Chapter - 2

Literatures and Research Studies

§ 2.1 Introduction:

The need of some theories / formulations in the field of statistics appeared in the works done under the project entitled "Probabilistic Forecasting of Time Series". Some of the needed theories / formulations are established ones that are available in many literatures. However, some needed formulations / theories are not established ones. Theses had to be developed due to necessity. In this chapter, a brief discussion on each of the theories / formulations that have been developed in this study have been presented below. The existing theories / formulations which have been applied in this study have also been outlined below.

§ 2.2 Definitions of Probability:

The three significant approaches, other than the two unscientific ones viz. **Intuitive Approach** (*Ref.* 70, 71, 111 &112) and **Subjective Approach** (*Ref.* 3) through which the theory of probability has been brought to the current stage of its development are

(1) **Classical Approach** introduced by *Bernoulli* (*Ref.* 5, 6, 35, 53 & 118),

(2) **Empirical Approach** developed by *von Mises* (*Ref.* 51, 65, 123, 124, 125 & 126 and by *Fisher* (*Ref.* 46 & 48)

and (3) **Axiomatic Approach** constructrd by *Bernstein* (*Ref.* 7 & 8) and by *Kolmogorov* (*Ref.* 67, 68 & 69).

The two definitions of probability that play the vital role in the current study are its classical definition due to *Bernoulli* and its empirical definition due to *von Mises* and *Fisher*.

§ 2.2.1 Classical Definition of Probability:

The classical definition of probability due to *Bernoulli* is as follows:

Definition (2.2.1): *If an experiment results in n possible exhaustive, mutually exclusive and equally likely outcomes and if out of these n outcomes m outcomes are favourable to an event E then the probability of the occurrence of the event E denoted by* P(*E*) *is defined by*

$$P(E) = \frac{\textit{Number of mutually exclusive and equally likely cases favorable to } E}{\textit{Number of mutually exclusive, equally likely and exhaustive cases of the experiment}} \qquad (2.2.1)$$

§ 2.2.2 Basic Properties of Probability:

Following properties of probability that are basic in nature follow from Definition (2.2.1).

Property (2.2.1) : For any event E ,

$$0 \leq P(E) \leq 1 .$$

Corollary (2.2.1): For any event E ,

$$P(\bar{E}) = 1 - P(E) \text{ i.e. } P(\bar{E}) + P(E) = 1$$

where $\bar{E}$ is the complementary event of E .

Corollary (2.2.2): For any event E ,

$$P(E) = \begin{cases} 0, & \text{iff } E \text{ is impossible event} \\ 1, & \text{iff } E \text{ is certain event} \end{cases}$$

Property (2.2.2) : If $\{E_1, E_2, \ldots\ldots\ldots\ldots, E_n\}$ is a set of mutually exclusive and exhaustive events of a random experiment,

$$P(E_1) + P(E_2) + \ldots\ldots\ldots\ldots + P(E_n) = 1.$$

Property (2.2.3) : If A and B are two mutually exclusive events,

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

In general, if $A_1, A_2, \ldots\ldots\ldots\ldots, A_n$ are n mutually exclusive events ,

$$P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = P(A_1) + P(A_2) + \ldots\ldots\ldots\ldots + P(A_n)$$

Property (2.2.4) : For any two events A and B,

$$P(A \cap B) = P(A)\, P(B|A) = P(B)\, P(A|B)$$

where $A|B$ is the conditional event of A given B and similar is the case of $B|A$.

Property (2.2.5) : If A and B are two independent events then

$$P(A \cap B) = P(A)\, P(B).$$

§ 2.2.3 Empirical Definition of Probability:

The empirical definition of probability due to *Fisher* is as follows:

Definition (2.2.2): *If the experiment is repeated N times under identical conditions and if out of N repetitions an event E_i occurs N_i times then* P(E_i), *the probability of*

R-44

occurrence of E_i, is a number towards which the ratio N_i N (called the relative frequency of occurrence of E_i) approaches as N becomes larger i.e.

$$\frac{N_i}{N} \rightarrow P(E_i) \quad \text{as} \quad N \rightarrow \infty$$

$$\text{i.e.} \quad P(E_i) = \underset{N\rightarrow\infty}{Lt} \frac{N_i}{N} \qquad (2.2.2)$$

The basic properties of probability described in § 2.2.2 can be established from this definition also.

§ 2.3 One Generalization of the Classical Definition of Probability:

The classical definition of probability holds good if and only if the sample points are equally likely. In the situation where the sample points are not equally likely, one cannot define the probability of an event with the help of sample space. To overcome this trouble an attempt, in this study, has been made to extend the classical definition of probability introduced by *Bernoulli* to the situation where the possible outcomes are not equally likely. This has been discussed below.

Let us consider a simple random experiment that results in the n possible outcomes (also known as (i) elementary events and (ii) as cases) viz.

$$e_1, e_2, \ldots\ldots\ldots\ldots\ldots, e_n \qquad (2.3.1)$$

These possible outcomes are obviously exhaustive and mutually exclusive. The set

$$S = \{ e_1, e_2, \ldots\ldots\ldots\ldots\ldots, e_n \} \qquad (2.3.2)$$

is nothing but the sample space of the experiment due to *von Mises*. This sample space is the set that contains all the possible outcomes of the experiment. Therefore, this sample space of the experiment can also be called the **outcome space** or the **elementary event space** of the experiment. If the possible outcomes of the experiment are equally likely, the probability of an event E denoted by P(E) can be defined with the help of the outcome space S given by (2.3.2) since

$$P(E)=\frac{\text{Number of mutually exclusive and equally likely cases favorable to } E}{\text{Number of mutually exclusive, equally likely and exhaustive cases of the experiment}} \qquad (2.3.3)$$

and since both of the numerator and the denominator of the expression in the right hand side of equation (2.3.3) can be obtained directly from the outcome space S. However, if the possible outcomes are not equally likely, the probability of an event E associated to the experiment cannot be defined with the help of its outcome space since in this situation the numerator and the denominator of the ratio in the right hand side of equation (2.3.3) cannot be obtained from the outcome space S.

Now,

$$e_1, e_2, \ldots\ldots\ldots\ldots, e_n \text{ are equally likely}$$

$$\Leftrightarrow \mathrm{P}(e_1) \quad \mathrm{P}(e_2) \quad \ldots\ldots\ldots\ldots \quad \mathrm{P}(e_n) \quad 1/n \qquad (2.3.4)$$

But if

$$e_1, e_2, \ldots\ldots\ldots\ldots, e_n$$

are not equally likely then the equality in the right hand side of the tautology given by (2.3.4) does not hold good.

Let us consider the fact that the outcomes $e_1, e_2, \ldots\ldots\ldots\ldots, e_n$ are not necessarily equally likely. Then $\mathrm{P}(e_1)$, $\mathrm{P}(e_2)$, $\ldots\ldots\ldots\ldots$, $\mathrm{P}(e_n)$ are not necessarily identical. Suppose,

$$\mathrm{P}(e_i) \quad f_i \ N, \ (i = 1, 2, 3, \ldots\ldots, n) \qquad (2.3.5)$$

Then

$$\left.\begin{array}{ll} & \sum\limits_{i=1}^{n} f_i \quad N \\ \\ \text{since} & \sum\limits_{i=1}^{n} \mathrm{P}(e_i) = 1 \end{array}\right\} \qquad (2.3.6)$$

Here the outcome space is nothing but the set S, given by equation (2.3.2), containing n cases which does not admit to define $\mathrm{P}(e_i)$ as well as to define $\mathrm{P}(E)$ for any other event E.

Now,

$$\mathrm{P}(e_i) = f_i \ N$$

$$(\text{for } \ i = 1, 2, 3, \ldots\ldots, n)$$

means that had there been n mutually exclusive, equally likely and exhaustive cases, f_i cases would have been favourable to the elementary event e_i $(i = 1, 2, 3, \ldots, n)$. This means that the outcome space would have been S' where

$$S' \quad \{e_1, e_1, \ldots\ldots\ldots, e_1, e_2, e_2, \ldots\ldots\ldots\ldots, e_2, \ldots\ldots\ldots\ldots\ldots, e_n, e_n, \ldots\ldots\ldots, e_n\} \qquad (2.3.7)$$

containing N mutually exclusive, equally likely and exhaustive cases, treating each of them to be distinct cases, among which the outcome e_i appears f_i times $(i = 1, 2, \ldots\ldots, n)$. In other words, the outcome space would have been

$$S' = \{g_1, g_2, \ldots\ldots\ldots\ldots, g_N\} \tag{2.3.8}$$

containing N mutually exclusive, equally likely and exhaustive cases where

$$g_i = \begin{cases} e_1 & \text{for } i = 1, 2, \ldots\ldots\ldots, f_1 \\ e_2 & \text{for } i = f_1+1, f_1+2, \ldots\ldots\ldots\ldots, f_1+f_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ e_n & \text{for } i = \sum_{i=1}^{n} f_i+1, \sum_{i=1}^{n} f_i+2, \ldots\ldots\ldots \sum_{i=1}^{n} f_i \end{cases} \tag{2.3.9}$$

The space S' enables one to define $P(e_i)$ by the classical approach yielding

$$P(e_i) = f_i / N \ , \qquad (i = 1, 2, 3, \ldots\ldots\ldots, n)$$

Also, the probability of any other event E associated to the experiment can be defined by the classical approach with the help of the set S' since the requirements in defining $P(E)$ by the classical approach are fulfilled by S'. Thus, if an experiment results in n possible outcomes viz.

$$e_1, e_2, \ldots\ldots\ldots\ldots\ldots, e_n$$

which are not necessarily equally likely with

$$P(e_i) = f_i / N \qquad \text{for } i = 1, 2, 3, \ldots\ldots\ldots, n.$$

is conducted then its outcome space can be thought of as equivalent to the space S' given by the expression (2.3.7) or equivalently by the expression (2.3.8) which enables one to define the probability of an event as per the logic/philosophy behind the classical approach to probability. This space S' is thus the **probability definable space** of the experiment. Thus, one can define the probability definable space and the probability as follows.

Definition (2.3.1):

Definition ofProbability Definable Space

Suppose, the outcome space of an experiment is

$$S(e) = \{ e_1, e_2, \ldots\ldots\ldots\ldots\ldots, e_k \}$$

with

$$P(e_i) = n_i / N \ , \qquad (i = 1, 2, \ldots\ldots, k)$$

Then, the set $S(e')$ given by

$$S(e') = \{ e_1', e_2', e_3', \ldots\ldots\ldots\ldots\ldots, e_N' \}$$

where e_x' is a function of x with the domain

$$I_N = \{1, 2, 3, \ldots\ldots, N\} \tag{2.3.10}$$

and with the counter domain $S(e)$ with e_x' defined by

$$e_x \quad \begin{cases} e_1 & \text{for } x = 1, 2, \ldots\ldots\ldots\ldots, n_1 \\ e_2 & \text{for } x = n_1+1, n_1+2, \ldots\ldots\ldots, n_1+n_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ e_k & \text{for } x \ \ \sum_{i=1}^{k-1} n_i+1,\ \sum_{i=1}^{k-1} n_i+2, \ldots\ldots, \sum_{i=1}^{k-1} n_i = N \end{cases} \qquad (2.3.11)$$

can be called the probability definable space of the experiment .

