On Moments of Order Statistics

CHITRA S

(16PMA001)

Thesis Submitted to

Avinashilingam Institute for Home Science and Higher Education for Women

Coimbatore-641 043

In Partial Fulfilment of the Requirements for the Degree of

Master of Science in Mathematics

April 2018

On Moments of Order Statistics

CHITRA S

(16PMA001)

Thesis Submitted to

Avinashilingam Institute for Home Science and Higher Education for Women

Coimbatore-641 043

In Partial Fulfilment of the Requirements for the Degree of

Master of Science in Mathematics

April 2018

Signature of the Head of the Department **Signature of the Supervisor**

ACKNOWLEDGMENT

First and foremost, I extremely thankful to the **LORD ALMIGHTY** for his graces and blessings showered on me.

I take immense pleasure in thanking **Dr. P. R. KRISHNAKUMAR,** Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the infrastructure for the conduct of the research study.

I would like to thank **Dr.(Tmt.) PREMAVATHY VIJAYAN**, Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for the encouragement and for providing the opportunity to develop and establish my skills.

I extend my heartfelt thanks to **Dr. (Tmt.) S. KOWSALYA,** Registrar, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for the encouragement given during the investigation.

I would like to thank **Dr. (Tmt.) G.P.JEYANTHI**, Director (Research and Consultancy), Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for all the encouragement and for providing the opportunity.

I express my heartfelt thanks to **Dr. (Tmt.) A. PARVATHI,** Professor, Dean Faculty of Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her excellent support, unflinching encouragement and guidance during the course of the investigation

I express my heartfelt thanks to **Dr. (Tmt.) K.UDAYA CHANDRIKA,** Professor and Head of the Department of Mathematics, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her support and guidance during the course of the investigation.

I deeply indebted to my thesis supervisor **Dr.(Tmt.)S.MUTHULAKSHMI,** Professor, Department of Mathematics, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her inspiring guidance, innovative ideas, critical suggestions and constant encouragement throughout this work.

I would like to express my sincere thanks to all the **STAFF MEMBERS OF THE DEPARTMENT OF MATHEMATICS** who were responsible for the good finish of this dissertation.

I owe my special thanks to my **BELOVED PARENTS, LOVING SISTERS, BROTHERS, FRIENDS and also the GRACEFUL RELATIVES,** who helped me by providing full strength, support and encouragement to complete my project successfully.

CONTENT

INTRODUCTION

For the last two decades, research in the area of order statistics has been steadily and rapidly growing. Statistical inference theory has been developed for samples from populations having normal, binomial, Poisson, multinomial and other specified forms of distribution functions depending on one or more unknown population parameters. These developments fall into two main categories:

(i) statistical estimation,

(ii) the testing of statistical hypotheses.

The theory of statistical estimation deals with the problem of estimating values of the unknown parameters of distribution functions of specified form from random samples assumed to have been drawn from such populations. The testing of statistical hypotheses deals with the problem of testing, on the basis of a random sample, whether a population parameter has a specified value, or more generally whether one or more specified functional relationships exist among two or more population parameters.

There are many problems of statistical inference in which one is unable to assume the functional form of the population distribution. Many of these problems are such that the strongest assumption which can be reasonably made is continuity of the cumulative distribution function of the population. An increasing amount of attention is being devoted to statistical tests which hold for all populations having continuous cumulative distribution functions. Problems of this type in which the distribution function is arbitrary within a broad class are referred to as non-parametric problems of statistical inference. In nonparametric problems it is being found that order statistics, that is, the ordered set of values in a random sample from least to greatest, are playing a fundamental role.

There are both theoretical and practical reasons for this increased attention to nonparametric problems and order statistics. From a theoretical point of view it is obviously desirable to develop methods of statistical inference which are valid with respect to broad classes of population distribution functions. From a practical point of view it is desirable to make the statistical procedures themselves as simple and as broadly applicable as possible. This is indeed the case with statistical inference theory based on order statistics.

Order statistics also permit very simple solutions of some of the more important parametric problems of statistical estimation and testing of hypotheses.

The aim of this dissertation is to bring distributional properties of order statistics, based on any distribution. This dissertation includes

i. derivation of the expected values, moments, sample ranges and sample median of order statistics based on the binomial distribution.

ii. derivation of the expected values, moments, sample ranges and sample median of order statistics based on the exponential distribution.

iii. derivation of the expected values, moments, sample ranges and sample median of order statistics based on the logistic distribution.

SYNOPSIS

This dissertation is concerned with the development and evaluation of moments, sample median and sample range of order statistics from binomial, exponential and logistic distributions.

Chapter I - **Basic Concepts**, presents the basic definition of terms related to order statistics and its distributions.

Chapter II- **Binomial Distribution**, presents the derivation of probability mass function of Binomial distribution and its moments. The mean and variance of Extreme Order statistics with respect to Binomial distribution are derived and the tables are prepared. A Computer code using C language is presented to help the user in deriving the mean and variance of extreme order statistics.

Chapter III- **Exponential Distribution**, provides the probability mass function of Exponential distribution and its moments. The mean and variance of Order statistics with respect to Exponential distribution are derived and the tables are prepared. A Computer code using C language is presented to help the user in deriving the mean and variance of extreme order statistics.

Chapter IV- **Logistic Distribution**, gives the probability distribution function of Logistic distribution and its moments. The mean, variance, covariance of Order statistics with respect to Logistic distribution are derived and the tables are given.

REVIEW OF LITERATURE

A remarkably large body of literature is devoted on order statistics however, only the literature relevant for this dissertation is provided in this section.

Investigations on the sampling theory of order statistics was originated from **Galton Difference Problem** studied by Karl Pearson (1902). Tippett (1925) extended the work of Pearson and found the mean value of the sample range that is, the difference between the least and the greatest order statistics in a sample and tabulated for certain sample sizes ranging from 3 to 1000, with reference to the cumulative distribution function of the largest order statistic in a sample from a standard normal population.

Fisher and Tippett (1928), derived the limiting distributions of the largest and smallest order statistics under certain regularity conditions and also obtained the related asymptotic results when the sample size increases indefinitely using a method of functional equations.

Allen (1932) derived general expressions for the exact distribution functions of the median, quartiles, and range of a sample of size n. Gumbel (1935) derived the limiting distributions to study the extreme order statistics. Mises (1936) made a precise determination of regularity conditions. Gumbel (1941) attempted Various applications like flood flows and maximum time intervals between successive emissions of gamma rays from a given source.

Research work relating to the development of order statistics and their significance for a period of almost a quarter of a century have been summarized by Wilks (1948) in a survey paper. Hoeffding (1953) studied on the distribution of the expected values of order statistics.

Apart from the basic distribution theory and limit laws, attention has also been focused by various authors on problems involving order statistics in the theory of estimation, testing of hypotheses in multiple decision and comparison procedures. Most of these results are outlined in Gumbel (1958).

Sarhan and Greenberg (1962) studied extensively the exact distributions and properties of order statistics. Developments in the field of order statistics from the early 1960s are summarized in a book by Sarhan and Greenberg (1962).

Tanis(1964) developed linear forms in the order statistics with reference to exponential distribution. Generalised extreme-value distribution in estimating the extreme percentiles were discussed by Martiz and Munro(1967).

Applications of order statistics in tests of hypotheses and estimation methods based on censored samples from lifetime distributions of interest have been widely brought forward by Harter (1969).

Galambos (1978) focused on the asymptotic theory of extreme order statistics. David (1981) gave an exciting encyclopedic representation of order statistics. Nagaraja (1986) derived structures of discrete order statistics.

Balakrishnan and Cohen (1991) provided the estimation methods relating to order statistics. Nagaraja (1992) studied order statistics of discrete distributions.

Balakrishnan and Chan (1992) studied Order statistics from extreme order distribution and computed the tables of mean, variance and covariance. The theory, methods and applications of exponential distribution was brought forward by Balakrishana(1996).

Harter and Balakrishnan (1996) prepared a handbook of tables for the use of order statistics in estimation. Harter and Balakrishnan (1997) provided tables for the use of range and studentized range in tests of hypotheses.

Jones (2004) studied on the families of distributions arising from distributions of order statistics. Ahsanullah and Hamedani (2010) studied extensively the exponential distribution.

Rupert Jr. (2012) presented simultaneous inference methods with order statistics. An introductory level of order statistics was prepared by Ahsanullah et al. (2013).

CHAPTER I

BASIC CONCEPTS

Basic concepts include the definitions of related terms of order statistics, basic distribution theory of order statistics, distribution of sample median and range of order statistics, underlying basic distributions and the role of order statistics along with notations.

ORDER STATISTICS

Let $X_1, X_2, X_3, \dots X_n \dots$ be some random variables and let $X_{1,n} \le X_{2,n} \le X_{3,n}, \dots \dots \le X_{n,n}$ denote the corresponding variational series based on random variables $X_1, X_2, X_3, \dots \dots X_n$ .Elements $X_{k,n}$, $1 \le k \le n$, are called Order Statistics. Observed values of $X_{1,n}, X_{2,n}, X_{3,n}, \dots X_{n,n}$, are denoted by $x_{1,n}, x_{2,n}, x_{3,n}, \dots x_{n,n}$ and call a realization of order statistics and

$$X_{1,n} = \min\{X_1, X_2, X_3, \dots X_n\} \text{ and } X_{n,n} = \max\{X_1, X_2, X_3, \dots X_n\}, n = 1,2, \dots ..$$

MEAN

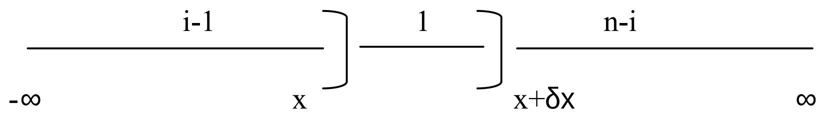
The mean or average is used to measure the central tendency of data. It is determined by adding values of all the data points in a population and then dividing the total by the number of points. The mean of i^{th} order statistics is denoted by $\mu_{i:n}$.

VARIANCE

Variance is the expectation of squared deviation of random variable from its mean. The variance of i^{th} order statistics is denoted by $\sigma_{i,i:n}$.

