NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

CHARACTERIZATION AND IDENTIFICATION OF PROBABILITY DISTRIBUTIONS

BY

MD. SHAFIQUR RAHMAN

@

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in
Statistics
at
Dalhousie University
Halifax, Nova Scotia
August, 1991

ISBN 0-315-71552-9

TO MY MOTHER AND THE MEMORY OF MY FATHER

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Let $P(x; \theta)$ be the probability mass function or probability density function of a random variable X where $\theta \varepsilon \Re^p$, p being finite. Using the first k ($k \geq p$) raw or central moments of this distribution we eliminate the p parameters in θ and obtain a moment relation in k moments. We derive the raw and the central moment relations for a number of discrete and continuous distributions. These moment relations are used as criteria to characterize a distribution. In general the present method is effective. But there are some special situations, where the moment relations of two or more distributions are same or one particular moment function takes same value for two or more distributions. In such a situation we propose two moment ratios as extra criteria for deciding among them. These ratios are also useful in approximating the Neyman type A and the Generalized Poisson distribution by the Negative Binomial distribution. We can identify a distribution by using the ratios of the co-efficients of the recurrence relations obtained from its generating function.

Subsequently, a special class of the Exponential family of distributions named the family of Transformed Chi-square distributions is defined. Explicit expressions for the MVUE with MV of a function of the parameter of this family are given. The critical region and the power function for various tests of hypotheses for the parameter of this family are also obtained. An identification procedure with probability of correct identification is discussed in detail.

LIST OF ABBREVIATIONS AND SYMBOLS

c.d.f. = Cumulative distribution function

GNB = Generalized Negative Binomial

G.P. = Generalized Poisson

iff = If and only if

iid = Independently and identically distributed

m.g.f. = Moment generating function

NBD = Negative Binomial distribution

MV = Minimum variance

MVB = Minimum variance bound

MVBUE = Minimum variance bound unbiased estimator

NB = Negative Binomial

p.d.f. = Probability density function

p.g.f. = Probability generating function

p.m.f. = Probability mass function

r.v. = Random variable

UMP = Uniformly most powerful

UMVUE = Uniformly minimum variance unbiased estimator

μ = Population mean

m = Sample mean

σ = Population standard deviation

s = Sample standard deviation

ACKNOWLEDGEMENT

I express my deepest sense of gratitude to my thesis supervisor, Professor R. P. Gupta, for his continual assistance and encouragement throughout the writing of this thesis. It was a privilege and a pleasure for me to work with him.

I would like to thank Dr. Lise Manchester and Dr. George Gabor for their careful reading of the thesis and invaluable suggestions.

I had the privilege to discuss academic problems with several members of the Department and friends. I have learnt a lot from these discussions. I am grateful to all of them. I have enjoyed working in a department where one can expect a friendly "hello" from everybody.

I wish to acknowledge with gratitude the Killam Trust and the Faculty of Graduate Studies, for awarding me Killam Scholarship and Dalhousie Graduate Fellowship without which this study would have been impossible.

Finally, deepest appreciation and many thanks are due to all of my family members for their love, inspiration and sacrifice during the period of study.

Chapter 1

Introduction

At present a large inventory of discrete as well as continuous probability distributions is available. Most of the probability distributions and their properties can be found in the recent works of Rothschild et al. (1985), Patel et al. (1976), Johnson et al. (1969, 1970) and Patil et al. (1968). On account of the wide variety of available probability distributions, researchers in applied fields have begun to wonder which distribution will be the most appropriate one in a particular case and how to choose it ? One conventional method in this respect is to use Chi-square goodness of fit test. But the Chi-square goodness of fit test may give insignificant results for two or more suspected distributions for a particular data set. In such a situation, it is a difficult task to make a choice. We can only say that one distribution gives a better fit than others, but statistically we can not reject the possibility that the data set is from some other distribution(s). Thus, we need some criteria for making a choice. One such criterion is the method of moment relations. Let us consider an arbitrary distribution with p parameters. Since the moments of any distribution are functions of its parameters, by using the first k ($k \geq p$) raw or central moments of this distribution it is possible to eliminate its p parameters and obtain a moment relation in k moments. This moment relation can be used to characterize that distribution. Ferguson(1967) and Khan et al. (1986, 1987) have characterized some probability distributions through the conditional moments of order statistics with single or higher order gaps. Lin (1987, 1988, 1989) has used several recurrence relations and identities for product moments of order statistics to characterize some probability distributions. Other recent works on this topic are those of Govindarajulu (1966), Gupta (1984), Hwang (1975), Hwang et al. (1984) and Kirmani (1984). But in this thesis we have used ordinary moment relations to characterize probability distributions. Our method is equivalent to the method developed by Lukacs(1981) to characterize a distribution by zero regression of certain statistics.

Characterization of distributions by means of zero regression has been thoroughly discussed by Lukacs et al. (1964) and Kagan et al. (1973). Other recent works on this topic are those of Bar-Lev et al. (1986, 1987), Gordon (1973), Heller (1979, 1983, 1984), Jorgensen (1987), Kushner et al. (1981), Kushner (1987), Lukacs (1963), Richards (1984), Seshadri (1983) and Tweedie (1984). In Chapter 2, moment relations for a number of discrete and continuous distributions have been derived and their uses have been discussed thoroughly with suitable examples.

In general the method of moment relations is effective in characterizing a distribution. But there are some special situations, where one particular moment function takes same value for two or more distributions or the moment relations for two or more distributions are same. In such situations we need another criterion for making a choice. One such criterion is the method of moment ratios. If the moment relations for two or more distributions are same then we shall use the moment ratios $d_1 = \mu\,\mu_3/\mu_2^2$ and $d_2 = (\mu_2\mu_4 - \mu_3^2)/\mu_2^3$ as extra criteria for deciding among them. This is discussed in chapter 3. With the help of these ratios Generalized Poisson (Borel-Tanner) and Neyman type A distributions can be approximated by the Negative Binomial distribution. Comparison of exact and approximate distributions have been studied.

The Compound Poisson distribution was first considered by Greenwood and Yule (1920). Let X be a Poisson random variable having p.d.f. $P(x;\theta) = \theta^x e^{-\theta}/x!$, $x = 0,1,2, \ldots$, where the parameter θ (> 0) gives the expected number of 'events'. If different individuals of a population are associated with different values of θ, and if θ is distributed as a random variable with distribution function $F(\theta)$, the probability of x events in the total population will be given by

$$p(x) = \int_0^\infty \frac{\theta^x e^{-\theta}}{x!}\, dF(\theta) \qquad (1.1)$$

Following Greenwood et al. (1920) we shall refer to (1.1) as a family of Compound Poisson distributions. Negative Binomial, Hermite, Borel-Tanner, Neyman type-A etc. are distributions belonging to this family. These types of distributions have been successfully applied by many authors such as Neyman (1939), Palm

(1937), Lundberg (1940), Greenwood et al. (1920) and Eggenberger et al. (1923,1924) to problems of accident statistics, telephone traffic, fire damage, sickness-insurance, life-insurance, risk theory, and even in engineering. Following Feller (1965), the probability generating function (p.g.f.) of a Compound Poisson distribution can be written as

$$P(s)= \Sigma p_j s^j = \exp[a_0(s-1) + a_1(s^2-1)/2 + a_2(s^3-1)/3 + \ldots] \qquad (1.2)$$

A recurrence relation to calculate successive p_j's obtained by differentiating (1.2) with respect to s and equating the coefficients of s^j is given by

$$p_{j+1}(j+1) = a_0 p_j + a_1 p_{j-1} + \ldots + a_j p_0$$

$j=0,1,2, \ldots, \quad p_j = 0$ if $j<0$.

The cumulants of the distribution can be obtained by taking the logarithm of both sides of (1.2) and expanding them .

The i-th cumulant so obtained is

$$\begin{aligned} \kappa_i &= a_0 + 2^{i-1} a_1 + 3^{i-1} a_2 + 4^{i-1} a_3 + \ldots \\ &= \Sigma_r (r+1)^{i-1} a_r , \quad \ldots . \qquad (1.3) \\ & \qquad i=1,2,3, \ldots \end{aligned}$$

Thus the cumulants of any Compound Poisson distribution can be computed from (1.3). Hinz et al. (1967) have suggested that the plots of the sample values η_j and γ_j against j may be used in discriminating among the Negative Binomial, Neyman type A, Poisson Pascal and Poisson Binomial distributions. Here $\eta_0(\kappa) = \kappa_{(1)}$, $\eta_j(\kappa) = \kappa_{(j+1)}/\kappa_{(j)}$ (j=1,2,3, . . .), where $\kappa_{(j)}$ is the jth factorial cumulant and $\gamma_0 = \mu_1'$, $\gamma_j = \tau_{j+1}(R)/\tau_j(R)$ (j=1,2,3, . . .), where $R_j = p_j/p_0$ and p_j is the probability of the random variable having the value j, and the $\tau_j(R)$ are defined by

$$\log G(z) = \log p_0 + \sum_{j=1}^{\infty} \frac{\tau_j(R)\, z^j}{j!}$$

In particular $\tau_1 = R_1$, $\tau_2 = 2R_2 - R_1^2$, $\tau_3 = 6(R_3 - R_1 R_2) + 2R_1^3$, $\tau_4 = 24(R_4 - R_1 R_3 + R_1^2 R_2) - 12R_2^2 - 6R_1^4$, where for convenience τ_j is written in place of $\tau_j(R)$. These can be obtained generally through the recurrence relation

$$\tau_j = j!\, R_j - \sum_{i=1}^{j-1} \frac{(j-1)!}{(j-i-1)!} R_i\, \tau_{j-i}$$

Earlier, Ottested (1939) used the ratio $\mu_{(j+1)}/\mu_{(j)}$, where $\mu_{(j)}$ is the jth factorial moment against j, to discriminate among the Binomial, Poisson and Negative Binomial distributions. One can use the corresponding sample values in these criteria to find out the possible form of the underlying distribution. Because of the sampling fluctuations, a particular criterion may not provide reliable information to draw sound conclusions. In fact whenever it is possible, more than one criterion should be used and other characteristics should be verified to ascertain a distribution. Here we suggest use of the ratios of a_j's rather than the ratios of moments for identifying certain Compound Poisson distributions, especially those listed in Table 4.1. It may be noted that the cumulants do not necessarily identify a distribution while within the Compound Poisson family the a_j's do, hence the use of a_j's in place of the cumulants has some merit. We have discussed these in chapter 4. Characteristics and applications of some distributions belonging to the Compound Poisson family are also discussed in this chapter.

Following Kotz and Johnson (1982) there are two main types of parametric families of distributions in Statistics such as the transformation (or group) families and the Exponential families. The families of distributions generated from a single probability measure by a group of transformations on the sample space are called transformation families. For example, any location-scale family $\beta^{-1}f\{(x-\alpha)/\beta\}$, where f is a known p.d.f. of the random variable X. Following Barndorff-Nielsen (1978), the Exponential families are characterized by having p.m.f. or p.d.f. of the form

$$f(x;\theta) = \exp[a(x)b(\theta)+c(\theta)+h(x)] \tag{1.4}$$

A large number of commonly occurring families of distributions belong to the exponential family. Examples of such families are the Binomial, Poisson, Geometric, Normal, Gamma etc. Both Exponential and the transformation families of distributions possess particularly nice properties. Their general structures have been studied by many authors. The Cramér-Rao inequality, the Bhattacharrya inequality, the Lehmann-Scheffé theorem etc. play important roles in minimum variance unbiased estimation. To obtain the MVUE of any parameter or any function of the parameters by conventional methods we need to apply the method separately for each individual distribution. In testing any statistical hypothesis we need to choose an appropriate test-statistic. To find the critical region and the power of the test, we need to know the distribution of the test statistic. The distribution of the test statistic depends on the parent population. It varies from population to population. We have developed some general results for different types of estimators, the test statistic, the critical region and the power function of the test of hypothesis regarding some parameter(s). These results are true for a special family of distributions named the Transformed Chi-square family, which is a sub-family of the large Exponential family of distributions. By observing the p.m.f. or p.d.f. of any probability distribution belonging to the Transformed Chi-square family, one can easily obtain without much derivation various estimators, critical regions and power functions of the test concerning the parameter(s). These are discussed in detail with suitable examples in chapter 5. To summarise here, let $\underset{\sim}{X} = (X_1, X_2 , \ldots, X_n)$ be a random sample of size n drawn from a distribution that belongs to the family of Transformed Chi-squares having density of the form (1.4), then

(i) the distribution of $-2\Sigma a(X_i)b(\theta)$ is Central Chi-square with nk degrees of freedom,

(ii) $\Sigma a(X_i)$ is a sufficient statistic for θ or, any one to one function of θ,

(iii) $\Sigma a(X_i)/n$ is the MLE and UMVUE of $[-k/\{2b(\theta)\}]$ with MV $[k/\{2nb^2(\theta)\}]$,

(iv) $\Sigma a(X_i)/n$ is the MVBUE of $[-k/\{2b(\theta)\}]$ with MV $[k/\{2nb^2(\theta)\}]$ provided Cramér-Rao regularity conditions hold,

(v) $\left(\dfrac{\chi^2_{nk,\alpha_1}}{2\Sigma a(X_i)}, \dfrac{\chi^2_{nk,(1-\alpha_2)}}{2\Sigma a(X_i)} \right)$ is a $100\{1- (\alpha_1 + \alpha_2)\}\%$ confidence interval of $\{-b(\theta)\}$,

(vi) an α level UMP test for testing the hypothesis

$H_0 : \theta \leq \theta_0$ against

$H_1 : \theta > \theta_0$ is $\phi(\underline{x}) = 1$, if $\Sigma a(x_i) \geq [\chi^2_{nk,(1-\alpha)}]/\{-2b(\theta_0)\}$, provided $b(\theta)$ is strictly increasing in θ. Also the power function of the test is $P_\theta\{\chi^2_{nk} \geq [b(\theta)\, \chi^2_{nk,(1-\alpha)}\}/b(\theta_0)]\}$.

Obviously this is a time saving device in estimation and tests of hypotheses. In some practical situations to which conventional tests of homogeneity are applied, such as the F-test for the equality of p population means, the tests (whether or not they yield statistically significant results) do not supply the information that the experimenter seeks. For example, let the p populations be the populations of p different cities or counties. Then, possibly the hypothesis that the different cities or counties have the same average income is an unrealistic one since it is likely that if the cities or counties are different, the average incomes will also be different, and a sufficiently large sample will establish this fact at any preassigned level of significance. Moreover, the experimenter's problems usually begin after obtaining a significant result. After establishing that the average incomes are different, the experimenter usually desires to select the one which is best. The best city or county can be defined as the one having the highest average income of the people. A general identification procedure and the probability of correct identification of the best population or subset of populations have been discussed in section 5.5.

Chapter 2

Characterization by Moment Relations

2.1 Introduction

In this chapter a method for the construction of moment relations is presented. We derive the raw and the central moment relations for a number of discrete and continuous distributions. These moment relations are used as criteria to characterize a distribution. The present method is equivalent to characterizing the distribution by zero regression of certain statistics [Lukacs (1981)]. We shall need the following definition.

Definition 2.1.1 Constant Regression and Zero Regression :

Let U and V be two random variables. Suppose that the expected value of V i.e., E(V) exists. Then V is said to have constant regression on U if the conditional expectation of V, given U, equals the unconditional expectation of V, that is, if the relation $E(V|U) = E(V)$ holds almost everywhere. If $E(V) = 0$, then we say that V has zero regression on U. Thus if V has zero regression on U, then $E(V|U) = 0$.

Let us state the following lemma which is very important in deriving some statistics to characterize probability distributions.

Lemma 2.1.1 The random variable V has constant regression on U, iff the relation $E(Ve^{tU}) = E(V)\,E(e^{tU})$ holds for all t. Thus if V has zero regression on U, then $E(Ve^{tU}) = 0$.

Proof : Let the random variable V have constant regression on U i.e., $E(V|U) = E(V)$.

Therefore, $E(Ve^{tU}) = E\{e^{tU}\,E(V|U)\} = E\{e^{tU}\,E(V)\} = E(V)\,E(e^{tU})$.

Conversely, let the relation $E(Ve^{tU}) = E(V)\,E(e^{tU})$ holds for all t.

Then

$$E[e^{tU}\{V - E(V)\}] = E(Ve^{tU}) - E(V)\,E(e^{tU}) = 0$$

or, $E[e^{tU}\,E(\{V - E(V)\}|U)] = 0$

or, $$\int_{-\infty}^{\infty} e^{tU}\,E[\{V-E(V)\}|U]\,dF_1(u) = 0$$

Here, $F_1(u)$ is the marginal distribution of the random variable U. We introduce the probability function $P_U(A)$ of the random variable U instead of the distribution function $F_1(u)$. This is a set function, defined on all Borel sets of R_1. The preceding equation then becomes

$$\int_{R_1} e^{tU}\,E[\{V-E(V)\}|U]\,dp_U = 0$$

Let $\mu(A) = \int_A E[\{V-E(V)\}|U]\, dp_U$.

This is a function of bounded variation which is defined on all Borel sets A of R_1 and we see that

$$\int_{R_1} e^{tU}\, d\mu = 0.$$

This implies that $\mu(A) = \mu(R_1) = 0$. This is only possible if

$E[\{V - E(V)\}|U] = 0$

or, $E(V|U) = E(V)$.

Therefore, V has constant regression on U.

Again, if V has zero regression on U, then $E(V|U) = E(V) = 0$.

Thus, $E(Ve^{tU}) = 0$.

Hence the lemma is proved.

Let X be a r.v. having p.m.f or p.d.f. given by

$$P(x;\theta)\ ,\quad \theta \in \Re^{P} \tag{2.1.1}$$

We are interested in characterizing the distribution of X. Our approach is subject to the following assumptions :

(i) The distribution function to be characterized depends only on a finite number of parameters.

(ii) The existence of the moment generating function of the distribution is necessary.

Let $M \equiv M(t)$ be the moment generating function of X about the origin. Using the first k ($k \geq p$) raw or central moments of this distribution we can eliminate the p parameters in θ. In order to eliminate the p parameters in θ let us differentiate M successively 'r' times with respect to t. Let the jth derivative of M be written as

$$d^j M/dt^j = M^{(j)} = g_j(t;\ \theta) \tag{2.1.2}$$

where j=1,2,. . .,r and r can be chosen in such a way that we can eliminate not only the p parameters but also all expressions which contain the arbitrary variable t explicitly from the set of equations (2.1.2). By this method we obtain an ordinary differential equation which involves the moment generating function M and its derivatives. Let this differential equation have the following form

$$\sum_{S_1}\sum_{S_2}\cdots\sum_{S_k} b_{S_1,S_2,\ldots,S_k} M^{(S_1)} M^{(S_2)} \ldots M^{(S_k)} = 0 \tag{2.1.3}$$

Let $f=f(t)=E\{\exp(itX)\}$ be the characteristic function. Now using f and proceeding as before we get,

$$\sum_{S_1}\sum_{S_2}\cdots\sum_{S_k} a_{S_1,S_2,\ldots,S_k} f^{(S_1)} f^{(S_2)} \ldots f^{(S_k)} = 0 \tag{2.1.4}$$

Let $(X_1, X_2, \ldots, X_n)$ be a random sample of size n from a population having p.m.f. or p.d.f. (2.1.1). Also let $\Delta = \Sigma X_j$ and $S = S(X_1, X_2, \ldots, X_n)$ be two sample statistics such that S has zero regression on Δ. Then

$$E(Se^{\Delta t}) = E(S)\, E(e^{\Delta t}) = 0 \tag{2.1.5}$$

Lukacs (1981) has pointed out that the statistic S characterizing the function (2.1.1) by zero regression can be constructed by using the differential equation (2.1.4). In section 2.2 we show that the statistic S can also be constructed by using the relations of the first k ($k \geq p$) moments. Raw and central moment relations for a number of distributions are presented in Table 2.3.A and 2.3.B. These moment relations will be used to characterize a particular distribution.

2.2 Construction of 'S' Statistics by Moment Relations

We have $M = M(t) = E(e^{tX})$

$$M^{(r)} = E(X^r e^{tX}) \tag{2.2.1}$$

$$E(e^{\Delta t}) = (M)^n \tag{2.2.2}$$

$$M^{(r)} = \{E(X^r e^{\Delta t})\}/(M)^{n-1} \tag{2.2.3}$$

Let s_j (j= 1,2, . . .,k) be integers such that $r \geq s_j \geq 0$ (j= 1,2, . . .,k) and suppose that $X_j \neq X_l$ for $j \neq l$. Then

$$E(X_1^{s_1} X_2^{s_2} \ldots X_k^{s_k} e^{t\Delta}) = \{M(t)\}^{n-k} \prod_{j=1}^{k} E(X_j^{s_j} e^{tX_j}) \tag{2.2.4}$$

It follows from (2.2.1) that

$M^{(s_1)} M^{(s_2)} \ldots M^{(s_k)} (M)^{n-k} = E(X_1^{s_1} X_2^{s_2} \ldots X_k^{s_k} e^{t\Delta})$. Therefore,

$$\sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} b_{s_1,s_2,\ldots,s_k} M^{(s_1)} M^{(s_2)} \ldots M^{(s_k)} (M)^{n-k}$$

$$= \sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} b_{s_1,s_2,\ldots,s_k} E(X_1^{s_1} X_2^{s_2} \ldots X_k^{s_k} e^{t\Delta})$$

$= E(Se^{\Delta t}) = 0$ [Using (2.1.5)],

where, $S = \sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} b_{s_1,s_2,\ldots,s_k} X_1^{s_1} X_2^{s_2} \ldots X_k^{s_k}$

Thus,

$$\sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} b_{s_1,s_2,\ldots,s_k} M^{(s_1)} M^{(s_2)} \ldots M^{(s_k)} (M)^{n-k} = 0$$

or, $$\sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} b_{s_1,s_2,\ldots,s_k} M^{(s_1)} M^{(s_2)} \ldots M^{(s_k)} = 0$$

Let μ, μ_r' and μ_r be respectively the mean, rth raw and rth central moments of the distribution of X having p.m.f. or p.d.f. (2.1.1). Let

$$\mu_r' = E(X^r) \qquad (2.2.5)$$

$$\mu_r = E(X-\mu)^r \qquad (2.2.6)$$

where, r = 1,2,3, ... We want to characterize the distribution of X. Taking the first k equations from (2.2.5) or (2.2.6) and eliminating the p parameters in θ we obtain an equation of the form

$$\sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} c_{s_1,s_2,\ldots,s_k}(\mu)^{s_1}(\mu_2')^{s_2}(\mu_3')^{s_3}\cdots(\mu_k')^{s_k} = 0 \qquad (2.2.7)$$

$$\text{or}, \sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} d_{s_1,s_2,\ldots,s_k}(\mu)^{s_1}(\mu_2)^{s_2}(\mu_3)^{s_3}\cdots(\mu_k)^{s_k} = 0 \qquad (2.2.8)$$

where $s_j \geq 0$, $j=1,2,\ldots,k$. Evidently, an unbiased estimator of the expression on the left side of (2.2.7) is

$T = T(X_1, X_2, \ldots, X_n)$

$$= \sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} c_{s_1,s_2,\ldots,s_k}(X_1X_2\cdots X_{s_1})(X_{s_1+1}X_{s_1+2}\cdots X_{s_1+s_2})^2\cdots$$

$$(X_{s_1+s_2+\cdots+s_{k-1}+1}X_{s_1+s_2+\cdots+s_{k-1}+2}\cdots X_{s_1+s_2+\cdots+s_k})^k \qquad (2.2.9)$$

Let T have zero regression on $\Delta = \Sigma X_j$, so that $E(Te^{t\Delta}) = 0$. Therefore, using (2.2.9) we may write,

$$\sum_{s_1}\sum_{s_2}\cdots\sum_{s_k} c_{s_1,s_2,\ldots,s_k}\{M^{(k)}\}^{s_k}\{M^{(k-1)}\}^{s_{k-1}}\cdots\{M^{(1)}\}^{s_1}\{M\}^{n-b_k} = 0 \qquad (2.2.10)$$

where $b_k = s_1+s_2+\ldots+s_k$. Evidently, the expressions (2.2.4) and (2.2.10) are the same under the assumption that T has zero regression on Δ. This differential equation (2.2.10) can be solved for M(t) corresponding to (2.1.1). On the other hand differentiating M(t) successively with respect to t and eliminating the parameters we can derive (2.2.10). Then using (2.2.1), (2.2.2) and (2.2.3) we get $E(Te^{\Delta t}) = 0 = E(T)$ which implies that (2.2.7) and (2.2.8) hold. Thus (2.2.7) and (2.2.8) imply that T has zero regression on Δ. Hence we can use (2.2.7) or, (2.2.8) to characterize a distribution. Now using the value of T in (2.2.9) and the sample $(X_1, X_2, \ldots, X_n)$ we can easily construct the sample statistic S characterizing the distribution (2.1.1) by zero regression on Δ i.e., satisfying $E(Se^{\Delta t}) = 0$. This is given by

$$S = \sum_i \sum_j \ldots \sum_z \sum_{s_1} \sum_{s_2} \ldots \sum_{s_k} C_{s_1, s_2, \ldots, s_k} \frac{X_i^k X_j^k \ldots X_z^k}{n(n-1)(n-2)\ldots(n-s_1-s_2-\cdots-s_{j+1})} \qquad (2.2.11)$$

where the summations go over all subscripts i,j, . . ., which are all different and vary from 1 to n. S statistics for a number of distributions have been constructed and presented in Table 2.5.

2.3 Discussion

Consider the Poisson distribution having p.m.f.

