.

Two laws of Probability

1. Multiplicative Law:

The probability of the intersection of two events A and B is given by

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

If A and B are independent events, then $P(A \cap B) = P(A)P(B)$.

2. Additive Law:

The probability of the union of two events A and B is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Note that if A and B are independent events, then $P(A \cup B) = P(A) + P(B)$. Further, one can verify $P(A) = 1 - P(\overline{A})$, where $\overline{A}$ is the compliment of event A. This result is useful when $P(\overline{A})$ is easier to calculate compared to $P(A)$.

Bayes' Rule

Suppose $\{B_1, B_2, ..B_k\}$ **is a partition of** $\mathcal{S}$ **(that is** $\mathcal{S}$ **=**$B_1 \cup B_2 \cup, ..., \cup B_k$ **and** $B_i \cap B_k = \phi$ **for** $i \neq j = 1, 2, .., k$), **such that** $P(B_i) > 0$, **for** $i = 1, 2, ..., k$. **Then, for any event** A

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

RandomVariables :

A random variable is a real-valued function for which the domain is a sample space.

Discrete Random Variable:

A random variable Y is said to be discrete if it can assume only a finite or countable infinite number of distinct values.

The probability that Y takes on the value y, $P(Y = y)$, (for notation simplicity this is denoted by $p(y)$) is defined as the sum of the probabilities of all sample points in $\mathcal{S}$ that are assigned the value y. **Probability Distribution for a Discrete random variable.** The probability distribution for a discrete random variable Y can be represented by a formula, a table, or a graph, which provides $p(y) = P(Y = y)$ for all y.

Expected Value of Y or $g(Y)$

Let Y be a discrete random variable with the probability function $p(y)$. Then the expected value of Y, $E(Y)$ is defined to be (provided it exists; see page 88 of your text)

$$E(Y) = \sum_y yp(y),$$

where the sum is taken over all y. Sometimes, $E(Y)$ is denoted by μ, the distribution(population) mean.

Let Y be a random variable with probability function $p(y)$ and $g(y)$ be a real-valued function of Y. Then the expected value of $g(Y)$ is given by

$$E[g(y)] = \sum_y g(y)p(y).$$

Variance of a random variable.

The variance of a random variable Y , denoted by $V(Y)$ is defined to be the expected value of $[Y - E(Y)]^2$. That is,

$$V(Y) = E[\{Y - E(Y)\}^2] = E(Y - \mu)^2 = \sum_y (y - \mu)^2 p(y)$$

The standard deviation of Y is the positive square root of $V(Y)$. Often, $V(Y)$ is abbreviated by σ^2.

Some Results on Expectation:

1. Let Y be a discrete random variable with probability function $p(y)$ and c be a constant. Then $E(c) = c$.

2. Let Y be a discrete random variable with probability function $p(y), g(Y)$ be a function of Y, and let c be a constant. Then

$$\begin{aligned} a). \quad E[cg(Y)] &= cE[g(Y)] \\ b). \quad E[c + g(Y)] &= c + E[g(Y)]. \end{aligned}$$

3. Let Y be a discrete random variable with probability function $p(y)$ and $g_1(Y), g_2(Y), ..., g_k(Y)$ be k functions of Y. Then

$$E[g_1(Y)+g_2(Y)+...+g_k(Y)] = E[\sum_{i=1}^{k} g_i(Y)] = \sum_{i=1}^{k} E[g_i(Y)] = E[g_1(Y)]+$$

4. Let Y be a discrete random variable with probability function $p(y)$, then

$$V(Y) = \sigma^2 = E[(Y-\mu)^2] = E(Y^2) - \mu^2, \text{ where } \mu = E(Y).$$

Tchebysheff's Theorem:

Let Y be a random variable with mean μ and finite variance σ^2. Then, for any constant $k > 0$,

$$P(|Y-\mu| < k\sigma) \geq 1 - \frac{1}{k^2} \text{ or}$$
$$P(|Y-\mu| \geq k\sigma) < \frac{1}{k^2}.$$

THE BINOMIAL PROBABILITY DISTRIBUTION

A binomial experiment possesses the following properties:

1. The experiment consists of a fixed number, n, of identical trials.

2. Each trial results in one of two outcomes. We will call one outcome success, S, and the other failure, F.

3. The probability of success on a single trial is equal to some value p and remains the same from trial to trial. The probability of a failure is equal to $q = 1 - p$.

4. The trials are independent.

5. The random variable of interest is Y, the number of successes observed during the n trials.

A random variable Y is said to have a binomial distribution based on n trials with success probability p if and only if

$$p(y) = \binom{n}{y} p^y (1-p)^{n-y}; \;\; y = 0, 1, 2, 3, ..., n \text{ and } 0 \leq p \leq 1.$$

We write $Y \sim Bin(n, p)$ to indicate that the random variable Y is distributed as a binomial with n trials and p as success probability p in a single trial.

SOME RESULTS ON BINOMIAL DISTRIBUTION.

$$
\begin{aligned}
E(Y) &= np \\
V(Y) &= np(1-p)
\end{aligned}
$$

Moment generating function: $M_Y(t) = E(e^{tY}) = [pe^t + (1-p)]^n$