DISTRIBUTION OF AN ORDER STATISTIC

Let us assume that $X_1, X_2, X_3, \dots X_n$ is a random sample from an absolutely continuous population with probability density function f(x) and cumulative distribution function F(x); let $X_{1,n} \le X_{2,n} \le X_{3,n}, \dots \dots \le X_{n,n}$ be the order statistics obtained by arranging the preceding random sample in increasing order of magnitude. Then, the event $x < X_{i,n} \le x + \delta x$ is essentially same as the following event:



$X_r \leq x$ for i-1 of the X_r's, $x < X_r \leq x + \delta x$ for exactly one of the X_r's and $X_r > x + \delta x$ for the remaining n-i of the X_r's. By considering δx to be small, we may write,

$$P\left(x \leq X_{i,n} \leq x + \delta x\right) = \frac{n!}{(i-1)!(n-i)!}\{F(x)\}^{i-1}\{1 - F(x + \delta x)\}^{n-i}$$

$$\times \{F(x + \delta x) - F(x)\} + O((\delta x)^2) \qquad (1.1)$$

where $O((\delta x)^2)$, a term of order $(\delta x)^2$ is the probability corresponding to the event of having more than one X_r in the interval $(x, x + \delta x]$. From (1.1), we may derive the density function of $X_{i,n}$ $(1 \leq i \leq n)$ to be

$$f_{i:n}(x) = \lim_{\delta x \to 0}\left\{\frac{P(x < X_{i,n} \leq x + \delta x)}{\delta x}\right\}$$

$$= \frac{n!}{(i-1)!\,(n-i)!}\{F(x)\}^{i-1}\{1 - F(x)\}^{n-i} f(x) \qquad -\infty < x < \infty \qquad (1.2)$$

The above expression of the pdf of $X_{i,n}$ can also be derived directly from the joint density function of all n order statistics. To show this we first of all need to note that, given that realizations of the n-order statistics to be $x_{1,n} < x_{2,n} < x_{3,n}, \dots\dots < x_{n,n}$, the original variables X_i are restrained to take on the values $x_{i,n}$ $(i = 1,2,\dots n)$, which by symmetry assigns equal probability for each of the $n!$ permutations of (1,2,...n). Hence we have the joint density function of all n-order statistics to be

$$f_{1,2,\dots,n:n}\left(x_{1,n} < x_{2,n} < x_{3,n}, \dots\dots < x_{n,n}\right) = n!\prod_{r=1}^{n} f(x_r), \quad -\infty < x_1 < x_2 < \infty \quad (1.3)$$

By considering the joint density function of all n order statistics in (1.3) and integrating out the variables $(X_{1:n}, \dots\dots X_{i-1:n})$ and $(X_{i+1:n}, \dots\dots X_{n:n})$ we derive the marginal density function of $X_{i,n}$ $(1 \leq i \leq n)$ to be ,

$$f_{i:n}(x) = n!\, f(x) \left\{ \int_{-\infty}^{x} \dots \int_{-\infty}^{x_2} f(x_1) \dots f(x_{i-1}) dx_1 \dots dx_{i-1} \right\}$$

$$\times \left\{ \int_{x}^{\infty} \dots \int_{x}^{x_{i+2}} f(x_{i+1}) \dots f(x_n) dx_{i+1} \dots dx_n \right\} \qquad (1.4)$$

Direct integration yields,

$$\int_{-\infty}^{x} \dots \int_{-\infty}^{x_3} \int_{-\infty}^{x_2} f(x_1) \dots f(x_{i-1}) dx_1 \dots dx_{i-1} = \frac{\{F(x)\}^{i-1}}{(i-1)!} \qquad (1.5)$$

and $$\int_{x}^{\infty} \dots \int_{x}^{x_{i+3}} \int_{x}^{x_{i+2}} f(x_{i+1}) \dots f(x_n) dx_{i+1} \dots dx_n = \frac{\{1-F(x)\}^{n-i}}{(n-i)!} \qquad (1.6)$$

Substitution of the eqn (1.5) and (1.6) for the two sets of integrals in eqn (1.4) gives the pdf of $X_{i,n}$ $(1 \le i \le n)$ to be exactly the similar to expression as in (1.2).

The pdfs of the smallest and largest order statistics follow from (1.2) (when i=1 and i=n) to be

$$f_{1:n}(x) = n\{1 - F(x)\}^{n-1} f(x), \quad -\infty < x < \infty \qquad (1.7)$$

and $$f_{n:n}(x) = n\{F(x)\}^{n-1} f(x), \quad -\infty < x < \infty \qquad (1.8)$$

respectively.

The distribution functions of the smallest and largest order statistics are easily derived, by integrating the pdfs in (1.7) and (1.8), to be

$$F_{1:n}(x) = 1 - \{1 - F(x)\}^n, \quad -\infty < x < \infty \qquad (1.9)$$

and $$F_{n:n}(x) = \{F(x)\}^n, \quad -\infty < x < \infty \qquad (1.10)$$

In general, the cdf of $X_{i,n}$ may be obtained by integrating the pdf of $X_{i,n}$ in (1.2). It may also be derived without much difficulty by realizing that

$$F_{i:n}(x) = P(X_{i,n} \le x)$$

$$= P(\text{at least } i \text{ of } X_1, X_2, X_3, \dots X_n \text{ are at most } x)$$

$$= \sum_{r=i}^{n} P(\text{exactly } r \text{ of } X_1, X_2, X_3, \dots X_n \text{ are atmost } x)$$

$$= \sum_{r=i}^{n} \binom{n}{r} \{F(x)\}^r \{1 - F(x)\}^{n-r} \qquad -\infty < x < \infty \qquad (1.11)$$

Thus, we find that the cdf of $X_{i,n}$ $(1 \le i \le n)$ is simply the tail probability of a binomial distribution with F(x) as the probability of success and n as the number of trials. Furthermore, by using the identity that

$$\sum_{r=i}^{n} \binom{n}{r} \{p\}^r \{1 - p\}^{n-r} = \int_0^p \frac{n!}{(i-1)!\,(n-i)!} t^{i-1}\{1-t\}^{n-i}\, dt, \qquad 0 < p < 1 \qquad (1.12)$$

we can write the cdf of $X_{i,n}$ from (1.11) equivalently as ,

$$F_{i:n}(x) = \int_0^{F(x)} \frac{n!}{(i-1)!\,(n-i)!} t^{i-1}\{1-t\}^{n-i}\, dt$$

$$= I_{F(x)}(i, n-i+1), \qquad -\infty < x < \infty \qquad (1.13)$$

which is incomplete Beta function. It should be pointed out here that the expression of $F_{i:n}(x)$ in (1.13) holds for any arbitrary population whether continuous or discrete. However, under the assumption that the population is absolutely continuous, we may differentiate the expression for the cdf in (1.13) and derive the pdf of $X_{i,n}$ $(1 \le i \le n)$ to be exactly similar expression in (1.2).

It is important to mention here that one can write the cdf of $X_{i,n}$ in terms of negative binomial probabilities instead of the binomial form given in (1.11). Let us write

$$F_{i:n}(x) = P(X_{i,n} \le x)$$

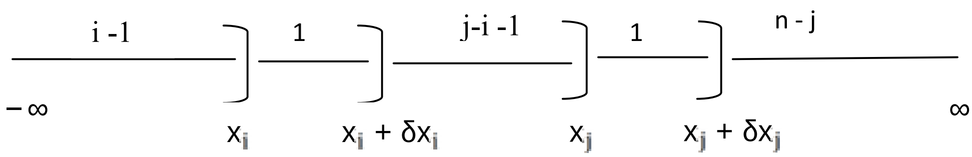
= P(reaching i successes in the course of atmost n trials with probability of success F(x))

$$= \binom{i-1}{i-1}\{F(x)\}^{i}\{1-F(x)\}^{0} + \binom{i}{i-1}\{F(x)\}^{i}\{1-F(x)\}^{i} + \cdots + \binom{n-1}{i-1}\{F(x)\}^{i}\{1-F(x)\}^{n-i}$$

$$= \sum_{r=0}^{n-i}\binom{n-1-r}{i-1}\{F(x)\}^{i}\{1-F(x)\}^{n-i-r} \qquad -\infty < x < \infty$$

JOINT DISTRIBUTION OF TWO ORDER STATISTICS

In order to derive the joint density function of two order statistics $X_{i,n}$ and $X_{j,n}$ $(1 \le i < j \le n)$, let us first visualise the event $(x_i < X_{i,n} \le x_i + \delta x_i\ ,\ x_j < X_{j:n} \le x_j + \delta x_j)$as follows



$X_r \le x_i$ for i-1 of the X_r's , $x_i < X_r \le x_i + \delta x_i$ for exactly one of the X_r's , $x_i + \delta x_i < X_r < x_j$ for j-i-1 of the X_r's , and $X_r > x_j + \delta x_j$ for the remaining n-j of the X_r's .By considering δx_i and δx_j to be both small, we may write

$$P(x_i < X_{i:n} \le x_i + \delta x_i, x_j < X_{j:n} \le x_j + \delta x_j)$$

$$= \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}\{F(x_i)\}^{i-1}\{F(x_j) - F(x_i + \delta x_i)\}^{j-i-1}$$

$$\times \{1 - F(x_j + \delta x_j)\}^{n-j} \times \{F(x_i + \delta x_i) - F(x_i)\}\{F(x_j + \delta x_j) - F(x_j)\}$$

$$+O((\delta x_i)^2 \delta x_j) + O((\delta x_j)^2 \delta x_i) \tag{1.14}$$

here $O((\delta x_i)^2 \delta x_j)$ and $O((\delta x_j)^2 \delta x_i)$ are higher-order terms which correspond to the probabilities of the event of having more than one X_r in the interval$(x_i, x_i + \delta x_i]$ and atleast one X_r in the interval $(x_j, x_j + \delta x_j]$, and of the event of having one X_r in $(x_i, x_i + \delta x_i]$ and more than one X_r in $(x_j, x_j + \delta x_j]$, respectively.

From (1.14), we may then derive the joint density function of $X_{i,n}$ and $X_{j,n}$ $(1 \le i < j \le n)$ to be

$$f_{i,j:n}(x_i, x_j) = \lim_{\delta x_i \to 0\, \delta x_j \to 0} \left\{ \frac{P(x_i < X_{i,n} \le x_i + \delta x_i\,,\; x_j < X_{j:n} \le x_j + \delta x_j)}{\delta x_i \delta x_j} \right\}$$

$$= \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}$$

$$\times \{F(x_i)\}^{i-1}\{F(x_j) - F(x_i)\}^{j-i-1}\{1 - F(x_j)\}^{n-j} f(x_i) f(x_j), \; -\infty < x_i < x_j < \infty$$

(1.15)

The joint density function of $X_{i:n}$ and $X_{j:n}$ $(1 \le i < j \le n)$ given in (1.15) can also be derived directly from the joint density function of all n order statistics. By considering the joint density function of all n order statistics in (1.3) and then integrating out the variables $(X_{1:n}, \dots \dots X_{i-1:n})$, $(X_{i+1:n}, \dots \dots X_{j-1:n})$ and $(X_{j+1:n}, \dots \dots X_{n:n})$,we derive the joint density function of $X_{i:n}$ and $X_{j:n}$ $(1 \le i < j \le n)$ to be

$$f_{i,j=n}(x_i, x_j) = n!\, f(x_i) f(x_j) \left\{ \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_2} f(x_1) \dots f(x_{i-1}) dx_1 \dots dx_{i-1} \right\}$$

$$\times \left\{ \int_{x_i}^{x_j} \dots \int_{x_i}^{x_{i+2}} f(x_{i+1}) \dots f(x_{j-1}) dx_{i+1} \dots dx_{j-1} \right\}$$

$$\times \left\{ \int_{x_j}^{\infty} \dots \int_{x_j}^{x_{j+2}} f(x_{j+1}) \dots f(x_n) dx_{j+1} \dots dx_n \right\} \qquad (1.16)$$