$P_x(\theta) = e^{-\theta}\theta^x/x!$, $x = 0,1,2, \ldots$.

The moment generating function is $M(t) = \exp(-\theta+\theta e^t)$.

The first two raw moments are

$\mu_1' = \theta = \mu$ and $\mu_2' = \theta^2 + \theta$.

Now eliminating θ from these two equations we get,

$$\mu_2' - \mu^2 - \mu = 0 . \qquad (2.3.1)$$

Here, $\mu_2 = \theta$. Thus the central moment relation is $\mu_2 - \mu = 0$. Let

$$T=X_1^2-X_2X_3-X_4 \ . \qquad (2.3.2)$$

We observe that T is an unbiased estimator of $\mu_2' - \mu^2 - \mu$, where $\{X_1, X_2, \ldots\}$ are any i.i.d. random variables whose 2nd moments exist. If T has zero regression on $\Delta = X_1+X_2+\ldots+X_n$, then $E(Te^{\Delta t}) = 0$. That is,

$$E(Te^{\Delta t}) = E[(X_1^2-X_2X_3-X_4)\exp\{t(X_1+X_2+\ldots+X_n)\}]$$
$$=M''M^{n-1}-(M')^2M^{n-2}-M'M^{n-1}$$

$$\text{or,} \quad MM''-(M')^2-M'M=0 \qquad (2.3.3)$$

This is the differential equation in terms of the moment generating function for zero regression. To solve (2.3.3) for M(t), we may write

$$M''/M' = M'/M + 1$$

Integrating both sides of this equation with respect to t we get $\ln M' = \ln M + t + c$, where c is a constant of integration.

or, $\ln(M'/M) = t + c$. or, $M'/M = c_1e^t$, ($c_1 = e^c$).

Integrating again with respect to t we get

$$\ln M = c_1e^t + c_2 \ .$$

As $M(0) = 1$, $c_1 + c_2 = 0$. Let $c_1 = \theta$, thus $c_2 = -\theta$. So,

$$M(t) = \exp(-\theta+\theta e^t) \qquad (2.3.4)$$

Therefore, if T has zero regression on Δ, then $X_1, X_2, \ldots, X_n$ are from the Poisson distribution. Now if we start from (2.3.4) and eliminate θ from M, M',M" we obtain (2.3.3), which can be written as

$$E[(X_1^2-X_2X_3-X_4)\exp\{t(X_1+X_2+\ldots+X_n)\}] = E(Te^{\Delta t}) = 0.$$

This implies that T has zero regression on Δ. Again substituting

$t=0$ in (2.3.3) and remembering that $M(0) = 1$, $M'(0) = \mu$, $M''(0) = \mu_2'$ we get (2.3.2) and (2.3.1). Therefore, these moment relations characterize the Poisson distribution. Using T and the sample $(X_1, X_2, \ldots, X_n)$ we can easily construct the statistic S (Table 2.5.A and 2.5.B).

Now consider the Negative Binomial distribution having p.m.f

$$P(x;\theta) = \binom{N+x-1}{x} \theta^x (1-\theta)^N, \quad x = 0,1,2,\ldots, \quad N, \theta > 0 .$$

The moment generating function is $M(t) = (1-\theta)^N (1-\theta e^t)^{-N}$. The first three raw moments are

$\mu = N\theta/(1-\theta)$, $\mu_2' = N\theta(1+N\theta)/(1-\theta)^2$ and

$\mu_3' = N\theta(N^2\theta^2+3N\theta+\theta+1)/(1-\theta)^3$.

Eliminating N and θ we get,

$$\mu_3'\mu - 2(\mu_2')^2 + \mu\mu_2' - \mu^3 + \mu^2\mu_2' = 0 \qquad (2.3.5)$$

This is the raw moment relation.The second and third central moments are $\mu_2 = N\theta/(1-\theta)^2$ and $\mu_3 = N\theta(\theta+1)/(1-\theta)^3$

Thus the central moment relation is

$$\mu_3\mu - 2\mu_2^2 + \mu\mu_2 = 0 \qquad (2.3.6)$$

Let $T = X_1X_2^3 - 2(X_1X_2)^2 + X_1X_2^2 - X_1X_2X_3 + X_1X_2X_3^2$ be an unbiased estimator of $\mu_3'\mu - 2(\mu_2')^2 + \mu\mu_2' - \mu^3 + \mu^2\mu_2'$, where X_1, X_2, X_3 are i.i.d. random variables whose third moments are finite. If T has zero regression on Δ, then $E(Te^{t\Delta}) = 0$. That is,

$$M'M'''M^{n-2}-2(M'')^2M^{n-2}+M'M''M^{n-2}-(M')^3M^{n-3}+(M')^2M''M^{n-3}= 0$$

$$\text{or, } M'M'''M -2(M'')^2M + M'M''M - (M')^3 + (M')^2M'' = 0 \qquad (2.3.7)$$

This is the differential equation in the moment generating function for zero regression. This can be written as

$$(M')^2D\{MM'' - (M')^2 - MM'\} - \{MM'' - (M')^2 - MM'\}D(M')^2 = 0$$

$$\text{or, } D\left[(M')^2/\{MM'' - (M')^2 - MM''\}\right] = 0$$

or, $(M')^2/\{MM'' - (M')^2 - MM'\} = N$, a constant.

$$\text{or, } (M'/M)^2/[\{MM'' - (M')^2\}/M^2 - M'/M] = N$$

$$\text{or, } (D\ln M)^2/(D^2\ln M - D\ln M) = N$$

where, D is the differential operator, i.e., $D = d/dt$.
This is the simplified form of the differential equation (2.3.7).
Let $Z = D\ln M$. Thus the above equation can be written as
$DZ - Z = Z^2/N$. Let $W= 1/Z$. Then we may write,

$$DW + W + 1/N = 0$$

Here the integrating factor is $\exp\{\int dt\} = e^t$. Now multiplying both sides by the integrating factor we get,

$$e^t DW + We^t = -1/N \; e^t$$

$$\text{or, } D(W e^t) = -1/N \; e^t$$

$$\text{or, } We^t = -1/N \; e^t + c$$

or, $Z = N\theta e^t/(1-\theta e^t)$ [Where $c=1/N\theta$]

or, $D\ln M = (-N)\{(-\theta e^t)/(1-\theta e^t)\}$.

or, $\ln M = -N\ln(1-\theta e^t) + \ln c_1$.

As $M(0) = 1$, thus $c_1 = (1-\theta)^N$. Hence,

$$M(t) = (1-\theta)^N (1-\theta e^t)^{-N} \tag{2.3.8}$$

Thus if T has zero regression on Δ then $X_1, X_2, \ldots, X_n$ are from the Negative Binomial distribution. Now if we start from (2.3.8) and eliminate N and θ from M, M', M" and M"' we obtain (2.3.7), which can be written as

$$E\{(X_1 X_2^3 - 2 X_1^2 X_2^2 + X_1 X_2^2 - X_1 X_2 X_3^2)e^{t\Delta}\} = E(Te^{t\Delta}) = 0.$$

This implies that T has zero regression on $\Delta = X_1 + X_2 + \ldots + X_n$.

Substituting t=0 in (2.3.7) and remembering that $M(0) = 1$, $M'(0) = \mu$, $M''(0) = \mu_2'$, $M'''(0) = \mu_3$, we get (2.3.5) and (2.3.6). This implies that these moment relations characterize the Negative Binomial distribution.

Using T and the sample $(X_1, X_2, \ldots, X_n)$ we can easily construct the statistic S (Table 2.5).

The p.d.f. of the Generalized Negative Binomial distribution (Jain and Consul, 1971) is

$$P_x(N,\theta) = \frac{N(N+\alpha x-1)!}{x!\,(N+\alpha x-x)!}\,\theta^x(1-\theta)^{N+\alpha x-x}, \qquad x=0,1,2,\ldots\ldots.$$
$$N>0,\ \alpha\geq 1,\ 0<\theta<1,\ 0<\alpha\theta<1 .$$

The first four moments are $\mu = N\theta/(1-\alpha\theta)$, $\mu_2 = N\theta(1-\theta)/(1-\alpha\theta)^3$,

$$\mu_3 = N\theta(1-\theta)\{1-2\theta+\alpha\theta(2-\theta)\}/(1-\alpha\theta)^5 ,$$

$$\mu_4 = 3\mu_2{}^2 + [N\theta(1-\theta)\{1-6\theta+6\theta^2+2\alpha\theta(4-9\theta+4\theta^2)$$
$$+(\alpha\theta)^2(6-6\theta+\theta^2)\}]/(1-\alpha\theta)^7.$$

The moment relations of this distribution can be obtained by using the following equations

$$(2-\theta)^2/(1-\theta) = b = (3\mu_2{}^2 - \mu\mu_3')^2/\mu\,\mu_2{}^3$$

and $(\mu_4 - 3\mu_2{}^2)/\mu_2{}^2 = 15\mu_2/\mu^2 + (b+2)/\mu - 10(b\mu_2/\mu^3)^{1/2}.$

The derivation of moment relations for other distributions is straightforward. The p.m.f. or p.d.f., central or raw moments, central and raw moment relations for some important discrete and continuous distributions are given in Table 2.3A and Table 2.3B. To derive the differential equation for Generalized Poisson (Consul and Jain,1973) or the Borel Tanner distribution we use the differential equation $G\{GG''-(G')^2\} = G'(G+G')^2$

$$\text{or, } \{GG''-(G')^2\}/GG' = (1+G'/G)^2$$

where $G'=DG$, $G''=D^2G$ and $G=G(t)$ is its moment generating function. Here $G(t)=\{M(t)\}^{1/N}$, $\{GG''-(G')^2\}/GG' = \{M''M - (M')^2\}/MM'$, where $M = M(t)$ is the moment generating function for the Generalized Poisson distribution.

Differentiating the equation $(1+G'/G)^2 = \{M''M - (M')^2\}/MM'$ with respect to t and simplifying we get

$$9\{M''M - (M')^2\}^4+(M')^2\{M'''M^2-3MM'M''+2(M')^3\}^2-6M'\{M''M - (M')^2\}^2$$
$$\{M'''M^2-3MM'M''+2(M')^3\} - 4MM'\{M''M - (M')^2\}^3 = 0$$

Differential equations for other distributions are straightforward. Differential equations for zero regression that satisfies $E(Se^{t\Delta}) = 0$ for some discrete and continuous distributions are given in Table 2.4.A and 2.4.B. Differential equations in the moment generating function to identify some distributions are also given in the fourth column of Table 2.4.A. If we integrate these equations with respect to t we get the moment generating function for the corresponding distribution, while the differentiation gives further differential equations as listed in the third column of Table 2.4.A. Hence these differential equations also characterize a particular distribution.

2.4 Applications of the Moment Relations

The moment relations are used as criteria for discriminating one distribution from another. In particular consider the four distributions, the Negative Binomial, the Neyman type A, the Hermite and the Generalized Poisson. The Negative Binomial distribution results if the Poisson distribution is generalized by the Logarithmic distribution. The Neyman type A distribution is obtained if the Poisson distribution is generalized by another Poisson distribution. If X_1 and X_2 be two independent Poisson variables then the random variable $X_1 + 2X_2$ has the Hermite distribution. The Generalized Poisson distribution results if the Poisson distribution is generalized by the Borel distribution. The moments function $\mu\mu_3 - 2\mu_2^2 + \mu\mu_2$ is zero for the Negative Binomial distribution, negative for the Neyman type A and also for the Hermite distribution and positive for the Generalized Poisson distribution. The function $(\mu\mu_3 - \mu_2^2 - \mu\mu_2 + \mu^2)$ is zero for the Neyman type A, negative for the Hermite distribution and positive for the Generalized Poisson and Negative Binomial distributions. The moment relation for the Generalized Poisson distribution is $\{3/2 - \mu\mu_3/(2\mu_2^2)\}^2 = \mu/\mu_2$ which can be used in deciding which of the following kinds of distribution to use, if one of them is appropriate.

Let $a = \{3/2 - \mu\mu_3/(2\mu_2^2)\}^2$ and $b = \mu/\mu_2$. Then for the Poisson distribution $a = b = 1$, for the Generalized Poisson $0 < a = b < 1$, for the Neyman type A $49/64 < a > b < 1$, for the Negative Binomial $1/4 < a > b < 1$ and for the Generalized Negative Binomial $0 < b < a$.

In Table 2.1 we consider Bortkewitch's data on the "Number of deaths caused by horsekicks in the Prussian Army Corps". Fisher (1958) fitted the data with the Poisson distribution, Jain and Consul (1971) with the Generalized Negative Binomial (GNB) and Consul and Jain (1973) with the Generalized Poisson distribution. The calculated values of the Chi-square goodness of fit test statistics are insignificant in all cases. From the data we get, $m = 0.61$, $m_2 = 0.6109548$, $m_3 = 0.590562$, $m_4 = 1.643373$ and

(i) $m_2 - m = 0.0009548$ for Poisson,

(ii) $\{3/2 - m_3 m/(2m_2^2)\}^2 - m/m_2 = 0.03675632$ for Generalized Poisson and

(iii) $15m_2^4 + 2mm_2^3 + (mm_3 - 3m_2^2)^2 - m^2 m_2(m_4 - 3m_2^2) + 10(mm_3 - 3m_2^2)m_2^2 = -0.009132426$ for GNB distribution.

Here (i) is very close to zero compared to (ii) and (iii).

Here, $\hat{a} = \{3/2 - mm_3/(2m_2^2)\}^2 \approx 1$ and $\hat{b} = m/m_2 \approx 1$.

Thus $\hat{a} \approx \hat{b} \approx 1$.

Consider the moment ratios $d_1 = \mu\mu_3/\mu_2^2$ and

$$d_2 = \{\mu_2\mu_4 - \mu_3^2\} / \mu_2^3.$$

Then for the Poisson distribution $d_1 = 1$ and $d_2 = 3$, for the GNB $d_1 < 3$ and $d_2 \geq 1$, for the G. Poisson $1 \leq d_1 < 3$ and $d_2 \geq 3$.

Here $\hat{d}_1 \approx 1$ and $\hat{d}_2 \approx 3$.

Therefore, the distribution is more likely to be Poisson.

For assessing statistical accuracy we use the bootstrap method discussed by Efron (1982). This is a computer based method. Theoretically it is difficult to find the exact form of the distribution of a function of the sample moments. The bootstrap can routinely give us approximate distribution. There are two types of bootstrap: parametric and nonparametric. Here we used the nonparametric bootstrap method for Bortkewitch's data given in Table 2.1. First we drew a sample of size 200 from the given 200 observations with replacement and calculated the mean, the variance, the third, and the fourth central moments. Then we calculated the values of the moments functions for the Poisson, Generalized Poisson, and Generalized Negative Binomial distributions. We repeated this experiment first 1000 times, then 1500 times, and finally 2000 times in three stages and obtained three sets of 1000, 1500, and 2000 values for the three moments functions. Then we computed the mean and the standard deviation for each set and are given in table 2.2A. To do this we used a Fortran program given in Appendix1. Now we are interested in testing the following hypotheses:

(i) $H_{o1} : \mu_2 - \mu = 0$ against

$H_{a1} : \mu_2 - \mu \neq 0$

(ii) $H_{o2} : \{3/2 - \mu_3\mu/(2\mu_2^2)\}^2 - \mu/\mu_2 = 0$ against

$H_{a2} : \{3/2 - \mu_3\mu/(2\mu_2^2)\}^2 - \mu/\mu_2 \neq 0$

(iii) $H_{o3} : 15\mu_2^4 + 2\mu\mu_2^3 + 10(\mu\mu_3 - 3\mu_2^2)\mu_2^2 - \mu^2\mu_2(\mu_4 - 3\mu_2^2) + (\mu\mu_3 - 3\mu_2^2)^2 = 0$ against

$H_{a3} : 15\mu_2^4 + 2\mu\mu_2^3 + 10(\mu\mu_3 - 3\mu_2^2)\mu_2^2 - \mu^2\mu_2(\mu_4 - 3\mu_2^2) + (\mu\mu_3 - 3\mu_2^2)^2 \neq 0$

To test the above hypotheses we consider a test statistic which is of the following general form

$Z = [\,H(m, m_2, m_3, \ldots) - 0]/S.E.[H(m, m_2, m_3, \ldots)]$,

where $H(m, m_2, m_3, \ldots)$ is the mean of bootstrap sample.

Let us assume that in some neighbourhood of the point $m = \mu$, $m_i = \mu_i$, $(i=1,2, \ldots)$ the function H is continuous and has continuous derivatives of the first and second order with respect to the arguments m and m_i. According to the central limit theorem the test statistic Z follows the N(0,1) distribution at least approximately. The values of the test statistic, P-values, and the corresponding conclusion are given in Table 2.2B. In all cases the P-values are very large for the Poisson distribution and nearly zero for the Generalized Negative Binomial and Generalized Poisson distributions. Thus there is no evidence against the hypothesis that the data set is from a Poisson population.

For validation, we have considered a random sample of size 200 from a Poisson population having mean 0.61 and calculated the first four sample moments given by $m = 0.68$, $m_2 = 0.680856$, $m_3 = 0.714$, $m_4 = 2.223$. We have used these values in the moment functions for the Poisson, Generalized Poisson and Generalized Negative Binomial distributions and obtained the following results,

(a) $m_2 - m = 0.000856$ for Poisson,

(b) $\{3/2 - m_3 m/(2m_2^2)\}^2 - m/m_2 = -0.0455435$ for Generalized Poisson and

(c) $15m_2^4 + 2mm_2^3 + (mm_3 - 3m_2^2)^2 - m^2 m_2(m_4 - 3m_2^2) + 10(mm_3 - 3m_2^2)m_2^2 = 0.0135369$ for GNB distribution.

Evidently, (a) is close to zero compared to (b) and (c).

Also, $\hat{d}_1 = mm_3/m_2^2 \approx 1$ and $\hat{d}_2 = \{m_2 m_4 - m_3^2\}/m_2^3 \approx 3$.

Subsequently, we have applied nonparametric bootstrap method discussed earlier to the randomly chosen sample obtained from the Poisson population with mean 0.61 for testing the hypotheses (i), (ii) and (iii). The results are given in Table 2.2C and 2.2D. Thus there are similarities between the results obtained from a Poisson population and those from Bortkewitch's data.

Table 2.1 : DEATHS DUE TO HORSE-KICKS IN THE PRUSSIAN ARMY, BORTKEWITCH'S DATA.

Number of deaths	Observed frequency	Poisson	GNB	Generalized Poisson
0	109	108.67	109.12	108.72
1	65	66.29	65.27	66.22
2	22	20.22	20.74	20.22
3	3	4.11	4.27	4.12
4 or more	1	0.71	0.60	0.72
Total	200	200	200	200
χ^2		0.322	0.230	0.330

Source : Fisher(1958)

TABLE 2.2 A : SUMMARY OF THE OUTCOMES OBTAINED FROM THE FORTRAN PROGRAM (BASED ON BORTKEWITCH'S DATA)

Moment relation estimate for	Bootstrap samples	Mean	Standard deviation
Poisson	1000	0.00044	0.06169
G. Poisson	1000	0.05570	0.09769
GNB	1000	-0.06606	0.31427
Poisson	1500	-0.00001	0.06132
G. Poisson	1500	0.05522	0.10093
GNB	1500	-0.00635	0.03033
Poisson	2000	-0.00066	0.06042
G. Poisson	2000	0.05614	0.10022
GNB	2000	-0.00682	0.02954

G. = Generalized

GNB = Generalized Negative Binomial

TABLE 2.2 B : TEST RESULTS (BASED ON BORTKEWITCH'S DATA)

Hypotheses	Bootstrap samples	Values of the test statistic	P-value	Conclusion
H_{o1}	1000	0.225547	0.8216	Do not reject
H_{o2}	1000	18.030389	0	Reject
H_{o3}	1000	-6.647153	0	Reject
H_{o1}	1500	-0.006316	0.995	Do not reject
H_{o2}	1500	21.18955	0	Reject
H_{o3}	1500	-8.10862	0	Reject
H_{o1}	2000	-0.488515	0.6252	Do not reject
H_{o2}	2000	25.05145	0	Reject
H_{o3}	2000	-10.32497	0	Reject

TABLE 2.2 C : SUMMARY OF THE OUTCOMES OBTAINED FROM THE FORTRAN PROGRAM (BASED ON POISSON SAMPLE).

Moment relation estimate for	Bootstrap samples	Mean	Standard deviation
Poisson	1000	0.00079	0.05984
G. Poisson	1000	-0.72855	0.07686
GNB	1000	0.00491	0.03250
Poisson	1500	-0.00125	0.05949
G. Poisson	1500	-0.72225	0.07507
GNB	1500	-0.00707	0.02893
Poisson	2000	-0.00087	0.05889
G. Poisson	2000	-0.72664	0.07472
GNB	2000	-0.00724	0.02885

TABLE 2.2 D : TEST RESULTS (BASED ON POISSON SAMPLE).

Hypotheses	Bootstrap samples	Values of the test statistic	P-value	Conclusion
H_{o1}	1000	0.41748	0.6764	Do not reject
H_{o2}	1000	-299.74986	0	Reject
H_{o3}	1000	4.77747	0	Reject
H_{o1}	1500	-0.813789	0.4158	Do not reject
H_{o2}	1500	-372.620517	0	Reject
H_{o3}	1500	-9.464913	0	Reject
H_{o1}	2000	-0.66068	0.5088	Do not reject
H_{o2}	2000	-434.908	0	Reject
H_{o3}	2000	-11.22297	0	Reject

TABLE 2.3.A MOMENT RELATIONS CHARACTERIZING CERTAIN DISCRETE DISTRIBUTIONS.

Distribution with prob. mass function	Central or raw moments	Central moment relations	Raw moment relations
Geometric $(1-\theta)\theta^x$	$\mu = \theta/(1-\theta)$, $\mu_2 = \theta/(1-\theta)^2$	$\mu_2 - \mu^2 - \mu = 0$	$\mu_2' - 2\mu^2 - \mu = 0$
Poisson $e^{-\theta}\theta^x/x!$	$\mu = \theta = \mu_2$	$\mu_2 - \mu = 0$	$\mu_2' - \mu^2 - \mu = 0$
Negative Binomial $\Gamma(N+x)\theta^x(1-\theta)^N/(x!\Gamma N)$	$\mu = N\theta/(1-\theta)$, $\mu_2 = N\theta/(1-\theta)^2$, $\mu_3 = N\theta(1+\theta)/(1-\theta)^3$	$\mu\mu_3 - 2\mu_2^2 + \mu\mu_2 = 0$	$\mu\mu_3' - 2(\mu_2')^2 + \mu\mu_2' - (\mu)^3 + \mu^2\mu_2' = 0$
Borel $\theta^x(1+x)^{x-1}e^{-\theta(1+x)}/x!$	$\mu = \theta/(1-\theta)$, $\mu_2 = \theta/(1-\theta)^3$,	$\mu_2 - \mu(1+\mu)^2 = 0$	$\mu_2' - \mu - 3\mu^2 - \mu^3 = 0$
Borel Tanner or Generalized Poisson $\theta^x N(N+x)^{x-1}e^{-\theta(N+x)}/x!$	$\mu = N\theta/(1-\theta)$, $\mu_2 = N\theta/(1-\theta)^3$, $\mu_3 = N\theta(1+2\theta)/(1-\theta)^5$	$\{3/2 - \mu_3\mu/(2\mu_2^2)\}^2 - \mu/\mu_2 = 0$	$9(\mu_2')^4 - 18\mu^2(\mu_2')^2 - 21\mu^4(\mu_2')^2 + \mu^8 - 6\mu^6\mu_2' + \mu^2(\mu_3')^2 + 6\mu^3\mu_2'\mu_3' - 2\mu^5\mu_3' - 6\mu(\mu_2')^2\mu_3' - 4\mu(\mu_2')^3 + 12\mu^3(\mu_2')^2 - 12\mu^5\mu_2' + 4\mu^7 = 0$
Hermite $\sum_{j=0}^{[x/2]} \frac{\theta_1^{x-2j}\theta_2^{j}e^{-\theta_1-\theta_2}}{j!\,(x-2j)!}$	$\mu = \theta_1 + 2\theta_2$, $\mu_2 = \theta_1 + 4\theta_2$, $\mu_3 = \theta_1 + 8\theta_2$	$\mu_3 - 3\mu_2 + 2\mu = 0$	$\mu_3' - 3\mu\mu_2' + 2\mu^3 - 3\mu_2' + 3\mu^2 + 2\mu = 0$

$x = 0,1,2,\ldots$ for all cases.