Definition (2.3.2):

Definition ofProbability

The probability of an event E associated to a random experiment can be defined by

$$\mathrm{P}(E) \quad \frac{\text{Number of case in the probability definable space favorable to } E}{\text{Number of cases in the probability definable space of the experiment}} \qquad (2.3.12)$$

The basic properties of probability described in § 2.2.2, which follow from both of the definitions viz. the classical definition and the empirical definition can be derived from this definition also.

§ 2.4 Link between the Classical Definition and the Empirical Definition of Probability:

In this section, a brief discussion on an attempt that has been made to search for if there exists any link between the two definitions of probability viz. the classical definition and the empirical definition has been presented.

As earlier suppose, an experiment results in n possible outcomes viz.

$$e_1,\ e_2,\ \ldots\ldots\ldots\ldots,\ e_n$$

These are obviously mutually exclusive and exhaustive. They may be equally likely or may not be equally likely. If the experiment is repeated N times under identical conditions and if out of N repetitions the event e_i occurs N_i times then by the empirical definition of probability, the probability of occurrence of e_i denoted by $\mathrm{P}(e_i)$ is a number p_i towards which the ratio (known as the relative frequency of occurrence of e_i) N_i/N approaches as N becomes larger i.e. $\mathrm{P}(e_i) = p_i$ is a number such that

$$\frac{N_i}{N} \rightarrow p_i = \mathrm{P}(e_i) \text{ as } N \rightarrow \infty$$

i.e. $$P(e_i) = p_i = \underset{N \to \infty}{Lt} \frac{N_i}{N} \qquad (2.4.1)$$

This can be expressed as

$$P(e_i) = \frac{N_i}{N} + V_i(N) \qquad (2.4.2)$$

where $V_i(N) \to 0$ as $N \to \infty$

Here, $V_i(N)$ is the difference between the probability of occurrence of e_i and the relative frequency of occurrence of e_i among N repetitions that moves towards zero as the number of repetitions N becomes larger. Hence, $V_i(N)$ can be called the **vanishing term** of the event e_i in N repetitions of the experiment.

Basic Properties of P(e_i)

The basic properties of $P(e_i)$ are discussed below.

Property (2.4.1.): For any event e_i ,

$$0 \le P(e_i) \le 1 .$$

Proof: This follows from the fact that the number of occurrences of e_i in any number of repetitions of the experiment is non-negative and cannot exceed the number of repetitions.

Property (2.4.2): The sum of probabilities of the elementary events is unity i.e

$$\sum_{i=1}^{n} P(e_i) = 1$$

Proof :

$$\sum_{i=1}^{n} P(e_i)$$

$$= \sum_{i=1}^{n} \left(\underset{N \to \infty}{Lt} N_i/N \right) \quad \text{, by equation (2.4.1)}$$

$$= \underset{N \to \infty}{Lt} \left(\sum_{i=1}^{n} \frac{N_i}{N} \right) \quad \text{, by additive property of limit}$$

$$= 1 .$$

Property (2.4.3.): For two elementary events e_i *and* e_j ,

$$P(e_i \cup e_j) = P(e_i) + P(e_j)$$

Proof: If e_j occurs N_j times and e_i occurs N_i times out of N repetitions of the experiment, $e_i \cup e_j$ occurs $(N_i + N_j)$ times out of the same N repetitions because of the fact that e_i and e_j are mutually exclusive. Hence

$$P(e_i \cup e_j) = \underset{N \to \infty}{Lt} \left(\frac{N_i + N_j}{N} \right), \text{ by equation (2.4.1)}$$

$$= \underset{N \to \infty}{Lt} \frac{N_i}{N} + \underset{N \to \infty}{Lt} \frac{N_j}{N}$$

$$= P(e_i) + P(e_j), \quad \text{by equation (2.4.2).}$$

Generalization: Generalizing it , one can obtain that

$$P\left(\bigcup_{i=1}^{m} e_{ii} \right) = \sum_{i=1}^{m} P(e_{ii}), \quad m \leq n$$

Corollary (2.4.1): $P\left(\bigcup_{i=1}^{n} e_i \right) = 1$

i.e. probability of occurrence of either of the elementary events of a random experiment is unity.

Proof:

$$P\left(\bigcup_{i=1}^{n} e_i \right)$$

$$= \sum_{i=1}^{n} P(e_i), \text{ by the property (2.2.3)}$$

$$= \sum_{i=1}^{n} \underset{N \to \infty}{Lt} \frac{N_i}{N}, \text{ by equation (2.4.1)}$$

$$\underset{N \to \infty}{Lt} \sum_{i=1}^{n} \frac{N_i}{N}, \text{ by additive property of limit}$$

$$1, \text{ since } \sum_{i=1}^{n} N_i \quad N .$$

Property (2.4.4): $P(e_i) \quad 1 \quad P(e_i)$

where $\overline{e}_i$ is the complementary event of e_i.

Proof: The event $\overline{e_i}$ is the non-occurrence of e_i. Thus if e_i occurs N_i times out of N repetitions, e_i does not occur $N - N_i$ times out of N repetitions. Hence

$$P(\overline{e_i})$$

$$= \underset{N \to \infty}{Lt} \frac{N - N_i}{N} \quad , \quad \text{by equation (2.4.1)}$$

$$-- 1 - P(e_i)$$

This property can also be obtained from the earlier property also as follows.

$\overline{e_i}$ is the non-occurrence of e_i. This means that $\overline{e_i}$ is the occurrence of either of the elementary events e_j ($j = 1, 2, \ldots, n$ & $j \neq i$). Hence

$$P(\overline{e_i})$$

$$P(\bigcup_{i \neq l\ j \neq i}^{n} e_i)$$

$$= \sum_{j\ l\ j \neq i}^{n} P(e_j) \text{, by the generalization of property (2.4.3)}$$

$$= \sum_{j=l}^{n} P(e_j) - P(e_i)$$

$$= 1 - P(e_i) \text{, by the corollary of property (2.4.1).}$$

Basic Properties of $V_i(N)$

The following two properties of $V_i(N)$ play the vital role in the case of empirical probability.

Property (2.4.5): $V_i(N) \to 0$ as $N \to \infty$

$$\text{i.e. } \underset{N \to \infty}{Lt} V_i(N) \quad 0$$

Proof: This is obvious from equation (2.4.2)

Property (2.4.6):

$$\sum_{i\ l}^{n} V_i(N) = 0$$

i.e. the sum of the vanishing terms corresponding to the elementary events in any number of repetitions vanishes.

Proof: We have

$$\sum_{i\ l}^{n} P(e_i) \quad 1 , \qquad \text{by property (2.4.2)}$$

Hence,

$$\sum_{i=1}^{n} \left\{ \frac{N_i}{N} + V_i(N) \right\} = 1, \qquad \text{by equation (2.4.2)}$$

Hence, $$\sum_{i=1}^{n} V_i(N) = 0, \quad \text{since} \quad \sum_{i=1}^{n} N_i = N$$

Definition (2.4.1):

Definition of Empirical Probability in terms of $V_i(N)$

With the help of the above two properties of $V_i(N)$ viz. Property (2.4.5) & Property (2.4.6) and equation (2.4.2), the empirical definition of probability can also be expressed as follows:

If an experiment, having the elementary events $e_1, e_2, \ldots\ldots, e_n$, is repeated N times under identical conditions and if out of these N repetitions an event e_i occurs N_i times then the probability of occurrence of e_i denoted by $P(e_i)$ *can be defined by*

$$P(e_i) = \frac{N_i}{N} + V_i(N)$$

where, (i) $V_i(N) \to 0$ *as* $N \to \infty$

& (ii) $\sum_{i=1}^{n} V_i(N) = 0$

§ 2.4.1 Random Events and Empirical Probability:

A random (or contingent) event associated to a random experiment is a combination of some or all of the elementary events of the experiment. For a random event associated to a random experiment a lemma that can be proved by constructing events is stated now without proof.

Lemma (2.4.1): *For any random event associated to a random experiment there exists a set of mutually exclusive and exhaustive events associated to the experiment that contains the said event. This set is not necessarily unique.*

Let E be a random event associated to the experiment. Then by the above lemma, there exists a set of mutually exclusive and exhaustive events, say,

$$A_1, A_2, \ldots\ldots\ldots\ldots\ldots\ldots, A_r$$

associated to the experiment that contains E. Thus, if E is the event A_j for some j $(1 \le j \le r)$ then by the empirical definition of probability due to *Fisher* ,

$$P(E) = P(A_i) = \underset{N \to \infty}{Lt} \quad \frac{N(A_i)}{N}$$

where $N(A_i)$ is the number of times the event A_i occurs out of N repetitions of the trial This means,

$$P(E) = P(A_i) = \frac{N(A_i)}{N} \quad | \quad \underset{N \to \infty}{Lt} \quad V(A_i : N)$$

where $V(A_i : N) \to 0$ as $N \to \infty$.

Here $V(A_i : N)$ can be called the vanishing term, as described in § 2.4, corresponding to the event A_i in the N repetitions.

Basic Properties of $P(A_i)$ and of $V(A_i : N)$

Applying similar logic as in the cases of $P(e_i)$ and $V_i\,(N)$ described in § 2.4, one can obtain the following properties of $P(A_i)$ and $V(A_i : N)$.

Property (2.4.7): For a random event A_i

$$0 \le P(A_i) \le 1$$

Property (2.4.8) : $\sum_{i=1}^{r} P(A_i) = 1$

i.e. the sum of probabilities of the events in a set of mutually exclusive and exhaustive events is unit.

Property (2.4.9) : For the events $A_1, A_2, \ldots\ldots\ldots\ldots, A_q \; (q \le r)$

$$P\Big(\bigcup_{i=1}^{q} A_i\Big) = \sum_{i=1}^{q} P(A_i)$$

Corollary (2.4.2): For the events $A_1, A_2, \ldots\ldots\ldots\ldots, A_r$ we have

$$P\Big(\bigcup_{i=1}^{r} A_i\Big) = 1$$

i.e. the probability of occurrence of either of the events $A_1, A_2, \ldots\ldots\ldots\ldots, A_r$ is unity.