By direct integration we obtain,

$$\int_{-\infty}^{x_i} \dots \int_{-\infty}^{x_3} \int_{-\infty}^{x_2} f(x_1) f(x_2) \dots f(x_{i-1})\, dx_1 dx_2 \dots dx_{i-1} = \{F(x_i)\}^{i-1}/(i-1)! \qquad (1.17)$$

$$\int_{x_i}^{x_j} \cdots \int_{x_i}^{x_{i+3}} \int_{x_i}^{x_{i+2}} f(x_{i+1})f(x_{i+2}) \ldots f(x_{j-1})\, dx_{i+1}dx_{i+2} \ldots dx_{j-1} = \{F(x_j) - F(x_i)\}^{j-i-1}/(j-i-1)! \tag{1.18}$$

and

$$\int_{x_j}^{\infty} \cdots \int_{x_j}^{x_{i+3}} \int_{x_j}^{x_{i+2}} f(x_{j+1})f(x_{j+2}) \ldots f(x_n)\, dx_{j+1}dx_{j+2} \ldots dx_n = \{1 - F(x_j)\}^{n-j}/(n-j)! \tag{1.19}$$

Upon substituting the expressions (1.17)-(1.19) for the three sets of integrals in equation (1.16), we obtain the joint density function of $X_{1:n}$ and $X_{j:n}$ $(1 \le i < j \le n)$ to be exactly the same expressions as derived in (1.17) .

In particular, by setting i=1 and j=n in (1.15), we obtain the joint density function of the smallest and largest order statistics to be

$$f_{1,n:n}(x_1, x_n) = n(n-1)\{F(x_n) - F(x_1)\}^{n-2} f(x_1)f(x_n) \quad -\infty < x_1 < x_n < \infty \tag{1.20}$$

similarly, by setting j=i+1 in (1.17), we obtain the joint density function of two contiguous order statistics, $X_{1:n}$ and $X_{i+1:n}$ $(1 \le i \le n-1)$, to be

$$f_{i,i+1}(x_i, x_{i+1}) = \frac{n!}{(i-1)!\,(n-i-1)!} \times \{F(x_i)\}^{i-1}\{1 - F(x_{i+1})\}^{n-i-1} f(x_i)f(x_{i+1})$$

$$-\infty < x_i < x_{i+1} < \infty \tag{1.21}$$

The joint cumulative distribution function of $X_{1:n}$ and $X_{j:n}$ can, in principle, be obtained through double integration of the joint density function of $X_{1:n}$ and $X_{j:n}$ in (1.15) . It may also be written as

$$F_{i,j:n}(x_i, x_j) = F_{j:n}(x_j) \qquad \text{for } x_i \ge x_j,$$

and $F_{i,j:n}(x_i, x_j) = P(X_{i:n} \le x_i, X_{j:n} \le x_j)$ for $x_i < x_j$,

$= P($at least i of $X_1, X_2, \ldots X_n$ are at most x_i and atleast j of $X_1\ X_2, \ldots X_n$ are at most x_j)

$$= \sum_{s=j}^{n} \sum_{r=i}^{s} P(\text{exactly } r \text{ of } X_1, X_2, \ldots X_n \text{ are at most } x_i \text{ and exactly } s \text{ of } X_1\ X_2, \ldots X_n \text{ are at most } x_j)$$

$$= \sum_{s=j}^{\infty} \sum_{r=i}^{\infty} \frac{n!}{r!\,(s-r)!\,(n-s)!}$$

$$\times \{F(x_i)\}^r \{F(x_j) - F(x_i)\}^{s-r} \{1 - F(x_j)\}^{n-s} \qquad (1.22)$$

Thus, we find that the joint cdf of $X_{1:n}$ and $X_{j:n}$ $(1 \le i \le n-1)$ is the tail probability of a bivariate binomial distribution. By using the identity that

$$\sum_{s=j}^{\infty} \sum_{r=i}^{\infty} \frac{n!}{r!\,(s-r)!\,(n-s)!} P_1^{\,r} (P_2 - P_1)^{s-r} (1 - P_2)^{n-s}$$

$$= \int_0^{P_1} \int_{t_1}^{P_2} \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!} \times t_1^{\,i-1} (t_2 - t_1)^{j-i-1} (1 - t_2)^{n-j} dt_2\, dt_1$$

$$0 < P_1 < P_2 < 1 \qquad (1.23)$$

We can write the joint cdf of $X_{1:n}$ and $X_{j:n}$ in (1.22), equivalently as,

$$F_{i,j:n}(x_i, x_j) = \int_0^{F(x_i)} \int_{t_1}^{F(x_j)} \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!} \times t_1^{\,i-1} (t_2 - t_1)^{j-i-1} (-t_2)^{n-j} dt_2\, dt_1,$$

$$-\infty < x_i < x_j < \infty \qquad (1.24)$$

which may be noted to be incomplete Bivariate Beta function. The expression of $F_{i,j:n}(x_i, x_j)$ in (1.24) holds for any arbitrary population. when the population is continuous the joint density function of $X_{i:n}$ and $X_{j:n} (1 \le i \le n)$ in (1.15) may be derived from (1.24) by differentiating with respect to both x_i and x_j.

DISTRIBUTIONS OF SAMPLE MEDIAN AND SAMPLE RANGE

SAMPLE MEDIAN

The sample median, denoted by $\tilde{X}_n$, is a number such that approximately one half of the observations are less than $\tilde{X}_n$ and one half are greater. In terms of Order Statistics, $\tilde{X}_n$ is defined by

$$\tilde{X}_n = \begin{cases} X_{(n+1)/2):n} & \text{if n is odd} \\ \dfrac{(X_{(\frac{n}{2}):n} + X_{(\frac{n}{2}+1):n})}{2} & \text{if n is even} \end{cases}$$

Consider the sample size n to be odd. Then from equation (1.2) we have the pdf of the sample median $\tilde{X}_n = X_{(n+1)/2):n}$ to be

$$f_{\tilde{X}_n}(x) = \frac{n!}{\{|\frac{n-1}{2}|!\}^2}\{F(x)(1-F(x))\}^{\frac{n-1}{2}} f(x), \quad -\infty < x < \infty, \quad (1.25)$$

In particular for standard uniform population, the pdf of the sample median given in (1.25) becomes

$$f_{\tilde{U}_n}(u) = \frac{n!}{\{|\frac{n-1}{2}|!\}^2} u^{(n-1)/2}(1-u)^{(n-1)/2}, \quad 0 \le u < 1, \quad (1.26)$$

from which the m^th^ moment of $\tilde{U}_n$ is obtained as

$$E(\tilde{U}_n) = \frac{n!}{(n+m)!}\frac{|\frac{n-1}{2}+m|!}{(\frac{n-1}{2})!}, \quad m = 1,2,\ldots.. \quad (1.27)$$

Then the mean and variance of $\tilde{U}_n$ are

$$E(\tilde{U}_n) = \frac{1}{2}\ ; \ var(\tilde{U}_n) = \frac{1}{4(n+2)}$$

where $\tilde{U}_n$ is the sample median of standard uniform distribution.

SAMPLE RANGE

The sample range ,$W_n = X_{n:n} - X_{1:n}$ is the distance between the smallest and largest observation.

The pdf of sample range W_n is

$$f_{W_n}(w) = n(n-1)\int_{-\infty}^{\infty}\{F(x_1+w)-F(x_1)\}^{n-2}f(x_1)f(x_1+w)dx_1, \quad 0 < w < \infty \quad (1.28)$$

The cdf of sample range W_n may be derived as

$$F_{W_n}(w) = P(W_n \le w)$$

$$= n \int_{-\infty}^{\infty} \{F(x_1 + w) - F(x_1)\}^{n-1} f(x_1)dx_1, \quad 0 < w < \infty \quad (1.29)$$

In particular in standard uniform distribution, since $f(x_1 + w) \equiv 0$ when $x_1 + w > 1$, we obtain from equation(1.28) the pdf of sample range W_n as

$$f_{W_n}(w) = n(n-1) \int_0^{1-w} w^{n-2}\, dx_1$$

$$= n(n-1)w^{n-2}(1-w) \qquad 0 < w < 1 \qquad (1.30)$$

From equation (1.30), we obtain the cdf of the sample range W_n as

$$F_{W_n}(w) = n(n-1) \int_0^{w} w^{n-2}\, dx_1$$

$$= nw^{n-1} - (n-1)w^n, \qquad 0 < w < 1 \qquad (1.31)$$

BINOMIAL DISTRIBUTION

The probability mass function is given by

$$p(x) = \binom{n}{x} p^x q^{n-x} \ , \ x=0,1,2......,n \ ; \ p+q = 1$$

and cumulative distribution function is given by

$$F(k) = Pr(X \le k)$$

$$= \sum_{x=0}^{k} \binom{n}{x} p^x q^{n-x}$$

EXPONENTIAL DISTRIBUTION

The probability density function is given by

$$f(x, \theta) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & \text{Otherwise} \end{cases}$$

and cumulative distribution function is given by

$$F_x(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-\lambda x} & \text{if } x \ge 0 \end{cases}$$

LOGISTIC DISTRIBUTION

The probability density function of logistic distribution is

$$f_X(x) = \frac{e^{-\frac{(x-\mu)}{\sigma}}}{\sigma[1 + e^{-\frac{(x-\mu)}{\sigma}}]^2}, \qquad -\infty < x < \infty, \sigma > 0, -\infty < \mu < \infty$$

The cumulative distribution function for this distribution is

$$F_x(x) = \frac{1}{1 + e^{-\frac{(x-\mu)}{\sigma}}}, \qquad \sigma > 0, -\infty < \mu < \infty$$

COMPLETE BETA FUNCTION

A generalization of the complete beta function is defined by

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}\, dt, \qquad a, b > 0$$

INCOMPLETE BETA FUNCTION

Given a > 0 , b > 0 and $0 \leq p \leq 1$

$$I_p(a, b) = \frac{\int_0^p t^{a-1}(1-t)^{b-1}\, dt}{\int_0^1 t^{a-1}(1-t)^{b-1}\, dt}$$

is the incomplete beta function.