Distribution with mass function	Central or Raw moments	Central moment relations	Raw moment relations
Logarithmic Series	$\mu = c\theta/(1-\theta)$,	$\mu\mu_3 - \mu^2\mu_2 - \mu^4$	$\mu\mu_3' + \mu\mu_2' - 2(\mu_2')^2 = 0$
$c\theta^x/x$, $x=1,2,3,\ldots$	$\mu_2' = \mu/(1-\theta)$,	$-2(\mu_2)^2 + \mu\mu_2 + \mu^3 = 0$	
$c = [-\ln(1-\theta)]^{-1}$	$\mu_3' = \mu(1+\theta)/(1-\theta)^2$		
Neyman type A	$\mu = \lambda\theta$,	$\mu\mu_3 - \mu\mu_2 - (\mu_2)^2$	$\mu\mu_3' - \mu^2\mu_2' + \mu^4 + \mu^3$
$\exp[-\lambda(1-e^{-\theta})]$, $x=0$	$\mu_2 = \lambda\theta(1+\theta)$,	$+ \mu^2 = 0$	$-(\mu_2')^2 - \mu\mu_2' + \mu^2 = 0$
$\sum_{j=1}^{x} e^{-\lambda}\left(\frac{\lambda^j}{j!}\right) e^{-j\theta}\frac{(j\theta)^x}{x!}$,	$\mu_3 = \lambda\theta(1+3\theta+\theta^2)$		
Generalized Geometric	$\mu = \theta/(1-\alpha\theta)$,	$4\mu\mu_2^3 - 9\mu_2^4 + \mu^4\mu_2^2$	$4\mu(\mu_2')^3 - 12\mu^3(\mu_2')^2$
$\frac{\Gamma(1+\alpha x)\,\theta^x(1-\theta)^{1+\alpha x-x}}{x!\,\Gamma(\alpha x-x+2)}$	$\mu_2 = \mu^3(1-\theta)/\theta^2$,	$+6\mu\mu_2^2\mu_3 - \mu^2\mu_3^2$	$+24\mu^5\mu_2' - 9(\mu_2')^4 + 18\mu^2(\mu_2')^3$
$x = 0,1,2,\ldots$	$\mu_3 = \mu^2\mu_2[3(1-\theta) -$	$= 0$	$-14\mu^4(\mu_2')^2 - 18\mu^6\mu_2' + 4\mu^8$
	$(1-\alpha\theta)(2-\theta)]/\theta^2$		$-\mu^2(\mu_3')^2 - 4\mu^6 - 6\mu^3\mu_2'\mu_3' - 4\mu^7$
			$-4\mu^4\mu_3' + 6\mu(\mu_2')^2\mu_3' = 0$
Generalized	$\mu = N\theta/(1-\alpha\theta)$,	$15\mu_2^4 + 2\mu\mu_2^3$	$6[\mu^5\mu_2' - (\mu_2')^4 - \mu^3(\mu_2')^2]$
Negative Binomial	$\mu_2 = N\theta(1-\theta)/(1-\alpha\theta)^3$,	$+10(\mu\mu_3 - 3\mu_2^2)\mu_2^2$	$+2[\mu^6\mu_2' + \mu(\mu_2')^3 - \mu^7]$
$\frac{N\,\Gamma(N+\alpha x)\,\theta^x(1-\theta)^{N+\alpha x-x}}{x!\,\Gamma(N+\alpha x-x+1)}$	$\mu_3 = N\theta(1-\theta)\{1-2\theta+$	$+(\mu\mu_3 - 3\mu_2^2)^2$	$+15[\mu^2(\mu_2')^3 - \mu^4(\mu_2')^2] + \mu^4\mu_4'$
$x = 0,1,2,\ldots$	$+\alpha\theta(2-\theta)\}/(1-\alpha\theta)^5$	$-\mu^2\mu_2(\mu_4 - 3\mu_2^2)$	$+4[\mu^5\mu_3' + \mu(\mu_2')^3\mu_3'] + \mu^2\mu_3'$
	$\mu_4 = 3\mu_2^2 + [N\theta(1-\theta)\{1-6\theta$	$= 0$	$-10\mu^3\mu_2'\mu_3' - \mu^2\mu_2'\mu_4' = 0$
	$+6\theta^2 + 2\alpha\theta(4-9\theta+4\theta^2)+$		
	$\alpha^2\theta^2(6-6\theta+\theta^2)\}]/(1-\alpha\theta)^7$		

TABLE 2.3.B MOMENT RELATIONS FOR SOME CONTINUOUS DISTRIBUTIONS.

Distribution with prob. density function	Central or raw moments	Central moment relations	Raw moment relations
Gaussian or Normal $\exp\{-(x-\theta)^2/2\sigma^2\}/(\sigma\sqrt{2\pi})$, $-\infty < x < \infty$	$\mu = \theta$, $\mu_2 = \sigma^2$, $\mu_3 = 0$, $\mu_4 = 3\sigma^4$	$\mu_4 - 3\mu_2^2 = 0$	$\mu_4' - 4\mu\mu_3' + 12\mu^2\mu_2' - 3(\mu_2')^2$ $- 6\mu^4 = 0$
Inverse Gaussian $\sqrt{(\lambda/2\pi x^3)}[\exp\{-\lambda(x-\theta)^2$ $/(2x\theta^2)\}]$, $x > 0$.	$\mu = \theta$, $\mu_2 = \theta^3/\lambda$ $\mu_3 = 3\mu_2^2/\mu$	$\mu\mu_3 - 3\mu_2^2 = 0$	$\mu\mu_3' + 3\mu^2\mu_2' - \mu^4 - 3(\mu_2')^2 = 0$
Gamma $\theta^\alpha x^{\alpha-1}e^{-\theta x}/\Gamma\alpha$, $x > 0$.	$\mu = \alpha/\theta$, $\mu_2' = \alpha(\alpha+1)/\theta^2$, $\mu_3' = \alpha(\alpha+1)(\alpha+2)/\theta^3$	$\mu\mu_3 - 2\mu_2^2 = 0$	$\mu\mu_3' + \mu^2\mu_2' - 2(\mu_2')^2 = 0$
Exponential $(1/\theta)\exp(-x/\theta)$, $x>0$.	$\mu = \theta$, $\mu_2 = \theta^2$	$\mu_2 - \mu^2 = 0$	$\mu_2' - 2\mu^2 = 0$
Maxwell $\sqrt{(2\theta^3/\pi)}\, x^2\exp(-\theta x^2/2)$,	$\mu = 2\sqrt{(2/\pi\theta)}$, $\mu_2 = (3-8/\pi)/\theta$	$\mu_2 - (3\pi/8 - 1)\mu^2 = 0$	$\mu_2' - (3\pi/8)\mu^2 = 0$
Chi-square $e^{-x/2}(2x)^{\nu/2}/x\Gamma(\nu/2)$,	$\mu = \nu$, $\mu_2 = 2\nu$	$\mu_2 - 2\mu = 0$	$\mu_2' - \mu^2 - 2\mu = 0$
Laplace $\exp(-\|x-\theta\|/\alpha)/2\alpha$ $-\infty < x < \infty$.	$\mu = \theta$ $\mu_2 = 2\alpha^2$, $\mu_3 = 0$, $\mu_4 = 24\alpha^4$	$\mu_4 - 6\mu_2^2 = 0$	$\mu_4' - 4\mu\mu_3' + 18\mu^2\mu_2' - 6\mu_2'^2$ $- 9\mu^4 = 0$

TABLE 2.4 .A DIFFERENTIAL EQUATIONS IN MOMENT GENERATING FUNCTIONS FOR ZERO REGRESSION AND IDENTIFYING CERTAIN DISTRIBUTIONS.

Distribution	Moment generating functions, $M=M(t)$	Differential equations satisfying $E(Se^{\Delta t}) = 0$	Differential equations identifying distributions.
Geometric	$(1-\theta)/(1-\theta e^{t})$	$MM''-2(M')^2-MM' = 0$	$M/(1+D\ln M) = c,\ 0< c <1.$
Poisson	$\exp(-\theta+\theta e^{t})$	$MM''-(M')^2-MM' = 0$	$D\ln M - \ln M = c,\ \ c > 0.$
Negative Binomial	$(1-\theta)^N/(1-\theta e^{t})^N$	$MM'M'''+M''(M')^2-2M(M'')^2 - (M')^3+MM'M'' = 0$	$(D\ln M)^2/(D^2\ln M-D\ln M) = c,\ c > 0.$
Borel	$\exp\{\theta(Me^{t}-1)\}$	$M^2M''-M(M')^2-M'(M'+M)^2 = 0$	$D\ln M/(1+D\ln M)-\ln M= c,\ 0< c<1$
Borel Tanner	$\exp\{N\theta M^{1/N}e^{t}-N\theta\}$	$9(MM''-M'^2)^4+(M^2M'''+2M'^3 -3MM'M'')\{M'^2-6M'(MM''-M'^2)^2\} - 4MM'(MM''-M'^2)^3 = 0$	$G(GG''-G'^2) = G'(G+G')^2,$ where $G = M^{1/N}.$
Hermite	$\exp\{-\theta_1(1-e^{t})-\theta_2(1-e^{2t})\}$	$M^2(M'''-3M''+2M')+2M'^3 -3MM'(M''-M') = 0$	$D^2\ln M-3D\ln M+2\ln M = c,\ c > 0.$
Logarithmic	$\ln(1-\theta e^{t})/\ln(1-\theta)$	$M'M'''+M'M''-2(M'')^2 = 0$	$(DM)^2/(D^2M-DM) = c,\ c > 0.$
Neyman type A	$\exp[-\lambda\{1-\exp -\theta(1-e^{t})\}]$	$M^2M'(M'''-M''+M')-M^2M''^2 -MM''M'^2+M'^4+MM'^3 = 0$	$(D\ln M)^2/(D^2\ln M-D\ln M) - \ln M = c,\ c > 0.$

TABLE 2.4 .B DIFFERENTIAL EQUATIONS IN MOMENT GENERATING FUNCTIONS FOR ZERO REGRESSION .

Distribution	Moment generating functions	Differential equations satisfying $E(Se^{\Delta t})=0$
Normal	$\exp(\theta t + t^2\sigma^2/2)$	$M''''M^3-4M'M'''M^2+12MM''M'^2-3M^2(M'')^2-6(M')^4=0$
Inverse Gaussian	$\exp\{\lambda[1-\sqrt{(1+2\theta^2 t/\lambda)}]/\theta\}$	$M^2M'M''' + 3MM''M'^2 - (M')^4 - 3M^2(M'')^2 = 0$
Gamma	$(1 - t/\theta)^{-\alpha}$	$MM'M''' + M''(M')^2 - 2M(M'')^2 = 0$
Exponential	$(1 - t\theta)^{-1}$	$MM'' - 2(M')^2 = 0$
Maxwell	$[1+2t^2\{1-\varphi(-t/\sqrt{\theta})\}/\theta]\exp(-t^2/2\theta)$ $+ t\sqrt{(8/\theta\pi)}$	$MM'' - (3\pi/8)(M')^2 = 0$
Chi-square	$(1-2t)^{-v/2}$	$MM'' - M'^2 - 2MM' = 0$
Laplace	$e^{\theta t}(1-\alpha^2t^2)^{-1}$	$M^2\{MM''''-4M'M'''-6(M'')^2\})$ $+ 9M'^2(2MM''-M'^2) = 0$

TABLE 2.5.A SAMPLE STATISTICS S WITH ZERO REGRESSION ON $\Delta = X_1 + X_2 + \ldots + X_n$

Distribution	S Statistic
Geometric	$(n-1)\Sigma X_i (X_i - 1) - 2\Sigma X_i X_j$
Poisson	$(n-1)\Sigma X_i (X_i - 1) - \Sigma X_i X_j$
Negative Binomial	$(n-2)\Sigma X_i X_j^2 (X_j - 2X_i + 1) + \Sigma X_i X_j X_k (X_k - 1)$
Borel	$(n-1)(n-2)\Sigma X_i (X_i - 1) - 3(n-2)\Sigma X_i X_j - \Sigma X_i X_j X_k$
Borel Tanner	$(n-4)(n-5)(n-6)(n-7)\Sigma X_i X_j X_k^2 X_l^2 (9X_i X_j - 18X_k X_l - 6X_j X_l - 4X_j) + 6(n-5)$ $(n-6)(n-7)\Sigma X_i X_j X_k X_l^2 X_m^2 (X_m + 2) - (n-6)(n-7)\Sigma X_i X_j X_k X_l X_m X_p^2 (21X_m +$ $2X_p + 12) - 2(n-7)\Sigma X_i X_j X_k X_l X_m X_p X_o (3X_o - 2) + \Sigma X_i X_j X_k X_l X_m X_p X_o X_q$
Hermite	$(n-1)(n-2)\Sigma X_i (X_i^2 - 3X_i + 2) - 3(n-2)\Sigma X_i X_j (X_j - 1) + 2\Sigma X_i X_j X_k$
Logarithmic Series	$\Sigma X_i X_j^2 (X_j - 2X_i + 1)$
Neyman Type A	$(n-2)(n-3)\Sigma X_i X_j (X_j^2 - X_i X_j - X_j + 1) - (n-3)\Sigma X_i X_j X_k (X_k - 1) + \Sigma X_i X_j X_k X_l$
Generalized Geometric	$(n-4)(n-5)(n-6)(n-7)\Sigma X_i X_j X_k^2 X_l^2 (4X_j - 9X_i + 6X_j X_l - X_k X_l) - 2(n-5)(n-6)(n-7)$ $\Sigma X_i X_j X_k X_l X_m^2 (6X_l - 9X_k + 3X_l X_m + 2X_m) - 2(n-6)(n-7) \Sigma X_i X_j X_k X_l X_m X_p (7X_m X_p$ $- 12X_p + 2 - 3X_p^2) - 4(n-7)\Sigma X_i X_j X_k X_l X_m X_p X_o (2X_o + 1) + 4\Sigma X_i X_j X_k X_l X_m X_p X_o X_q$
Generalized Negative Binomial	$(n-4)(n-5)(n-6)\Sigma X_i X_j X_k^2 X_l^2 (2X_j - 6X_i X_j - X_l^2 + X_k X_l) + (n-5)(n-6)\Sigma X_i X_j X_k X_l$ $X_m^2 (15X_k X_l - 6X_l - 10X_l X_m + 4X_j X_k X_l X_m + X_m^2) + 2\Sigma X_i X_j X_k X_l X_m X_p X_o (X_o - 1)$ $+ (n-6)\Sigma X_i X_j X_k X_l X_m X_p^2 (6 - 15X_m + 4X_p)$

TABLE 2.5.B SAMPLE STATISTICS S WITH ZERO REGRESSION ON $\Delta = X_1 + X_2 + \ldots + X_n$

Distribution	S Statistic
Normal or Gaussian	$(n-2)(n-3)\Sigma X_i^2 \{(n-1)X_i^2 - 4X_i X_j - 3X_j^2\} + 6\Sigma X_i X_j X_k\{2(n-3)X_i - X_l\}$
Inverse Gaussian	$(n-2)(n-3)\Sigma X_i X_j^2 (X_j - 3X_i) + (n-3)\Sigma X_i X_j X_k^2 - \Sigma X_i X_j X_k X_l$
Gamma	$(n-2)\Sigma X_i X_j^2 (X_j - 2X_i) + \Sigma X_i X_j X_k^2$
Exponential	$(n-1)\Sigma X_i^2 - 2\Sigma X_i X_j$
Maxwell	$(n-1)\Sigma X_i^2 - (3\pi/8)\Sigma X_i X_j$
Chi-square	$(n-1)\Sigma X_i (X_i - 2) - \Sigma X_i X_j$
Laplace	$(n-2)(n-3)\Sigma X_i^2 \{(n-1)X_i^2 - 4X_i X_j - 6X_j^2\} + 9\Sigma X_i X_j X_k\{2(n-3)X_i - X_l\}$

Chapter 3

Identification by Moment Ratios

3.1. Introduction

In this chapter we have proposed two moment ratios. These ratios are useful in identifying different members of a class of discrete or continuous distributions. These ratios are also useful in approximating the Neyman type A and the Generalized Poisson distributions by the Negative Binomial distribution. The impact of using approximate distributions instead of the exact distributions is studied.

We know that the values of the pair (β_1 , β_2) of moment ratios $\beta_1=\mu_3^2/\mu_2^3$ and $\beta_2=\mu_4/\mu_2^2$, where μ_i is the ith central moment of the distribution, are (0,3) for the Normal distribution. Hence a comparison of the point (β_1, β_2) of any given distribution with (0,3) will give an idea about the departure from the shape of the Normal distribution. For discrete distributions Jain and Gupta (1980) have defined the moment ratios

$b_1 = (\mu\mu_3 - \mu_2^2)/\mu_2^2$

and

$b_2 = (\mu\mu_3 - 3\mu_2^2)^2/\mu\,\mu_2^3 = \mu_2(b_1 - 2)^2/\mu,$

where μ is the mean. The value of the point (b_1, b_2) is (0,4) for the Poisson distribution. Therefore, it provides useful information on how similar a discrete distribution is to the Poisson distribution. They have also given approximations of the Generalized Poisson (Borel-Tanner) and Neyman type A distributions by the Negative Binomial distribution by equating the μ and b_1 values of these distributions. In this chapter we introduce moment ratios

$$d_1 = \mu\mu_3/\mu_2^2 \quad \text{and}$$

$$d_2 = (\mu_2\mu_4 - \mu_3^2)/\mu_2^3$$

based on the first four moments. The value of this pair (d_1, d_2) is (1, 3) for the Poisson distribution. The p.m.f. of the Generalized Poisson distribution is

$$P(Y=j) = \lambda(\lambda+\theta j)^{j-1}e^{-(\lambda+\theta j)}/j! \quad .$$

It is shown in Section 3.2 that the use of (μ, d_2) provides better approximations of Generalized Poisson and Neyman type A distributions by the Negative Binomial distribution. Moreover, by the use of (μ, b_1) the approximations for the Generalized Poisson distribution are valid only for $\theta < 0.5$, but the use of (μ, d_2) gives quite satisfactory approximations even for $\theta < 0.8$. The approximations are, however, not satisfactory for the following cases :

(i) when $\theta > 0.2$, probability sums for $r=0$;

(ii) when $\theta > 0.4$, probability sums for $r<10$;

(iii) when $\theta \geq 0.7$ and $\lambda > 2$;

(iv) when $\theta = 0.8$ and $\lambda > 1$;

(v) when $\theta > 0.8$.

Taking random samples from Generalized Poisson, Neyman type A and their approximate distributions [say, Modified Generalized Poisson and Modified Neyman type A], it is shown in Section 3.3 that the estimator of the population mean remains unbiased in each case. But the variance is overestimated in case of Generalized Poisson and underestimated in case of Neyman type A if we use the Modified distributions instead of the exact distributions. We know from the discussion in chapter 2 that moment relations can also be used to identify a distribution. But in some cases moment relations of two or more distributions are same. In such a situation the moment ratios (d_1,d_2) can be used to identify a distribution. Because of the sampling fluctuations, a particular criterion may not provide reliable information to draw sound conclusions. So, it is better to use moment ratios with moment relations to identify a distribution.

3.2 Moment Ratios and their Uses

Let the moment ratios based on the first four moments be defined by $d_1 = \mu\mu_3/\mu_2^2$ and $d_2 = (\mu_2\mu_4 - \mu_3^2)/\mu_2^3$. The exact expressions of the moment ratios d_1 and d_2 as a function of parameters for some discrete and continuous distributions are given in Table 3.1 and Table 3.2 respectively. The ranges of d_1 and d_2 for all distributions in Table 3.1 and Table 3.2 are given in Table 3.3. We know that the Generalized Negative Binomial (GNB) distribution is the generalization of the Binomial, Negative Binomial, Geometric and the Generalized Geometric distributions. Let X be a GNB variable having probability mass function

$N\Gamma(N + \alpha x)\ \theta^{x}(1-\theta)^{N+\alpha x - x}/\{x!\Gamma(N + \alpha x - x + 1)\}$.

If $\alpha = 0$, the distribution of X is Binomial, if $\alpha = 1$, the distribution of X is Negative Binomial, if $N = 1$, the distribution of X is Generalized Geometric and if $N = 1$ and $\alpha = 1$, the distribution of X is Geometric. For the Binomial distribution $d_1 < 1$ and $1 \le d_2 < 3$, for the Negative Binomial distribution $1 < d_1 < 2$ and $3 < d_2 \le 5$, for the Generalized Geometric distribution $d_1 < 3$ and $d_2 \ge 5$, for the Geometric distribution $1 < d_1 < 2$ and $d_2 = 5$ and for the GNB distribution $d_1 < 3$ and $d_2 \ge 1$. The Gamma distribution is the general form of Chi-square and Exponential distribution. Let X be a Gamma variable having density $\theta^{p}x^{p-1}e^{-\theta x}/\Gamma(p)$. If $p = \nu/2$ and $\theta = 1/2$, the distribution of X is Central Chi-square with ν degrees of freedom and if $p = 1$, the distribution of X is Exponential with mean $1/\theta$. For the Chi-square distribution $d_1 = 2$ and $3 < d_2 < 7$, for the Exponential distribution $d_1 = 2$ and $d_2 = 5$ and for the Gamma distribution $d_1 = 2$ and $d_2 > 3$. Let X be a Weibull variable having density $\theta p x^{p-1}\exp(-\theta x^{p})$. If $p = 1$, the distribution of X is Exponential and if $p = 2$, the distribution of X is Rayleigh. For the Rayleigh distribution $d_1 = 2.2074$ and $d_2 = 2.8468$, for the Exponential distribution $d_1 = 2$ and $d_2 = 5$ and for the Weibull distribution $d_1 > 1$ and $d_2 > 1$.

The Generalized Hermite distribution is the generalization of the Hermite and Poisson distributions. The p.m.f. of the Generalized

Hermite distribution is $e^{-(\alpha+\beta)} H_{j,\theta}(\alpha, \beta)/j!$, j=0,1,2,....

$$\text{where, } H_{j,\theta}(\alpha,\beta) = \sum_{k=0}^{[j/\theta]} \beta^k \alpha^{j-\theta k} j! /\{k! (j-\theta k)!\}$$

If $\theta = 2$, it reduces to the Hermite distribution and if $\theta = 2$ and $\beta=0$ it reduces to the Poisson distribution. For the Poisson distribution $d_1 = 1$ and $d_2 = 3$; for the Hermite distribution $1 \leq d_1 < 2$ and $3 \leq d_2 < 4$; for the Generalized Hermite distribution $1 \leq d_1 < \infty$ and $3 \leq d_2 < \infty$. Therefore, the values of d_1 and d_2 could be used to identify a distribution.

The Negative Binomial, Neyman type A and the Generalized Poisson distributions are contagious. Tables are available for computing c.d.f. of the Negative Binomial distribution only. One can also compute the c.d.f. of the Negative Binomial distribution by using the Binomial probability Tables and the incomplete Beta function Tables.

In Table 3.4 we approximate the Neyman type A probabilities $\Sigma f_N(j;\lambda,\theta)$ by the Negative Binomial probabilities $\Sigma f_{NB}(j;N,\alpha)$ by equating the means and the d_2 values of these distributions with

$\alpha = \lambda\theta/ (\lambda\theta + N)$, $N = 2\lambda(1+\theta)^3/(2+2\theta + \theta^2)$

and $\Sigma f_N(j; \lambda,\theta) = \Sigma f_{NB}(j; N,\alpha)$.

In parentheses we consider the corresponding values taken from Jain and Gupta (1980) based on the first three moments. The probability values based on the first four moments seem to provide better approximations.

Similarly by equating the means and the d_2 values of the Generalized Poisson and the Negative Binomial distributions in Table 3.5 we have approximated the Generalized Poisson sums by the Negative Binomial sums with $\alpha = \theta(\theta+2)/(1+2\theta^2)$,
$N = \lambda(1-\theta)/\{\theta(\theta+2)\}$ and $\Sigma f_{GP}(j;\ \lambda,\theta) = \Sigma f_{NB}(j;\ N,\alpha)$.

We know from the discussion in chapter 2 that the moment relations can be used to identify a distribution. The moments function $\mu\mu_3 - 2\mu_2^2 + \mu\mu_2$ is zero for the Binomial, Poisson and the Negative Binomial distributions. But for the Binomial distribution $d_1 < 1$ and $1 \le d_2 < 2$; for the Negative Binomial distribution $1 < d_1 < 2$ and $3 < d_2 \le 5$ and for the Poisson distribution $d_1 = 1$ and $d_2 = 3$. The function $\mu_3 - 3\mu_2 + 2\mu$ is zero for the Hermite and the Poisson distributions. But for the Hermite distribution $1 \le d_1 < 2$ and $3 \le d_2 < 4$; for the Poisson distribution $d_1 = 1$ and $d_2 = 3$. The function $\mu\mu_3 - \mu_2^2 - \mu\mu_2 + \mu^2$ is zero for the Poisson and the Neyman type A distributions. But for the Poisson distribution $d_1 = 1$ and $d_2=3$; for the Neyman type A distribution, $1 < d_1 < 1.25$ and $3 < d_2 < \infty$.

3.3 Comparison between Exact and Approximate Distributions

3.3.1. Generalized Poisson and its Approximate Distributions

The p.m.f. of the Generalized Poisson distribution is
$P(Y=j) = f_{GP}(j;\ \lambda,\theta) = \lambda(\lambda+\theta j)^{j-1} e^{-(\lambda+\theta j)}/j!$
and that of its approximate distribution is

$P(X=j) = f_{NB}(j;\ N,\alpha) = \Gamma(N+j)\alpha^j(1-\alpha)^N/(j!\ \Gamma N)$,

where $N = \lambda(1-\theta)/\{\theta(\theta+2)\}$, $\alpha = \theta(\theta+2)/(1+2\theta^2)$ and $j = 0,1,\ldots.$

Let the name of this approximate distribution be Modified Generalized Poisson. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from Generalized Poisson distribution and $\bar{Y}=\Sigma Y_i/n$ be the sample mean. Then $E(\bar{Y}) = \lambda/(1-\theta) = \mu$ (say), where μ is the population mean and $V(\bar{Y}) = \lambda/\{n(1-\theta)^3\}$. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Modified Generalized Poisson and $\bar{X} = \Sigma X_i/n$ be the sample mean. Then $E(\bar{X}) = \lambda/(1-\theta) = \mu$ and $V(\bar{X}) = \lambda(1+2\theta^2)/\{n(1-\theta)^3\}$.

Here, $V(\bar{Y}) - V(\bar{X}) = -2\lambda^2/\{n(1-\theta)^3\}$, which is negative.

Thus $V(\bar{Y}) < V(\bar{X})$.