Property (2.4.10) : For an event A_i

$$P(\overline{A}_i) \quad 1 \quad P(A_i)$$

where $\overline{A}_i$ is the complementary event of A_i.

Property (2.4.11) : For an event A_i

$$\underset{N \to \infty}{Lt} V(A_i : N) = 0$$

Property (2.4.12) : $\sum_{i=1}^{r} V(A_i : N) = 0$

i.e. the sum of the vanishing terms in any repetition corresponding to the events in a set of mutually exclusive and exhaustive events vanishes.

Definition (2.4.2):

Definition of Empirical Probability in terms of Vanishing Term

With the help of the above two properties of $V(A_i : N)$ viz. Property (2.4.11) and Property (2.4.12), the empirical probability of a random event can be defined in the following pattern also.

If an experiment is repeated N times under identical conditions and if out of N repetitions a random (contingent) event E, which is an element Ai (say) of a set of mutually exclusive and exhaustive events viz. $\{A_1, A_2, \ldots\ldots, A_r\}$ *associated to the experiment, occurs* $N(A_i)$ *number of times then the probability of occurrence of E denoted by* $P(E)$ *is defined by*

$$P(E) = P(A_i) = \frac{N(A_i)}{N} + V(A_i : N)$$

$$\textit{where (i)} \quad \underset{N \to \infty}{Lt} V(A_i : N) = 0$$

$$\& \quad (ii) \quad \sum_{i=1}^{r} V(A_i : N) = 0 .$$

Here $V(A_i : N)$ is the vanishing term in N repetitions corresponding to the event A_i.

Note: It can be shown that this definition implies the definition of empirical probability due to *Fisher* and vice versa. Hence, the two definitions are equivalent.

§ 2.4.2 Classical Probability As a Special Case:

Let the random experiment be such that the possible outcomes i.e. the elementary events

$$e_1, e_2, \ldots\ldots\ldots\ldots, e_n$$

are equally likely i.e. each of them has equal probability of being occurred. Now,

$e_1, e_2, \ldots\ldots\ldots\ldots, e_n$ are equally likely

$\Leftrightarrow \quad P(e_1) = P(e_2) = \ldots\ldots\ldots = P(e_n)$

$\Leftrightarrow \quad P(e_i) = 1/n$, for all i.

(since $\sum_{i=1}^{n} P(e_i) = 1$, by property (2.2.2)

Let A be an event, not necessarily elementary, associated to the experiment. Then A is a combination of some of the elementary events $e_1, e_2, \ldots\ldots, e_n$. Without loss of generality let A be a combination of $e_1, e_2, \ldots\ldots, e_m$ ($m \leq n$). Then

$$A = \bigcup_{i=1}^{n} e_i$$

Using Property (2.4.3) one can obtain that

$$P(A) = \sum_{i=1}^{m} P(e_i)$$

$$= \frac{m}{n} \quad , \quad \text{since } P(e_i) = 1/n \text{, for all } i$$

$$\text{i.e. } P(A) = \frac{\text{Number of elementary events favourable to A}}{\text{Number of mutually exclusive, equally likely and exhaustive elementary events}}$$

This is the classical definition of P(A) the probability of occurrence of the event. Thus the classical definition of probability can be obtained from its empirical definition when the elementary events are equally likely.

<u>**Some Remarks**</u>

Remark (2.4.1): The vital properties of probability discussed in § **2.2.1** can be derived from this definition also.

Remark (2.4.2) : The set

$$\{ e_1, e_2, \ldots\ldots\ldots\ldots, e_n \}$$

is nothing but the elementary event set (or space), abbreviated as EES, of the experiment while the set

$$\{ e_1, e_2, \ldots\ldots\ldots\ldots, e_m \}$$

is nothing but the favourable elementary event set (or space), abbreviated as FEES, in the EES of the event A so that

n = the size of the EES of the experiment

and m = the size of the FEES in the EES of A.

Therefore, P(A) can be summarized as

$$P(A) = \frac{\text{The size of the FEES, in the ESS, of } A}{\text{The size of the EES of the experiment}}$$

§ 2.5 A Theoretical Definition of Probability:

Here, a brief discussion on an attempt that has been made to search for an approach of defining probability that can be a basis of searching for method / methods of determining the exact value of the probability of an event has been presented

Consider an experiment that has two possible outcomes viz.

$$e_1 \ \& \ e_2$$

Now and onwards, performing of an experiment (i.e. an experimentation) will be called perfect if it does not contain any error (even the random error too). If the experiment is performed once, any one of the two possible outcomes may occur. But, none of them is sure to be occurred. This is due to the fact that in one trial only one of the two possible outcomes can occur. Since e_1 & e_2 are possible outcomes, each of them has a chance (and hence a probability which is nothing but a measure of chance) of occurring in a trial. By the statement

" the event e_1 has a chance of occurring in a trial"

we mean that the event e_1 will occur sometimes if the trial is repeated a sufficient number of times and if the experimentation is a perfect one. Similar is the case of e_2. Now, e_1 & e_2 may have equal chances or different chances. If they have equal chances, they will occur equal number of times if the trial is repeated even number of times and if the experimentation is perfect. If the chance of occurrence of e_1 is more than that of e_2, under similar repetitions of the trial e_1 will occur more times than that of e_2 and vice versa. Thus, if

$$P(e_1) = P(e_2)$$

then each of e_1 & e_2 will occur

once out of 2 repetitions of the trial ,

twice out of 4 repetitions of the trial ,

....

n times out of $2n$ repetitions of the trial ,

.... ,

provided the experimentation is a perfect one.

Conversely, if the above picture of occurrences of e_1 & e_2 happens then this means that e_1 & e_2 have equal chances of occurrence i.e.

$$P(e_1) = P(e_2).$$

Similarly, if

$$P(e_1) = m.P(e_2)$$

then

e_1 will occur m times and e_2 will occur once out of $m+1$ trials ,

e_1 will occur $2m$ times and e_2 will occur 2 times out of $2(m+1)$ trials ,

....,

e_1 will occur nm times and e_2 will occur n times out of $n(m+1)$ trials ,

....,

provided the experimentation is perfect.

Conversely, if the above picture of occurrences of e_1 & e_2 happens then this means that the chance of occurrence of e_1 is m times than that of e_2 i.e.

$$P(e_1) = m\,P(e_2) .$$

Now, to define the probability the following three facts are to be considered:

(1) Probability has been defined in terms of relative frequency in the empirical approach due to *Fisher* and *von Mises*.

(2) The classical definition of probability is a special case of its empirical definition (discussed in § 2.4.2).

(3) The axioms or the postulates by which the probability is defined in the axiomatic approach by *Bernstein* and *Kolmogorov* are nothing but the theorems derived from its classical definition due to *Bernoulli*.

These three facts together establish that the definition of probability in terms of relative frequency can be treated as a basic definition of probability. Now, in the first picture of e_1 & e_2, the relative frequencies of occurrences of e_1 are

$$½,\ 2/4 = ½,\ 3/6 = ½,\ \ldots\ldots\ldots\ldots ,\ m/(2m) = ½ ,\ \ldots\ldots$$

that are constant and equal to ½ . Hence,

$$P(e_1) = ½ .$$

Similarly, for the outcome e_2 ,

$$P(e_2) = ½ .$$

Applying similar logic as in the case of the first picture of occurrences of e_1 & e_2 to the second picture of occurrences of e_1 & e_2 , one can obtain that

$$P(e_1) = \frac{m}{m+1}$$

$$\text{and}\qquad P(e_2) = \frac{1}{m+1}$$

Thus it has been found that

(1) if $P(e_1) = ½ = P(e_2)$ then the first picture of occurrences of e_1 & e_2 happens and vice versa

and (2) if $P(e_1) = \frac{m}{m+1}$ & $P(e_2) = \frac{1}{m+1}$

then the second picture of occurrences of e_1 & e_2 happens and vice versa .

In a similar manner one can arrive at the fact that if

$$P(e_1) \quad \frac{m}{M}$$

then e_1 will occur

m times out of M repetitions of the trial ,
$2m$ times out of $2M$ repetitions of the trial,
.... ,
nm times out of nM repetitions of the trial ,
....

provided the experimentation is a perfect one and conversely if the above picture of occurrences of e_1 & e_2 happens in a perfect experimentation then

$$P(e_1) \quad \frac{m}{M} .$$

In general, if a random experiment has k possible outcomes viz.

$$e_1, e_2, \ldots\ldots\ldots\ldots, e_k$$

then by applying similar logic one can arrive at the fact that if

$$P(e_i) = m_i / M , \quad (i = 1, 2, \ldots\ldots, k)$$

then e_i will occur

m_i times out of M repetitions of the trial ,
$2m_i$ times out of $2M$ repetitions of the trial ,
.... ,
nm_i times out of nM repetitions of the trial ,
.... ,

provided the experimentation is a perfect one and vice versa

Note : From the above theoretical facts it is clear that an arbitrary event E occurs nr times out of nR repetitions of the trial (n = 1, 2, ...) for unknown r and R . Thus the probability can be defined theoretically as follows:

Definition (2.5.1): *If a trial is repeated in such a way that the experimentation is a perfect one and if an event E occurs nm times out of nM repetitions of the trials for n =* 1, 2, 3, ... *then the probability of occurrence of the event E, denoted by* P(E) , *is defined by*

$$P(E) \quad \frac{m}{M} .$$

Definition (2.5.2): *The probability of an event E associated to a random experiment is a rational number (m/M) such that if the trial is repeated in such a way that the experiment is a perfect one, the event E will occur nm times out of nM repetitions of the trial for n =* 1, 2, 3, ...

Combining these two definitions (or equivalently from the facts mentioned above) one can obtain the following fundamental proposition:

Proposition (2.5.1): *If E is an event associated to a random experiment then*

$$P(E) \quad \frac{m}{M}$$

if and only if E occurs nm times out of nM repetitions of the trial for n = 1, 2, 3, 4, ..., *provided the experimentation is a perfect one.*

<u>Basic Properties</u>

The basic properties of probability discussed in § 2.2.1 can be established from this definition also.

§ 2.6 A Method of Determining the Exact Value of the Probability of an Event:

Now, a brief discussion on an attempt of searching for a method of determining the value of the probability of an event has been outlined below.

As earlier here also we consider a random experiment having possible outcomes

$$e_1, e_2, \ldots\ldots\ldots\ldots, e_k$$

with respective probabilities

$$p_1, p_2, \ldots\ldots\ldots, p_i$$

Suppose, the associated trial is repeated a sufficiently large number of times (say, N times) under identical condition and the event e_i occurs $f(e_i : N)$ times out of these N repetitions. Then $R(e_i : N)$, given by,

$$R(e_i : N) = \frac{f(e_i : N)}{N}$$

is the relative frequency of occurrence of e_i out of N repetitions of the trial.