Assumptions to define order statistic

In the classical theory of order statistics, it is assumed that the X's are independent and identically distributed random variables. But the definition of order statistics need not require that

i. $X_i^{\prime}s$ to be identically distributed

ii. $X_i^{\prime}s$ to be idependent

iii. associated distributions to be continuous

iv. associated densities need to exist

ROLE OF ORDER STATISTICS

Order statistics and related theory have many interesting and important applications. Certain significant applications are listed.

i. **Robust location estimates**. Suppose that n independent measurements are available and we wish to estimate their assumed common mean. It has long been recognize that the sample mean suffers from an extreme sensitivity to outliers and model variations. Estimates based on the median or the average of central order statistics are less sensitive to model assumptions. A particular application of this observation is the accepted practice of using trimmed means (ignoring highest and lowest scores) especially in evaluating Olympic skating performances.

ii. **Detection of outliers**. If one is confronted with a set of measurements and is concerned with determining whether some have been incorrectly made or reported, attention naturally focuses on certain order statistics of the sample. Usually the largest one (or two) and/or the smallest one (or two) are deemed most likely to be outliers.

iii. **Natural disaster**. Disastrous floods and destructive earthquakes recur throughout history. Dam construction has long been focused on so called 100-year flood. Presumably the dams are built big enough and strong enough to handle any water flow to be encountered except for a level expected to occur only once every 100 years. Whether one agrees or not with the 100-year disaster philosophy, it is obvious that designers of dams and skycrapers, and even doghouses, should be concerned with the distribution of large order statistics from a possibly dependent, but possibly not identically distributed sequence.

iv. **Strength of materials**. The adage that a chain is no longer than its weakest link underlines much of the theory of strength of materials, whether they are threads, sheets, or blocks. By considering failure potential in infinitely small sections of the material, quickly lead to strength distributions associated with limits of distributions of sample minima. Of course, if we stick to the finite chain with n links, its strength would be the minimum of the strengths of its n component links, again an order statistic.

v. **Reliability**. The example of a cord composed of n threads can be extended to lead us to reliability applications of order statistics. It may be that failure of one thread will cause the cord to break (the weakest link), but more likely the cord will function as long as r (a number less than n) of the threads remains unbroken, as such it is an example of a r out of n system commonly discussed in reliability settings. With regard to tire failure in automobile, is often an example of a 4 out 5 system (remember the spare).

Borrowing on terminology from electrical systems, the n out of n system is known as a series system, any component failure is disastrous. The 1 out of n system is known as a parallel system, it will function as long as any of the component survives. The life of the r out of n system is clearly$X_{n-r+1:n}$, the (n - r + 1)th largest observation of the component lifetimes, or equivalently, the time until less than r components are functioning. The study of system lifetime will necessarily involve distributions of order statistics.

vi. **Quality control**. Here we use example of production of snickers candy bars passing through a conveyor belt. Each candy bar should weigh 2.1 ounces. No matter how well the pouring machine functions, minor fluctuations will occur, and potentially major aberrations might be encountered. We must be alert for correctable malfunctions causing unreasonable variation in the candy bar weight. In quality control, a sample of candy bars is weighted every hour, and close attention is paid to the order statistics of the weights so obtained. If the median or the mean is far from the target value, we must shut down the line. Attention is also focused on the sample range, if it is too large, the process is out of control, and the widely fluctuating candy bar weights will probably cause problems further down the line. Hence, quality control clearly involve order statistics.

vii. **Censored sampling.** Consider life-testing experiments, in which a fixed number n of items are placed on test and the experiment is terminated as soon as a prescribed number r have failed. The observed lifetimes are thus $X_{1:n} \le X_{2:n} \le \cdots .. \le X_{r:n}$ whereas the lifetimes $X_{r+1:n} \le X_{r+2:n} \le \cdots .. \le X_{n:n}$ remain unobserved.

viii. **Characterizations and goodness of fit.** The exponential distribution is famous for its so-called lack of memory. The usual model involves an electronic device. The argument goes that a light bulb that has been in service 20 hours is no more and no less likely to fail in the next minute than one that has been in service for, say, 5 hours, or even, than a brand new bulb. Such a curious distributional situation is reflected by the order statistics from exponential samples.

NOTATIONS

X	=	population random variable
pdf	=	probability density function
pmf	=	probability mass function
cdf	=	cumulative distribution function
i.i.d	=	independent and identically distributed
f(x)	=	probability density function
F(x)	=	cumulative distribution function , Pr $(X \leq x)$
$X_{i,n}$	=	i th order statistic in a sample of size n
$f_{i:n}(x)$	=	probability mass function of $X_{1,n}, X_{2,n}, \ldots \ldots X_{n,n}$
$f_{i,j:n}(x_i, x_j)$	=	joint probability mass function of $X_{i,n}$ and $X_{j,n}$ for $i < j$
$F_{i:n}(x)$	=	cumulative distribution function of $X_{i,n}$
$F_{i,j:n}(x_i, x_j)$	=	joint cumulative distribution function of $X_{i,n}$ and $X_{j,n}$ for $i < j$
$M_{i:n}(t)$	=	moment generating function
$x_{i,n}$	=	realisation of X
$\sigma_{i,i:n}$	=	variance of $X_{i,n}$ or var $(X_{i,n})$
$\sigma_{i,\,j:n}$	=	covariance between $X_{i,n}$ and $X_{j,n}$ or cov $(X_{i,n}, X_{j,n})$
$\tilde{X}_n$	=	sample median
$\tilde{U}_n$	=	sample median of standard uniform distribution
W_n	=	sample range

CHAPTER II
BINOMIAL DISTRIBUTION

This chapter presents the Binomial distribution and its moments. The mean and variance of extreme order statistics with respect to binomial distribution are derived and tables are prepared for their values. A computer code using C language is presented to help the user in deriving mean and variance of extreme order statistics.

James Bernoulli discovered Binomial distribution in the year 1700 and first published posthumously in 1713, eight years after his death. Binomial distribution is applicable under the following situations

i. random experiment to be performed repeatedly, each repetition being called a trial.

ii. each trial turns in only in one of the two outcomes, the occurrence of an event in a trial be called a success and its non-occurrence is a failure.

iii. the probability of occurrence of an event for each trial remains the same.

Consider a set of n independent Bernoullian trials in which the probability p of success in any trial is constant for each trial, then $q=1-p$, is the probability of failure in any trial.

The probability of x successes and consequently (n-x) failures in n independent trials, in a specified order SSFSFFFS....FSF (where S represents success and F represents failure) is given by the compound probability theorem

$$P(SSFSFFFS...FSF) = P(S)\ P(S)\ P(F)\ P(S)\ P(F)\ P(F)\ P(F)P(S)........P(F)\ P(S)\ P(F)$$

$$= p.p.q.p.q.q.q.p..........q.p.q$$

$$= p.p.p......p\ q.q.q.......q$$

$$= p^{x}\ .q^{n-x}$$

But x succcesses in n trails can occur in $\binom{n}{x}$ ways and the probability for each of these ways is same, $p^x q^{n-x}$. Hence the probability of x-successes in n-trials in any order is given by the addition theorem of probability by the expression,

$$\binom{n}{x} p^x q^{n-x}$$

A random variable X is said to follow binomial distribution if it assumes only non-negative values and its probability mass function is given by

$$P(X = x) = p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} , & x = 0,1,2 \ldots n \quad q = 1 - p \\ 0 , & \text{otherwise} \end{cases}$$

The two independent constants n and p in the distribution are known as the parameters of the distribution. n is also known as the degree of the binomial distribution.

X~B(n,p) denotes that the random variable X follows binomial distribution with parameters n and p. The graphical representations of pdf and cdf given in Fig.(2.1) and Fig.(2.2) respectively.

Moments of Binomial distribution

The first four moments about the origin of binomial distribution are obtained using the method of mathematical expectation.

The first moment, mean of binomial distribution is

$$\mu_1' = E(x)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$

$$= np \sum_{x=0}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= np(q + p)^{n-1}$$

$$= np$$

The second moment is

$$\mu_2' = E(X^2)$$

$$= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \{x(x-1) + x\} \frac{n(n-1)}{x(x-1)} \binom{n-2}{x-2} p^x q^{n-x}$$

$$= n(n-1)p^2 + np$$

The third moment is

$$\mu_3' = E(X^3)$$

$$= \sum_{x=0}^{n} x^3\, p(x)$$

$$= \sum_{x=0}^{n} \{x(x-1)(x-2) + 3x(x-1) + x\} \binom{n}{x} p^x q^{n-x}$$

$$= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np$$

The fourth moment is

$\mu_4' = E(X^4)$ which may be derived as

$$= n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3$$

$$+7n(n-1)p^2 + np$$

Central Moments of Binomial Distribution

$$\mu_1 = \mu_1' - \mu_1' = 0$$

$$\mu_2 = npq$$

$$\mu_3 = npq(q-p)$$

$$\mu_4 = npq\{1 + 3(n-2)pq\}$$

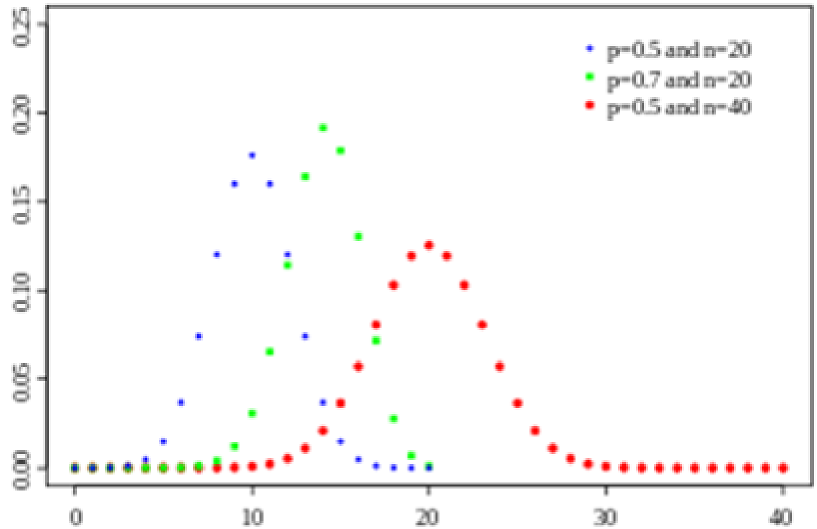


Fig. (2.1) The pmf of Binomial distribution

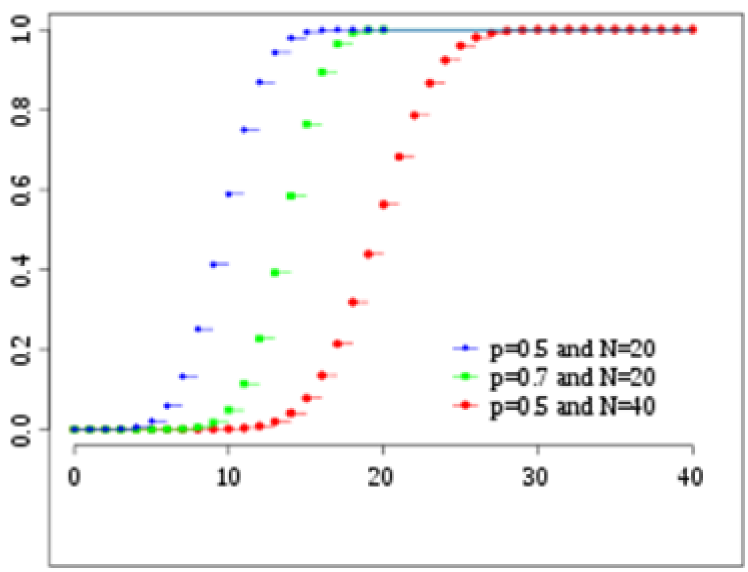


Fig. (2.2) The cdf of Binomial distribution

ORDER STATISTICS FOR BINOMIAL DISTRIBUTION

Consider the binomial population with pmf,

$$P(X = x) = \binom{N}{x} p^{x}(1-p)^{N-x} \quad , \qquad x = 0,1,2,\dots\dots N \qquad (2.1)$$

and cdf

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

$$= \sum_{r=0}^{x} \binom{N}{r} p^{r}(1-p)^{N-r} \quad , \quad x = 0,1,2,\dots\dots N \qquad (2.2)$$