This implies that if we use Modified Generalized Poisson distribution instead of Generalized Poisson distribution then the estimator of the population mean (μ) will remain unbiased but the variance will be overestimated.

3.3.2. Neyman Type A and its Approximate Distributions

The p.m.f. of a Neyman type A distribution is

$P(Y=j) = f_N(j;\ \lambda,\ \theta) = \Sigma e^{-(\lambda+k\theta)}\ \lambda^k(k\theta)^j/(j!\ k!)$ and that of its approximate distribution is

$P(X=j) = f_{NB}(j;\ N,\alpha) = \Gamma(N+j)\ \alpha^j(1-\alpha)^N/(j!\ \Gamma N)$,

where, $N = 2\lambda(1+\theta)^3/(2+2\theta+\theta^2)$, $\alpha = \lambda\theta/(\lambda\theta+N)$ and $j = 0,1,2,\ldots$

Let the name of this approximate distribution be Modified Neyman type A distribution. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a Neyman type A distribution and $\bar{Y} = \Sigma Y_i/n$ be the sample mean. Then $E(\bar{Y}) = \lambda\theta = \mu_1'$ (say), where μ_1' is the population mean and $V(\bar{Y}) = \lambda\theta(1+\theta)/n$. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Modified Neyman type A and $\bar{X} = \Sigma X_i/n$ be the sample mean.

Then $E(\bar{X}) = \lambda\theta = \mu_1'$ and $V(\bar{X}) = \lambda\theta^2(2+2\theta+\theta^2)/\{2n(1+\theta)^3\} + \lambda\theta/n$.

Here, $V(\bar{Y}) - V(\bar{X}) = \lambda\theta^3(4 + 5\theta + 2\theta^2)/\{2n(1+\theta)^3\}$, which is positive.

So $V(\bar{Y}) > V(\bar{X})$.

This implies that if we use Modified Neyman type-A distribution instead of Neyman type-A distribution then the estimator of the population mean will remain unbiased but the variance will be underestimated.

TABLE 3.1: d_1 AND d_2 FUNCTIONS FOR SOME DISCRETE DISTRIBUTIONS.

Name of distribution with p.m.f.	d_1	d_2
Binomial ${}^{N}C_j\,(\theta)^j(1-\theta)^{N-j}$, $0<\theta<1$	$(1-2\theta)/(1-\theta)$	$3-2/N$
Negative Binomial $\Gamma(N+j)\ (\alpha)^j(1-\alpha)^N/(j!\ \Gamma N)$	$1+\alpha$	$3+2/N$
Generalized Negative Binomial $N\Gamma(N+\alpha j)\theta^j(1-\theta)^{N+\alpha j-j}/\{j!\ \Gamma(N+\alpha j-j+1)\}$	$1+\theta\alpha+\theta(\alpha-1)/(1-\theta)$	$3+2\{(2-2\theta+\alpha\theta)\alpha-1\}/\{N(1-\alpha\theta)\}$
Geometric $(1-\theta)\theta^j$, $j=0,1,2,\ldots\ldots$	$1+\theta$	5
Generalized Geometric $\Gamma(1+\alpha j)\theta^j(1-\theta)^{1+\alpha j-j}/\{j!\ \Gamma(\alpha j-j+2)\}$	$1+\theta\alpha+\theta(\alpha-1)/(1-\theta)$	$3+2\{(2-2\theta+\alpha\theta)\alpha-1\}/(1-\alpha\theta)$
Neyman type A $\exp\{-\lambda(1-e^{-\theta})\}$ for $j=0$ $\sum_{k=0}^{j} e^{-(\lambda+k\theta)}\lambda^k\,(k\theta)^j/(j!\ k!)$, $j=1,2,3,\ldots\ldots$	$1+\theta/(1+\theta)^2$	$3+(2+2\theta+\theta^2)/\lambda(1+\theta)^3$
Borel $e^{-\theta(1+j)}(1+j)^{j-1}\theta^j/j!$,$j=0,1,2,\ldots$	$1+2\theta$	$1+6/(1-\theta)$
Poisson $e^{-\lambda}(\lambda)^j/j!$, $0<\lambda<1$	1	3
Generalized Poisson $\lambda e^{-(\lambda+\theta j)}(\lambda+\theta j)^{j-1}/j!$,	$1+2\theta$	$3+2\theta(\theta+2)/\{\lambda(1-\theta)\}$
Hermite $\sum_{k=0}^{[j/2]}\theta_1^{j-2k}\theta_2^j e^{-\theta_1-\theta_2}/\{k!\ (j-2k)!\}$, $j=0,1,2,\ldots\ldots$	$1+2\theta_1\theta_2/(\theta_1+4\theta_2)^2$	$3+4\theta_1\theta_2/(\theta_1+4\theta_2)^3$
Generalized Hermite $e^{-(\alpha+\beta)}H_{j,m}(\alpha,\beta)/j!$, where, $H_{j,m}(\alpha,\beta)=\sum_{k=0}^{[j/m]}\beta^k\alpha^{j-mk}j!/\{k!\ (j-mk)!\}$ $j=0,1,2,\ldots.$	$1+m(m-1)^2\alpha\beta/(\alpha+m^2\beta)^2$	$3+m^2(m-1)^2\alpha\beta/(\alpha+m^2\beta)^3$

TABLE 3.2 : d_1 AND d_2 FUNCTIONS FOR SOME CONTINUOUS DISTRIBUTIONS.

Name of distribution with p.d.f.	d_1	d_2
Normal $\exp\{-(x-\theta)^2/2\sigma^2\}/(\sigma\sqrt{2\pi})$	0	3
Student's t $1/\{\sqrt{p}\ \beta(p/2,1/2)(1+t^2/p)^{(p+1)/2}\}$	0	$3 + 6/(p-4)$
Gamma $(\theta)^p x^{p-1} e^{-x\theta}/\Gamma p$	2	$3 + 2/p$
Exponential $\theta e^{-x\theta}$	2	5
Chi-square $2^{-\nu/2} e^{-x/2} x^{\nu/2-1}/\Gamma(\nu/2)$	2	$3 + 4/\nu$
Rayleigh $2\theta x \exp(-\theta x^2)$	$2\pi(\pi-3)/(4-\pi)^2$	$(128 - 50\pi + \pi^3)/(4-\pi)^3$
Weibull $p\theta x^{p-1}\exp(-\theta x^p)$ where, $a_r = \Gamma(1+r/p)$	$a_1(a_3-3a_1a_2+2a_1^2)/$ $(a_2-a_1)^2$	$\{(a_4-4a_1a_3+6a_2a_1^2-$ $3a_1^4)^2(a_2-a_1)-(a_3-$ $3a_1a_2+2a_1^2)^2/(a_2-a_1)^3$
Gumbel $u/\theta\, e^{-u}$, where $u = e^{-(x-a)/\theta}$	$0.8881(a/\theta)$ -0.4874	4.10265
Logistic $e^{-(x-a)/b}\{1+e^{-(x-a)/b}\}^{-2}/b$	0	4.2
Uniform $1/(b-a)$	0	1.8
Laplace $(1/2ab)\exp\{-\lvert x-a\rvert/b\}$	0	6
Lognormal $\exp\{-(\ln x-\theta)^2/2\sigma^2\}/(x\sigma\sqrt{2\pi})$	$\exp(\sigma^2)+2$	$\exp(4\sigma^2)+\exp(3\sigma^2)+1$
Inverse Gaussian $\lambda\exp\{-\lambda(x-\theta)^2/(2x\theta^2)\}/(2\pi x^3)$	3	$6\theta/\lambda$

TABLE 3.3 : RANGES OF d_1 AND d_2 FOR SOME DISCRETE AND CONTINUOUS DISTRIBUTIONS.

DISTRIBUTION	d_1	d_2	DISTRIBUTION	d_1	d_2
Binomial	$(-\infty , 1]$	$[1 , 3)$	Normal	0	3
N. Binomial	$(1 , 2)$	$(3 , 5]$	Student's t	0	$(3 , 9)$
G. N. Binomial	$(-\infty , 3)$	$[1 , \infty)$	Gamma	2	$(3 , \infty)$
Geometric	$(1 , 2)$	5	Exponential	2	5
G. Geometric	$(-\infty , 3)$	$[5 , \infty)$	Chi-square	2	$(3 , 7)$
Neyman type A	$(1 , 1.25)$	$(3 , \infty)$	Rayleigh	1.2074	2.8468
Borel	$(1 , 3)$	$[7 , \infty)$	Weibull	$(1 , \infty)$	$(1 , \infty)$
Poisson	1	3	Logistic	- 1	4.2
G. Poisson	$[1 , 3]$	$[3 , \infty)$	Uniform	- 1	1.8
Hermite	$[1 , 2)$	$[3 , 4)$	Gumbel	$(-\infty, \infty)$	4.10265
G. Hermite	$[1 , \infty)$	$[3 , \infty)$	Laplace	0	6
Log Normal	$(3 , \infty)$	$(3 , \infty)$	Inverse Gaussian	3	$(0 , \infty)$

G.= Generalized and N. = Negative.

TABLE 3.4 : EXACT AND APPROXIMATE VALUES FOR NEYMAN'S TYPE-A SUMS.

λ	θ	α	N	r	Exact $\Sigma f_N(j;\ \lambda, \theta)$	Approx. by $\Sigma f_{NB}(j;\ N,\alpha)$
1	0.1	0.0767	1.205	0	0.90922	0.90840(0.90869)
1	0.1	0.0767	1.205	1	0.99150	0.99228(0.99205)
1	0.1	0.0767	1.205	5	0.99999	0.99999(0.99242)
1	0.2	0.1237	1.464	0	0.83421	0.82938(0.83076)
1	0.2	0.1237	1.464	1	0.97081	0.97473(0.97383)
1	0.2	0.1237	1.464	5	0.99999	0.99999(0.97614)
1	0.3	0.1552	1.633	0	0.77168	0.75925(0.76213)
1	0.3	0.1552	1.633	1	0.94319	0.95189(0.95019)
1	0.3	0.1552	1.633	5	0.99993	0.99995(0.95647)
1	0.5	0.1940	2.077	0	0.67471	0.63890(0.64417)
1	0.5	0.1940	2.077	1	0.87933	0.89636(0.89467)
1	0.5	0.1940	2.077	5	0.99917	0.99965(0.91519)
2	0.1	0.0767	2.409	0	0.82669	0.82520(0.82578)
2	0.1	0.0767	2.409	1	0.97629	0.97759(0.98485)
2	0.1	0.0767	2.409	5	0.99999	0.99999(0.98490)
2	0.2	0.1237	2.833	0	0.69591	0.68786(0.69016)
2	0.2	0.1237	2.833	1	0.92381	0.9286990.92787)
2	0.2	0.1237	2.833	5	0.99992	0.99993(0.95284)
8	0.1	0.0767	9.636	0	0.46706	0.46370(0.46487)
8	0.1	0.0767	9.636	1	0.80515	0.80622(0.80604)
8	0.1	0.0767	9.636	5	0.99949	0.99953(0.94094)
8	0.2	0.1237	11.331	0	0.23453	0.22388(0.22688)
8	0.2	0.1237	11.331	1	0.54177	0.54177(0.53946)
8	0.2	0.1237	11.331	5	0.98767	0.98938(0.82395)
8	0.3	0.1552	13.068	0	0.12575	0.11043(0.11382)
8	0.3	0.1552	13.068	1	0.34933	0.33435(0.33850)
8	0.3	0.1552	13.068	5	0.94310	0.95074(0.89839)
20	0.1	0.0767	24.091	0	0.14908	0.14642(0.14734)
20	0.1	0.0767	24.091	1	0.41887	0.41680(0.41768)
20	0.1	0.0767	24.091	5	0.97809	0.97892(0.85429)

TABLE 3.5 : EXACT AND APPROXIMATE VALUES FOR G. P. (BOREL-TANNER) SUMS.

θ	λ	α	N	r	Exact $\Sigma f_{GP}(j;\ \lambda, \theta)$	Approx. by $\Sigma f_{NB}(j;\ N,\alpha)$
0.05	1	0.102	9.26	0	0.36788	0.36897
0.05	1	0.102	9.26	5	0.99820	0.99816
0.05	1	0.102	9.26	10	0.99999	0.99999
0.05	2	0.102	18.34	0	0.13534	0.13614
0.05	2	0.102	18.34	5	0.97279	0.97248
0.05	2	0.102	18.34	10	0.99994	0.99994
0.05	5	0.102	46.34	0	0.00674	0.00684
0.05	5	0.102	46.34	5	0.57240	0.57245
0.05	5	0.102	46.34	10	.97482	0.97453
0.1	1	0.206	4.29	0	0.36788	0.37234
0.1	1	0.206	4.29	5	0.99584	0.99552
0.1	1	0.206	4.29	10	0.99999	0.99999
0.1	2	0.206	8.57	0	0.13534	0.13863
0.1	2	0.206	8.57	5	0.95875	0.95729
0.1	2	0.206	8.57	10	0.99972	0.99969
0.1	5	0.206	21.43	0	0.00674	0.00716
0.1	5	0.206	21.43	5	0.52977	0.53047
0.1	5	0.206	21.43	10	0.95778	0.95629
0.2	1	0.407	1.82	0	0.36788	0.38622
0.2	1	0.407	1.82	5	0.98605	0.98330
0.2	1	0.407	1.82	10	0.99980	0.99972
0.2	2	0.407	3.63	0	0.13534	0.14916
0.2	2	0.407	3.63	5	0.92048	0.91376
0.2	2	0.407	3.63	10	0.99756	0.99678
0.2	5	0.407	9.09	0	0.00674	0.00860
0.2	5	0.407	9.09	5	0.44885	0.45515
0.2	5	0.407	9.09	10	0.90412	0.89667
0.3	1	0.585	1.014	0	0.36788	0.40999
0.3	1	0.585	1.014	5	0.96780	0.95901
0.3	1	0.585	1.014	10	0.99833	0.99718
0.3	2	0.585	2.029	0	0.13534	0.16810
0.3	2	0.585	2.029	5	0.87005	0.85673
0.3	2	0.585	2.029	10	0.98916	0.98418
0.3	5	0.585	5.072	0	0.00674	0.01159
0.3	5	0.585	5.072	5	0.37552	0.39582
0.3	5	0.585	5.072	10	0.82408	0.80960
0.4	1	0.727	0.625	0	0.36788	0.44395
0.4	1	0.727	0.625	5	0.94063	0.92452
0.4	1	0.727	0.625	10	0.99280	0.98716
0.4	2	0.727	1.25	0	0.13533	0.19709
0.4	2	0.727	1.25	5	0.81056	0.79576
0.4	2	0.727	1.25	10	0.96835	0.95350

EXACT AND APPROXIMATE VALUES FOR G. P. (BOREL-TANNER) SUMS.

θ	λ	α	N	r	Exact $\Sigma f_{GP}(j;\ \lambda, \theta)$	Approx. by $\Sigma f_{NB}(j;\ N, \alpha)$
0.4	5	0.727	3.125	0	0.00674	0.01724
0.4	5	0.727	3.125	5	0.31094	0.35390
0.4	5	0.727	3.125	10	0.72335	0.71064
0.5	1	0.833	0.400	0	0.36788	0.48836
0.5	1	0.833	0.400	5	0.90562	0.88714
0.5	1	0.833	0.400	10	0.97934	0.96516
0.5	2	0.833	0.800	0	0.13533	0.23849
0.5	2	0.833	0.800	5	0.74578	0.74165
0.5	2	0.833	0.800	10	0.93026	0.90447
0.5	5	0.833	2.000	0	0.00674	0.02778
0.5	5	0.833	2.000	5	0.25536	0.33020
0.5	5	0.833	2.000	10	0.61200	0.61867
0.6	1	0.907	0.256	0	0.36788	0.54392
0.6	1	0.907	0.256	5	0.86485	0.85598
0.6	1	0.907	0.256	10	0.95485	0.93389
0.6	2	0.907	0.513	0	0.13534	0.29585
0.6	2	0.907	0.513	5	0.67934	0.70346
0.6	2	0.907	0.513	10	0.87390	0.84808
0.6	5	0.907	1.282	0	0.00674	0.04761
0.6	5	0.907	1.282	5	0.20846	0.32518
0.6	5	0.907	1.282	10	0.50073	0.54769
0.7	1	0.955	0.159	0	0.36788	0.61223
0.7	1	0.955	0.159	5	0.82064	0.83918
0.7	1	0.955	0.159	10	0.91857	0.90357
0.7	2	0.955	0.317	0	0.13533	0.37483
0.7	2	0.955	0.317	5	0.61425	0.68856
0.7	2	0.955	0.317	10	0.80261	0.80098
0.7	5	0.955	0.794	0	0.06738	0.08602
0.7	5	0.955	0.794	5	0.16948	0.34253
0.7	5	0.955	0.794	10	0.39805	0.50695

Chapter 4

Family of Compound Poisson Distributions

4.1 Introduction

In this chapter we have used some ratios of the co-efficients of a recurrence relation obtained from the generating function of a Compound Poisson distribution to identify different members of the Compound Poisson family. Moments of some distributions belonging to the Compound Poisson family are also presented.

A Compound Poisson distribution can be defined as a family of distributions having the following probability generating function (p.g.f.) [Feller 1965, p. 271]

$$P(s)= \Sigma p_j s^j = \exp[a_0(s-1) + a_1(s^2 -1)/2 + a_2(s^3 -1)/3 + . . .] \quad (4.1.1)$$

This represents the model for cumulative effects of singlets, doublets etc., each with a Poisson law. The Poisson, Hermite, Negative Binomial, Neyman type A etc. are distributions belonging to this family and can be obtained by suitable choices of the coefficients $a_0, a_1, . . .$ Table 4.1 gives a few Compound Poisson distributions , their p.g.f.s and a_i-values.

To obtain a recurrence relation for calculating successive p_j 's let us differentiate (4.1.1) with respect to s and get

$$P'(s) = \Sigma j p_j s^{j-1} = [a_0 + a_1 s + a_2 s^2 + \ldots] \exp[a_0(s-1) + a_1(s^2-1)/2 + \ldots]$$
$$= P(s)\,[a_0 + a_1 s + a_2 s^2 + \ldots]$$
$$= [p_0 + p_1 s + p_2 s^2 + \ldots]\,[a_0 + a_1 s + a_2 s^2 + \ldots]$$

Now equating the coefficients of s^j we can write the recurrence relation as follows

$$p_{j+1}(j+1) = a_0 p_j + a_1 p_{j-1} + \ldots + a_j p_0 \qquad (4.1.2)$$
$$j=0,1,2,\ \ldots\ ,\quad p_j = 0 \text{ if } j<0.$$

The cumulants of (4.1.1) can be obtained by taking the logarithm of both sides and expanding.

$$\ln P(s) = [a_0(s-1) + a_1(s^2-1)/2 + a_2(s^3-1)/3 + \ldots]$$

Thus the cumulant generating function is

$$\kappa(t) = a_0(e^t-1) + a_1(e^{2t}-1)/2 + a_2(e^{3t}-1)/3 + \ldots$$

The co-efficient of $t^i/i!$ is the i-th (i=1,2,3, . . .) cumulant

$$\kappa_i = a_0 + 2^{i-1} a_1 + 3^{i-1} a_2 + 4^{i-1} a_3 + \ldots = \Sigma_r (r+1)^{i-1} a_r \qquad (4.1.3)$$

Thus the cumulants of any Compound Poisson distribution can be computed from (4.1.3) . Hinz and Gurland (1967) have suggested that the plots of the sample values of the cumulant ratios $\eta_j = \kappa_{(j+1)}/\kappa_{(j)}$, where $\kappa_{(j)}$ is the jth factorial cumulant, against j may be used in discriminating among certain Compound Poisson distributions. Earlier Ottested (1939) used the ratio $\mu_{(j+1)}/\mu_{(j)}$, where $\mu_{(j)}$ is the jth factorial moment against j, to discriminate among the Binomial, Poisson and the Negative Binomial distributions. One can use the corresponding sample values in these criteria to find out the possible form of the underlying distribution. Because of the sampling fluctuations, a particular criterion may not provide reliable information to draw sound conclusions. In fact whenever it is possible, more than one criterion should be used and other characteristics should be verified to ascertain a distribution. Here we suggest a use of the ratios of a_j's, rather than the ratios of moments, identifying certain Compound Poisson distribution especially those listed in Table 4.1. It may be noted that the cumulants do not necessarily identify a distribution while the a_j's do identify them, hence the use of a_j's in place of the cumulants has some merit.

4.2. Identification in the Compound Poisson Family

The coefficients $a_0, a_1, a_1, \ldots$, can be obtained systematically from (4.1.2) and can be written in matrix notation as

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_j \end{bmatrix} = \begin{bmatrix} p_0 & 0 & 0 \cdots & 0 \\ p_1 & p_0 & 0 \cdots & 0 \\ p_2 & p_1 & p_0 \cdots & 0 \\ & \vdots & & \\ p_j & p_{j-1} & p_{j-2} \cdots & p_0 \end{bmatrix}^{-1} \begin{bmatrix} p_1 \\ 2p_2 \\ 3p_3 \\ \cdot \\ \cdot \\ \cdot \\ (j+1)p_{j+1} \end{bmatrix}, j = 0,1,2, \ldots \quad (4.2.1)$$

Let

$$P_j = \begin{bmatrix} p_0 & 0 & 0 & \dots & 0 \\ p_1 & p_0 & 0 & \dots & 0 \\ p_2 & p_1 & p_0 & \dots & 0 \\ & & \vdots & & \\ p_j & p_{j-1} & p_{j-2} & \dots & p_0 \end{bmatrix}$$

and $p_{(j)} = (p_1, p_2, \dots, p_j)'$.

Then (4.2.1) can be written as

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_j \end{bmatrix} = \begin{bmatrix} \frac{1}{p_0} & 0 \\ -\frac{P_{j-1}^{-1} p_{(j)}}{p_0} & P_{j-1}^{-1} \end{bmatrix} \begin{bmatrix} p_1 \\ 2p_2 \\ 3p_3 \\ \vdots \\ (j+1)p_{j+1} \end{bmatrix}, \quad j = 0,1,2,\dots \qquad (4.2.2)$$

In particular using (4.2.2) for j=0,1,2, . . . we obtain a_0, a_1, a_2, a_3 etc. as follows

$a_0 = p_1 / p_0$

$a_1 = 2p_2 / p_0 - (p_1 / p_0)^2$

$a_2 = 3p_3 / p_0 - 3p_1p_2 / p_0^2 + (p_1 / p_0)^3$

$a_3 = 4p_4 / p_0 - 2(p_2^2 + 2p_1p_3) / p_0^2 + 4p_1^2p_2 / p_0^3 - (p_1 / p_0)^4$

$a_4 = 5p_5/p_0 - 5p_1p_4/p_0^2 - 5p_2p_3/p_0^2 + 5p_1^2p_3/p_0^3 + 5p_1p_2^2/p_0^3$
$- 5p_1^3p_2/p_0^4 + (p_1/p_0)^5$, etc.

These coefficients can also be obtained by using the recurrence relations

$a_0 = p_1/p_0$

$a_j = p_{j+1}(j+1)/p_0 - a_0\, p_j/p_0 - a_1 p_{j-1}/p_0 - \ldots - a_{j-1}p_1/p_0$, $j=1,2,\ldots$

For discriminating among the compound distributions of table 4.1 we define the following ratios :

$n_i = a_{i+1}/a_i$, $i=0,1,2,\ldots$

and $t_i = (1+i)a_{i+1}/a_i$, $i=0,1,2,\ldots$

Table 4.1 gives the values of n_i's. In table 4.2 we state the behaviour of n_i's and t_i's for these distributions. The corresponding sample values may, therefore, be useful in discriminating among these distributions. Figures 4.1 and 4.2 give the plots of (1) n_i against i and (2) n_i against t_i for $i=0,1,2,\ldots$

Because of the sampling fluctuations values of the estimates of n_i for $i \geq 3$ may not be reliable and conclusions may have to be based on n_0, n_1 and n_2 only. Furthermore if these ratios give some indication of a particular form of an underlying distribution, it may be advantageous to verify other criteria and characteristics of the distribution. For example for the Binomial, Poisson or the Negative Binomial distributions it is known that κ_2/κ_1 is less than, equal to or greater than one respectively. It is also known that the quantity, $a_j = p_{j+1}(j+1)/jp_j - p_1/j\,p_0$, $j=1,2,3,\ldots$, should be $-p/(1-p)$, 0 or p for the Binomial, Poisson or the Negative Binomial distributions respectively. Hence the sample counterparts of a_j should be approximately a negative constant, zero or a positive constant for these distributions.

As an example we consider the Bell Telephone Company data in table 4.3 regarding the lost articles found in the Telephone and Telegraph building, New York city. The sample values of n_i are given in the table 4.3 and indicate that the distribution may be the Negative Binomial. For this data set m=1.03783, m_2=1.27044, m_3 =1.75591, m_4 = 7.90823, moment function for the Negative Binomial distribution $mm_2-2(m_2)^2+mm_2 \approx 0$, moment ratios $\hat{d}_1$ = 1.12373, $\hat{d}_2$ = 3.38359 (for the Negative Binomial distribution $1 < d_1 < 2$ and $3 < d_2 \leq 5$). Therefore, the distribution is more likely to be the Negative Binomial.