Suppose, E is an arbitrary event (elementary or other) associated to the random experiment with

$$P(E) = \frac{m}{M}$$

where m & M are unknown. Of course, m and M are positive integers with $m \leq M$. The aim here is to determine the value of $P(E)$. If m/M becomes equal to some relative frequency / frequencies then it can be equal to the relative frequencies

$$R(E : r, M),\ R(E : r, 2M), \ldots\ldots\ldots\ldots,\ R(E : r, nM), \ldots\ldots\ldots\ \text{etc.}$$

(for $r = 1, 2, 3, \ldots\ldots$) where $R(E : r, M)$ denotes the relative frequency of occurrence of E out of M successive trials treating the r^{th} trial as the first one. It can never be equal to any relative frequency other than these.

Consider the relative frequencies

$$R(E : r, M) = \frac{f(E : r, M)}{M}, \quad (r = 1, 2, 3, \ldots\ldots),$$

where $f(E : r, M)$ denotes the frequency of occurrence of E out of M successive trials treating the r^{th} trial as the first one. The ratio m/M can be equal to $R(E : r, M)$ if the associated experimentation is a perfect one. In reality, however, it may not be equal to $R(E : r, M)$ due to the fact that no experimentation in reality is free from random error which is uncontrollable in any real situation and hence $f(E : r, M)$ is influenced by the presence of this error.

In a trial, the frequency of occurrence of E is either '0' or '1'. It is '1' if E occurs and '0' if it does not occur in the trial. When the true frequency of E is '1', due to the presence of random error the observed frequency must be

either '0' if the random error exists

or '1' if the random error does not exist

(since the only two possible frequencies are '0' & '1'). Thus in this case the possible values of the random error are

$$-1 \ \& \ 0 \ .$$

Similarly when the true frequency of E is '0', the possible values of the random error are

$$0 \ \& \ 1 \ .$$

Therefore, the frequency $f(E : r, M)$ suffers from an error

$$e(E : r, M) = \sum_{i=r}^{r+M-1} d(E : i)$$

where $d(E : i)$ assumes the values

(i) either '0' or '1' corresponding to each of the $f(E : r, M)$ trials that results in the occurrence of E

and (ii) either '-1' or '0' corresponding to each of the $M - f(E : r, M)$ trials that results in the non-occurrence of E.

Since d$(E : i)$ occurs due to the random error and since its possible values are

$$-1, 0 \ \& \ 1$$

therefore, for some value or values of r the term $e(E : r, M)$ vanishes . This means that

$$\frac{m}{M} = R(E : r, M) \, , \quad \text{for the same value or values of } r \, .$$

Similarly, the ratio viz.

$$\frac{m}{M}$$

will be equal to each of the following viz.

$R(E : r, 2M)$ for some value or values of r ,

$R(E : r, 3M)$ for some value or values of r ,

.... ,

$R(E : r, kM)$ for some value or values of r ,

....

Therefore, the ratio m/M, which is equal to the value of P(E), will be the relative frequency that corresponds to each of the following viz.

the M^{th} trial in one or more sets of the collection $\check{C}(E)$,

the $2M^{th}$ trial in one or more sets of the collection $\check{C}(E)$.

.... ,

the kM^{th} trial in one or more sets of the collection $\check{C}(E)$.

.... , where

$$\check{C}(E) = \{ \, S(E : r) : r = 1, 2, 3, \ldots\ldots \}$$

with

$$S(E : r) = \{ \, R(E : r, n) : \; n = 1,2, \ldots\ldots, M, \ldots\ldots, 2M, \ldots\ldots\ldots, kM, \ldots\ldots, N \},$$
$$(r = 1, 2, 3, \ldots\ldots) .$$

Thus, P(E) is a fixed relative frequency that corresponds to the T_i^{th} (i = 1, 2, 3, ...) trials in one or more sets in the collection $\check{C}(E)$ where T_{i+1} is a multiple of T_i . Thus the following theorem has been obtained.

Theorem (2.6.1): *If E is an event associated to a random experiment and if the associated trial is repeated a sufficiently large number of times under identical condition then* P(E), *the probability of occurrence of the event E, is a fixed relative frequency that corresponds to each of the $(mT)^{th}$ trials (m = 1, 2, 3, ...) in one or more sets in the collection $\check{C}(E)$ given by*

$$\check{C}(E) = \{ \, S(E : r) : r = 1, 2, 3, \ldots\ldots \}$$

with

$$S(E : r) = \{ \, R(E : r, n) : \; n = 1,2, \ldots\ldots, M, \ldots\ldots, 2M, \ldots\ldots\ldots, kM, \ldots\ldots, N \}$$

where $R(E : r, n)$ is the relative frequency of occurrence of E out of n trials treating the r^{th} trial as the first one and ignoring the earlier r-1 trials .

Note: In the theorem, the term "sufficiently large number of repetitions of the trial" is used to mean that the trial is repeated such a number of times so that it is found that a fixed.relative frequency corresponds to each of the $(mT)^{th}$ trials (m = 1, 2, 3, ...) at least in one set belonging to the collection $\check{C}(E)$.

§ 2.6.1 Method of Determining P(E):

In Theorem (2.6.1), one can observe that the method of determining the value of P(E) consists of

(1) performing the associated trial a sufficiently large number of times under identical condition,

(2) constructing the sets

$$S(E : r) = \{ \, R(E : r, n) : \; n = 1,2, \ldots\ldots, M, \ldots\ldots, 2M, \ldots\ldots\ldots, kM, \ldots\ldots, N \}$$

for r = 1, r = 2, r = 3, etc. (as many as required)

to obtain the collection

$$\check{C}(E) = \{ \, S(E : r) : r = 1, 2, 3, \ldots\ldots \}$$

and (3) to observe if in the collection $C(E)$ there exists a fixed relative frequency that corresponds to each of the $(mT)^{th}$ trials ($m = 1, 2, 3, \ldots.$) at least in one of the sets belonging to $C(E)$.

Thus, in determining the value of P(E) one needs to proceed with the following steps:

Steps in the Method

From the above theorem, it is clear that to determine the value of P(E) one needs to proceed with the following steps:

Step (1): Repeat the trial a large number of times (say, N times) under identical condition and note down the frequency of occurrence of E at each repetition. The frequency of occurrence of E at any repetition is '1' if E occurs at the repetition and '0' otherwise.

Step (2): Calculate $f(E : n, 1)$, the frequency of occurrence of E out of n repetitions starting from the 1st trial, for $n = 1, 2, \ldots\ldots, N$ by the formula

$$f(E : n, 1) = \sum_{i=1}^{n} b(E : i)$$

$$\text{where} \qquad b(E : i) \begin{cases} 1, & \text{if } E \text{ occurs} \\ 0, & \text{otherwise}. \end{cases}$$

Step (3): Construct the set

$$S(E : 1) = \{ R(E : n, 1) : n = 1, 2, \ldots\ldots, N \}$$

calculating $R(E : n)$, the relative frequency of occurrence of E out of n repetitions , by the formula

$$R(E : n, 1) = f(E : n, 1)/n .$$

Step (4): If in the set $S(E : 1)$ it is found that there exists a fixed relative frequency that corresponds to each of the $(m_1 T_1)^{th}$ trials ($m_1 = 1, 2, \ldots, N_1$) for some m_1 & T_1 where N_1 is the greatest multiple of T_1 less than or equal to N then that relative frequency is the true/exact value of P(E) .

Step (5): If no such fixed relative frequency, as mentioned in Step (4), is found then calculate $f(E : n, 2)$, the frequency of occurrence of E out of n repetitions treating the 2nd trial as the first one ignoring the 1st trial, by the formula

$$f(E : n\,,\,1) \;=\; \sum_{i=2}^{n+1} b(E : i)$$

for $n = 1, 2, \ldots\ldots, N-1$

and then construct the set

$$S(E : 2) = \{\; R(E : n, 2) : n = 1, 2, \ldots\ldots\ldots, \mathrm{N}-1 \;\}\,.$$

Step (6): If in the collection

$$\check{C}_2\,(E) = \{\; S(E : r) : r = 1, 2 \;\}$$

it is found that there exists a fixed relative frequency that corresponds to each of the ($m_2 T_2$)th trials (m_2 = 1, 2, ... , N_2) for some m_2 & T_2 where N_2 is the greatest multiple of T_2 less than or equal to N at least once in at least one set belonging to $\check{C}_2\,(E)$ then that relative frequency is the value of $P(E)$.

Step (7): If no such fixed relative frequency, as mentioned in Step (6), is found then calculate $f(E : n\,,\,3)$, the frequency of occurrence of E out of n repetitions treating the 3rd trial as the first one ignoring the earlier two trials, by the formula

$$f(E : n\,,\,1) = \sum_{i=3}^{n+2} b(E : i)$$

for $n = 1, 2, \ldots\ldots, N-2$

and then construct the set

$$S(E : 3) = \{\; R(E : n, 3) : n = 1, 2, \ldots\ldots\ldots, N-2 \;\}\,.$$

Step (8): If in the collection

$$\check{C}_3\,(E) = \{\; S(E : r) : r = 1, 2, 3 \;\}$$

it is found that there exists a fixed relative frequency that corresponds to each of the ($m_3 T_3$)th trials (m_3 = 1, 2, ... , N_3) for some m_3 & T_3 where N_3 is the greatest multiple of T_3 less than or equal to N at least once in at least one set belonging to $\check{C}_3$ (E) then that relative frequency is the value of P(E).

Step (9): *If no such relative frequency, as mentioned in* **Step (8), exists then repeat the process and stop at the sth step if at the sth step it is found that there exists a fixed relative** frequency that corresponds to each of the ($m_s T_s$)th trials (m_s = 1, 2, ... , N_s) for some m_s & T_s where N_s is the greatest multiple of T_s less than or equal to N at least once in at least one set belonging to

$$\check{C}_s\,(E) = \{\; S(E : r) : r = 1, 2, 3, \ldots\ldots\ldots, s \;\}\,.$$

This fixed relative frequency is the value of P(E).

Step (10): If by the above steps the value of P(E) can not be determined, increase the number of repetitions and repeat the steps

Note: If by the above steps the value of P(E) can not be determined, it is to be understood that the number of repetitions of the trial is not sufficient. Therefore, in that situation one needs to increase the number of repetitions and repeat the steps. The value of P(E) can be determined by this method if the trial is repeated under identical condition.

§ 2.7 Stability Property of Relative Frequency ---- Point Projection on the Total Population of Sub Regions of a Region:

Let a region R be consist of k sub regions viz.

$$R_1, R_2, \ldots\ldots\ldots, R_k .$$

Suppose,

$N(R_i : t)$ = Number of persons in the subregion R_i at time t.

