CHARACTERIZATION OF PROBABILITY DISTRIBUTIONS BY VARIANCE BOUNDS AND ITS APPLICATIONS

Thesis submitted to the
Cochin University of Science and Technology
for the Award of the Degree of

Doctor of Philosophy

under the Faculty of Science

By

SUDHEESH KUMAR KATTUMANNIL

Department of Statistics
Cochin University of Science and Technology
Cochin- 682 022

DECEMBER 2007

In Loving Memory of My Father

CERTIFICATE

Certified that the thesis entitled '***Characterization of Probability Distributions by Variance Bounds and its Applications***' is a bonafide record of work done by Shri. Sudheesh Kumar Kattumannil under my guidance in the Department of Statistics, Cochin University of Science and Technology and that no part of it has been included anywhere previously for the award of any degree or title.

Cochin- 22,
31st December 2007.

Dr. N Unnikrishnan Nair

Emeritus Fellow,
Department of Statistics,
Cochin University of
Science and Technology.

DECLARATION

This thesis contains no material which has been accepted for the award of any other Degree or Diploma in any University and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due references are made in the text of the thesis.

Cochin- 22

31st December 2007.

Sudheesh Kumar Kattumannil

ACKNOWLEDGEMENT

With immense pleasure, I express my sincere gratitude to my honourable research guide, Dr. N. Unnikrishnan Nair, Emeritus Fellow, Department of Statistics, Cochin University of Science and Technology, Cochin- 682022, for his valuable guidance, constant encouragement, patience and affectionate attitude without which I would not have been able to complete this work satisfactorily.

I am thankful to Dr K R Muraleedharan Nair, Head, Department of Statistics and Dr. V. K. Ramachandran Nair, Former Head, Department of Statistics for their support and motivation. I also express my gratefulness to all the faculty members of the Department of Statistics for their sustained assistance. My sincere thanks are also due to the entire office staff of the Department.

The support and encouragement received from all my fellow research scholars of the Department of Statistics is deeply acknowledged. I also extend my warm appreciation to all my friends and well-wishers for their advice, help and encouragement.

I express my indebtedness to my mother Smt. Padmavathi Amma and other members of my family for their prayers for my success.

I also thankfully acknowledge the financial assistance accorded to me by the University Grants Commission, Government of India, for carrying out this research work.

"Paramathmani Niveditam"

CONTENTS

Chapter 1
Introduction

Computation of bounds on variance of functions of random variables, conditions for attainment and the form of probability distributions admitting such bounds have been always fascinating problems in statistics and probability. Some classical examples in this direction are those representing the lower bound to the variance of an unbiased estimator proposed in the well known Cramer-Rao, Bhattacharya and the Chapman- Robbins inequalities. One stream of research belonging to this category stimulated by a simple inequality giving the upper bound to the variance of a standard normal variate proposed by Chernoff in 1981 in connection with the solution of a variational problem, has produced a vast amount of literature. Chernoff's (1981) result is that when X is $N(0,1)$ and $c(x)$ is absolutely continuous function with derivative $c'(x)$, under appropriate condition on the moments,

$$V\big(c(X)\big) \le E\big(c'(X)\big)^2 \tag{1.1}$$

with equality sign holding good if and only if $c(x)$ is linear. This was followed by a plethora of extensions initially to higher dimensions in the normal case and subsequently to inequalities for other probability distributions including discrete ones. While Chernoff (1981) made use of Hermite polynomials as the main fabric to arrive at (1.1), other methods of proof like the use of Schwarz inequality, Legranges identity etc were proposed to obtain what was later known as Chernoff- type inequalities as well as lower bounds to the variance of functions of random variables and vectors. Another notable turn in this research area was the exploration of the possibilities to characterize discrete and continuous distributions in terms of the upper and lower bounds. These characterizations were further consolidated by prescribing a general measure- theoretic framework that encompassed the discrete and continuous cases and also extended to cover singular distributions. The impact of the theoretical developments on the usefulness of the bounds and covariance identities proved in the course of establishing various results were felt in other areas like probability theory, statistical estimation, isoperimetric problems, reliability modeling, etc. In view of the considerable scope this area of research has, to open up new theoretical developments as well as potential to encourage building up meaningful models and analysis of random phenomena, an attempt is made in this thesis to study a different aspect of the work induced by Chernoff– type inequalities.