Then from Equation (2.1) and (2.2) the pmf of $X_{i:n}$ $(1 \le i \le n)$ may be written as

$$f_{i:n}(x) = P(X_{i:n} = x)$$

$$= \sum_{r=i}^{n} \binom{n}{r} [\{F(x)\}^{r}\{1-F(x)\}^{n-r} - \{F(x-1)\}^{r}\{1-F(x-1)\}^{n-r}]$$

$$x = 0,1,2,\dots N \qquad (2.3)$$

with $F(-1)=0$, The pmf of i^{th} order statistics in beta integral form is

$$f_{i;n}(x) = C(i;n) \int_{F(x-)}^{F(x)} u^{i-1}(1-u)^{n-i} du \qquad \text{where } C(i;n) = \frac{n!}{(i-1)!\,(n-i)!}$$

which can be expressed as

$$f_{i;n}(x) = I_{F(x)}(i, n-i+1) - I_{F(x-1)}(i, n-i+1), \qquad x = 0,1,2,\dots\dots N \qquad (2.4)$$

where $I_{\alpha}(a,b)$ is the incomplete beta function defined by

$$I_{\alpha}(a,b) = \frac{1}{B(a,b)} \int_{0}^{\alpha} t^{a-1}(1-t)^{b-1}\, dt \qquad (2.5)$$

The cdf of $X_{i:n}$ $(1 \le i \le n)$ is given as

$$F_{i:n}(x) = P(X_{i:n} \le x)$$

$$= \sum_{r=i}^{n} \binom{n}{r} \{F(x)\}^r \{1 - F(x)\}^{n-r}, \qquad x = 0,1,2, \ldots . N \qquad (2.6)$$

The first two moments of $X_{i:n}$ are

$$\mu_{i:n} = \sum_{x=0}^{\infty} \{1 - F_{i:n}(x)\}$$

$$\mu_{i:n}^{(2)} = 2 \sum_{x=0}^{\infty} x \{1 - F_{i:n}(x)\} + \mu_{i:n}$$

From Eqn(2.6) we obtain the cdf of $X_{i:n}$ and $X_{n:n}$ as

$$F_{1:n}(x) = \sum_{r=1}^{n} \binom{n}{r} \{F(x)\}^r \{1 - F(x)\}^{n-r}$$

$$= 1 - \{1 - F(x)\}^n, \qquad x = 0,1,2, \ldots . N \qquad (2.7)$$

and $$F_{n:n}(x) = \{F(x)\}^n, \qquad x = 0,1,2, \ldots . N \qquad (2.8)$$

Therefore the first two moments of order statistics particularly for Binomial distribution are,

$$\mu_{i:n} = \sum_{x=0}^{N-1} \{1 - F_{i:n}(x)\} \qquad 1 \le i \le n \qquad (2.9)$$

and $$\mu_{i:n}^{(2)} = 2 \sum_{x=0}^{N-1} x \{1 - F_{i:n}(x)\} + \mu_{i:n} \qquad 1 \le i \le n \qquad (2.10)$$

By using equations (2.7) and (2.8) in (2.9) and (2.10) we obtain

$$\mu_{1:n} = \sum_{x=0}^{N-1} \{1 - F(x)\}^n \qquad (2.11)$$

$$\mu_{n:n} = \sum_{x=0}^{N-1} \{1 - \{F(x)\}^n\} \qquad (2.12)$$

$$\mu_{1:n}^{(2)} = 2 \sum_{x=0}^{N-1} x \{1 - F(x)\}^n + \mu_{i:n} \qquad (2.13)$$

and

$$\mu_{n:n}^{(2)} = 2 \sum_{x=0}^{N-1} x\{1 - \{F(x)\}^{n}\} + \mu_{i:n} \qquad (2.14)$$

By making use of the expressions in (2.11) to (2.14), the mean and variance of $X_{1:n}$ and $X_{n:n}$ for various choices of n,N and p may be computed.

By considering N = 5, 10, 15, p = 0.1, 0.2, 0.3, 0.4, 0.5 and n = 5, 10 the values of mean and Variance for extreme order statistics are computed and are presented in Table 2.1. By considering N = 5, 10, 15, p = 0.1, 0.2, 0.3, 0.4, 0.5 and n = 15, 20 the values of Mean and Variance for extreme order statistics are computed and presented in Table 2.2

Table 2.1 Mean and variance of Extreme Order Statistics for the Binomial population.

		n=5				n=10			
N	P	$\mu_{1:n}$	$\mu_{n:n}$	$\sigma_{1,1:n}$	$\sigma_{n,n:n}$	$\mu_{1:n}$	$\mu_{n:n}$	$\sigma_{1,1:n}$	$\sigma_{n,n:n}$
5	0.1	0.0115	1.3188	0.0114	0.4544	0.0001	1.6544	0.0001	0.4202
	0.2	0.1386	2.0711	0.1219	0.5895	0.0189	2.4703	0.0185	0.4873
	0.3	0.4220	2.7049	0.2911	0.6286	0.1594	3.1222	0.1351	0.4993
	0.4	0.7985	3.2634	0.4301	0.4301	0.4615	3.6736	0.2814	0.4595
	0.5	1.2388	3.7612	0.5388	0.5388	0.8543	4.1457	0.3791	0.3791
10	0.1	0.1185	2.1550	0.1070	0.7460	0.0137	2.6089	0.0136	0.6478
	0.2	0.6649	3.5140	0.4263	1.0281	0.3301	4.0586	0.2391	0.8521
	0.3	1.4070	4.7089	0.7090	1.1628	0.9574	5.2929	0.4700	0.9279
	0.4	2.2560	5.8014	0.9397	1.1812	1.7317	6.3919	0.6672	0.9104
	0.5	3.1886	6.7114	1.1019	1.1019	2.6183	7.3817	0.8186	0.8186
15	0.1	0.3347	2.9116	0.2608	1.0085	0.1001	3.4480	0.0908	0.8699
	0.2	1.3214	4.8506	0.7264	1.4494	0.8658	5.5052	0.4626	1.1925
	0.3	2.4200	6.5926	1.1332	1.6815	1.9398	7.3026	0.7873	1.3396
	0.4	3.8449	8.2110	1.4526	1.7449	3.1838	8.9364	1.0570	1.3514
	0.5	5.2702	9.7298	1.6625	1.6625	4.5617	10.4383	1.2512	1.2512

Table 2.2 Mean and variance of Extreme Order Statistics for the Binomial population.

		n=15				n=20			
N	P	$\mu_{1:n}$	$\mu_{n:n}$	$\sigma_{1,1:n}$	$\sigma_{n,n:n}$	$\mu_{1:n}$	$\mu_{1:n}$	$\sigma_{1,1:n}$	$\sigma_{n,n::n}$
5	0.1	0.0000	1.84807	0.0000	0.3999	0.0000	1.9845	0.0000	0.3690
	0.2	0.0026	2.6819	0.0026	0.4491	0.0003	2.8271	0.0003	0.4254
	0.3	0.0633	3.3409	0.0593	0.4351	0.0252	3.4839	0.0246	0.4016
	0.4	0.2990	3.8847	0.2138	0.3947	0.1983	4.0237	0.1596	0.3488
	0.5	0.6656	4.3346	0.3115	0.3115	0.5457	4.4543	0.2794	0.2794
10	0.1	0.0016	2.8565	0.0016	0.6058	0.0002	3.0269	0.0001	0.5788
	0.2	0.1828	4.3514	0.15111	0.7711	0.1032	4.5489	0.0927	0.7199
	0.3	0.7398	5.6019	0.3722	0.8229	0.6032	5.8084	0.3184	0.7597
	0.4	1.4686	6.6997	0.5542	0.7928	1.2985	6.9036	0.4895	0.7233
	0.5	2.3259	7.6740	0.6985	0.6985	2.1342	7.8658	0.6285	0.6285
15	0.1	0.0315	3.7401	0.0305	0.7954	0.0099	3.93919	0.0098	0.7535
	0.2	0.6491	5.8665	0.3586	1.1039	0.5142	6.0953	0.3014	1.0054
	0.3	1.6511	7.6747	0.9212	1.1679	1.4645	7.9288	0.5634	1.0701
	0.4	2.8463	9.3146	0.8961	1.1776	2.6258	9.5669	0.8014	1.0763
	0.5	3.1947	10.8053	3.6860	1.0754	2.9527	11.0473	6.8783	0.9723

Computer code using C language to enable the user to compute mean and variance of extreme order statistics relating to binomial distribution

```
#include<stdio.h>
#include<conio.h>
#include<math.h>
void main()
{
int N,n,i,Nr;
long int fact(int);
float p,Sig1n,Signn,F[20],m1,m2,mx,n1,n2,s1,s2,s11,t1,t2,t11;
clrscr();
printf("Enter the values of N, p, n:");
scanf("%d %f %d",&N,&p,&n);
m1=0,mx=0,n1=0,s1=0,t1=0;
for(i=0;i<=N;i++)
{
Nr=N-i;
F[i]=(fact(N)/fact(i)*fact(Nr))*pow(p,i)*pow((1-p),Nr));
mx=mx+F[i];
printf("The values of F[%d] is : %6f \n",i,F[i]);
m2=pow((1-mx),n);
m1=m1+m2;
n2=1.0-(pow(mx,n));
```

```
n1=n1+n2;
s2=(float)(i*m2);
s1=s1+s2;
t2=(float)(i*n2);
t1=t1+t2;
}
printf("The value of mu1n=%6f \n",m1);
printf("The value of munn= %6f \n",n1);
s11=(2*s1)+m1;
sig1n=s11-(m1*m1);
printf("The value of sig1n=%6f \n ",sig1n);
t11=(2*t1)+n1;
signn=t11-(n1*n1);
printf("The value of signn =%6f \n",signn);
getch();
}
long int fact(int N)
{
long int f=1;
int r;
for(r=1;r<=N;r++)
f=f*r;
return(f); }
```

Chapter III

EXPONENTIAL DISTRIBUTION

This chapter is devoted to the study of Exponential distribution and its moments. The mean and variance of order statistics with respect to exponential distribution are derived and the tables are constructed. As a ready reckoner Computer code using C language is presented to help the user in deriving mean and variance of i^{th} order statistics.