This notation automatically implies that $N(R, t)$ denotes the number of persons in the region R at time t.

Consider a trial of selecting a person in the region R at random. If a person from among the persons in R is selected at random, the selected person may belong to any one of the k subregions $R_1, R_2, \ldots\ldots\ldots, R_k$ and hence each of the subregions can be associated to a probability that the selected person belong to it.. Let E_i denote the event that the selected person belongs to R_i and P(E_i) the probability of occurrence of E_i (i.e. P(E_i) denotes the probability that the selected person belongs to R_i).

The empirical definition of probability due to *Fisher* (*Ref.* 46 & 48) that is based on the idea of statistical regularity says that if the trial is repeated N times under identical condition and if out of these repetitions the event E_i occurs $N(E_i)$ times then

$$P(E_i) = \underset{N \to \infty}{Lt} \frac{N(E_i)}{N} \qquad (2.7.1)$$

The empirical definition of probability due to *von Mises* (*Ref.* 51, 65, 123, 124, 125 & 126) says that if the trial is repeated N times under identical condition and if the event E_i occurs $N(E_i)$ times out of the N repetitions then

$$N(E_i) \to N.\ \mathrm{P}(E_i) \text{ as } N \to \infty \qquad (2.7.2)$$

This implies that

$$N(\mathrm{R}_i : t) \to N(\mathrm{R} : t).\mathrm{P}(E_i) \text{ as } N(\mathrm{R} : t) \to \infty \qquad (2.7.3)$$

which implies that

$$N(\mathrm{R}_i : t) \approx N(\mathrm{R} : t).\mathrm{P}(E_i) \text{ for large } N(\mathrm{R} : t) \qquad (2.7.4)$$

Thus the estimate/projection on $N(\mathrm{R}_i : t)$ can be determined from the estimate/projection of $N(\mathrm{R} : t)$ if the value of $\mathrm{P}(E_i)$ is known. To obtain the value of $\mathrm{P}(E_i)$, equation (2.6.1) can be applied. Equation (2.7.1) implies that

$$\mathrm{P}(E_i) = \underset{N(\mathrm{R} : t) \to \infty}{Lt} \frac{N(\mathrm{R}_i : t)}{N(\mathrm{R} : t)} \qquad (2.7.5)$$

which implies that

$$\mathrm{P}(E_i) \approx \frac{N(\mathrm{R}_i : t)}{N(\mathrm{R}: t)} \text{ for large } N(\mathrm{R} : t) \qquad (2.7.6)$$

Thus, equation (2.7.6) can be used to obtain an approximate value of $\mathrm{P}(E_i)$ with the observed values of $N(\mathrm{R}_i : t)$ and $N(\mathrm{R}_i : t)$ provided the observed values of $N(\mathrm{R} : t)$ are sufficiently large.

Note: Similar theory can be developed for number of persons of specified sex and specified age in a subregion. If

$N(\mathrm{R}_i, \mathrm{S}, \mathrm{A}: t)$ number of persons of the sex 'S' & of the age 'A' in the subregion R_i

at time t,

$E(\mathrm{R}_i, \mathrm{S}, \mathrm{A})$ = the event that the selected person belongs to all of R_i, S & A

and $\mathrm{P}\{E(\mathrm{R}_i, \mathrm{S}, \mathrm{A})\}$ = the probability that the event $E(\mathrm{R}_i, \mathrm{S}, \mathrm{A})$ occurs

then the formulae for $N(\mathrm{R}_i, \mathrm{S}, \mathrm{A}: t)$ and $\mathrm{P}\{E(\mathrm{R}_i, \mathrm{S}, \mathrm{A})\}$ will respectively be

$$N(\mathrm{R}_i, \mathrm{S}, \mathrm{A}: t) \approx N(\mathrm{R} : t).\ \mathrm{P}\{E(\mathrm{R}_i, \mathrm{S}, \mathrm{A})\} \text{ for large } N(\mathrm{R} : t) \qquad (2.7.7)$$

and $$\mathrm{P}\{E(\mathrm{R}_i, \mathrm{S}, \mathrm{A})\} = \underset{N(\mathrm{R} : t) \to \infty}{Lt} \frac{N(\mathrm{R}_i, \mathrm{S}, \mathrm{A}: t)}{N(\mathrm{R} : t)} \qquad (2.7.8)$$

$$\text{i.e} \quad P\{E(R_t, S, A)\} \approx \frac{N(R_t, S, A : t)}{N(R : t)} \quad \text{for large } N(R : t) \qquad (2.7.9)$$

§ 2.8 Arithmetic Progression and Point Projection on Total Population:

The theory of population due to *Malthus* (*Ref.* 85) states that population increases in geometric progression while subsistence increases in arithmetic progression. Though this theory suffers from some shortcomings, one should consider it to have significant contribution on the researchers in the respective field since it is the root source of further thinking for the researchers. In the current study, it has been thought of that the changes in total population over intervals (of some length) of time can be represented by an arithmetic progression.

Let

$N(t)$ Size of the population under study at time t.

Let us divide the interval $(t, t+h)$ into n subintervals of equal length δh viz.

$(t, t + \delta h)$, $(t + \delta h, t + 2\delta h)$, $(t + 2\delta h, t + 3\delta h)$, ..., $(t + (n-1)\delta h, t + n\delta h)$
(2.8.1)

Suppose that c_i is the amount of change in the number of persons in the i^{th} subinterval i.e.

$$c_i = N(t + i\delta h) - N(t + (i-1)\delta h) \qquad (2.8.2)$$

If it is assumed that the total population changes in an arithmetic progression then there exists some positive integer n such that

$$c_1, c_2, c_3, \ldots\ldots\ldots\ldots, c_n \qquad (2.8.3)$$

forms an arithmetic progression. Let the first term and the common difference of the arithmetic

progression be 'a' and 'd' respectively. Then

$$c_1 = a$$
$$\text{and } c_i - c_{i-1} = d \text{ for all } i \qquad (2.8.4)$$

There are three parameters of this representation of the change in total population. They are 'a', 'd' and 'n'. The problem here is to estimate the parameters on the basis of the observed data.

Let

$$N(t_0),\ N(t_1),\ N(t_2),\ \ldots\ldots\ldots\ldots,\ N(t_N) \qquad (2.8.5)$$

be the observed data on $N(t)$ at times

$$t_0,\ t_1,\ t_2,\ \ldots\ldots\ldots\ldots,\ t_N \qquad (2.8.6)$$

respectively where $t_0,\ t_1,\ t_2,\ \ldots\ldots\ldots\ldots,\ t_N$ are equally spaced i.e.

$$t_i - t_{i-1} = h\ ,\quad \text{(say) for all } i. \qquad (2.8.7)$$

Let us select three points

$$N(t_{m-2}),\ N(t_{m-1}),\ N(t_m) \qquad (2.8.8)$$

in such a way that

(i) $t_m - t_{m-1} = t_{m-1} - t_{m-2}$

and (ii) these three points can cover the given data.

Let us divide the interval and $(t_{m-2},\ t_{m-1})$ into n subintervals viz.

$$(t_{m-2},\ t_{m-2}+\delta h),\ (t_{m-2}+\delta h,\ t_{m-2}+2\delta h),\ \ldots\ldots\ldots,\ [t_{m-2}+(n-2)\delta h,\ t_{m-2}+(n-1)\delta h = t_{m-1}] \qquad (2.8.9)$$

each of length δh.

Similarly, let us divide the interval $(t_{m-1},\ t_m)$ into n subintervals viz.

$$(t_{m-1},\ t_{m-1}+\delta h),\ (t_{m-1}+\delta h,\ t_{m-1}+2\delta h),\ \ldots\ldots\ldots,\ [t_{m-1}+(n-2)\delta h,\ t_{m-1}+(n-1)\delta h = t_m] \qquad (2.8.10)$$

each of length δh.

Thus, there are $2n$ subintervals viz.

$$(t_{m-2},\ t_{m-2}+\delta h),\ (t_{m-2}+\delta h,\ t_{m-2}+2\delta h),\ \ldots\ldots\ldots,\ [t_{m-2}+(n-2)\delta h,\ t_{m-2}+(n-1)\delta h = t_{m-1}],\ (t_{m-1},\ t_{m-1}+\delta h),\ (t_{m-1}+\delta h,\ t_{m-1}+2\delta h),\ \ldots\ldots\ldots,\ [t_{m-1}+(n-2)\delta h,\ t_{m-1}+(n-1)\delta h = t_m] \qquad (2.8.11)$$

(arranged in ascending order)

each of length δh in the interval $(t_{m-2},\ t_m)$.

If the change in $N(t)$ in the interval $(t_{m-2},\ t_m)$ occurs in arithmetic progression then the amount of changes in these subintervals will be

$$a,\ a+d,\ a+2d,\ \ldots\ldots,\ a+(n-1)d,\ a+nd,\ a+(n+1)d,\ a+(n+2)d,\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ a+(2n-1)d \qquad (2.8.12)$$

respectively.

Automatically,

$$a + (a+d) + (a+2d) + \ldots\ldots\ldots\ldots + \{a+(n-1)d\} = N(t_m) - N(t_{m-1}) \qquad (2.8.13)$$

$$\&\ \{a+(n+1)d\} + \{a+(n+2)d\} + \{a+(2n-1)d\} = N(t_{m-1}) - N(t_{m-2}) \qquad (2.8.14)$$

Solving these one can obtain that

$$d = \Delta^2 N(t_{m-2})/n^2 \qquad (2.8.15)$$

and
$$c = \frac{2.n.\Delta\, N(t_{m-1}) - (n-1)\, \Delta^2 N(t_{m-2})}{2.n^2} \qquad (2.8.16)$$

Thus
$$N(t_{m+k}) = N(t_m) + [\, 2(c + 2nd) + (n - 1)d].(n/2) \qquad (2.8.17)$$

Estimates of 'd ', 'c ' and 'n' can be obtained from these equations (2.8.1) and (2.8.2). With the help of the estimates of 'd ', 'c ' and 'n' obtained, the projected value of $N(t_{m+k})$ can be obtained by the equation (2.8.3).

The basic problem here is to determine three parameters 'd ', 'c ' and 'n' from two equations given by (2.8.15) and (2.8.16). The method of determining these parameters has been discussed below.

§ 2.8.1 Method of Determining 'd ', 'c ' and 'n':

Two equations are available to determine three parameters 'd ', 'c ' and 'n'. Also we have information that each of these three parameters is a positive integer. Thus to determine these three parameters we are to proceed with the following steps:

Step (1): Start with taking a small value of n, say n_0.

Step (2): Compute the values of d and c by the formula (2.8.15) and (2.8.16) respectively.