4.3. Some Distributions belonging to the Compound Poisson Family

In nature the individuals of many species (e.g. plants, insects) have the tendency to cluster together. The variance of an observational series in such a situation will exceed its mean. A few distributions have been developed in recent years. One such distribution is the Neyman type A distribution, which assumes that the clusters are randomly dispersed over a given area according to the Poisson law, while the number of individuals within a cluster are also distributed randomly according to another Poisson law. Neyman (1939) used this model to fit the observed distribution of larvae in a randomly chosen area on a field. Thomas (1949) considered a modified form of the Neyman type A distribution by including the parent as well in the count for each cluster, and applied the distribution to fit the observed distribution of plants (Armeria martima and Plantago martima) per quadrant. There are situations where the hypothesis of the Poisson distribution of clusters may be reasonably justifiable, but the assumption of the Poisson distribution of the counts of a cluster may not be justifiable. Jain and Plunkett (1977) consider one such model by assuming that the clusters are randomly distributed according to the Poisson law with mean θ_1, and that the cluster size '1+i' (i=0,1,2,...) has the Borel distribution having probability mass function (p.m.f)

$$f_{1+i} = \theta_2^{i} (1+i)^{i-1} \exp\{-\theta_2(1+i)\}, \; i = 0,1,2, \ldots, \quad 0 < \theta_2 < 1 \qquad (4.3.1)$$

with the probability generating function (p.g.f.)

$$G(s) = \Sigma f_{1+i} s^{i+1} = sH(s), \qquad (4.3.2)$$

where H(s) is given by the functional relation

$$H(s) = \exp[-\theta_2\{1 - sH(s)\}] \qquad (4.3.3)$$

The distribution of the total count by mixing the Poisson and the Borel distributions can be shown to have the Borel-Tanner distribution or, the Generalized Poisson distribution given by the p.m.f.

$$t_j = \theta_1(\theta_1 + \theta_2 j)^{j-1} \exp\{-(\theta_1 + \theta_2 j)\}/j! \,, \quad j = 0,1,2,\ldots\ldots \qquad (4.3.4)$$

with the p.g.f.

$$T(s) = \exp[-\theta_1\{1 - sH(s)\}] \qquad (4.3.5)$$

The mean and variance of the cluster size as obtained from the Borel distribution (4.3.1) are respectively $1/(1-\theta_2)$ and $\theta_2/(1-\theta_2)^3$. The mean will, therefore, be smaller than, equal to or, greater than the variance depending on whether the value of θ_2 is greater than, equal to or, smaller than $(3-\sqrt{5})/2= 0.38197$. The Borel-Tanner distribution (4.3.4) may, therefore, provide a good fit to many situations in Entomology and Bacteriology where the mean and the variance of the cluster count are not necessarily equal. All of these distributions belong to the family of the Compound Poisson distribution. A Compound Poisson distribution is one which is formed by a mixture of any two or more distributions. The computation of the moments of a Compound distribution in terms of those of the mixtures are presented in section 4.4. In subsequent sections the mixtures of Poisson and Borel-Tanner, Borel and Poisson, and Borel and Borel distributions are discussed.

4.4. Moments of a Compound Distribution

Suppose that clusters are dispersed according to an arbitrary distribution a_i $(i=0,1,2,\ldots)$ with p.g.f. $A(s)$ and that the cluster sizes, including the parent, are distributed according to another arbitrary distribution b_{1+j} $(j=0,1,2,\ldots)$ with p.g.f. $B(s) = sC(s)$, where $b_{1+j} = c_j$ $(j=0,1,2,\ldots)$ and $C(s) = \sum c_j s^j$. Then the compound distribution, say g_j $(j=0,1,2,\ldots)$, of the total counts is given by

$$g_j = \sum_{i=1}^{j} a_i\, c_{j-i}^{(i)}\ ,\ j=1,2,\ldots,\ g_0=a_0\ ,\ \text{with p.g.f. } G(s) = A\{sC(s)\} \quad (4.4.1)$$

and where $c_j^{(i)}$ is given by the expansion

$$[C(s)]^i = \Sigma\, c_j^{(i)}\, s^j \qquad (4.4.2)$$

The moments of the distribution 'g_j' can easily be obtained by differentiating or expanding (4.4.1). Let $\psi_{(k)}$, $\phi_{(k)}$ and $\mu_{(k)}$ be the kth factorial moments, ψ_k , ϕ_k and μ_k be the kth central moments of the distributions a_j , c_j and g_j respectively. Then it can be shown that

$$\mu_{(1)} = \psi_{(1)}(1+ \phi_{(1)})$$
$$\mu_{(2)} = \psi_{(2)}[1+ \phi_{(1)}]^2 + \psi_{(1)}(2\phi_{(1)} + \phi_{(2)})$$
$$\mu_{(3)} = \psi_{(3)}[1+ \phi_{(1)}]^3 + 3\psi_{(2)} (1+ \phi_{(1)})(2\phi_{(1)} + \phi_{(2)}) +\psi_{(1)}(3\phi_{(2)} + \phi_{(3)})$$
$$\mu_{(4)} = \psi_{(4)}[1+ \phi_{(1)}]^4 + 6\psi_{(3)} [1+ \phi_{(1)}]^2(2\phi_{(1)} + \phi_{(2)})$$
$$+ 4\psi_{(2)} (1+ \phi_{(1)})(3\phi_{(2)} + \phi_{(3)}) + 3\psi_{(2)}[2\phi_{(1)} + \phi_{(2)}]^2 + \psi_{(1)}(4\phi_{(3)}$$
$$+ \phi_{(4)}) \qquad (4.4.3)$$

and

$$\mu_2 = \psi_2[1+ \phi_{(1)}]^2 + \psi_{(1)}\phi_2$$
$$\mu_3 = \psi_3[1+ \phi_{(1)}]^3 + 3(1+ \phi_{(1)})\phi_2\, \psi_2 + \psi_{(1)}\phi_3$$
$$\mu_4 = \psi_4[1+ \phi_{(1)}]^4 + 6\phi_2[\psi_3+ \psi_{(1)}\, \psi_2]\, [1+ \phi_{(1)}]^2$$
$$+\psi_2(4\phi_3 +3\phi_2{}^2 +4\, \phi_1\phi_3] + 3\psi_1{}^2\phi_2{}^2 + \psi_1(\phi_4 -3\, \phi_2{}^2)$$
$$(4.4.4)$$

Thus choosing $A(s)=\exp\{-\theta(1-s)\}$ and $C(s)=\exp\{-\lambda(1-s)\}$ gives the p.g.f. of the Thomas distribution as $G(s)= \exp[-\theta(1-s\ \exp\{-\lambda(1-s)\}]$. The mean and the other three central moments of the Thomas distribution are

$$\mu_{(1)} = \theta(1+ \lambda)$$

$\mu_2 = \theta(1+3\lambda+\lambda^2)$

$\mu_3 = \theta(1+7\lambda+6\lambda^2+\lambda^3)$

$\mu_4 = 3\theta^2(1+3\lambda+\lambda^2)^2 + \theta(1+15\lambda+25\lambda^2+10\lambda^3+\lambda^4)$

as given in Johnson and Kotz [1969].

Remark :-

For the compound distribution having the p.g.f. of the form $G(s)=A[C(s)]$, the probability distribution of the total count is

$g_j = \Sigma a_i c_j^{(i)}$, where $c_j^{(i)}$ is defined in (4.4.2).

The factorial and central moments can be shown to be as follows :-

$\mu_{(1)} = \psi_{(1)}\phi_{(1)}$,

$\mu_{(2)} = \psi_{(2)}\phi_{(1)}^2 + \psi_{(1)}\phi_{(2)}$

$\mu_{(3)} = \psi_{(3)}\phi_{(1)}^3 + 3\psi_{(2)}\phi_{(1)}\phi_{(2)}) + \psi_{(1)}\phi_{(3)}$

$\mu_{(4)} = \psi_{(4)}\phi_{(1)}^4 + 6\psi_{(3)}\phi_{(1)}^2\phi_{(2)} + 4\psi_{(2)}\phi_{(1)}\phi_{(3)} + 3\psi_{(2)}\phi_{(2)}^2 + \psi_{(1)}\phi_{(4)}$

$\mu_2 = \psi_2\,\phi_{(1)}^2 + \psi_{(1)}\phi_2$

$\mu_3 = \psi_3\,\phi_{(1)}^3 + 3\phi_{(1)})\phi_2\,\phi_3 + \psi_{(1)}\phi_3$

$\mu_4 = \psi_4\,\phi_{(1)}^4 + 6\phi_2\,\phi_{(1)}^2\,[\psi_3 + \psi_{(1)}\,\psi_2] + 4\psi_2\,\phi_{(1)}\phi_3 + 3\psi_{(1)}^2\phi_2^2$

$\quad + \psi_{(1)}(\phi_4 - 3\,\phi_2^2)$

A choice of $A(s)=\exp\{-\theta(1-s)\}$ and $C(s)=\exp\{-\lambda(1-s)\}$ gives the p.g.f. of the Neyman type A distribution $G(s)= \exp[-\theta(1-\exp\{-\lambda(1-s)\}]$, with the moments as

$\mu_{(1)} = \theta\lambda$, $\quad \mu_2 = \theta\lambda(1+\lambda)$

$\mu_3 = \theta\lambda(1+3\lambda+\lambda^2)$

$\mu_4 = \theta\lambda(1+7\lambda+6\lambda^2+\lambda^3) + 3\lambda^2\theta^2(1+\lambda)^2$

4.5. Mixtures of the Borel Distribution

Among the various mixtures which can be obtained by compounding a Borel distribution we consider mixtures of Poisson and Borel and the double Borel distributions.

(i) The Poisson-Borel-Tanner Distribution

A somewhat generalized model, to describe situations where the mean and the variance of the cluster counts are different, can be obtained by assuming that cluster size 1+i, is distributed according to the Borel-Tanner distribution,

$q_{1+i} = \theta_1(\theta_1 + \theta_2 i)^{i-1} \exp\{-(\theta_1 + \theta_2 i)\}/i!$, $i = 0,1,2,\ldots\ldots$

with the p.g.f. $Q(s) = \Sigma q_{1+i}\, s^{i+1} = sT(s)$, where T(s) is defined in (4.3.5). A further assumption of the Poisson distribution of clusters with mean θ gives the distribution of the total count 'j' as

$$g_j = e^{-\theta}\sum_{i=1}^{j} \theta^i(\theta_1 i)[\theta_1 i + \theta_2(j-i)]^{j-i-1} \exp\{-[\theta_1 i + \theta_2(j-1)]\} /\{i!\,(j-i)!\} \qquad (4.5.1)$$

$$j = 1,2,3,\ldots\ldots$$

$$= \exp(-\theta), \quad j=0$$

with the p.g.f.

$G(s) = \exp\{-\theta[1-sT(s)]\} = \exp\{-\theta[1-s\ \exp\{-\theta_1(1-sH(s))\}]\}$,

where H(s) is given by (4.3.3). It is evident that the distribution (4.5.1) reduces to that of the Borel-Tanner distribution if $\theta_1 = \theta_2$ and to that of the Thomas distribution if $\theta_2 = 0$.

The mean and the second, third and fourth central moments are found to be

$\mu_{(1)} = \theta(1+ \lambda_1')$

$\mu_2 = \theta[(1+\lambda_1')^2 + \lambda_2)]$

$\mu_3 = \theta[(1+ \lambda_1')^3 + 3\ (1+\lambda_1')\lambda_2 + \lambda_3\ ]$

$\mu_4 = \theta[(1+ \lambda_1')^4 + 6(1+\lambda_1')^2 \lambda_2 + 4(1+\lambda_1')\lambda_3 + \lambda_4] + 3\theta^2[(1+ \lambda_1')^2 + \lambda_2]^2$

where,

$\lambda_1' = \theta_1/(1- \theta_2)$

$\lambda_2 = \theta_1/(1- \theta_2)^3$

$\lambda_3 = \theta_1(1+2\theta_2)/(1- \theta_2)^5$

$\lambda_4 = 3\theta_1{}^2/(1- \theta_2)^6 + \theta_1(1+8\theta_2+6\theta_2{}^2)/(1- \theta_2)^7$

(4.5.2)

Therefore,

$\beta_1 = (1/\theta)[1+3\lambda_2/(1+\lambda_1')^2 + \lambda_3/(1+\lambda_1')^3][1+ \lambda_2/(1+\lambda_1')^2]^{-3}$

$\beta_2 = 3+ (1/\theta)[1+6\lambda_2/(1+\lambda_1')^2 + 4\lambda_3/(1+\lambda_1')^3 + \lambda_4/(1+\lambda_1')^4]$

$[1+ \lambda_2/(1+\lambda_1')^2]^{-2}$.

Hence the distribution is positively skewed.

Solving the moment equations for parameters also gives

$9(\theta\mu_2 - \mu_{(1)}{}^2)^4 + (\theta - \mu_{(1)})^2[\theta^2\mu_3 - \mu_{(1)}{}^3 - 3\mu_{(1)}(\theta\mu_2 - \mu_{(1)}{}^2)]^2$

$- 4\theta(\theta\mu_2 - \mu_{(1)}{}^2)^3(\mu_{(1)} - \theta) - 6(\theta\mu_2 - \mu_{(1)}{}^2)[\theta^2\mu_3 - \mu_{(1)}{}^3 - 3\mu_{(1)}(\theta\mu_2 -$

$\mu_{(1)}{}^2)](\mu_{(1)} - \theta) = 0,$

$\theta_1 = (\mu_{(1)} - \theta)^{3/2}/[\theta(\theta\mu_2 - \mu_{(1)}{}^2)]^{1/2},$

$\theta_2 = 1 - \{\theta(\mu_{(1)} - \theta)/[\theta\mu_2 - \mu_{(1)}{}^2]\}^{1/2},$

(ii) The Borel-Poisson Distribution

For a situation where the dispersion of the clusters is 'Non Poisson', we consider here a particular one by assuming that the clusters are dispersed according to the Borel distribution

$$b_r = \alpha^r (1+r)^{r-1} \exp\{-\alpha(1+r)\}/r! , \qquad r=0,1,2, \ldots , \quad \alpha < 1. \qquad (4.5.3)$$

If the cluster size, including the parent, is assumed to be distributed as Poisson with mean θ , then the probability distribution of the total count is given by

$$g_j = \Sigma \, \alpha^i (1+i)^{i-1} \exp\{-\alpha(1+i)-\theta i\} (\theta i)^{j-1}/\{i! \ (j-i)!\} , \qquad (4.5.4)$$

$$j=1,2,3, \ldots , \qquad \alpha<1, \qquad \theta>0$$

$$= \exp\{-\alpha\} , \quad j=0,$$

which we call the Borel-Poisson distribution.

The mean and the second, third and fourth central moments of (4.5.4) can be obtained by using the moments of the Borel distribution [as in (4.5.2) with $\theta_1 = \theta_2 = \alpha$] and the moments of the Poisson distribution [as in (4.5.2) with $\theta_1 = \theta$, $\theta_2 = 0$] in (4.4.4).

Thus, the mean and the variance of (4.5.4) are

$$\mu_{(1)} = \alpha(1+ \theta)/(1- \alpha)$$

$$\mu_2 = \alpha(1+ \theta)^2/(1- \alpha)^3 + \alpha\theta/(1- \alpha) ,$$

from which the parameters θ and α of (5.5.4) can be computed as

$$\theta^2\mu_{(1)} - \theta(\mu_2 - 3\mu_{(1)} - 2\mu_{(1)}^2) + \mu_{(1)}^3 + 2\mu_{(1)}^2 + \mu_{(1)} - \mu_2 = 0 \quad (4.5.5)$$

and

$$\alpha = (2+2\theta+ \mu_{(1)})/(1+ \theta +\mu_{(1)}) \qquad (4.5.6)$$

(iii) The Double Borel Distribution

If the clusters are assumed to be dispersed according to the Borel distribution (4.5.3) with parameter α and the cluster size is also assumed to be distributed according to the Borel distribution (4.5.3) with parameter β, then the total count has the following distribution,

$$g_j = \sum \alpha^i (1+i)^{i-1} \exp\{-\alpha(1+i)-\beta j\}\, i\, j^{j-i-1} \beta^{j-1}/\{i!\,(j-i)!\},$$

$$j=1,2,3,\ldots, \qquad \alpha>0, \qquad \beta<1$$

$$= \exp\{-\alpha\}, \quad j=0,$$

which we call the Double Borel distribution. The mean, the variance and the parametric relations of the Double Borel distribution are found to be

$$\mu_{(1)} = \alpha/(1-\alpha)(1-\beta),$$

$$\mu_2 = \alpha[1-\alpha\beta(2-\alpha)]/(1-\alpha)^3(1-\beta)^3,$$

$$\alpha^3\{\mu_2-(\mu_1')^3-(\mu_1')^2\} + \alpha^2\{3(\mu_1')^3+2(\mu_1')^2-\mu_2\} - 3\alpha(\mu_1')^3 + (\mu_1')^3 = 0$$

respectively and $\beta = 1-\alpha/\{\mu_1'(1-\alpha)\}$.

By computing the probabilities for $j = 0,1,2,\ldots$, and for different values of the parameters the behaviour of these distributions can be studied. In Table 4.4 and Figure 4.3 we exhibit the probabilities corresponding to these distributions for specific values of their parameters.

TABLE 4.1: PGF, VALUES OF a_i AND n_i FOR SOME DISTRIBUTIONS.

Distribution	PGF	Values of a_i and n_i
Poisson $a_0{}^n\{\exp(-a_0)\}/j!$	$\exp[a_0(s-1)]$	$a_0 > 0,\ a_i = 0,\ i=1,2,3,\ldots$ $n_i = 0,\ i = 0,1,2,\ldots$
Binomial ${}^nC_j\, p^j q^{n-j}$	$(q+ps)^n$	$a_i = n(p/q)^{i+1}(-1)^i,\ q=1-p$ $n_i = -q/p,\ i = 0,1,2,.$
Poisson Binomial $e^{-\lambda}(\frac{p}{q})^j (j!)^{-1} \sum_{k\geq j/n} \frac{(nk)^j(\lambda q^n)^k}{k!}$	$\exp[\lambda\{(q+ps)^n - 1\}]$	$a_i = \lambda\, {}^nC_{1+i}(p/q)^{i+1}q^n(1+i)$, $i=0,1,2,\ldots,n-1$ $n_i = p(n-i-1)/\{q(i+1)\},\ i=0,1,2,\ldots$
Negative Binomial ${}^{n+j-1}C_j\, p^j q^n$	$\{(1-p)/(1-ps)\}^n$	$a_i = np^{i+1},\ i = 0,1,2,\ldots$ $n_i = p,\ i = 0,1,2,\ldots$
Poisson-Negative Binomial $\sum_i e^{-\lambda} \frac{\lambda^i(1-p)^{ni}}{i!}\binom{ni+j-1}{j}p^j$	$\exp[-\lambda\{1-\{q/(1-ps)\}^n\}]$	$a_i = {}^{n+i}C_i\, n\lambda p^{i+1}q^n$, $n_i = p(n+i+1)/(i+1)$ $i=0,1,2,\ldots$
Generalized Poisson $\lambda_1(\lambda_1 + \lambda j)^{j-1}e^{-(\lambda_1+\lambda j)}/j!$	$\exp[\lambda_1\{\Sigma(\lambda j)^{j-1}e^{-js}s^j/j! - 1\}]$	$a_i = (1+i)g_{i+1},\ i=0,1,2,\ldots$ where, $g_i = (\lambda i)^{i-1}e^{-\lambda i}/i!$, $n_i = \{1+1/(i+1)\}^{i+1}\lambda e^{-\lambda},\ i=1,2,\ldots$
Hermite $\exp[-(a_0+\frac{a_1}{2})]\sum_{k=0}^{[j/2]}\frac{a_1^k a_0^{j-2k}}{k!2^k(j-2k)!}$	$\exp[a_0(s-1)+a_1(s^2-1)]$	$a_0, a_1 > 0,\ a_i = 0,\ i=1,2,3,\ldots$ $n_0 = \frac{a_1}{a_0},\ n_i = 0,\ i=1,2,\ldots$
Neyman Type A $\sum_k e^{-\lambda}(\frac{\lambda^k}{k!})e^{-k\theta}\frac{(k\theta)^j}{j!}$	$\exp[-\lambda\{1-\exp\{-\theta(1-s)\}\}]$	$a_i = \lambda e^{-\theta}\theta^{i+1}/i!,\ i=0,1,2,\ldots$ $n_i = \frac{\theta}{i+1},\ i=0,1,2,\ldots$

TABLE 4.2: IDENTIFICATION OF THE MEMBERS OF COMPOUND POISSON FAMILY.

Poisson	$n_i = 0,\ i = 0,1,2,\ldots,$	$t_i = 0,\ i = 0,1,2,\ldots$
Binomial	n_i = a constant $i = 0,1,2,\ldots$	t_i decreases with i such that $t_{i+1} - t_i = n_i$, a constant, $i = 0,1,2\ldots$
Poisson Binomial	n_i decreases to zero and $i = n-1$	t_i decreases to zero and $i = n-1$
Negative Binomial	n_i = a constant $i = 0,1,2,\ldots$	t_i increases with i such that $t_{i+1} - t_i = n_i$, a constant $i = 0,1,2\ldots$
Poisson Negative Binomial	n_i decreases with i and tends to a constant	t_i increases with i
Generalized Poisson or, Borel-Tanner	n_i increases slowly with i and tends to a constant	t_i increases with i
Hermite	n_0 = a constant $n_i = 0,\ i = 1,2,3,\ldots$	$t_0 = n_0$ $t_i = 0,\ i = 1,2,\ldots$
Neyman type A	n_i decreases slowly to zero, $i = 0,1,2,\ldots$	t_i = a constant $i = 0,1,2,\ldots$

TABLE 4.3: LOST ARTICLES FOUND IN THE TELEPHONE AND TELEGRAPH BUILDING, NEW YORK CITY.

No. of articles lost	No. of days	n_i	Expected Negative Binomial
0	169	0.31	166.02
1	134	0.32	140.42
2	74	0.25	72.37
3	32		29.33
4	11		10.27
5	2		3.26
6	0		0.96
7	1		0.37
Total	423		423.00

Source : Thorndike (1926)

TABLE 4.4 : PROBABILITIES OF POISSON-BOREL-TANNER, BOREL-TANNER AND DOUBLE BOREL DISTRIBUTIONS FOR SPECIFIC VALUES OF THEIR PARAMETERS.

j	PROBABILITIES g_j		
	Poisson-Borel -Tanner	Borel-Poisson	Double Borel
0	0.548812	0.740818	0.670320
1	0.551819	0.081760	0.133148
2	0.274488	0.050974	0.069263
3	0.141212	0.025996	0.041508
4	0.074103	0.013933	0.026511
5	0.039442	0.007818	0.017581
6	0.021227	0.004524	0.011960
7	0.011528	0.002682	0.008292
8	0.006308	0.001621	0.005834
9	0.003475	0.000996	0.004155
10	0.001925	0.000618	0.002990
11	0.001971	0.000389	0.002170
12	0.000599	0.000247	0.001587
13	0.000336	0.000158	0.001169
14	0.000189	0.000102	0.000865
15	0.000107	0.000066	0.000644
	$\theta = .6$, $\theta_1 = .4$ $\theta_2 = .3$	$\theta = .4$, $\alpha = .3$	$\alpha = .4$, $\beta = .4$

FIGURE 4.1 GRAPHS OBTAINED BY PLOTTING THE RATIOS n_i AGAINST i FOR DIFFERENT DISTRIBUTIONS.

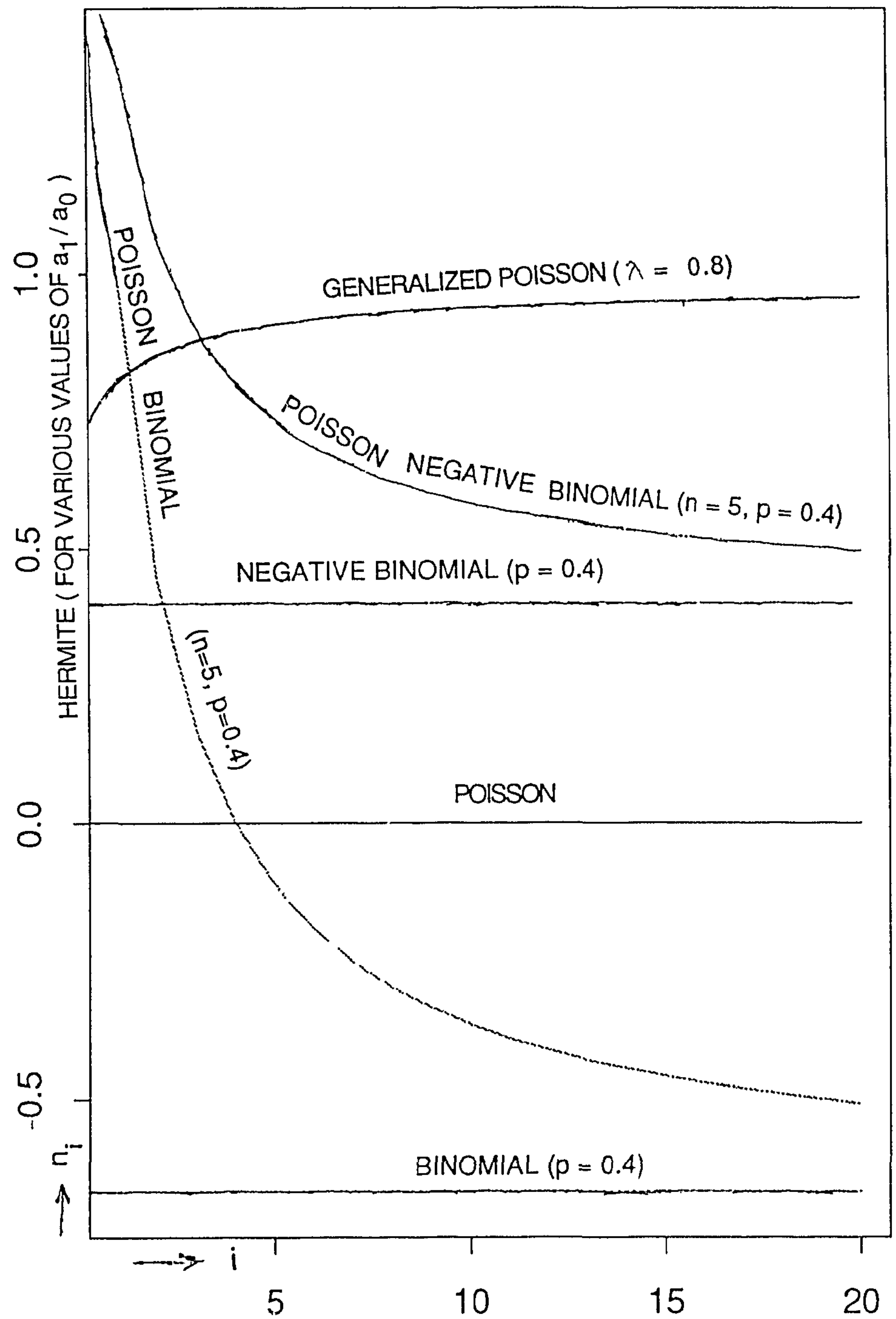


FIGURE 4.2 GRAPHS OBTAINED BY PLOTING THE RATIOS n_i AGAINST THE RATIOS t_i FOR DIFFERENT DISTRIBUTIONS.