The exponential distribution is a model widely used in reliability theory and survival analysis. Properties of order statistics and the use of derived results in estimating parameters of exponential distribution has been studied by Sarhan and Greenberg (1962), David (1981) and Balakrishnan and Cohen (1991).

Exponential Distribution refers to a statistical distribution used to model the time between independent events that happen at a constant average rate.

A random variable X is said to have an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$$f(x,\theta) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{Otherwise} \end{cases}$$

and cumulative distribution function is given by

$$F_x(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-\lambda x} & \text{if } x \geq 0 \end{cases}$$

$X \sim Exp(\lambda)$ denotes that X has an exponential distribution with its parameter λ, where λ is the average number of events within a given time period. The graphical representations of pdf and cdf are given in Fig. (3.1) and Fig.(3.2) respectively.

Moments of exponential distribution are derived through moment generating function.

Moment generating function of exponential distribution is

$$M_x(t) = E(e^{tx})$$

$$= \lambda \int_0^\infty e^{tx} e^{-\lambda x} \, dx$$

$$= \lambda \int_0^\infty \exp\{-(\lambda - t)x\} \, dx$$

$$= \left(1 - \frac{t}{\lambda}\right)^{-1}, \qquad \text{where } \left(\frac{t}{\lambda}\right) < 1$$

r^{th} order raw moment is $\mu_r' = \text{co. efficient of } \frac{t^r}{r!} \text{ in } M_X(t)$

$$= \frac{r!}{\lambda^r}$$

MEAN

$$E(X) = \int_0^\infty x \lambda e^{-\lambda x} \, dx$$

$$= \frac{1}{\lambda}$$

$$E(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx$$

$$= \frac{2}{\lambda^2}$$

VARIANCE

$$\text{Variance}(X) = E(X^2) - (E(X))^2$$

$$= \frac{1}{\lambda^2}$$

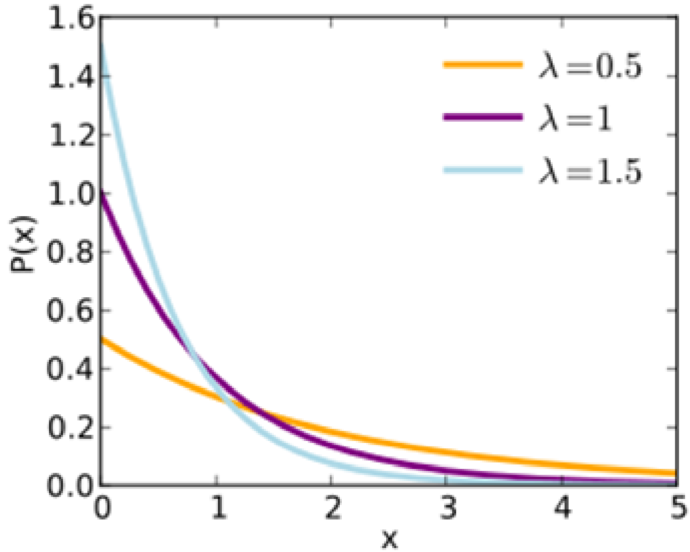


Fig.(3.1) The pdf of exponential distribution

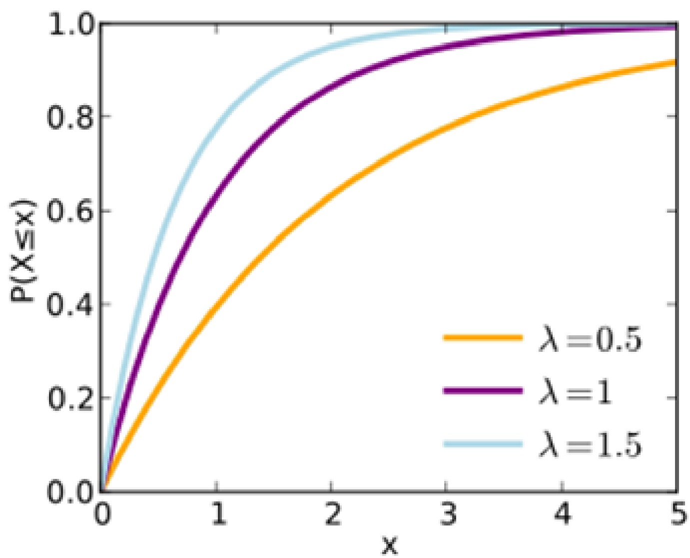


Fig.(3.2) The cdf of exponential distribution

ORDER STATISTICS FROM EXPONENTIAL DISTRIBUTION

The standard exponential population with p.d.f

$$f(x) = e^{-x}, \qquad 0 \le x \le \infty \qquad (3.1)$$

The joint density function of all n order statistics is

$$f_{1,2,\ldots\ldots n:n}(x_1, x_2, \ldots\ldots x_n) = n! \prod_{r=1}^{n} f(x_r) \qquad -\infty < x_1 < x_2, \ldots\ldots\ldots < x_n < \infty$$

Therefore the joint density function of $X_{1:n}, X_{2:n}, \ldots\ldots X_{2:n}$ with reference to exponential population is

$$f_{1,2,\ldots\ldots n:n}(x_1, x_2, \ldots\ldots x_n) = n!\, e^{-\sum_{i=1}^{n} x_i}, \qquad 0 \le x_1 < x_2, \ldots\ldots . < x_n < \infty \qquad (3.2)$$

Considered the transformation,

$$Z_1 = nX_{1:n}, Z_2 = (n-1)(X_{2:n} - X_{1:n}), \ldots., Z_n = X_{n:n} - X_{n-1:n}$$

The equivalent transformation is

$$X_{1:n} = \frac{Z_1}{n}, X_{2:n} = \frac{Z_1}{n} + \frac{Z_2}{n-1} + \cdots X_{n:n} = \frac{Z_1}{n} + \frac{Z_2}{n-1} + \cdots \ldots + Z_n \qquad (3.3)$$

The Jacobian of this transformation is $\frac{1}{n!}$ and that

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} (n-i+1)(x_i - x_{i-1}) = \sum_{i=1}^{n} Z_i$$

From (3.2) the joint density function of $z_1, z_2, \ldots\ldots z_n$ is obtained to be

$$f_{z_1,z_2,\ldots\ldots z_n}(z_1, z_2, \ldots\ldots z_n) = e^{-\sum_{i=1}^{n} z_i}, \qquad 0 \le z_1 < z_2, \ldots < z_n < \infty \qquad (3.4)$$

By using the factorization theorem, it is seen from (3.4) that the variables $z_1, z_2, \ldots\ldots z_n$ are independent and identically distributed standard exponential random variables.

Then the random variables $z_1, z_2, \ldots .. z_n$ where

$$Z_i = (n - i + 1)(X_{i:n} - X_{i-1:n}), \qquad i = 1,2, \ldots .n$$

with $X_{0:n} \equiv 0$, are all statistically independent and also have exponential distributions.

Further, from (3.3) we obtain

$$X_{i:n} \stackrel{d}{=} \sum_{r=1}^{i} \frac{Z_r}{(n - r + 1)}, \qquad i = 1,2, \ldots \ldots n \qquad (3.5)$$

The expression (3.5) shows that the i^th order statistics in a sample of size n from the standard exponential distribution is a linear combination of i independent standard exponential random variables. The representation of $X_{i:n}$ in (3.5) will enable us to derive the means, variances and covariances of exponential order statistics.

$$\mu_{i:n} = E(X_{i:n})$$

$$= \sum_{r=1}^{i} \frac{E(Z_r)}{(n - r + 1)}$$

$$= \sum_{r=1}^{i} \frac{1}{(n - r + 1)}, \qquad 1 \le i \le n \qquad (3.6)$$

$$\sigma_{i,i:n} = Var(X_{i:n})$$

$$= \sum_{r=1}^{i} \frac{Var(Z_r)}{(n - r + 1)^2}$$

$$= \sum_{r=1}^{i} \frac{1}{(n - r + 1)^2}, \qquad 1 \le i \le n \qquad (3.7)$$

and

$$\sigma_{i,j:n} = Cov\left(X_{i:n}, X_{j:n}\right)$$

$$= \sum_{r=1}^{i} \frac{Var(Z_r)}{(n - r + 1)^2} \qquad \text{for } 1 \le i < j \le n$$

$$\sigma_{i,i:n} = \sum_{r=1}^{i} \frac{1}{(n-r+1)^2} \qquad (3.8)$$

As an alternative method, by making use of the fact that $f(x) = 1 - F(x), (x \geq 0)$ for the standard exponential distribution, one can derive recurrence relations for the computations of single and product moments of all order statistics.

With the help of the above expressions (3.6),(3.7) and (3.8) the mean and variance for $X_{i:n}$ for various choices of n may be computed.

By considering n=1 to 15 the values of mean and variance for i^{th} order statistics are computed. The computed values are presented in Tables 3.1 and 3.2

A Computer code using C- language is presented to enable the user to compute mean and variance of i^{th} order statistics relating to Exponential distribution

```
#include <stdio.h>

#include<conio.h>

#include<math.h>

void main()

{
 int i,j,n;

float T,S,p,T1,S1;

clrscr();

printf("Enter the value of n:");

scanf("%d",&n);

S=0,S1=0;

for(i=1;i<=n;i++)

{

T=1.0/(float)(n-i+1);

S=S+T;

p=(float)((n-i+1)*(n-i+1));

T1=1.0/p;

S1=S1+T1;

printf("The value of i=%d \t mu=%6f \t variance=%6f \n",i,S,S1);

} getch(); }
```

Table 3.1 Mean and Variance of Exponential Order Statistics for n upto 10

n	i	$\mu_{i:n}$	$\sigma_{i,i:n}$	n	i	$\mu_{i:n}$	$\sigma_{i,i:n}$
1	1	1.0000	1.0000	7	7	2.5928	1.5117
2	1	0.5000	0.2500	8	1	0.1250	0.0156
	2	1.5000	1.2500		2	02678	0.0360
					3	0.4345	0.0638
3	1	0.3333	0.1111		4	0.6345	0.1038
	2	0.8333	0.3611				
	3	1.8333	1.3611		5	0.8845	0.1663
					6	1.2718	0.2724
4	1	0.2500	0.0625		7	1.7178	0.5274
	2	0.5833	0.1736		8	2.7178	1.5274
	3	1.0833	0.4236				
	4	2.0833	1.4236	9	1	0.1111	0.0123
					2	0.2361	0.0279
5	1	0.2000	0.0400		3	0.3789	0.0483
	2	0.4500	0.1025		4	0.5456	0.0761
	3	0.7833	0.2136				
	4	1.2833	0.4636		5	0.7456	0.1161
	5	2.2833	1.4636		6	0.9956	0.1786
					7	1.3289	0.2897
6	1	0.1666	0.0277		8	1.8289	0.5397
	2	0.3666	0.0677		9	2.8289	1.5397
	3	0.6166	0.1302				
	4	0.9500	0.2413	10	1	0.1000	0.0100
	5	1.4500	0.4913		2	0.2111	0.0223
	6	2.4500	1.4913		3	0.3361	0.0379
					4	0.4789	0.0583
7	1	0.1428	0.0204		5	0.6456	0.0861
	2	0.3095	0.0481				
	3	0.5095	0.0881		6	0.8456	0.1261
	4	0.7595	0.1506		7	1.0956	0.1886
	5	1.0928	0.2617		8	1.4289	0.2997
	6	1.5928	0.5117		9	1.9289	0.5497
					10	2.9289	1.5497