Step (3): Observe whether the values of d and c obtained in Step (2) are positive integers.

Step (4): If both of the values of d and c obtained in Step (2) are found to be positive integers then these values of d and c are their required values and the corresponding value n_0 of n is its required value.

Step (5): If both of the values of d and c obtained in Step (2) are not found to be positive integers then take $(n_0 + 1)$ as the value of n and repeat Step (2) and Step (3).

Step (6): If both of the values of d and c obtained in Step (5) are found to be positive integers then these values of d and c are their required values and the corresponding value $(n_0 + 1)$ of n is its required value.

Step (7): If both of the values of d and c obtained in Step (5) are not found to be positive integers then take $(n_0 + 1 + 1)$ as the value of n and repeat Step (2) and Step (3).

Step (8): Continue the process taking an increased value of n (increased by 1 at each step) until both of the values of d and c are not found to be positive integers. Stop at the step where both of the values of d and c are found to be positive integers. These integral values of d and c are their required values and the corresponding value of n is its required value.

§ 2.9 Interval Projection on The Total Population of a Region:

A method of determining interval projection on the total population of a region has been innovated by *Chakrabarty* and *Baruah* in 1993 (*Ref.* 23). It has been found that the method of interval population projection. innovated by *Chakrabarty* and *Baruah*, has yielded acceptable results in the context of total population of India (*Ref.* 23). The method is outlined below.

The principle behind the method of interval population projection due to *Chakrabarty* and *Baruah* is to determine an interval of length as small as possible within which the actual total population lies. This can be achieved if one can find out two sets of estimates: one of underestimates and the other of overestimates, the two sets being as large as possible. The interval with the maximum of the underestimates as the lower bound and the minimum of the overestimates as the upper bound is an interval that satisfies this principle. The method has been discussed below:

Let $N(t_0)$, $N(t_1)$, $N(t_2)$, ... $P(t_n)$ be the observed sizes of the population of a region at times t_0, t_1, t_2... t_n respectively where t_0, t_1, t_2... t_n are equally spaced with

$$t_0 - t_1 = t_2 - t_1 = t_3 - t_2 = \cdots\cdots = t_n - t_{n-1} = h, \text{ say.}$$

The aim is to project on $N(t_{n+k})$ on the basis of these observed data. Let $\{u_0, u_1, u_2, \ldots\ldots\ldots, u_p\}$ be a set of underestimate and $\{U_0, U_1, U_2, \ldots\ldots\ldots, U_q\}$ a set of overestimates of $N(t_{n+k})$ obtained by analyzing the trends of various measures of population growth. Then each interval (u_i, U_j) will contain $N(t_{n+k})$. (u_i, U_j) is the interval containing $N(t_{n+k})$ with the smallest length among all the pq intervals. Therefore, by the above principle behind the method of interval projection, one can treat the interval $(u_i$, $U_j)$ as a projected interval value of $N(t_{n+k})$.

The basic problem in this method is to determine underestimates and overestimates of $N(t_{n+k})$. *Chakrabarty* and *Baruah*, after searching, have found out the following techniques of determining underestimates and/or overestimates of $N(t_{n+k})$.

Technique A

$N(t_i)$ satisfies the relationship

$$\Delta\{N(t_i) = N(t_{i+1}) - N(t_i) \qquad (2.9.1)$$

where Δ is the forward difference operator.

Here $\Delta\{N(t_i)$ is the amount of growth during the interval (t_i , t_{i+1}) . If the trend of $\Delta\{ N(t_i)$ is constant then

$$\hat{N}(t_{n+k}) = N(t_n) + \frac{k}{h} N(t_{n-1}) \qquad (2.9.2)$$

would be an exact estimate of $N(t_{n+k})$. On the other hand, $\hat{N}(t_{n+k})$ given by the above expression would be an underestimate of $N(t_{n+k})$ if the trend of $\Delta\{$ $N(t_i)$ is increasing. Similarly, it would be an overestimate if the trend is decreasing.

Technique B

$N(t)$ satisfies the relationship

$$N(t_i) = N(t_{i-1}) + \Delta\{N(t_{i-2})\} + \Delta^2\{N(t_{i-2})\} \qquad (2.9.3)$$

Here $\Delta^2\{N(t_{i-2})\}$is the difference of the growths occurred during the periods (t_i , t_{i+1}) and (t_{i+1} , t_{i+2}) respectively. An exact estimate of underestimate or overestimate of $N(t_{n+k})$ can be obtained from this relationship observing the trend of $\Delta^2\{N(t_{i-2})\}$. The formula for $\hat{N}(t_{n+k})$ is given by

$$\hat{N}(t_{n+k}) = N(t_n) + \frac{k}{h} \Delta N(t_{n-1}) + \Delta^2 N(t_{n-2}) , \; k \leq h \qquad (2.9.4)$$

Technique C

$\Delta\{N(t_0)/N(t_0)\}$, $\Delta\{N(t_1)/N(t_1)\}$, $\Delta\{N(t_2)/N(t_2)\}$, ..., $\Delta\{N(t_{n-1})/N(t_{n-1})\}$are the relative growths in the intervals (t_0 , t_1), (t_1 , t_2), ..., (t_{n-1}, t_n) respectively. These are known. In this case,

$$N(t_{n+k}) = \frac{k}{h}\,\frac{\Delta N(t_{n-1})}{N(t_{n-1})}\, N(t_n) + N(t_n) \qquad (2.9.5)$$

$$(k \leq h)$$

would be an exact estimate or underestimate or overestimate of $P(t_{N+k})$ depending on the trend of relative growth.

Technique D

$\Delta\{N(t_{i+1})\}/\Delta\{N(t_i)\}$ is the ratio of growth during (t_{i+1}', t_{i-2}) and (t_i , t_{i+1}). Here the estimate of $N(t_{n+k})$ is provided by

$$\underset{k<h}{\hat{N}(t_{n+k})} = \frac{k}{h}\,\frac{\{\Delta N(t_{n-1})\}^2}{\Delta N(t_{n-2})}\,.N(t_n)\,, \qquad (2.9.6)$$

Technique E

An exact estimate or underestimate or overestimate of $N(t_{n+k})$ can also be obtained from the nature of the trend of the ratio $N(t_{i+1})/N(t_i)$. Here the estimate $P(t_{N+K})$ of $N(t_{n+k})$ is given by

$$\hat{N}(t_{n+k}) = \begin{cases} \dfrac{N(t_{n-1}) + (k/h)\,\Delta\,N(t_{N-1})}{N(t_{n-1})}\,.N(tn) & \text{for } k<h \\[2ex] \dfrac{N(tn)}{N(t_{n-1})}\,.N(t_n\,) & \text{for } k=h \end{cases} \qquad (2.9.7)$$

Technique F

$\Delta N(t)$ is the change in total population during the interval $(t\,,\, t+h)$. So, $\Delta N(t)/h$ is the change in total population in unit time during the period $(t\,,\, t+h)$. Let us divide each interval $(t\,,\, t+h)$ into $(2m+1)$ equal subintervals. We assume that the amount of change $\Delta N(t)/(2m+1)$ occurs at the $(m+1)^{th}$ subinterval of each interval. To find out the changes in the first and the last subintervals of an interval one can use the average change in the interval i.e. the change in the $(m+1)^{th}$ subinterval. One can see that

$$\frac{m+1}{(2m+1)^2}\,\Delta N(t_i) + \frac{m}{(2m+1)^2}\,\Delta N(t_{i-1}) = c\left(t_i, t_i + \frac{h}{2m+1}\right) \qquad (2.9.8)$$

is the change in the subinterval (t_i, $t_i + \frac{h}{2m+1}$) and that

$$c\left(t_i + \frac{2mh}{2m+1}, t_i + h\right) = \frac{(m+1)}{(2m+1)^2}\,\Delta N(t) - \frac{(m+1)}{(2m+1)^2}\,\Delta N(t-h) \qquad (2.9.9)$$

is the change in the subinterval ($t_i + \frac{2mh}{(2m+1)h}$, t_i+h) .

Therefore ,

$$\hat{N}(t_{n+k}) = N(t_n) + \frac{h}{k}\left\{ \frac{3m+1}{2m+1} \Delta N(t-h) - \frac{m}{2m+1} \Delta N(t-2h) \right\} \qquad (2.9.10)$$

would be an exact estimate or underestimate or overestimate of $N(t_{n+k})$ depending on the trend of the changes.

§ 2.10 Population -- Point Projection From Interval Projection:

In the method of determining projected interval on the total population in a region innovated by *Chakrabarty* and *Baruah* (*Ref.* 23), projected interval is determined from the underestimates and overestimates of the same. The underestimates and overestimates, used there, suffer from errors. It has been thought of that by eliminating the errors involved there, it might be possible to obtain projected point value on the same. Hence, due to the necessity of some method of projecting point value on total population of the whole region, an attempt has been made to innovate a method of projecting point value on total population in the whole region based on underestimates and overestimates of the same. The method is outlined below.

The following notations have been used here:

$N(t)$ = Number of persons in a region at time t.

$N^T(t)$ = Number of persons in a region at time t.

$\hat{N}(t)$ = Point estimate of $N(t)$.

$\hat{N}_i^U(t)$ = Under estimate of $N(t)$ based on the method i (i = 1, 2, ... , m).

$\hat{N}_j^O(t)$ = Over estimate of $N(t)$ based on the method j (j = 1, 2, ..., n).

Automatically

$$N_i^U(t) = N^T(t) + \epsilon_i(t) \qquad (2.10.1)$$

for some $\epsilon_i(t)$

where $\epsilon_i(t) \geq 0$.

Similarly, $$\hat{N}_j^O(t) = N^T(t) - \epsilon_j'(t) \qquad (2.10.2)$$

for some $\epsilon_j'(t)$

where $\epsilon_j'(t) \geq 0$.

Now consider the sum

$$S = \sum_{i=1}^{m+n} \hat{N}_i^{a}(t) \qquad (2.10.3)$$

$$\text{where } \hat{N}_i^{a}(t) = \begin{cases} \hat{N}^{U}(t) & \text{for } i = 1, 2, 3, \ldots\ldots\ldots, m \\ \hat{N}^{O}(t) & \text{for } i = m+1, m+1, \ldots\ldots\ldots, m+n \end{cases}$$

Think of the averages

$$A_i = \frac{1}{m+n-1} \sum_{\substack{j=1 \\ j \neq i}}^{m+n} \hat{N}_j^{a}(t) \qquad (2.10.4)$$

for $i = 1, 2, \ldots\ldots\ldots, m+n$

Here, A_i is the average of ali estimates $N^a_j(t)$ except the estimate $N^a_i(t)$.