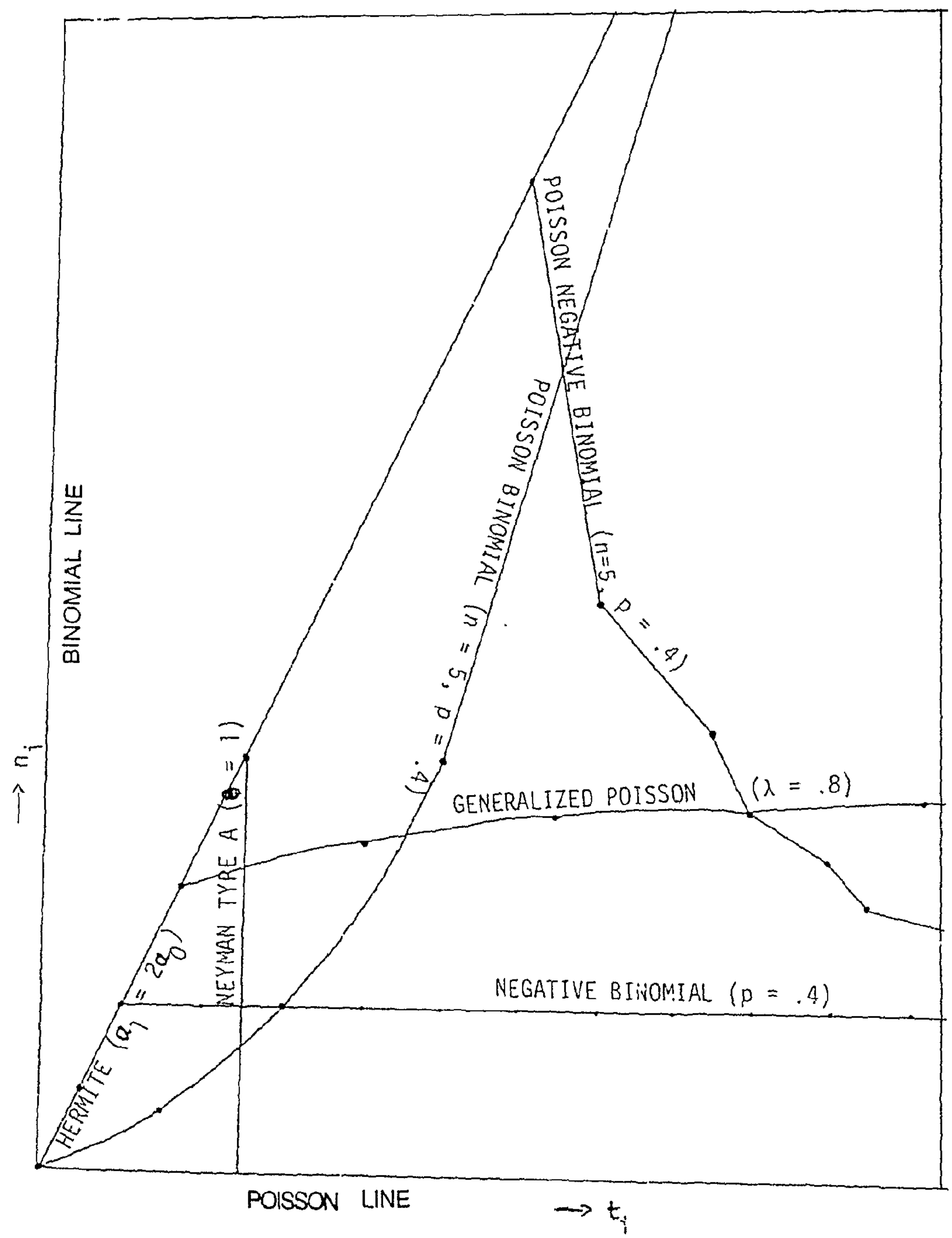
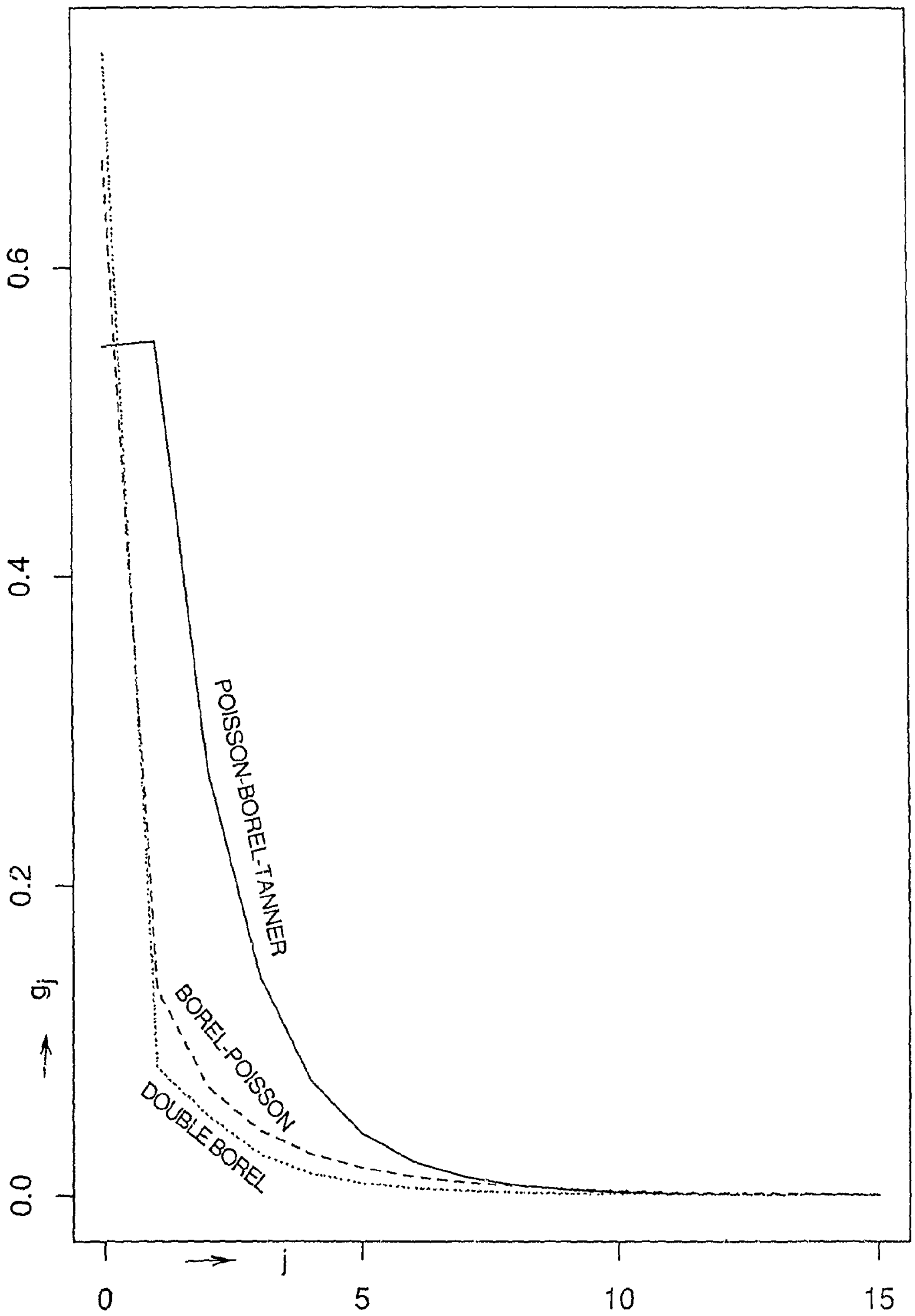


FIGURE 4.3 PROBABILITY CURVES FOR POISSON-BOREL-TANNER, BOREL POISSON AND DOUBLE BOREL DISTRIBUTIONS.



Chapter 5

Family of Transformed Chi-square distributions

5.1 Introduction

Family of Transformed Chi-square distributions is a sub-family of the Exponential family of distributions. Consider a random variable X whose probability mass function (p.m.f.), or probability density function (p.d.f.), $f(x;\theta)$ depends on a scalar parameter of interest θ. Let the distribution of X belong to the Exponential family [Barndorff-Nielsen (1978)], i.e. $f(x;\theta)$ given by,

$$f(x;\theta) = \exp[a(x)b(\theta)+c(\theta)+h(x)] \tag{5.1.1}$$

The Binomial, Poisson, Normal, Exponential, Gamma, Geometric, Rayleigh, etc. are distributions that belong to this family and can be obtained by suitable choices of $a(x)$, $b(\theta)$, $c(\theta)$ and $h(x)$. Here $b(\theta)$ is a non-trivial continuous function of θ for $\theta \,\varepsilon\, \Omega=(c_1, c_2)$ [Patel, Kapadia and Owen (1976)] where c_1 and c_2 are real numbers. The likelihood function of a sample of size n from (5.1.1) can be written as

$$L(\theta;\underset{\sim}{x}) = \exp[b(\theta)\Sigma a(x_i)+nc(\theta)+\Sigma h(x_i)] \quad (5.1.2)$$

where, $\underset{\sim}{x} = (x_1,x_2, \ldots ,x_n)$. According to the Neyman-Fisher Factorization theorem, $\Sigma a(X_i)$ is sufficient for θ or, for any one to one function of θ. It follows from theorem 1, page 142 of Lehmann (1986) that $\Sigma a(X_i)$ is also a complete statistic. Then according to the Lehmann-Scheffé (1950,1955) theorem, $\Sigma a(X_i)$ is the unique uniformly minimum variance unbiased estimator [UMVUE] of its expected value which is a function of θ. If the Cramér-Rao regularity conditions hold then $\Sigma a(X_i)$ is the minimum variance bound unbiased estimator [MVBUE] of

$$E\{\Sigma a(X_i)\} = \psi(\theta) \text{ (say)} \quad (5.1.3)$$

iff,

$$V\{\Sigma a(X_i)\} = \{\psi'(\theta)\}^2/V(\partial \ln L/\partial\theta) \quad (5.1.4)$$

General expressions for the exact form of (5.1.3) and (5.1.4) are not available. If $b(\theta)$ is strictly increasing in θ then $\Sigma a(X_i)$ is an optimal test statistic for testing $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ [Bickel & Doksum (1976)]. In order to calculate the critical region and the power of the test we need to find the exact distribution of $\Sigma a(X_i)$. This distribution is also needed to obtain a confidence interval for θ. The general form of the distribution of $\Sigma a(X_i)$ is not available.

We prove a theorem regarding some characteristics of a class of distributions in section 5.2 and then define a sub-family of the Exponential family of distributions called the Transformed Chi-square family. The Gamma, Rayleigh, Normal, Lognormal, Pareto, Exponential, etc. belong to this family. Irrespective of the form of the original distribution of a random variable X belonging to this family, the distribution of $-2a(X)b(\theta)$ follows a Central Chi-square distribution with appropriate degrees of freedom. This sub-family is thus named the Transformed Chi-square family. Without much derivation one can easily obtain MVBUE or UMVUE, and also an interval estimator, of θ or any function of θ; this is discussed in section 5.3. In Section 5.4 the critical region and the power of the tests concerning the parameter θ are given. A general selection procedure to identify the best population, or a subset of the best populations in the Transformed Chi-square family with probability of correct identification, has been discussed in section 5.5.

5.2 The Transformed Chi-square Family

The Transformed Chi-square family is a sub-family of the Exponential family of distributions. Let X be a continuous type random variable having p.d.f. of the form (5.1.1) then under certain conditions $-2a(X)b(\theta)$ will follow a Central Chi-square distribution with appropriate degrees of freedom. We prove this result in the following theorem.

Theorem 5.2.1. In a one parameter Exponential family of the form (5.1.1) iff

$$2c'(\theta)b(\theta)/b'(\theta) = k \tag{5.2.1}$$

where k is positive and free from θ, then $-2a(X)b(\theta)$ follows a Gamma distribution with parameters k/2 and 1/2 . In case k is an integer then $-2a(X)b(\theta)$ follows a Central Chi-square distribution with k degrees of freedom.

Proof :

Let $2c'(\theta)b(\theta)/b'(\theta) = k$.

Since $\exp[a(x)b(\theta)+c(\theta)+h(x)]$ is a p.d.f. we must have

$\int\exp[a(x)b(\theta)+c(\theta)+h(x)]\,dx = 1$,

or, $\int\exp[a(x)b(\theta)+h(x)]\,dx = \exp[-c(\theta)]$ (5.2.2)

We have from (5.2.1),

$c'(\theta) = (1/2)kb'(\theta)/b(\theta)$

Integrating both sides of this equation with respect to θ we get,

$c(\theta) = (1/2)k \ln b(\theta) + k_1$,

where k_1 is a constant of integration.

Thus (5.2.2) becomes

$$\int \exp[a(x)b(\theta)+h(x)]\,dx = \exp[-(1/2)k\ln b(\theta) - k_1]$$

Let $U = -2a(X)b(\theta)$. The characteristic function of U is

$\varphi_u(t) = E\{\exp(itU)\} = E[\exp\{-2it\ a(X)b(\theta)\}]$

$= \exp\{c(\theta)\}\int\exp[a(x)b(\theta)(1-2it)+h(x)]dx$

$= \exp\{c(\theta)\}\ \exp[(-1/2)k\ \ln\{b(\theta)(1-2it)\} - k_1] = (1-2it)^{-k/2}$

which is the characteristic function of a Gamma distribution with parameters k/2 and 1/2. As the characteristic function uniquely determines the distribution function, $-2a(X)b(\theta)$ follows a Gamma distribution with parameters k/2 and 1/2.

Conversely, let $-2a(X)b(\theta) = Y$ be a Gamma variate with parameters k/2 and 1/2. Then the p.d.f. of Y is

$\{\exp(-y/2)\}y^{k/2-1}/\{2^{k/2}\Gamma(k/2)\}$.

Thus, the p.d.f. of a(X) is

$[\exp\{a(x)b(\theta)\}]\{-2b(\theta)\}\{-2a(x)b(\theta)\}^{k/2\ -\ 1}/\{2^{k/2}\Gamma(k/2)\}$

$= [\exp\{a(x)b(\theta)\}]\{-b(\theta)\}^{k/2}\{a(x)\}^{k/2\ -\ 1}/\{\Gamma(k/2)\}$

$= \exp[a(x)b(\theta) + (k/2)\ln\{-b(\theta)\} + (k/2 -1)\ln\{a(x)\} - \ln\{\Gamma(k/2)\}]$,

which is of the form (5.1.1). This implies that the distribution of X belongs to the Exponential family.

Here, $c(\theta) = (k/2)\ln\{-b(\theta)\}$

or, $2c'(\theta)b(\theta)/b'(\theta) = k$.

It is also evident that if k is an integer then $-2a(X)b(\theta)$ follows a Central Chi-square distribution with k degrees of freedom.

Hence the theorem is proved.

Example 5.2.1. Let X be an Exponential variate with p.d.f.

$$f(x,\theta)= \theta\exp[-\theta x] .$$

Here, $a(X)=X$, $b(\theta)= -\theta$, $c(\theta)= \ln\theta$, $2c'(\theta)b(\theta)/b'(\theta) = 2$,

$$-2a(X)b(\theta) = 2X\theta .$$

Thus $2X\theta$ is distributed as a Central Chi-square with 2 d. f.

Table 5.1 gives the different expressions of the functions such as the p.d.f., $a(X)$, $b(\theta)$, $c(\theta)$, $-2a(X)b(\theta)$ and the values of $k= 2c'(\theta)b(\theta)/b'(\theta)$ for the Normal, Lognormal, Gamma, Exponential, Rayleigh, Pareto, Weibull, Erlang, Maxwell and Inverse Gaussian distributions.

DEFINITION 5.2.1. A sub-family of the one parameter Exponential family, having p.d.f. of the form (5.1.1) and satisfying (5.2.1), will be called a family of Transformed Chi-square distributions, provided that k is a positive integer.

The Normal, Lognormal, Gamma, Exponential, Rayleigh, Pareto, Weibull, Erlang, Maxwell, Inverse Gaussian, etc. are distributions that belong to the Transformed Chi-square family. However, all continuous distributions belonging to the Exponential family are not members of the Transformed Chi-square family. This may be seen from the following example.

Example 5.2.2. Let X be a random variable having density

$f(x,\theta) = \theta k^{-\theta} x^{\theta-1}$, $0 < x < k$, $\theta > 0$, k being known.

Clearly, the distribution of X belongs to the Exponential family with $a(X) = \ln X$, $b(\theta) = \theta - 1$ and $c(\theta) = \ln\theta - \theta \ln k$.

Here, $2c'(\theta)b(\theta)/b'(\theta) = 2(\theta-1)(1/\theta - \ln k)$, which is a function of θ. Therefore, this distribution does not belong to the Transformed Chi-square family.

5.2.1 Moments of the Distribution of a(X) in the Family of Transformed Chi-square Distributions

The characteristic function of $U = -2a(X)b(\theta)$ is

$\varphi_U(t) = (1-2it)^{-k/2}$.

Therefore, the moment generating function and the cumulant generating function are

$M_u(t) = (1-2t)^{-k/2}$ and

$\kappa(t) = \ln M_u(t) = -(k/2)\ln(1-2t)$ respectively.

$\kappa(t)$ *can* be written as

$\kappa(t) = kt/1! + 2kt^2/2! + 8kt^3/3! + 48kt^4/4! + \ldots\ldots.$

The rth cumulant k_r is the co-efficient of $t^r/r!$.

Thus $\kappa_1 = k$, $\kappa_2 = 2k$, $\kappa_3 = 8k$, $\kappa_4 = 48k$.

$E\{a(X)\} = -k/\{2b(\theta)\}$,

$V\{a(X)\} = k/\{2b^2(\theta)\} = \mu_2$, (5.2.3)

$\mu_3 = -k/b^3(\theta)$,

$\mu_4 = \{12k + 3k^2\}/\{4b^4(\theta)\}$,

$\beta_1 = 8/k$,

$\beta_2 = 3 + 12/k$.

Hence the distribution of a(X) is positively skewed and leptokurtic.

Example 5.2.3. Let a random variable X follow a Pareto distribution with p.d.f. $f(x,\theta) = \theta x^{-(\theta+1)}$, $x > 0$.

Here, $a(X) = \ln X$, $b(\theta) = -\theta$, $c(\theta) = \ln\theta$, $k = 2c'(\theta)b(\theta)/b'(\theta) = 2$.

Therefore,

$$E\{a(X)\} = -k/2b(\theta) = 1/\theta$$

$$V\{a(X)\} = k/\{2b^2(\theta)\} = 1/\theta^2$$

$$\mu_3 = -k/b^3(\theta) = 2/\theta^3$$

$$\mu_4 = \{12k + 3k^2\}/\{4b^4(\theta)\} = 9/\theta^4$$

$$\beta_1 = 8/k = 4$$

$$\beta_2 = 3 + 12/k = 9$$

Corollary 5.2.1. If X_j is distributed with p.d.f. $f(x_j,\theta_j) = \exp[a_j(x_j)b_j(\theta_j)+c_j(\theta_j)+h_j(x_j)]$ and satisfies the conditions $2c'_j(\theta_j)b_j(\theta_j)/b_j'(\theta_j) = k_j$, k_j being a positive integer, where $j=1,2,...,r$ and $X_1,X_2,...,X_r$ are independent, then $-2\Sigma a_j(X_j)b_j(\theta_j)$ follows a Central Chi-square distribution with Σk_j d.f.

Proof : Let $U_j = -2a_j(X_j)b_j(\theta_j)$.

It follows from theorem 5.2.1 that U_j is a Central Chi-square variate with k_j d.f. ($j= 1,2,...,r$) and the characteristic function of U_j is $(1-2it)^{-k_j/2}$.

Since $X_1,X_2,...,X_r$ are independent, $-2\Sigma a_j(X_j)b_j(\theta_j) = \Sigma U_j$ is the sum of r independent Chi-square variates. Hence by the additive property of Chi-square variates, $\Sigma U_j = -2\Sigma a_j(X_j)b_j(\theta_j)$ follows a Central Chi-square distribution with Σk_j degrees of freedom.

Example 5.2.4. Let X_1 and X_2 be independently distributed with p.d.f. $\theta_1\exp(-x_1\theta_1)$ and $\{\exp[-(\{x_2\}^2)/(2\{\theta_2\}^2)]/\{\theta_2\sqrt{(2\pi)}\}$ respectively.

Here, $a_1(X_1) = X_1$, $b_1(\theta_1) = -\theta_1, c_1(\theta_1) = \ln\theta_1$, $a_2(X_2) = \{X_2\}^2$,

$b_2(\theta_2) = -1/(2\{\theta_2\}^2)$, $c_2(\theta_2) = -\ln\theta_2$.

Thus $-2\Sigma a_i(X_i)b_i(\theta_i) = 2X_1\theta_1 + (X_2/\theta_2)^2$ follows a Central Chi-square distribution with 3 d.f.

Corollary 5.2.2. Let X be a random variable having density

$$f(x,\theta)=x^{r-1}[\exp\{-x^r/(r\theta^r)\}]/\theta^r, \quad x>0, \theta>0, \qquad (5.2.3.1)$$

r being a positive integer. Then, for any value of r, $2X^r/\{r\theta^r\}$

follows a Central Chi-square distribution with 2 d.f.

Proof : The characteristic function of $2X^r/(r\theta^r)$ is

$\varphi(t) = E[\exp(2itX^r/\{r\theta^r\})]$

$\quad = (1-2it)^{-2/2}$,

which is the characteristic function of a Central Chi-square distribution with 2 d.f. Hence, for any value of r, $2X^r/\{r\theta^r\}$ follows a Central Chi-square distribution with 2 d.f. °

Example 5.2.5. Let a random variable X follow a Rayleigh distribution with p.d.f. $f(x,\theta) = x[\exp\{-x^2/(2\theta^2)\}]/\theta^2$, $x > 0$.
Here, r=2 and $2X^r/(r\theta^r) = X^2/\theta^2$ follows a Central Chi-square distribution with 2 d.f.

5.2.2 Moments of the Distribution of X^r in (5.2.3.1)

Let $W=2X^r/(r\theta^r)$. The moment generating function of W is $(1-2t)^{-1}$ and the cumulant generating function is

$\kappa(t) = -\ln(1-2t) = 2t/1! + 4t^2/2! + 16t^3/3! + 96t^4/4! + \ldots\ldots$

Thus $E(X^r) = r\theta^r = \mu'_1$, $V(X^r) = (r\theta^r)^2 = \mu_2$,

$\mu_3 = 2r^3\theta^{3r}$, $\mu_4 = 9r^4\theta^{4r}$, $\beta_1 = 4$, $\beta_2 = 9$.

Hence the distribution of X^r is always positively skewed and leptokurtic.

Example 5.2.6. Let a random variable X follow a Rayleigh distribution with p.d.f. $f(x,\theta) = x[\exp\{-x^2/(2\theta^2)\}]/\theta^2$, $x > 0$.
Here, r=2 and $E(X^2)=2\theta^2$, $V(X^2)=4\theta^4$, $\mu_3=16\theta^6$, $\mu_4=144\theta^8$, $\beta_1=4$, $\beta_2=9$.

Corollary 5.2.4. If X is a random variable having p.d.f. of the form (5.2.3.1), then for any value of r, X^r is distributed as an Exponential Variate.

Proof: The characteristic function of X^r is

$$\varphi(t) = E[\exp(itX^r)]$$
$$= \int \exp(itx^r)\ x^{r-1}\ [\exp\{-x^r/(r\theta^r)\}]/\theta^r\ dx$$

Let $y = x^r/(r\theta^r)$

Therefore, $(x^{r-1}/\theta^r)dx = dy$ and

$$\varphi(t) = \int \exp(-y\ +ityr\theta^r)\ dy$$
$$= \int \exp\{-y(1 - itr\theta^r)\}\ dy$$
$$= (1 - itr\theta^r)^{-1}$$

Which is the characteristic function of an Exponential distribution with mean $r\theta^r$. Hence, X^r is distributed as an Exponential Variate with mean $r\theta^r$.