Table 3.2 Mean and Variance of Exponential Order Statistics for n from 11 to 15

n	i	$\mu_{i:n}$	$\sigma_{i,i:n}$	n	i	$\mu_{i:n}$	$\sigma_{i,i:n}$
11	1	0.09090	0.00826	13	11	1.68013	0.32089
	2	0.19090	0.01826		12	2.18013	0.57089
	3	0.30201	0.03060		13	3.18013	1.57089
	4	0.42701	0.04623				
				14	1	0.07142	0.00051
	5	0.56987	0.06664		2	0.14832	0.01101
	6	0.73654	0.09441		3	0.23168	0.01796
	7	0.93654	0.13441		4	0.32259	0.02622
	8	1.18654	0.19691				
					5	0.42259	0.03622
	9	1.51987	0.30803		6	0.53370	0.04857
	10	2.01987	0.55803		7	0.65870	0.06419
	11	3.01987	1.55803		8	0.80156	0.08460
12	1	0.08333	0.00694		9	0.96822	0.11238
	2	0.17424	0.01520		10	1.16822	0.15238
	3	0.27424	0.02520		11	1.47182	0.21488
	4	0.38535	0.03755		12	1.75156	0.32899
	5	0.51035	0.05317		13	2.25156	0.57599
	6	0.65321	0.07358		14	3.25156	1.57595
	7	0.81987	0.10136				
	8	1.01987	0.14136	15	1	0.06666	0.00444
					2	0.13809	0.00946
	9	1.26987	0.20386		3	0.21501	0.01546
	10	1.60320	0.31497		4	0.29835	0.02240
	11	2.10320	0.56497				
	12	3.10320	1.56497		5	0.38926	0.03067
					6	0.48926	0.04061
13	1	0.07692	0.00591		7	0.60037	0.05301
	2	0.16025	0.01286		8	0.72537	0.06864
	3	0.25116	0.02112				
	4	0.35116	0.03112		9	0.86822	0.08905
					10	1.03489	0.11682
	5	0.46227	0.04347				
	6	0.58727	0.05909		11	1.23489	0.15682
	7	0.73013	0.07950		12	1.48289	0.21932
	8	0.89680	0.10728				
					13	1.81822	0.33043
	9	1.09680	0.14728		14	2.31822	0.05804
	10	1.34680	0.20978		15	2.31822	1.55043

CHAPTER IV

LOGISTIC DISTRIBUTION

This chapter presents the Logistic distribution and its moments. The mean and variance of extreme order statistics with respect to logistic distribution are derived and the tables are prepared.

Pearl and Reed (1920) suggested the logistic growth function as a tool for use in demographic studies and particularly to estimate the growth of human population. The term logistic distribution function was developed by Reed and Berkson (1929). To study income distributions Fisk (1961) introduced the logistic distribution.

A detailed discussion of order statistics from the logistic distribution and some of their properties are presented in Gupta and Balakrishnan (1990). They derived the exact and explicit expressions for the single and product moments in terms of gamma function.

The probability density function of logistic distribution is

$$f_X(x) = \frac{e^{-\frac{(x-\mu)}{\sigma}}}{\sigma[1 + e^{-\frac{(x-\mu)}{\sigma}}]^2}, \qquad -\infty < x < \infty, \sigma > 0, -\infty < \mu < \infty$$

The cumulative distribution function for this distribution is

$$F_x(x) = \frac{1}{1 + e^{-\frac{(x-\mu)}{\sigma}}}, \qquad \sigma > 0, -\infty < \mu < \infty$$

The graphical representations of pdf and cdf are given in Fig.(4.1) and Fig.(4.2) respectively.

MEAN

$$E(X) = \mu$$

VARIANCE

$$V(X) = \frac{\sigma^2\pi^2}{3}$$

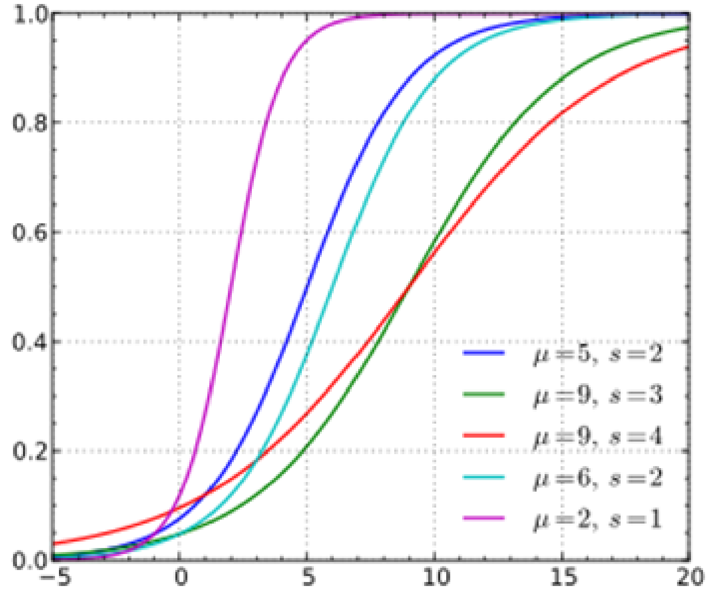


Fig.(4.1) The pdf of Logistic distribution

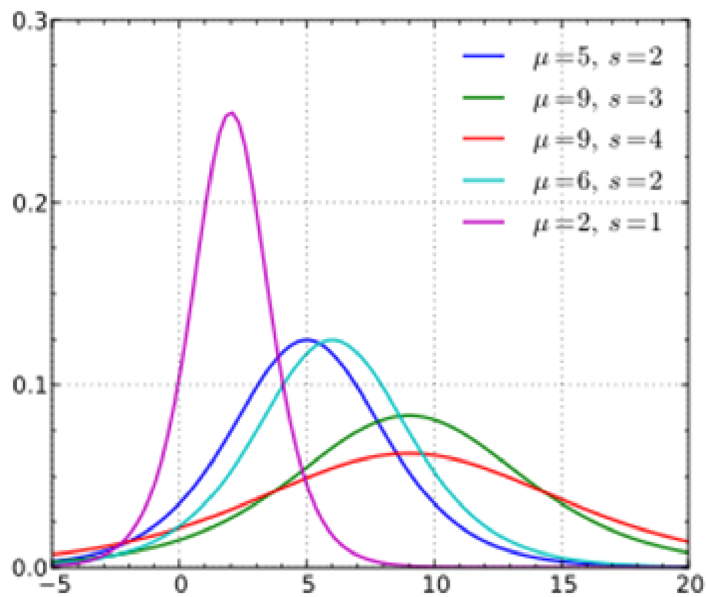


Fig.(4.2) The cdf of Logistic distribution

ORDER STATISTICS FROM LOGISTIC DISTRIBUTION

Consider the standard logistic population with pdf

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad -\infty < x < \infty \qquad (4.1)$$

and cdf

$$F(x) = \frac{1}{(1+e^{-x})}, \qquad -\infty < x < \infty \qquad (4.2)$$

By using $F^{-1}(u) = \log[\frac{u}{1-u}]$, we may write the moment generating function of $X_{i:n}$ from ,

$$f_{i:n}(x) = \lim_{\delta x \to 0} \left\{ \frac{P(x < X_{i,n} \le x + \delta x)}{\delta x} \right\}$$

$$= \frac{n!}{(i-1)!\,(n-i)!} \{F(x)\}^{i-1} \{1 - F(x)\}^{n-i} f(x), \qquad -\infty < x < \infty \quad \text{as}$$

Moment generating functions of logistic distribution is

$$M_{i:n}(t) = E(e^{tX_{i:n}})$$

$$= \frac{n!}{(i-1)!\,(n-i)!} \int_0^1 \left(\frac{u}{1-u}\right)^t u^{i-1} \{1-u\}^{n-i} \, du$$

$$= \frac{n!}{(i-1)!\,(n-i)!} B(i+t, n-i-t+1)$$

$$= \frac{\Gamma(i+t)\Gamma(n-i+1-t)}{\Gamma(i)\Gamma(n-i+1)} \qquad 1 \le i \le n \qquad (4.3)$$

where $\Gamma(.)$ is the complete Gamma function. From (4.3) we obtain the cumulant-generating function of $X_{i:n}$ as,

$$K_{i:n}(t) = \log M_{i:n}(t)$$

$$= \log\Gamma(i + t) + \log\Gamma(n - i + 1 - t) - \log\Gamma(i) - \log(n - i + 1),$$

$$1 \le i \le n \qquad (4.4)$$

From (4.4), we obtain the m^{th} cumulant of $X_{i:n}$ to be

$$\kappa_{i:n}^{(m)} = \frac{d^m}{dt^m} K_{i:n}(t)\Big|_{t=0}$$

$$= \frac{d^m}{dt^m} \log\Gamma(i + t)\Big|_{t=0} + \frac{d^m}{dt^m} \log\Gamma(n - i + 1 - t)\Big|_{t=0}$$

$$= \psi^{(m-1)}(i) + (-1)^m \psi^{(m-1)}(n - i + 1) \qquad (4.5)$$

where $\psi^{(0)}(z) \equiv \psi(z) = (\frac{d}{dz}) \log\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ is the digamma function, and $\psi^{(1)}(z), \psi^{(2)}(z), \ldots$ are the successive derivatives of $\psi(z)$, referred to as polygamma functions.From (4.5), we obtain in particular that

$$\mu_{i:n} = \kappa_{i:n}^{(1)} = \psi(i) - \psi(n - i + 1) \qquad (4.6)$$

and

$$\sigma_{i,i:n} = \kappa_{i:n}^{(2)} = \psi^{(1)}(i) + \psi^{(1)}(n - i + 1) \qquad (4.7)$$

where ψ and $\psi^{(1)}$ are the digamma and trigamma functions.

where $\psi^{(1)}(z) \equiv \psi(z) = (\frac{d^2}{d^2 z}) \log\Gamma(z)$ is trigamma function.Explicit expression for the covariance $\sigma_{i,j:n}$ in terms of digamma and trigamma functions.