The sum S can be expressed as

$$S = (m+n)\, N^T(t) + \left(\sum_{i=1}^{m} \epsilon_i - \sum_{i=m+1}^{m+n} \epsilon_i \right) \qquad (2.10.5)$$

$$\text{which} \Rightarrow A = N^T(t) + \left(\sum_{i=1}^{m} \epsilon_i - \sum_{i=m+1}^{n} \epsilon_i \right) / (m+n) \qquad (2.10.6)$$

where $A = \dfrac{S}{m+n}$, the average of all the estimates (under and over). Since each of the errors $\epsilon_1, \epsilon_2, \epsilon_3, \ldots\ldots \epsilon_{m+n}$ is positive, the 2nd expression in the right hand side of equation (2.10.6) may be very near to zero. Thus

$$N^{t}(t) \cong A . \qquad (2.10.7)$$

Though A is approximately equal to $N^T(t)$, it is to be noted that it may not be exactly equal to $N^T(t)$.

Thus the problem is to determine the true value of unknown $N^T(t)$. To do this let us deal with the averages

$$A_1, A_2, \ldots\ldots A_m, \ldots\ldots A_{m+n}$$

By the same logic as in the case of A,

$$A_i \simeq N^T(t)\ ,\ i=1,2,\ldots\ldots,m+n \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.10.8)$$

If $\qquad A_i = N^T(t) \quad , \ \forall\ i$

then $A_1, A_2, \ldots\ldots, A_m, \ldots\ldots A_{m+n}$ must be identical i.e

$$A_1 = A_2 = \ldots\ldots = A_m = A_{m+1} = \ldots\ \ldots\ldots = A_{m+n}$$

and vice versa.

Thus if it is found that all A_i s are equal then that common value of A_i will be the value of $N^T(t)$. However if not, the same process, may be repeated upon

$$A_1, A_2, \ldots\ldots, A_m, A_{m+1}, \ldots\ldots\ldots, A_{m+n}$$

treating them as estimates (some of which are underestimates and the others are overestimates) of $N^T(t)$ until stabilized value of the averages are obtained. The stabilized value will be the required value of $N^T(t)$.

Possible situations of stabilized value of average

Let

L = the minimum of over estimate $\hat{N}^T(t)$

and M = the maximum of underestimate $\hat{N}^T(t)$

Then the interval

$$(L, M)$$

is the projected interval for $N^T(t)$ (*Ref.* 23).

Now, there are three possible situations regarding the position of the stabilized value.

(1) Stabilized value falls below the lower limit of the projected interval.

(2) Stabilized value falls above the upper limit of the projected interval

(3) Stabilized value falls within the projected interval

It is obvious that the actual value (true point value) must fall within the corresponding projected interval and hence the corresponding projected value must fall within the same. Thus, one computed point value can be treated as an acceptable projection if it falls within the same. Thus in the third situation, the stabilized value can be treated as the corresponding projected value though may not be exactly equal to the corresponding actual value. The first possibility arises due to the effect of extreme value /values of the under estimates used. Similarly the second possibility arises due the effect of extreme value/values of the overestimates used. Thus in any of these two situations, the process may be repeated using the estimates excluding the associated extreme value till the third situation is obtained in which situation the stabilized value can be treated as the projected point value.

§ 2.11 Some Existing & Commonly Used Laws of Population Growth:

Here, two commonly used formulae for estimating / projecting the total population of a region have been mentioned with the respective methods of fitting of them to observed data.

A very satisfactory formula for estimating / projecting total population of a region is represented by the **logistic curve** (*Ref.* 63, 64, 99, 100, 106 & 110). The curve is of the form

$$N(\mathrm{R} : t) = \frac{L}{1 + \exp\{r.(\beta - t)\}} \qquad (2.11.1)$$

where

(i) $N(\mathrm{R} : t)$ is the total population of the region R under study at time t,

(ii) L is the upper limit of $N(\mathrm{R} : t)$,

(iii) β is the value of t for which $N(\mathrm{R} : t)$ is $L/2$

and (iv) r is the value of

$$\frac{1}{N(\mathrm{R} : t)} \; \frac{d}{dt} N(\mathrm{R} : t)$$

when $N(\mathrm{R} : t) = L$.

Here L, r and β are the parameters of the curve which are to be determined on the basis of the observed data. Putting

$$Y_t = 1 / N(\mathrm{R} : t) \; ,$$

$$X_t = 1 / N(\mathrm{R} : t\text{-}1) \; ,$$

$$A = \{1 - \exp(r)\}$$

$$\& \; B = \exp(r)$$

the logistic curve can be written in the form

$$Y_t = A + B\,X_t \qquad (2.11.2)$$

Thus if (t, N_t), $(t = 1, 2, \ldots\ldots, n)$ are the observed data on $\{t, N(\mathrm{R} : t)\}$ then the two constants A and B can be estimated by the equations

$$B = \left[\left\{ \sum_{t=1}^{n-1} (Y_t - Y)^2 \right\} / \sum_{t=1}^{n-1} (X_t - X)^2 \right\}\right]^{1/2} \tag{2.11.3}$$

$$\& \; A = Y - B.X \tag{2.11.4}$$

From the estimates of the constants A and B, the estimates of the parameters L, r can be obtained. Finally, the parameter β can be estimated from the equation

$$\beta = (1/nr). \sum_{t=1}^{n-1} \log(z_t - 1) + (\mathrm{n}-1)/2 \tag{2.11.5}$$

where $\quad z_t = L / N_t$.

This method of estimating the parameters of the logistic curve is due to *Rhodes* (**Ref.** 110).

Another formula for estimating / projecting total population of a region is represented by the **exponential curve** (*Ref.* 63, 64 & 106) which is of the form

$$N(\mathrm{R} : t) = \mu.\exp(-\lambda.t), \tag{2.11.6}$$
$$\mu > 0, \; \lambda > 0$$

where μ and λ are the parameters which are to be determined on the basis of the observed data. Putting

$$y_t = \log N(\mathrm{R} : t)$$
$$\& \quad \mathrm{v} = \log \mu$$

the exponential curve can be written in the form

$$y_t = \mathrm{v} - \lambda.t \tag{2.11.7}$$

Thus if (t, N_t), $(t = 1, 2, \ldots\ldots, n)$ are the observed data on $\{t, N(\mathrm{R} : t)\}$ then the two constants v and λ can be estimated by the equations

$$\sum_{t=1}^{n} y_t = n\mathrm{v} - \lambda \sum_{t=1}^{n} t \tag{2.11.8}$$

$$\& \sum_{t=1}^{n} t\,y_t = \mathrm{v}\sum_{t=1}^{n} t - \lambda \sum_{t=1}^{n} t^2 \tag{2.11.9}$$

where $\quad y_t = \log N_t$

From the estimates of the constants v an estimate of the parameters μ can be obtained from the equation

$$v = \log \mu \qquad (2.11.10)$$

§ 2.12 Testing of Goodness of Fit:

A very powerful test for testing the significance of the discrepancy between theory and experiment, popularly known as **Chi-square Test of Goodness of Fit**, was given by *Professor Karl Pearson* in 1900 (*Ref.* 12, 91, 92, 127 & 129). It enables one to find if the deviation of the experiment from the theory is just by chance or is it really due to the inadequacy of the theory to fit the observed data.

If O_i, (i = 1, 2, 3, ..., n) is a set of observed (experimental) frequencies and E_i, (i = 1, 2, 3, ..., n) is the set of the corresponding expected (theoretical or hypothetical) frequencies then *Karl Pearson*'s chi-square (abbreviated as χ^2) given by

$$\chi^2 = [\sum_{i=1}^{n} \{(O_i - E_i)^2\}/E_i], \qquad (2.12.1)$$

$$(\sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i)$$

follows chi-square (χ^2) distribution with (n-1) *d.f.* (*Ref.* 91, 92 & 127).

Thus, the null hypothesis

" H_0 : The discrepancy between the observed frequency and the theoretical frequency

is not significant "

against the alternative hypothesis

" H_1 : The discrepancy between the observed frequency and the theoretical frequency

is significant "

is retained or rejected at the significance level α according to the calculated value of χ^2 for (n – 1) *d.f.* at the significance level α is less than or greater than the corresponding theoretical value of χ^2 for the same *d.f.* and at the same significance level.

§ 2.13 Probabilistic Forecasting of Temperature and Rainfall:

Temperature and rainfall are two of the major factors that determine the weather and the climate of a location. Temperature of a location on the earth surface is a variable, which changes over time. It forms a time series where the period of the smallest periodic component is a day and the period of the highest periodic component in normal situation is a year. It may or may not have cyclical component. Of course, random factor always effects upon it and hence it contains the random component. Similar is the case of rainfall also.

The following characteristics together can give a picture of temperature at a location in a month:

(i) Mean Maximum Temperature (monthly).

(ii) Highest Maximum Temperature (monthly).

(iii) Mean Minimum Temperature (monthly).

and (iv) Lowest Minimum Temperature (monthly).

Similarly, the following characteristics together can give a picture of rainfall at a location in a month:

(i) Total Rainfall (monthly).

(ii) Heaviest 24 Hours Rainfall (monthly).

and (iii) Number of Rainy Days (Monthly)

The discovery of normal probability distribution, discovered by *Gauss* (*Ref.* 50, 60, 115 & 117), is the most significant discovery in the theory of statistics. It has been thought of that it may be possible to apply the area property of normal distribution in developing the literature on how to know whether there exists any significance assignable cause in a region which forces the temperature of the region to be changed as well as on how to determine forecasted interval value with desired probability (i.e. with desired confidence). Similar is the case for rainfall also. This literature, developed, has been thrown below. The literature, thrown below, is for the mean maximum temperature only. The literatures for the other characteristics are similar with this.

§ 2.13.1 Method of Obtaining Forecasted Interval:

Let Y be the random variable that represents the mean maximum temperature at a location. The maximum temperature at a location in a particular day in a year should remain the same in every successive year provided there is no any assignable cause of variation. But, assignable cause/causes of variation like rainfall, winds, clouds etc. may appear in the same day of the successive years that influence upon the temperature. So, analysis of the daily data on temperature cannot yield valid results.

However, monthly data on such variable can estimate this type of causes of variation. Hence, it would be reasonable to analyses monthly data instead of daily data.

Let

Y_{ij} = the mean maximum temperature observed at a location in the month 'i' of the

year 'j' (i = 1, 2, ………,12 & j = 1, 2, …………., n).

For fixed i, the values of Y_{ij} (j = 1, 2, …………., n) should be constant if there exists no cause of variation in Y_{ij} over year. However, random cause of variation always exists. Thus if no assignable cause of variation exists in Y_{ij} over year then for fixed i we have

$$Y_{ij} = \mu + \varepsilon_{ij}$$

where μ = the true value of the mean maximum temperature

& ε_{ij} = the error associated to Y_{ij} due to random cause (i.e. chance cause) of variation.