Example 5.2.7 Let a random variable X follow a Rayleigh distribution with p.d.f. $f(x,\theta) = x[\exp\{-x^2/(2\theta^2)\}]/\theta^2$, $x > 0$.
Here, r=2 . Let us make the transformation, $w=x^2$.

The Jacobian of the transformation is $J = (1/2)w^{-1/2}$.

Therefore, the p.d.f. of w is $g(w,\theta) = (1/2\theta^2)\exp(w/2\theta^2)$.

=> $w = x^2$ follows Exponential distribution with mean $2\theta^2$.

5.3 Estimation of Parameters in the Transformed Chi-square Family

In general there are two types of estimation, Point estimation and Interval estimation. These are discussed below.

5.3.1 Point Estimation

There are various methods of point estimation. Most commonly used point estimators are MLE, MVUE and MVBUE.

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a population having p.d.f. (5.1.1) and satisfying (5.2.1). Then the likelihood function of the sample observations is given by

$$L = \exp\{b(\theta)\Sigma a(x_i) + nc(\theta) + \Sigma h(x_i)\}.$$

or, $\ln L = b(\theta)\Sigma a(x_i) + nc(\theta) + \Sigma h(x_i)$.

Differentiating partially with respect to θ and setting this partial derivative to zero we get,

$b'(\theta)\Sigma a(x_i) + nc'(\theta) = 0.$

or, $-k/\{2b(\theta)\} = \Sigma a(x_i)/n$.

Therefore, $\Sigma a(X_i)/n$ is the MLE of $-k/\{2b(\theta)\}$.

The MVBUE of a function of θ is given in the following theorem.

Theorem 5.3.1. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from (5.1.1) satisfying (5.2.1). Then under the Cramér-Rao regularity conditions $\Sigma a(X_i)/n$ is the MVBUE of $[-k/\{2b(\theta)\}]$ with MV $[k/\{2nb^2(\theta)\}]$.

Proof : The log-likelihood function of the sample observations is given by $\ln L = b(\theta)\Sigma a(x_i) + nc(\theta) + \Sigma h(x_i)$. Differentiating partially with respect to θ and θ^2, we get,

$\partial \ln L/\partial\theta = b'(\theta)\Sigma a(x_i) + nc'(\theta)$ and

$$\partial^2 \ln L/\partial\theta^2 = b''(\theta)\Sigma a(x_i) + nc''(\theta) \qquad (5.3.1)$$

Using (5.2.3),

$E\{\Sigma a(X_i)/n\} = -k/\{2b(\theta)\} = \psi(\theta)$ (say) and $V\{\Sigma a(X_i)/n\} = k/\{2nb^2(\theta)\}$.

Taking expectations on both sides of (5.3.1) and simplifying we get,

$-E(\partial^2 \ln L/\partial\theta^2) = n\{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)\}/b'(\theta)$

Hence, the Cramér-Rao lower bound (CRLB) for an unbiased estimator of $\psi(\theta)$ is $\{\psi'(\theta)\}^2/-E(\partial^2 \ln L/\partial\theta^2)$.

As $V\{a(X)\} = \{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)\}/\{b'(\theta)\}^3$

$= k/\{2b^2(\theta)\}$ [Dobson(1983)],

thus, $-E(\partial^2 \ln L/\partial\theta^2) = nk\{b'(\theta)\}^2/\{2b^2(\theta)\}$ and the CRLB is

$\{kb'(\theta)/[2b^2(\theta)]\}^2 \,.\, 2\{b(\theta)\}^2/[nk\{b'(\theta)\}^2] = k/\{2nb^2(\theta)\} = V\{\Sigma a(X_i)/n\}$.

Hence the theorem is proved.

The UMVUE of a function of θ is given in the following theorem.

Theorem 5.3.2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from (5.1.1) and satisfying (5.2.1). Then $\Sigma a(X_i)/n$ is the UMVUE of [-

$k/\{2b(\theta)\}]$ with MV $[k/\{2nb^2(\theta)\}]$.

Proof : The likelihood function of the sample observations is given by $L = \exp\{b(\theta)\Sigma a(x_i) + nc(\theta) + \Sigma h(x_i)\}$.

According to the Neyman-Fisher Factorization theorem, $\Sigma a(X_i)$ is a sufficient statistic. It is evident from theorem 1, page 142 of Lehmann (1986) that $\Sigma a(X_i)$ is also a complete statistic. Then according to the Lehmann-Scheffé (1950,1955) theorem, $\Sigma a(X_i)$ is the unique uniformly minimum variance unbiased estimator [UMVUE] of its expected value.

Since $E\{\Sigma a(X_i)/n\} = -k/\{2b(\theta)\}$ and $V\{\Sigma a(X_i)/n\} = k/\{2nb^2(\theta)\}$.

Therefore, $\Sigma a(X_i)/n$ is the unique UMVUE of $[-k/\{2b(\theta)\}]$ with MV $[k/\{2nb^2(\theta)\}]$.

Example 5.3.1. For a random sample $X_1, X_2, \ldots, X_n$ of size n from a Rayleigh density, the likelihood function is

$L(\theta,\underline{x}) = \exp[-\Sigma x_i^2/(2\theta^2) - 2n\ln\theta + \Sigma \ln x_i]$.

Here, $a(X_i) = X_i^2$, $b(\theta) = -1/(2\theta^2)$,

$k = 2$, $-k/\{2b(\theta)\} = 2\theta^2$ and $k/\{2nb^2(\theta)\} = 4\theta^4/n$.

Thus, $\Sigma X_i^2/2n$ is the MLE, MVBUE and UMVUE of θ^2 with MV θ^4/n.

5.3.2 Interval Estimation

A general method of constructing an ordinary confidence interval and the shortest confidence set is given below.

(a) Confidence interval by pivotal method :

For a random sample $X_1, X_2, \ldots, X_n$ of size n from (5.1.1) satisfying (5.2.1), we consider $-2b(\theta)\Sigma a(X_i)$ as a pivot, the distribution of which is independent of θ . Let k be an integer. Choose two values $\chi^2_{\alpha_1}$ and $\chi^2_{1-\alpha_2}$ such that

$$P\{\chi^2_{\alpha_1} < -2\sum_{i=1}^{n} a(X_i)b(\theta) < \chi^2_{1-\alpha_2}\} = 1-(\alpha_1+\alpha_2)$$

or, $P\{\, t_1(\underset{\sim}{X}) < \theta < t_2(\underset{\sim}{X})\,\} = 1-\alpha$, where $\alpha_1+\alpha_2=\alpha$.

Hence $\{\, t_1(\underset{\sim}{X})\ ,\ t_2(\underset{\sim}{X})\,\}$ is a $100(1-\alpha)\%$ confidence interval for θ .

Example 5.3.2. Let $X_1, X_2, \ldots, X_n$ be a random sample from an Exponential distribution with mean $1/\theta$. Then the likelihood function is given by $L(\theta,\underset{\sim}{x}) = \exp[-\Sigma x_i/\theta - n\ln\theta]$ and

$-2\Sigma a(X_i)b(\theta) = 2\Sigma X_i/\theta$ is a Central Chi-square variate with 2n d.f. Thus,

$$P\{\chi^2_{2n,\alpha_1} < \frac{2\Sigma X_i}{\theta} < \chi^2_{2n,(1-\alpha_2)}\} = 1-\alpha, \quad \text{where } \alpha_1+\alpha_2=\alpha.$$

$$\text{or, } P\left\{\frac{2\Sigma X_i}{\chi^2_{2n,(1-\alpha_2)}} < \theta < \frac{2\Sigma X_i}{\chi^2_{2n,\alpha_1}}\right\} = 1-\alpha$$

$$\text{Hence, } \left(\frac{2\Sigma X_i}{\chi^2_{2n,(1-\alpha_2)}}\ ,\ \frac{2\Sigma X_i}{\chi^2_{2n,\alpha_1}}\right)$$

is a $100(1-\alpha)\%$ confidence interval for θ .

(b) Shortest confidence set :

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from (5.1.1) and satisfying (5.2.1). Let $b(\theta)$ be strictly increasing in θ and let k be an integer. Then an α level UMP test for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ exists with the critical region

$$W_\alpha = \{\chi^2 : \Sigma a(X_i) \geq [\chi^2_{nk,(1-\alpha)}/\{-2b(\theta_0)\}]\}.$$

Let $W_\alpha{}^c$ be the region complementary to W_α.

Then by the result 7b.2.1 of Rao (1973) the 100(1- α)% shortest confidence set for θ is $I(\chi^2) = \{\theta_0 : \chi^2 \varepsilon W_\alpha{}^c\}$

$$\text{or,} \quad I(\chi^2) = \{\theta_0 : \Sigma a(X_i) \leq [\chi^2_{nk,(1-\alpha)}/\{-2b(\theta_0)\}]\}.$$

Similarly, an α level UMP test for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ exists and the 100(1- α)% shortest confidence set for θ is of the form $J(\chi^2) = \{\theta_0 : \Sigma a(X_i) \geq [\chi^2_{nk,\alpha}/\{-2b(\theta_0)\}]\}.$

According to lemma 1 of Lehmann (1986, page-135), there also exits an α level UMP unbiased test for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ given by

$\phi(\underline{x}) = 1$, if $\Sigma a(x_i) < c_1$ and $\Sigma a(x_i) > c_2$,

$\phi(\underline{x}) = 0$, otherwise,

where the constants c_1 and c_2 are determined by

$$E_{\theta_0}\{\phi(X)\} = \alpha \quad \text{and}$$

$$E_{\theta_0}\{\phi(X)\sum_{i=1}^{n} a(X_i)\} = E_{\theta_0}\{\phi(X)\}\,\alpha$$

Here, the critical region is $W_\alpha = \{\chi^2 : \Sigma a(X_i) < c_1,\ \Sigma a(X_i) > c_2\}$.

Let $W_\alpha^{\ c}$ be the region complementary to W_α.

Then by the result 7b.2.1 of Rao (1973) the $100(1-\alpha)\%$ shortest confidence set for θ is $I_1(\chi^2) = \{\theta_0 : \chi^2 \varepsilon W_\alpha^{\ c}\}$.

5.4 Tests of Hypotheses in the Transformed Chi-square Family

Let $X_1, X_2, \ldots, X_n$ be a random sample from (5.1.1) satisfying (5.2.1). Let $b(\theta)$ be strictly increasing in θ and let k be an integer. Then by theorem 6.2.1 of Bickel & Doksum(1976), $\Sigma a(X_i)$ is an optimal test statistic for testing

$H_0 : \theta \leq \theta_0$ against

$H_1 : \theta > \theta_0$

and an α level test is

$\phi(\underset{\sim}{x}) = 1$, if $\Sigma a(x_i) \geq c$ and

$\phi(\underset{\sim}{x}) = 0$, otherwise,

where c is determined by

$E_\theta\{\phi(\underset{\sim}{X})\} = \alpha$, or, $P\{-2b(\theta_0)\Sigma a(X_i) \geq -2b(\theta_0)c\} = \alpha$

or, $P\{\chi^2_{nk} \geq -2b(\theta_0)c\} = \alpha$.

Let $-2b(\theta_0)c = \chi^2_{nk,(1-\alpha)}$.

Thus $c = [\chi^2_{nk,(1-\alpha)}]/\{-2b(\theta_0)\}$ and the power function is

$E_\theta\{\phi(X)\} = P_\theta\{\Sigma a(X_i) \geq c\} = P_\theta\{\chi^2_{nk} \geq [b(\theta)\,\chi^2_{nk,(1-\alpha)}\}/b(\theta_0)]\}$.

Here, the word optimal is used in the sense of UMP. Other UMP tests are given in section 5.3.

Example 5.4.1. Let $X_1, X_2, \ldots, X_n$ be a random sample from $G(p,\theta)$. Then the likelihood function is

$$L(\theta,\underline{x}) = [(x_1 x_2 \ldots x_n)^{p-1} \exp\{-\Sigma x_i/\theta\}]/[\theta^{np}\{(p-1)!\}^n],$$

where $\theta>0$, $x_i>0$ and p is known. We want to test,

$H_0 : \theta \leq \theta_0$ against

$H_1 : \theta > \theta_0$.

Here $b(\theta) = -1/\theta$ is strictly increasing in θ. Hence an α level optimal test is

$\phi(\underline{x}) = 1$, if $\Sigma x_i \geq c$ and

$\phi(\underline{x}) = 0$, otherwise.

Here, $k = 2p$, $b(\theta) = -1/\theta$, $c = (1/2)\theta_0\,\chi^2_{2np,(1-\alpha)}$, $b(\theta)/b(\theta_0) = \theta_0/\theta$ and $2\Sigma X_i/\theta$ is distributed as a Central Chi-square with 2np d.f. Therefore, the power function is $P\{\chi^2_{2np} \geq (\theta_0/\theta)\,\chi^2_{2np,(1-\alpha)}\}$.

If $\theta = \theta_0$, then power is α and

if $\theta > \theta_0$, then power $\geq \alpha$.

This implies that the test is also unbiased.

5.5 Identification of the Best Population in the Transformed Chi-square Family

Bechhofer (1954) and Bechhofer et al. (1955) introduced a single sample multiple decision procedure for ranking means and variances of normal populations respectively. Similar procedures are discussed by many authors for other populations. A general identification procedure and the probability of correct identification of the best population or subset of populations in Transformed Chi-square family are given below.

Let X_{ij} $(i=1,2,...,p\ ;\ j=1,2,...,N_i;\ \Sigma N_i=N)$ be independently distributed with p.d.f. $f(x_{ij},\theta_i)=\exp[a(x_{ij})b(\theta_i)+c(\theta_i)+h(x_{ij})]$ and satisfying the conditions $2c'(\theta_i)b(\theta_i)/b'(\theta_i) = k_i$, where θ_i are unknown. Let $\theta_{[1]} \leq \theta_{[2]} \leq \theta_{[3]} \leq ... \leq \theta_{[p]}$ be the ranked θ_i ; it is assumed that we do not know which population is associated with $\theta_{[i]}$, i=1,2,. . ,p. Let us assume that a population is characterized by the value of the parameter θ, with the 'best' population being the one having the largest θ , the 'second best' being the one having the second largest θ, and so on. On the other hand, we may define the 'best' population as being the one having the smallest θ, the 'second best' being the one having the second smallest θ, and so on.

However, the mathematical theory is the same for both cases. The p populations may be the populations of p different cities or counties and $\psi_i(\theta)$ may be the average income of the people of the ith city or county, where $\psi_i(\theta)$ is a one to one function of θ. For example, if the income follows a Pareto distribution, then the average income will be of the form $\psi(\theta) = \theta/(1+\theta)$. We are interested in identifying a population having the largest average income, which is equivalent to identifying a population with the largest θ. The p populations may be p different telephone exchanges and $\psi_i(\theta)$ may be the average time interval between two successive calls at the ith exchange. The time interval between two successive calls follows the Exponential distribution with mean θ. The telephone company may be interested in identifying the telephone exchange earning the maximum profit or the minimum profit. We would like, on the basis of a sample of $N=\Sigma N_i$ independent observations, to make some inferences about the true ranking of the populations. Our inferences will be based on the sample estimates of some function of θ. The MVBUE of $-k_i/2b(\theta_i)$ is $\Sigma a(x_{ij})/N_i = a_i$ (say) for the ith population. Let the sample estimate and sample size associated with the population having population parameter $\theta_{[i]}$ be denoted by $a_{(i)}$ and $N_{(i)}$ respectively, i=1,2, . . ., p; that is the expected value of $a_{(i)}$ is $\psi(\theta_{[i]})$.

The ranked a_i are denoted by

$$a_{[1]} < a_{[2]} < \dots < a_{[p]} \tag{5.5.1}$$

The event $a_i = a_j$ $(i \neq j)$ has probability zero and can be ignored in probability calculations. However, in practical situations this event can occur frequently because of the limitations of the measuring instrument of any experiment. If two or more a_i are equal, they should be ranked by using a randomized procedure which assigns equal probability to each ordering. Let us assume that

$$\alpha_{ij} = b(\theta_{[i]})/b(\theta_{[j]}), \ (i,j = 1,2, \ldots, p) .$$

Goals : Different goals are appropriate for different practical situations. In each situation it is the experimenter's responsibility to decide what the goal is before taking a sample.

For example, the goal may be to find any of the following :-

(i) The best population

(ii) The best two populations with regard to order

(iii) The best two populations without regard to order

(iv) The best three populations with regard to order

(v) The best three populations without regard to order, and so on.

The choice of a goal may depend on economic and other considerations outside the control of the statistician. These goals are the special cases of the following two representative goals.

Goal 1: To divide the p populations into two groups, the t best(unordered) and p-t worst(unordered) populations, where $1 \le t \le (p-1)$.

Goal 2: To divide the p populations into t+1 groups, the t best(ordered) and the p-t worst(unordered) populations, where $1 \le t \le (p-1)$.

It is obvious that, for Goal 1, the problem of choosing the t best is equivalent to choosing the p-t worst. On the other hand, for Goal 2, if $t=p-1$, then we need a complete ranking. The two goals coincide for $t=1$.

Assumptions: For goal 1 it is assumed that the experimenter can specify a smallest value of $\alpha_{t+1,t}$, say $\alpha^*_{t+1,t}$, that he desires to detect. He also must specify the smallest acceptable probability of achieving Goal 1 when $\alpha_{t+1,t} \geq \alpha^*_{t+1,t}$.

For Goal 2 it is assumed that the experimenter can specify a smallest value of each $\alpha_{i+1,i}$, say $\alpha^*_{i+1,i}$ (i = p-t+1, p-t+2, . . . , p) that he desires to detect. The experimenter also must specify the smallest acceptable probability of achieving Goal 2 when $\alpha_{i+1,i} \geq \alpha^*_{i+1,i}$ (i=p-t+1, p-t+2, . . ., p).

5.5.1 Identification Procedure

Having chosen a goal, the statistical procedure is elementary. We take a random sample of N_i observations from the ith population (i=1,2, . . ,p). Then we compute the p-statistics $a_1, a_2, \ldots, a_p$ and arrange them in ascending order of magnitude like (5.5.1). We then take the decision as follows :

If our goal is (i), the population associated with $a_{[p]}$ is the best population. If our goal is (ii), the populations associated with $a_{[p]}$ and $a_{[p-1]}$ are the best and second best populations respectively. If our goal is to find (iii), (iv), etc., we can make similar statements. In general, for Goal 1 the populations that give rise to the t largest a_i are the t best populations and the p-t remaining populations are the worst populations. For Goal 2, the populations that give rise to the largest, second largest, . . ., t-th largest a_i are the best, second best, . . ., t-th best populations respectively, and the remaining p-t populations are the worst populations.

5.5.2 Probability of Correct Identification

To calculate the probability of correct identification, we must first state our goal. A general goal can be expressed as follows :- To find the p_s best populations, the p_{s-1} second best populations, the p_{s-2} third best populations, etc., and finally the p_1 worst populations. Here $p_1, p_2, \ldots, p_s$ ($s \le p$) are positive integers such that $\Sigma p_i = p$. The probability of a correct ranking associated with this can be written as :

$$
\begin{aligned}
P[&\max\{a_1, a_2, \ldots, a_{p_1}\} < \min\{a_{p_1+1}, a_{p_1+2}, \ldots, a_{p_1+p_2}\},\\
&\max\{a_{p_1+1}, a_{p_1+2}, \ldots, a_{p_1+p_2}\} < \min\{a_{p_1+p_2+1}, a_{p_1+p_2+2}, \ldots, a_{p_1+p_2+p_3}\}, \ldots,\\
&\max\{a_{p-p_s-p_{s-1}+1}, \ldots, a_{p-p_s}\} < \min\{a_{p-p_s+1}, a_{p-p_s+2}, \ldots, a_p\}]
\end{aligned}
$$

If we assign particular values to s and p_i we obtain several special cases of interest. For example, for $s = 2$; $p_1 = p-t$, $p_2 = t$, we have

$$P[\max\{a_{(1)}, a_{(2)}, \ldots, a_{(p-t)}\} < \min\{a_{(p-t+1)}, a_{(p-t+2)}, \ldots, a_{(p)}\}] \qquad (5.5.2)$$

For $s = t+1$; $p_1 = p-t$ and $p_2 = p_3 = \ldots = p_{t+1} = 1$ we have

$$P[\max\{a_{(1)}, a_{(2)}, \ldots, a_{(p-t)}\} < a_{(p-t+1)} < a_{(p-t+2)} < \ldots < a_{(p)}\}] \qquad (5.5.3)$$

and for $s = p$; $p_1 = p_2 = \ldots = p_t = 1$, we have

$$P\{a_{(1)} < a_{(2)} < \ldots < a_{(p)}\}.$$

Thus (5.5.2) is the probability that the best population will yield the largest sample statistic a_i, then (5.5.2) for t=1, 2, 3 is the probability of a correct ranking associated with (i), (iii), (v) respectively. Also (5.5.3) for $t = 2, 3$ is the probability of a correct ranking associated with (ii), (iv) respectively. Evidently (5.5.2) and (5.5.3) represent the probabilities of correct ranking for Goals 1 and 2 respectively. The expression (5.5.2) can be written as

$$\sum_{j=1}^{p-t} P[\max\{a_{(1)},\ldots,a_{(j-1)},a_{(j+1)},\ldots,a_{(p-t)}\} < a_{(j)} < \min\{a_{(p-t+1)},\ldots,a_{(p)}\}]$$

$$= \sum_{j=1}^{p-t} P\left\{a_{(i)} < a_{(j)} < a_{(l)}\ ;\ \begin{array}{l} i=1,2,\ldots,j-1,j+1,\ldots,p-t \\ l=p-t+1,\,p-t+2,\ldots,p \end{array}\right\}$$

$$= \sum_{j=1}^{p-t} P\left\{\begin{array}{ll} a_{(i)} < a_{(j)} & ;\ i=1,2,\ldots,j-1,j+1,\ldots,p-t \\ a_{(l)} > a_{(j)} & ;\ l=p-t+1,p-t+2,\ldots,p \end{array}\right\} \qquad (5.5.4)$$

By theorem 5.2.1, $-2\Sigma a(X_{ij})\, b(\theta_{[i]}) = -2N_{(i)}\, a_{(i)}\, b(\theta_{[i]}) = U_{(i)}$ (say) is distributed as a Gamma variate with parameters $k_{(i)}/2$ and 1/2. Therefore, (5.5.4) can be written as

$$\sum_{j=1}^{p-t} P\left[\begin{array}{ll} U_{(i)} < \dfrac{N_{(i)}}{N_{(j)}}\alpha_{ij}\, U_{(j)} & ;\ i=1,2,\ldots,j-1,j+1,\ldots,p-t \\ U_{(l)} > \dfrac{N_{(l)}}{N_{(j)}}\alpha_{lj}\, U_{(j)} & ;\ l=p-t+1,p-t+2,\ldots,p \end{array}\right]$$

If for each j the above probability is evaluated for $U_{(j)}$ fixed (say at u), and the expectation is taken over u, then (5.5.4) can be written as

$$\sum_{j=1}^{p-t}\int_0^{\infty}\Big[\prod_{\substack{i=1\\ i\neq j}}^{p-t} F_j\Big(\frac{N_{(i)}}{N_{(j)}}\,\alpha_{ij}\,u\Big)\Big]\Big[\prod_{l=p-t+1}^{p}\Big\{1-F_j\Big(\frac{N_{(l)}}{N_{(j)}}\,\alpha_{lj}\,u\Big)\Big\}\Big]\,f_j(u)\,du \qquad (5.5.5)$$

where $f_j(u)$ and $F_j(u)$ are the probability density function and cumulative distribution function, respectively, of the Gamma variable U with parameters $k_{(j)}/2$ and 1/2. The probability (5.5.2) can be evaluated for arbitrary values of N_i and α_{ij} (i,j=1,2, . . . , p) using (5.5.5). If $N_1 = N_2 = \ldots = N_p = n$ (say) then (5.5.5) will be of the following form

$$\sum_{j=1}^{p-t}\int_0^{\infty}\Big[\prod_{\substack{i=1\\ i\neq j}}^{p-t} F_j(\alpha_{ij}\,u)\Big]\Big[\prod_{l=p-t+1}^{p}\Big\{1-F_j(\alpha_{lj}\,u)\Big\}\Big]\,f_j(u)\,du$$

The expression (5.5.3) can be written as

$$\sum_{j=1}^{p-t} P[\max\{a_{(1)},\ldots,a_{(j-1)},a_{(j+1)},\ldots,a_{(p-t)}\} < a_{(j)} < a_{(p-t+1)} < \ldots < a_{(p)}]$$

$$= \sum_{j=1}^{p-t} P[\,a_{(i)} < a_{(j)} < a_{(p-t+1)} < \ldots < a_{(p)}\ ;\ i=1,2,\ldots,j-1,j+1,\ldots,p-t]$$

$$= \sum_{j=1}^{p-t} P\begin{bmatrix} a_{(i)} < a_{(j)} & ; i = 1,2, \ldots, j-1, j+1, \ldots, p-t \\ a_{(p-t+1)} > a_{(j)} & \\ \vdots & \\ a_{(p)} > a_{(p-1)} & \end{bmatrix}$$

$$= \sum_{j=1}^{p-t} P\begin{bmatrix} U_{(i)} < \frac{N_{(i)}}{N_{(j)}}\alpha_{ij}\, U_{(j)} & ; i = 1,2, \ldots, j-1, j+1, \ldots, p-t \\ U_{(p-t+1)} > \frac{N_{(p-t+1)}}{N_{(j)}}\alpha_{p-t+1,j}\, U_{(j)} & \\ \vdots & \\ U_{(p)} > \frac{N_{(p)}}{N_{(p-1)}}\alpha_{p,p-1}\, U_{(p-1)} & \end{bmatrix} \qquad (5.5.6)$$

If $N_1 = N_2 = \ldots = N_p = n$ (say) then (5.5.6) can be written as

$$= \sum_{j=1}^{p-t} P\begin{bmatrix} U_{(i)} < \alpha_{ij}\, U_{(j)} & ; i=1,2, \ldots, j-1, j+1, \ldots, p-t \\ U_{(p-t+1)} > \alpha_{p-t+1,j}\, U_{(j)} & \\ \vdots & \\ U_{(p)} > \alpha_{p,p-1}\, U_{(p-1)} & \end{bmatrix}$$

On the other hand, if we define the 'best' population as being the one having the smallest θ, the 'second best' being the one having the second smallest θ and so on, then the probability of correct ranking for Goal 1 will be of the following form

$$P[\max\{a_{(1)}, a_{(2)},\ldots,a_{(t)}\} < \min\{a_{(t+1)}, a_{(t+2)},\ldots,a_{(p)}\}] \qquad (5.5.7)$$

Here s = 2, p_1 = t and p_2 = p-t.