From eqns (4.1) and (4.2), it is easy to note that

$$f(x) = F(x)\{1 - F(x)\} \qquad -\infty \le x \le \infty \qquad (4.8)$$

The relation in (4.8) has been utilized by shah(1966,1970) to establish some simple recurrence relations satisfied by the single and the product moments of order statistics.

For the standard logistic population with pdf and cdf as in (4.1) and (4.2) we have for m=1,2,.....

$$\mu_{i:n+1}^{(m)} = \mu_{1:n}^{(m)} - \frac{m}{n}\mu_{1:n}^{(m-1)}, \qquad n \geq 1 \tag{4.9}$$

and

$$\mu_{i+1:n+1}^{(m)} = \mu_{1:n}^{(m)} - \frac{m}{i}\mu_{1:n}^{(m-1)}, \qquad 1 \leq i \leq n \tag{4.10}$$

where $\mu_{1:n}^{(m)} \equiv 1$, for $1 \leq i \leq n$.

By making use of the above expressions, the mean and variance of $X_{1:n}$ and $X_{n:n}$ by various choices of n may be computed.

By considering n=1 to 10 the values of mean and variance for order statistics are computed. The computed values are presented in Tables 4.1, 4.2 and 4.3.

Table 4.1 Mean of Logistic Order Statistics

n	i	$\mu_{i:n}$	n	i	$\mu_{i:n}$
1	1	0.000000	7	7	2.450000
2	2	1.000000	8	5	0.250000
3	2	0.000000	8	6	0.783333
3	3	1.500000	8	7	1.450000
4	3	0.500000	8	8	2.592857
4	4	1.833333	9	5	0.000000
5	3	0.000000	9	6	0.450000
5	4	0.833333	9	7	0.950000
5	3	2.083333	9	8	1.592857
6	4	0.333333	9	9	2.717857
6	5	1.083333	10	6	0.200000
6	6	2.283333	10	7	0.616667
7	4	0.000000	10	8	1.092857
7	5	0.583333	10	9	1.717857
7	6	1.283333	10	10	2.828968

Table 4.2 Variance and Covariance of Logistic Order Statistics n upto 6

n	i	j	$\sigma_{i,\, j:n}$	n	i	j	$\sigma_{i,\, j:n}$
1	1	1	3.289868	5	1	5	0.269035
2	1	1	2.289868	5	2	2	0.928757
2	1	2	1.000000	5	2	3	0.590527
3	1	1	2.039868	5	2	4	0.433653
3	1	2	0.855066	5	3	3	0.789868
3	1	3	0.539868	6	1	1	1.826257
3	2	2	1.289868	6	1	2	0.738319
4	1	1	1.928757	6	1	3	0.458165
4	1	2	0.793465	6	1	4	0.331705
4	1	3	0.496403	6	1	5	0.259882
4	1	4	0.361111	6	1	6	0.213611
4	2	2	1.039868	6	2	2	0.866257
4	2	3	0.670264	6	2	3	0.546413
5	1	1	1.866257	6	2	4	0.399331
5	1	2	0.759642	6	2	5	0.314798
5	1	3	0.472872	6	3	3	0.678757
5	1	4	0.342976	6	3	4	0.502298

Table 4.3 Variance and Covariance of Logistic Order Statistics for n =7 to 10

n	i	j	$\sigma_{i,\ j:n}$	n	i	j	$\sigma_{i,\ j:n}$	n	i	j	$\sigma_{i,\ j:n}$
7	1	1	1.798479	8	2	7	0.199582	9	5	5	0.442646
7	1	2	0.723663	8	3	3	0.576257	10	1	1	1.750100
7	1	3	0.448112	8	3	4	0.4221171	10	1	2	0.698437
7	1	4	0.324031	8	3	5	0.333326	10	1	3	0.430922
7	1	5	0.253667	8	3	6	0.275479	10	1	4	0.310962
7	1	6	0.208385	8	4	4	0.505146	10	1	5	0.243115
7	1	7	0.176813	8	4	5	0.401428	10	1	6	0.199531
7	2	2	0.826257	9	1	1	1.762446	10	1	7	0.169184
7	2	3	0.518475	9	1	2	0.704838	10	1	8	0.146843
7	2	4	0.377749	9	1	3	0.435270	10	1	9	0.129711
7	2	5	0.297171	9	1	4	0.314261	10	1	10	0.116157
7	2	6	0.244966	9	1	5	0.245775	10	2	2	0.762446
7	3	3	0.616257	9	1	6	0.201761	10	2	3	0.474405
7	3	4	0.453287	9	1	7	0.171104	10	2	4	0.343956
7	3	5	0.358864	9	1	8	0.148528	10	2	5	0.269720
7	4	4	0.567646	9	1	9	0.131213	10	2	6	0.221830
8	1	1	1.778071	9	2	2	0.778071	10	2	7	0.188382
8	1	2	0.712975	9	2	3	0.485139	10	2	8	0.163699
8	1	3	0.440811	9	2	4	0.352157	10	2	9	0.144736
8	1	4	0.318472	9	2	5	0.276365	10	3	3	0.528071
8	1	5	0.249174	9	2	6	0.227421	10	3	4	0.384963
8	1	6	0.204612	9	2	7	0.193208	10	3	5	0.302949
8	1	7	0.173560	9	2	8	0.167946	10	3	6	0.249782
8	1	8	0.150686	9	3	3	0.548479	10	3	7	0.212513
8	2	2	0.798479	9	3	4	0.400684	10	3	8	0.184934
8	2	3	0.499214	9	3	5	0.315762	10	4	4	0.437368
8	2	4	0.362941	9	3	6	0.2602607	10	4	5	0.345659
8	2	5	0.285120	9	3	7	0.221890	10	4	6	0.285865
8	2	6	0.234796	9	4	4	0.465146	10	4	7	0.243769
				9	4	5	0.368453	10	5	5	0.402646
				9	4	6	0.305223	10	5	6	0.334261

SUMMARY AND CONCLUSION

The main work of this dissertation together with conclusions and recommendations for future work are presented in this part.

In Chapter I Various terms and definitions required for the preparation of dissertation are given.

In Chapter II the mean and variance of extreme order statistics with respect to binomial distribution are derived and the tables are prepared.

In Chapter III the mean and variance of order statistics with respect to exponential distribution are derived and the tables are prepared.

In Chapter IV the mean, variance and covariance of order statistics with respect to logistic distribution are derived and the tables are given.

Recommendations for Further Research

Mean and Variance of extreme order statistics may be computed for

i. Poisson Distribution

ii. Geometric Distribution and

iii. Normal Distribution

BIBILIOGRAPHY

Ahsanullah, M. and Hamedani, G. G. (2010), Exponential distribution, Nova Science Publication, New York.

Ahsanullah, M., Nevzorov, V. B., and Shakil, M. (2013), An introduction to order statistics, Atlantis Press, Paris.

Allen, T. C. (1932), On the distributions of certain statistics, Amer. J. Math., 54, pp.353-366.

Balakrishnan, N. and Cohen, A. (1991), Order statistics and inference: Estimation methods, Academic Press, San Diago.

Balakrishnan, N. and Chan, P. (1992), Order statistics from extreme value distribution, tables of means, variances and covariances. Communications in Statistics-Simulation and Computation, 21(4), pp.1199-1217.

Balakrishnan, K. (1996), Exponential distribution: theory, methods and applications, CRC press, Florida, US.

David, H. (1981), Order statistics, John Wiley Sons, New York.

Fisher, R. A. and Tippett, L. H. C. (1928), Limiting forms of the frequency distribution of the largest or smallest member of a sample,In Mathematical Proceedings of the Cambridge Philosophical Society, Cambridge University Press, England, 24, pp. 180-190,

Fisk, P. R. (1961), The graduation of income distributions. Econometrica: journal of the Econometric Society, 29(2),pp.171-185.

Galambos, J. (1978), The asymptotic theory of extreme order statistics, Volume 352, Wiley, New York.

Gumbel, E. J. (1935), Les valeurs extrmes des distributions statistiques, Annales del'Institut Henri Poincar, 4, pp.115-158.

Gumbel, E. J. (1941), The return period of flood flows, Ann. Math. Statist., 12, pp.163-190.

Gumbel, E. (1958), Statistics of extremes, Columbia University press, New York.

Gupta, S. S. and Balakrishnan, N. (1990), Logistic order statistics and their properties, Technical report, DTIC Document.

Harter, H. L. (1969), Order statistics and their use in testing and estimation, Volume 1, US Government Printing Office, Washington, DC.

Harter, H. L. and Balakrishnan, N. (1996), CRC handbook of tables for the use of order statistics in estimation, CRC press,Florida, US.

Harter, H. L. and Balakrishnan, N. (1997), Tables for the use of range and studentized range in tests of hypotheses, CRC Press,Florida,US.

Hoeffding, W. (1953), On the distribution of the expected values of the order statistics, Ann. Math. Statist., 24,pp. 93-100.

Jones, M. (2004), Families of distributions arising from distributions of order statistics,Test, 13(1), pp.1-43.

Karl Pearson, (1902), Note on Francis Gallon's problem, Biometrika, 1, pp.390-399.

Kong, L., L. C. and Sepanski, J. (2007), On the properties of beta-gamma distribution. Journal of Modern Applied Statistical Methods, (1), pp.187-211.

Maritz, J. and Munro, A. (1967), On the use of the generalised extreme-value distribution in estimating extreme percentiles, Biometrics,23(1), pp.79-103.

Mises, V. R. (1936), La distribution de la plus grande de n valeurs. Revue de l'Union Interbalkanique, 1, pp.1-20.

Nagaraja, H. (1986), Structure of discrete order statistics,Journal of statistical planning and inference, 13, pp.165-177.

Nagaraja, H. (1992), Order statistics from discrete distributions, Statistics: a journal of theoretical and applied statistics, 23(3), pp.189-216.

Pearl, R. and Reed, L. J. (1920), On the rate of growth of the population of the united states since 1790 and its mathematical representation, Proceedings of the National Academy of Sciences, 6(6), pp.275-288.

Reed, L. J. and Berkson, J. (1929), The application of the logistic function to experimental data, The Journal of Physical Chemistry, 33(5), pp.760-779.

Rupert Jr, G. (2012), Simultaneous statistical inference, Springer Science and Business Media, Germany.

Sarhan, A. E. and Greenberg, B. G. (1962), Contributions to order statistics, Wiley NewYork.

Shah, B. K. (1966), On the bivariate moments of order statistics from a logistic distribution, Ann. Math. Statist., 37,pp.1002-1010.

Shah, B. K. (1970), Note on moments of a logistic order statistics, Ann. Math. Statist., 41, pp.2151-2152.

Tanis, E. A. (1964), Linear forms in the order statistics from an exponential distribution, Ann. Math. Statist.,35, pp. 270-276.

Tippett, L. H. (1925), On the extreme individuals and the range of samples taken from a normal population, Biometrika, 17(3and4),pp. 364-387.

Wilks, S. S. (1948), Order statistics, Bulletin of the American Mathematical Society,54(1), pp.6-50.