The common assumption of ε_{ij} is that ε_{ij} s are independently and identically distributed $N\left(0,\sigma_\varepsilon^2\right)$ variates where the notation $N\left(0,\sigma_\varepsilon^2\right)$ represents normal variate with mean 0 and variance σ_ε^2.

Now,

since ε_{ij} is a $N\left(0,\sigma_\varepsilon^2\right)$ variate

therefore, Y_{ij} - μ is a $N\left(0,\sigma_\varepsilon^2\right)$ variate

which implies

(i) $P\{(Y_{ij} - \mu)/\sigma_\varepsilon \leq 1.96\} = 0.95$,

(ii) $P\{(Y_{ij} - \mu)/\sigma_\varepsilon \leq 2.58\} = 0.99$

& (iii) $P\{(Y_{ij} - \mu)/\sigma_\varepsilon \leq 3\} = 0.9973$

Therefore, the intervals

(i) $\mu - 1.96\sigma_\varepsilon \leq Y_{ij} \leq \mu + 1.96\sigma_\varepsilon$,

(ii) $\mu - 2.58\sigma_\varepsilon \leq Y_{ij} \leq \mu + 2.58\sigma_\varepsilon$

& (iii) $\mu - 3\sigma_\varepsilon \leq Y_{ij} \leq \mu + 3\sigma_\varepsilon$

are respectively the 95%, 99% & 99.73% confidence intervals of Y_{ij} , the mean maximum temperature at the location considered for the month 'i'. These mean that (i)

95% or more, (ii) 99% or more & (iii) 99.73% or more of the observations Y_{ij} (j = 1, 2, ..., n) will fall within (and consequently (i) 5% or less, (ii) 1% or less & (iii) 0.27% or less of the observations Y_{ij} (j = 1, 2, ..., n) will fall outside) the intervals

(i) $(\mu - 1.96\sigma_\varepsilon \;,\; \mu + 1.96\sigma_\varepsilon)$,

(ii) $(\mu - 2.58\sigma_\varepsilon \;,\; \mu + 2.58\sigma_\varepsilon)$

& (iii) $(\mu - 3\sigma_\varepsilon \;,\; \mu + 3\sigma_\varepsilon)$

respectively.

Conversely, if it is found that (i) 5% or less, (ii) 1% or less & (iii) 0.27% or less of the observations Y_{ij} (j = 1, 2, ..., n) fall outside (or equivalently (i) 95% or more, (ii) 99% or more & (iii) 99.73% or more of the observations Y_{ij} (j = 1, 2, ..., n) fall within) the intervals

(i) $(\mu - 1.96\sigma_\varepsilon \;,\; \mu + 1.96\sigma_\varepsilon)$,

(ii) $(\mu - 2.58\sigma_\varepsilon \;,\; \mu + 2.58\sigma_\varepsilon)$

& (iii) $(\mu - 3\sigma_\varepsilon \;,\; \mu + 3\sigma_\varepsilon)$

respectively where

μ = mean of Y_{ij} (j = 1, 2, ..., n)

& σ_ε^2 = variance of Y_{ij} (j = 1, 2, ..., n)

then

(i) $P\{(Y_{ij} - \mu)/\sigma_\varepsilon \leq 1.96\} = 0.95$,

(ii) $P\{(Y_{ij} - \mu)/\sigma_\varepsilon \leq 2.58\} = 0.99$

& (iii) $P\{(Y_{ij} - \mu)/\sigma_\varepsilon \leq 3\} = 0.9973$

These mean that the variable $(Y_{ij} - \mu)$ is a $N\left(0,\sigma_\varepsilon^2\right)$ variate which further means that the variable ε_{ij} where

$$\varepsilon_{ij} = Y_{ij} - \mu$$

is $N\left(0,\sigma_\varepsilon^2\right)$ variate. Hence Y_{ij} can be represented by

$$Y_{ij} = \mu + \varepsilon_{ij}$$

where μ = the mean of Y_{ij} (j = 1, 2, ..., n)

& ε_{ij} = the error associated to Y_{ij}

with the assumption that ε_{ij} s are independently and identically distributed

$N\left(0, \sigma_{\varepsilon}^{2}\right)$ variates.

This implies that there exists no assignable cause of variation in Y_{ij} (j = 1, 2, ..., n) over the years j (j = 1, 2, ..., n). Consequently, one can conclude that if the current condition prevails in future then these 95%, 99% & 99.73% confidence intervals of Y_{ij} will be the corresponding projected confidence intervals of the mean maximum temperature at the location for the month 'i'.

Some Notes

(i) Here,

$$\sigma_{\varepsilon}^{2} = \text{Variance of } \varepsilon_{ij}$$

$$= \text{Variance of } (Y_{ij} - \mu), \qquad \text{since } Y_{ij} - \mu = \varepsilon_{ij}$$

$$= \text{Variance } Y_{ij}$$

Thus σ_{ε}^{2} is the variance of Y_{ij} also.

(ii) Since for fixed i there is no more value of Y_{ij} besides the values of Y_{ij} (j = 1, 2, ..., n) if the period from the year '1' to the year 'n' is considered, the values of μ and σ_{ε}^{2} will be

$$\mu = \frac{1}{n}\sum_{j} Y_{ij}$$

$$\text{and } \sigma_{\varepsilon}^{2} = \frac{1}{n}\sum_{j} (Y_{ij} - \mu)^{2}$$

respectively.

§ 2.14 Analysis of Variance Technique and Analysis of Data on Temperature & Rainfall:

Here also the same characteristics of temperature and of rainfall mentioned above have been considered.

The analysis of variance, discovered by *Fisher* (*Ref.* 31, 49, 102, 103 & 113), is a statistical tool that can be used to know whether there exists any significance assignable cause in a region which forces the temperature of the region to be changed. It may be possible to apply this tool in determining forecasted interval value with desired probability. Similar is the case for rainfall also. An outline of the literature of this tool has been thrown below. The literature, thrown below, is for the mean maximum temperature only. The literatures for the other characteristics are similar with this

Let

Y_{ij} = the mean maximum temperature observed at a location in the month 'i' of the year 'j' (i = 1, 2, ...,12 & j = 1, 2, ..., n).

There are two sources of variation occurred in the data viz. (i) Month and (ii)Year.

(It is to be noted that these observations constitute the concerned population for the period from the year '1' to the year 'n'.)

The technique of analysis of variance, discovered by *Sir Ronald A. Fisher* (1890 –1962), can be applied to these data to test the significance of differences

(i) among the effects of different months over the mean maximum temperature

(ii) among the effects of the years 1, 2, ..., n over the same.

The mathematical model is

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where μ = the constant that would have been happened had there been no any other cause of variation,

α_i = the effect of causes of variation occurred in the month 'i',

β_j = the effect of causes of variation in the year 'j'

& ε_{ij} = the random effect.

The common assumption of ε_{ij} viz. the assumption that

$$\varepsilon_{ij} \text{ obeys } N\left(0, \sigma_\varepsilon^2\right) \text{ law}$$

is retained here.

The least squares estimates $\hat{\mu}$, $\hat{\alpha}_i$ & $\hat{\beta}_j$ of μ, α_i & β_j respectively are found to be

$$\hat{\mu} = \bar{y}_{00} \quad , \hat{\alpha}_i = \bar{y}_{i0} \quad \& \quad \hat{\beta}_j = \bar{y}_{0j}$$

The format of the table for performing analysis of variance is shown in the following table (Table-2.14.1).

Table-2.14.1

(Analysis of Variance Table)

Source	Degrees of freedom	Sum of squares	Mean squares	Calculated value of F statistic	Tabulated value F statistic
Month	11	$S_m^2 =$	$s_m^2 = S_m^2/11$	s_m^2 / s_e^2 with 11 & 11(n-1) degrees of freedom	
Year	n- 1	$S_y^2 =$	$s_y^2 = S_y^2/n$	s_y^2 / s_e^2 with n & 11(n-1) degrees of freedom	-
Error	11(n-1)	$S_e^2 =$	$s_e^2 = S_e^2/11(n-1)$		
Total	12n -1	$S_t^2 =$	$s_t^2 = S_t^2/(12n-1)$		

(It has been established that the error mean square is an unbiased estimate of σ_ε^2.)

Now, "the difference among the effects of causes of variations over the years is insignificant" implies that there is no assignable cause(s) of variations (in the data) that arise due to the change of years. Thus if the difference among the effects of causes of variations over the years is found to be insignificant then one can conclude that there is no assignable cause(s) of variations (in the data) that arise due to the change of years. In that case we can conclude that the current picture of the mean maximum temperature will prevail in future.

If the said difference is found to be significant then we can conclude that there exists assignable cause(s) of variation that occurs due to change in years. However, before drawing this conclusion, we are to test the difference between the effects of the two years in each pair of the years. This is necessary due to the reason that the analysis of variance may show significance of difference among the years due to the existence of the significance of difference between one pair of years only. If this situation is found, analysis of variance is to be carried out again on the observations excluding those that correspond to this pair of years. Of course, this is not to be repeated more than once. The method of testing the significance of the difference between two means has been outlined below.

§ 2.14.1 Test of Difference of Means:

A famous statistician *W. S. Gosset*, who wrote under pseudonym (prename) of Student defined a test statistic known as '*t*' and investigated its sampling distribution, somewhat empirically, in a paper entitled " **The Probable error of the Mean** " published in 1908 (***Ref.*** 119) while another statistician *Professor R. A. Fisher* defined the same statistic '*t*' in a more general way and gave a rigorous proof for its sampling distribution in 1926 (***Ref.*** 97). This statistic has a lot of applications one of which is the "**testing of the significance of difference between two means**" that has been outlined below.

Suppose, we want to test if two independent samples

$$\{x_1, x_2, \ldots\ldots\ldots, x_m\}$$

$$\& \ \{y_1, y_2, \ldots\ldots\ldots, y_n\}$$

of sizes m and n have been drawn from two normal populations with means μ_x and μ_y respectively. Then for testing the null hypothesis

$$H_0 : \ \mu_x = \mu_y$$

the test statistic t, under the assumption that the population variances are equal, is given by

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{S\{(1/m) + (1/n)\}^{1/2}}$$

where

$$\bar{x} = (1/m)\sum_{i=1}^{m} x_i$$

$$\bar{y} = (1/n)\sum_{i=1}^{n} y_i$$

$$\text{and } S^2 = 1/(m+n-2)\left\{ \sum_{i=1}^{m} (x_i - \bar{x})^2 + \sum_{i=1}^{n} (y_i - \bar{y})^2 \right\}$$

Here, t follows "t" distribution with $(m + n - 2)$ *d.f.* (***Ref.*** 97 & 119).