For s = t+1; $p_1 = p_2 = \ldots = p_t = 1$ and p_{t+1} = p-t we have

$$P[a_{(1)} < a_{(2)} < \ldots < a_{(t)} < \min\{a_{(t+1)}, a_{(t+2)}, \ldots, a_{(p)}\}] \qquad (5.5.8)$$

Thus (5.5.7) is the probability that the best population will yield the smallest sample statistic a_i, then (5.5.7) for t =1, 2, 3 is the probability of a correct ranking associated with (i), (iii), (v) respectively. Also (5.5.8) for t = 2, 3 is the probability of a correct ranking associated with (ii), (iv) respectively. Evidently (5.5.7) and (5.5.8) represent the probabilities of correct ranking for Goals 1 and 2 respectively. The probability expression (5.5.7) can be written as

$$\sum_{j=1}^{t} P[\max\{a_{(1)}, \ldots,a_{(j-1)},a_{(j+1)}, \ldots,a_{(t)}\} < a_{(j)} < \min\{a_{(t+1)}, \ldots,a_{(p)}\}]$$

$$= \sum_{j=1}^{t} P[\max\{a_{(1)}, \ldots,a_{(j-1)},a_{(j+1)}, \ldots,a_{(t)}\} < a_{(j)} < \min\{a_{(t+1)}, \ldots,a_{(p)}\}]$$

$$= \sum_{j=1}^{t} P\left\{a_{(i)} < a_{(j)} < a_{(l)} \; ; \begin{array}{l} i=1,2, \ldots,j-1,j+1, \ldots, t \\ l=t+1,t+2, \ldots, p \end{array}\right\}$$

$$= \sum_{j=1}^{t} P\left\{\begin{array}{ll} a_{(i)} < a_{(j)} & ; \; i=1,2, \ldots,j-1,j+1, \ldots, t \\ a_{(l)} > a_{(j)} & ; \; l=t+1,t+2, \ldots, p \end{array}\right\} \qquad (5.5.9)$$

By theorem 5.2.1, $-2\sum a(X_{ij})\, b(\theta_{[i]}) = -2N_{(i)}\, a_{(i)}\, b(\theta_{[i]}) = U_{(i)}$ (say) is distributed as a Gamma variate with parameters $k_{(i)}/2$ and 1/2.

Therefore, (5.5.9) can be written as

$$\sum_{j=1}^{t} P\left[\begin{array}{ll} U_{(i)} < \frac{N_{(i)}}{N_{(j)}}\alpha_{ij}\, U_{(j)} & ; i = 1,2,\ldots,j-1,\ j+1,\ldots, t \\ U_{(l)} > \frac{N_{(l)}}{N_{(j)}}\alpha_{lj}\, U_{(j)} & ; l = t+1, \quad t+2, \ldots, p \end{array}\right]$$

If for each j the above probability is evaluated for $U_{(j)}$ (fixed say at u), and the expectation is taken over u, then (5.5.9) can be written as

$$\sum_{j=1}^{t} \int_0^{\infty} \Big[\prod_{\substack{i=1 \\ i\neq j}}^{t} F_j\Big(\frac{N_{(i)}}{N_{(j)}}\alpha_{ij}u\Big)\Big]\Big[\prod_{l=t+1}^{p}\Big\{1 - F_j\Big(\frac{N_{(l)}}{N_{(j)}}\alpha_{lj}u\Big)\Big\}\Big]\, f_j(u)\, du \qquad (5.5.10)$$

where $f_j(u)$ and $F_j(u)$ are the probability density function and cumulative distribution function, respectively, of the Gamma variable U with parameters $k_{(j)}/2$ and 1/2.

If $N_1 = N_2 = \ldots = N_p = n$ (say) then (5.5.9) will be of the following form

$$\sum_{j=1}^{t} \int_0^{\infty} \Big[\prod_{\substack{i=1 \\ i\neq j}}^{t} F_j(\alpha_{ij}\, u)\Big]\Big[\prod_{l=t+1}^{p}\{1-F_j(\alpha_{lj}\, u)\}\Big]\, f_j(u)\, du$$

Example 5.5.1. Let X_{ij} be iid random variables distributed as $N(\mu_i, \sigma_i^2)$, $(i=1,2, \ldots, p;\ j=1,2, \ldots, N_i)$. We assume that the μ_i's are known and that the σ_i^2's are unknown. Let $\sigma_{[1]}^2 \leq \sigma_{[2]}^2 \leq \ldots \leq \sigma_{[p]}^2$ be the ranked σ_i^2. Suppose that it is not known which population is associated with $\sigma_{[i]}^2$. We further assume that for the ith population, the only parameter of interest is the population variance σ_i^2. The 'best' population being the one having the smallest variance, the 'second best' being the one having the second smallest variance, etc. The p populations may be p different measuring instruments and σ_i^2 may be the population variance of measurement of the ith instrument. This variance, which characterises the reproducibility of repeated measurements of the same quantity, can be used as an index of the precision of the measuring instrument. We would like, on the basis of a sample of $\Sigma N_i = N$ independent observations, to make some inferences about the true ranking of the populations.

The p.d.f. of X_{ij} is $\exp\{-(x_{ij} - \mu_i)^2/(2\sigma_i^2) - \ln \sigma_i - \ln\sqrt{(2\pi)}\}$.

Therefore, $a(x_{ij}) = (x_{ij} - \mu_i)^2$

$$b(\sigma_i) = -1/(2\sigma_i^2) \ ; \ c(\sigma_i) = -\ln \sigma_i \ ; \ \theta_i = \sigma_i$$

$$a_i = \Sigma a(x_{ij})/N_i = \Sigma(x_{ij} - \mu_i)^2/N_i$$

$$k_i = 2c'(\sigma_i)\, b(\sigma_i)/\, b'(\sigma_i) = 1$$

Evidently a_i is the MLE of σ_i^2. Let the ranked a_i be denoted by

$$a_{[1]} < a_{[2]} < \ldots . < a_{[p]} .$$

Thus the population associated with $a_{[1]}$ is the best population and the population associated with $a_{[2]}$ is the second best population, and so on. The probability of correct identification can be obtained by using the expression (5.5.10) and remembering that $k_{(i)} = 1$ and α_{ij} are the ordered variance ratios (i, j =1,2, . ., p).

Example 5.5.2. Let $X_{i\,j}$ be iid random variables distributed as Exponential distributions having density of the form

$$f(x_{i\,j}) = (1/\theta_i)\exp(-x_{i\,j}/\theta_i), \quad x_{i\,j} > 0\,,$$
$$i = 1,2, \ldots, p; \; j = 1,2, \ldots, N_i .$$

Here, θ_i are unknown parameters. Let $\theta_{[1]} \leq \theta_{[2]} \leq \theta_{[3]} \leq \ldots \leq \theta_{[p]}$ be the ranked θ_i. It is not known which population is associated with $\theta_{[i]}$, i=1,2,. . , p. The p populations may be p telephone exchanges and $X_{i\,j}$ may be the time interval between jth and (j+1)th calls of the ith telephone exchange. We are interested in identifying the telephone exchange which earns the minimum profit, so that we can take preventive measures or increase the facility to improve the situation.

A telephone exchange with the highest average time interval will produce the lowest profit. The average time interval of the ith telephone exchange is θ_i. Thus, to identify a telephone exchange with the highest average time interval is equivalent to identifying a population with the largest θ_i. That is, $\theta_{[p]}$ is associated with the telephone exchange earning the lowest profit.

Let $\alpha_{i\,j} = \theta_{[i]}/\ \theta_{[j]}$ be the ratios of the ordered average time intervals. We have for Exponential distribution,

$$a(x_{ij}) = x_{ij}\ ,\ b(\theta_i) = -1/\theta_i,\ c(\theta_i) = -\ln\ \theta_i.$$

Therefore,

$$a_i = \Sigma a(x_{ij})/N_i = \Sigma x_{ij}/N_i = \overline{x_i}$$

$$k_i = 2\ c'(\theta_i)\ b(\theta_i)/\ b'(\theta_i) = 2$$

$$-k_i/2b(\theta_i) = \theta_i$$

By theorem 5.3.1, $a_i = \overline{x_i}$ is the MVBUE of θ_i. Let the ranked a_i be denoted by

$$a_{[1]} < a_{[2]} < \ldots . < a_{[p]}\ .$$

Thus the telephone exchange associated with $a_{[p]}$ is the one earning the lowest profit and the telephone exchange associated with $a_{[p-1]}$ is the one earning the second lowest profit, and so on. The probability of correct identification can be obtained by using the expression 5.5.5 for given values of α_{ij} say α_{ij}^*.

TABLE 5.1: CHARACTERISTICS OF SOME DISTRIBUTIONS BELONGING TO THE TRANSFORMED CHI-SQUARE FAMILY.

Name of Distribution with p. d. f.	a(X),	b(θ),	c(θ),	-2a(X)b(θ),	k
Normal $\{\exp(-x^2/2\theta^2)\}/\{\theta\surd(2\pi)\}$	X^2,	$-1/2\theta^2$,	$-\ln\theta$,	X^2/θ^2,	1
Lognormal $\{\exp[-(\ln x)^2/2\theta^2]\}/\{\theta x\surd(2\pi)\}$,	$(\ln X)^2$,	$-1/2\theta^2$,	$-\ln\theta$,	$(\ln X)^2/\theta^2$,	1
Exponential $\{\exp(-x/\theta)\}/\theta$,	X,	$-1/\theta$,	$-\ln\theta$,	$2X/\theta$,	2
Gamma $x^{p-1}\{\exp(-x/\theta)\}/\{\theta^p(p-1)!\}$,	X,	$-1/\theta$,	$-p\ \ln\theta$,	$2X/\theta$,	2p
Rayleigh $x\{\exp(-x^2/2\theta^2)\}/\theta^2$,	X^2,	$-1/2\theta^2$,	$-2\ \ln\theta$,	X^2/θ^2,	2
Pareto $\theta x^{-(\theta+1)}$	$\ln X$,	$-\theta$,	$\ln\theta$,	$2\theta\ln X$,	2
Weibull $\theta p x^{p-1}\exp\{-\theta x^p\}$,	X,	$-\theta$,	$\ln\theta$,	$2\theta X^p$,	2
Erlang $(\theta p)^p x^{p-1}\{\exp(-p\theta x)\}/(p-1)!$,	X,	$-\theta p$,	$p\ln\theta$,	$2\theta p X$,	2p
Maxwell $\surd(2/\pi)\ \theta^{3/2}x^2\exp\{-(\theta/2)x^2\}$,	X^2,	$-\theta/2$,	$(3/2)\ln\theta$,	θX^2,	3
Inverse Gaussian $\surd\{\theta/2\pi x^3\}\exp\{-\theta(x-p)^2/2p^2x\}$,	$(X-p)^2/p^2X$,	$-\theta/2$,	$(\ln\theta)/2$,	$\theta(X-p)^2/p^2X$,	1

Appendix 1

Fortran Program

```
      INTEGER s1,s2,s3
      REAL u
      COMMON /DAT/ s1,s2,s3,u
      REAL  Y(200),YSTAR(200),BSTRAPV(2000,200)
      OPEN  (unit=5,file='horse.dat',form='formatted',status='old')
      OPEN  (unit=6,file='horseout.dat',form='formatted',status='new')
            N=200
            NBOOT=2000
            DO 10 I=1,N
            READ(5,*)  Y(I)
10          CONTINUE
C           Enter the three seeds
            s1  =  297+I
            s2  =  1907
            s3  =  859+2*I
            WRITE(6,104)
104         FORMAT('    SAMPLE#    ',' POISSON    ',' G.POISSON ',
            $' G.N.BINOMIAL  ')
```

```
      WRITE(6,*)  '          -----------------------
     $---------------------------'
      WRITE(6,*) '          '
      DO 20 I=1,NBOOT
      DO 30 J=1,N
      call  random
      II=INT(u*N) + 1
      YSTAR(J)=Y(II)
      BSTRAPV(I,J) = YSTAR(J)
30    CONTINUE
      THBAR=0
      DO 40 J=1,N
      THBAR=THBAR+BSTRAPV(I,J)
40    CONTINUE
      THBAR = THBAR/N
      THVAR=0
      DO 50 J=1,N
      THVAR=THVAR+(BSTRAPV(I,J)-THBAR)**2
50    CONTINUE
      THVAR = THVAR/(N-1)
      THTAR=0
      DO 60 J = 1,N
      THTAR = THTAR + (BSTRAPV(I,J) - THBAR)**3
60    CONTINUE
      THTAR = THTAR/N
      THFAR = 0
      DO 70 J = 1,N
```

```
          THFAR = THFAR + (BSTRAPV(I,J) - THBAR)**4
70        CONTINUE
          THFAR = THFAR/N
          THPO = THVAR - THBAR
          THGP = (3/2 - (THTAR)*(THBAR)/(2*(THVAR)**2))**2
          -THBAR/(THVAR)
          THGN1 = 15*(THVAR)**4 +2*(THBAR)*((THVAR)**3)
          THGN2 = THGN1+ (THBAR*THTAR-3*((THVAR)**2))**2
          THGN3 = THGN2
             $ - ((THBAR)**2)*(THVAR)*(THFAR-3*((THVAR)**2))
          THGNB = THGN3 + 10*(THBAR*THTAR
                $ -   3*((THVAR)**2))*((THVAR)**2)
          WRITE(6,102) I,THPO,THGP,THGNB
102       FORMAT(I10,4F13.5)
          WRITE(6,*) '        '
20        CONTINUE

          STOP
          END

          SUBROUTINE RANDOM
c         This random number generator appeared in the
          March, 1987 issue of Byte magazine.
c         The algorithm uses three 2 byte integer seeds
          s1, s2 and s3 to produce a real between 0 and 1.
c         The cycle length is around 7 E+12. That is, if
          1000 numbers are generated every second then the
```

```
      numbers will not repeat for 220 years.
      INTEGER s1,s2,s3
      REAL x,temp
      COMMON /DAT/ s1,s2,s3,x
c     First generator
      s1 = 171*MOD(s1,177) - 2*(s1/177)
      IF(s1.LT.0) s1 = s1 + 30269
c     Second generator
      s2 = 172*MOD(s2,176) - 35*(s2/176)
      IF(s2.LT.0) s2 = s2 + 30307
c     Third generator
      s3 = 170*MOD(s3,178) - 63*(s3/178)
      IF(s3.LT.0) s3 = s3 + 30323
c     Combine to give random number
      temp = s1/30269.0 + s2/30307.0 + s3/30323.0
      x = temp - INT(temp)
      RETURN
      END
```

BIBLIOGRAPHY

Bar-Lev, S. K., and Enis, P. (1986) : Reproducibility and natural Exponential families with power variance functions. Ann. Statist. 14, 1507-1522.

Bar-Lev, S. K., and Stramer, O. (1987) : Characterisations of natural exponential families with power variance functions by zero regression properties, Probab. Th. Rel. Fields, 76, 509-522.

Barndorff-Nielsen, O. (1978) : Information and exponential families in Statistical theory, Wiley, New York.

Bickel, Peter J. and Doksum, Kjell A. (1976) : Mathematical Statistics, Holden-Day, Inc. San Francisco.

Bechhofer, R. E. (1954) : A single-sample multiple decision procedure for ranking means of normal populations with known variances, Ann. Math. Statist. 25, 16-39.

Bechhofer, R. E. and Sobel, M. (1955) : A single-sample multiple decision procedure for ranking variances of normai populations, Ann. Math. Statist. 26, 273-289.

Bliss, C. I. (1953) : Fitting the negative binomial distribution to biological data. Biometrics, 9, 176-200.

Consul , P. C. and Jain , G. C. (1973) : On some interesting properties of the generalized Poisson distribution, Biom. J. 15(7), 495-500.

Consul, P. C. and Jain, G. C. (1973) : On a generalization of the Poisson distribution. Technometrics 15 (4), 791-799.

Dobson, Annette J. (1983) : Introduction to Statistical modelling, Chapman & Hall, New York.

Eggenberger, F. (1924) : Die Wahrscheinlichkeitsansteckung, Mitteilungen der Vereinigung Schweizerischer Versicherungs-Mathematiker, 31-144.

Eggenberger, F. and Polya, G. (1923) ; Uber die Statistik verketteter Vorgange, Zeitschrift fur Angewandte Mathematik and Mechanik, 1, 279-289.

Efron, B. (1982) ;The jackknife, the bootstrap, and other resampling plans. SIAM CBMS. NSF Monograph 38.

Feller, W. (1965) ; An introduction to probability theory and its applications. John Wiley and sons, Inc.

Ferguson, T. S. (1967) ; On characterizing distributions by properties of order statistics. Sankhya, Series A, 29, 265-278.

Fisher, R. A. (1922) ; On the mathematical foundation of theoretical statistics, Philos. Trans. Roy. Soc., 222 : 309-368.

Fisher, R. A. (1958) : Statistical Methods for Research Workers, Oliver and Boyd, London.

Gordon, F. S. (1973) : Characterisations of populations using regressions properties, Ann. Statist. 1, 114-126.

Govindarajulu, Z. (1966) : Characterization of normal and generalized truncated normal distributions using order statistics, Ann. Math. Statist. 37, 1011-1015.

Gupta, R. C. (1984) : Relationships between order statistics and record values and some Characterization results, J. Appl. Prob. 21, 425-430.

Gurland, John (1963) ; A method of estimation for some generalized Poisson distributions. Proceedings of the International Symposium on Classical and Contagious distributions. Pergamon Press,141-158.

Greenwood, M. and Yule, G. Udny (1920) ; An inquiry into the nature of frequency distribution representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents, J. Roy. Stat. Soc., 83, 255-279.

Heller, B. (1979) : Special functions and characterisations of probability distributions by constant regression of polynomial statistics on the mean. Unpublished Ph. D. dissertation, Department of Statistics, University of Chicago.

Heller, B. (1983) : Special functions and characterisations of probability distributions by zero regression properties. J. Multi. Anal. 13, 473-487.

Heller, B. (1984) : Use of the hyperbolic operator to find a scalar statistic which has constant regression on the mean of a sample of Wishart matrices. J. Multi. Anal. 14, 101-104.

Hinz, Paul and Gurland, John (1967) ; Simplified techniques for estimating parameters of some generalized Poisson distributions. Biometrika 54, 555-566.

Hwang, J. S. (1975) : Characterization of distributions by the expected values of the order statistics. Ann. Inst. Statist. Math. 27, 87-93.

Hwang, J. S., Lin, G. D. (1984a) : Characterisations of distributions by linear combinations of moments of order statistics. Bull. Inst. Math. Acad. Sinica 12, 179-202.

Hwang, J. S., Lin, G. D. (1984b) : Extensions of Muntz-Szasz theorem and applications. Analysis 4, 143-160.

Jain, G. C. and Consul, P. C. (1971) : A generalized negative binomial distribution. SIAM J. Appl. Math. 21 (4), 501-513.

Jain, G. C. and Plunkett, I. G. (1977) : A Borel-Hermite distribution and its applications. Biom. J. 19 (5), 347-354.

Jain, G.C. and Gupta, R.P. (1980) : On certain moment ratios of a class of discrete distributions, J. Statist. Comput. Simul.10, pp. 263-270.

Johnson, N.L. and Kotz, S.(1969); Discrete distributions, Houghton Mifflin Company, Boston.

Johnson, N. L. and Kotz, S. (1970); Continuous univariate distributions, Houghton Mifflin Company, Boston.

Jorgensen, B. (1987) : Exponential dispersion models (with discussion), J. Roy. Statist. Soc. Ser. B, 49,

Kagan, A., Linnik, Yu., and Rao, C. R. (1973) : Characterization Problems in Mathematical Statistics. Wiley, New York.

Khan, A. H. and Ali, M. Masoom (1987) ; Characterization of probability distributions through higher order gap. Commun. Statist. -Theory Meth., 16(5), 1281-1287.

Khan, A. H. and Beg, M I. (1987) ; Characterization of the Weibull distribution. Sankhya A, 49, 268-271.

Khan, A. H. and Khan, I. A. (1986) ; Characterization of the Pareto and the Power function distributions. J. Statist. Res.

Khan, A. H. and Khan, I. A. (1987) ; Moments of order statistics from Burr distribution and its Characterization. Metron, 45, 21-29.

Kirmani, S. N. U. A. and Beg, M. I. (1984) : On Characterization of distributions by expected records. Sankhyā, 46, A, 463-465.

Kotz, S. and Johnson, N. L. (19[illegible] : Encyclopaedia of Statistical Sciences, Vol. 2, John Wiley and Sons, Inc. New York.

Kushner, H. B., Lebow, A. and Meisner, M. (1981) : Eigenfunctions of expected value operators in the Wishart distribution, II. J. Multi. Anal. 14, 336-347.

Kushner, H. B. (1987) : Dimensions of spaces of homogeneous zero regression polynomials, J. Multi. Anal. 22, 245-250.

Kemp, C. D. and Kemp, A. W. (1965) : Some properties of the Hermite distribution. Biometrika 52, 381-394.

Lehmann, E. L. (1986) : Testing Statistical Hypotheses, John Wiley and sons, New York.

Lehmann, E. L. and Scheffé, H. (1950) ; Completeness, similar regions and unbiased estimation, Part1, Sankhyā, 10 : 305-340.

Lehmann, E. L. and Scheffé, H. (1955) ; Completeness, similar regions and unbiased estimation, Sankhyā, 15 : 219-236.

Lin, G. D. (1987) : On Characterisations of distributions via moments of record values, Probab. Th. Rel. Fields, 74, 479-483.

Lin, G. D. (1988) : Characterisations of distributions via relationships between two moments of order statistics. J. St. Pl. inf., 19, 73-80.

Lin, G. D. (1989) : The product moments of order statistics with applications to characterisations of distributions. J. St. Pl. inf., 21, 395-406.

Lukacs, E. (1963) : Characterization problems for discrete distributions. In: Patil, G. (ed.) Proceedings International Symposium on Classical and Contagious Discrete Distributions, pp. 65-73. Pergamon, Oxford, New York.

Lukacs, E. (1970) : Characteristic functions, 2nd edn. Hafner, New York.

Lukacs, E. (1981) : Construction of Characterization theorems. Analytical methods in probability theory, Proc. Conf., Oberwolfach 1980, Lect. Notes Math. 861, 121-127.

Lukacs, E. and Laha, R. G. (1964) : Applications of Characteristic Functions. Griffin, London.

Lundberg, O. (1940) ; On random processes and their application to sickness and accident statistics, Thesis, University of Stockholm.

Neyman, J. (1935) ; Sur un teorems concerente le cosidette Statistiche Sufficient. Inst. Ital. Atti. Giorn., 6 : 320-334.

Neyman, J. (1939) : On a new class of contagious distribution applicable in entomology and bacteriology. Ann. Math. Statist., 10, 35-37.

Ottestad, P. (1939) ; On the use of the factorial moments in the study of discontinuous frequency distributions, Skandinazrisk Aktuarietidskrift 22, 22-31.

Palm, C. (1937) ; Inhomogeneous telephone traffic in full-availability groups, Ericsson Technics, 1, 1-36.

Patil, G.P. and Joshi, S.W. (1968); A dictionary and bibliography of discrete distributions, Oliver and Boyd, London.

Patel, J. K., Kapadia, C. H. and Owen, D. B. (1976); Handbook of Statistical distributions, Marcel Dekker, New York.

Rao, C. R. (1973); Linear Statistical Inference and its applications, John Wiley and Sons, New York.

Richards, D. St. P. (1984) : Constant Regression Polynomials and the Wishart Distribution. Mimeo series No. 1552, Institute of Statistics, University of North Carolina.

Rothschild, V. and Logothetis, N. (1985); Probability Distributions, John Wiley and sons, New York.

Seshadri, V. (1983) : The inverse Gaussian distributions : some Properties and characterisations. Can. J. Statist. 11, 131-136.

Shumway, R. and Gurland, J. (1960) ; A fitting procedure for some generalized Poisson distributions. Skand. Akt. 43, 87-108.

Student (1907) ; On the error of counting with a haemocytometer. Biometrika, 5, 351-360.

Thomas, M. (1949); A generalization of Poisson's binomial limit for use in Ecology, Biometrika, 36, 18-25.

Thorndike, F. (1926) ; Application of Poisson's Probability summation, Bell System Tech. J. 5, 604-624.

Tweedie, M. C. K. (1984) : An index which distinguishes between some important exponential families. In: Ghosh, J. K., Roy, J. (eds.) Statistics : applications and new directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, pp. 579-604. Indian Statistical Institute, Calcutta.

Zacks, S. (1971) ; The theory of Statistical Inference, John Wiley and Sons,Inc. New York.