An important thought originated from (1.1) is to characterize the class probability models through lower bound on the variance of a function of random variables satisfying specific conditions. Characterizations based on various extensions of Chernoff inequality have been obtained by Borokov and Utev (1984), Cacoullos and Papathanasiou (1985, 1989, 1992, 1995, 1997), Srivastava and Sreehari (1987, 1990), Prakasa Rao and Sreehari (1986, 1987, 1997), Purkayastha and Bhandari (1990), Korwar (1991), Papathanasiou (1993), Alharbi and Shanbhag (1996) and Borzadaran and Shanbhag (1998). Among these Alharbi and Shanbhag (1996) point out the application of the results in characterizing life distributions through a result similar to Cox representation of the survival function in terms of failure rate and suggest cases of some continuous distribution as illustrations. A detailed review of these results can be found in Chapter 2. The results due to Alharbi and Shanbhag (1996) throws light in to the potential application of the variance bounds in reliability analysis, though the authors makes a brief mention of this aspect. This line of thought is further strengthened and encouraged by another stream of research in reliability modelling during the last three decades, which has conditions similar to those in the work on Chernoff type inequalities.

Associated with a non- negative random variable X with distribution function $F(x)$, the concept of failure rate, mean residual life, reversed failure rate and the reversed mean residual life are extensively used in modelling and analysis in reliability studies and for applications in other disciplines such as economics, actuarial science, survival analysis and biology. Relationship between failure rate(reversed failure rate) and mean residual life(reversed mean residual life) or left (right) truncated expectations of functions of X were found to be quite useful in studying comparative behaviour of these functions and in characterizing probability distributions. In many cases some of these functions do not have a simple closed form for analytic treatment and this necessitates such relationships for identifying the underlying distribution through characterization theorems. Of these, reversed failure rate, which is receiving considerable attention recently (see Block et. al. (1998), Gupta and Nanda (2001), Nanda and Gupta (2001), Nanda et. al. (2003), Nanda et. al. (2003), Nair and Asha (2004), Nair et. al. (2005), Gupta et. al. (2005) and Nanda and Sengupta (2005)) opens up some interesting extensions hitherto not discussed in earlier papers. Like the Chernoff inequality it seems that the relationships between conditional expectations and failure rate was first derived in the case of the normal distribution $N(\mu,\sigma^2)$ (Kotz and Shanbhag (1980)) in the form

$$E(X \mid X > x) = \mu + \sigma^2 k(x)$$

where $k(x)$ is the failure rate of X (see Chapter 2 for a review and detailed discussion). Since then similar results were proved for individual distributions and families in both discrete and continuous cases by various researchers like Osaki and Li (1981), Ahmed (1991), Nair and Sankaran (1991), Glanzel (1991), Koicheva (1993), Ruiz and Navarro (1994), Consul (1995), Ghitany et. al. (1995), Gupta and Bradley (2003), Navarro and Ruiz (2004), etc.

In this thesis, our aim is to link the results in these two streams of characterization through a general theorem that holds for most of the life time distributions used in practice. By doing so our expectation is that the new theorem will enable the introduction of an alternative criteria for modelling lifetime data through the relationships between conditional expectations and failure rates and also to infer the parameters contained in the model, from among a large class of distributions through the variance bounds that could be established in the process. The relaxation of the support of the random variable to the whole real line instead of positive reals contemplated in reliability characterizations will give provision for the inclusion of the large number of distributions into the new format. Further our aim is to bring the entire results in the two streams of characterization mentioned above under a uniform framework. The variance bound so obtained is applicable to estimation theory once the random variable becomes the estimators and the corresponding distributions become the sampling distributions. Thus a new methodology for unbiased estimation can be chalked out. When this happens it becomes mandatory to compare it with the existing results in classical theory such as the Cramer- Rao, Chapman Robbins bounds and evaluate its performance with the classical counterparts. These considerations form the core themes in the present investigation. The work done in this direction is organized in to five chapters in the present thesis.

After describing the origin and development of Chernoff- type inequalities an attempt is made in Chapter 2 to review the two areas of characterizations mentioned above. In Section 2.2, a detailed study is carried out to describe the various extension of (1.1) and its applications. Characterizations based on various extension of (1.1) are reviewed in Section 2.3 and those on the relationships between conditional expectations and failure rate or reversed failure rate are given in Section 2.4 so that the results given in this chapter form the background material for the deliberations in the succeeding chapters.

The work done in the Chapter 3 revolves around a general theorem linking the variance bounds and relationship between conditional expectation and failure rates. In Section 3.2, we present the main result that obtains the identities connecting (reversed) failure rates and (right) left truncated means as necessary and sufficient conditions for the existence of lower bound to the variance. Since the families of distributions are most useful in modelling and cover a large number of potential individual distributions, the expression for the lower bound is calculated for the distributions belonging to the exponential, Pearson, generalized Pearson families, which cover most of the lifetime models used in practice. Section 3.3 explains the application of the results in unbiased estimation of parametric functions in the models involved in the above families. It is shown that the bound obtained in Section 3.2 contains the Cramer-Rao and Chapman-Robbins inequalities as particular cases and further that it compares favourably the Cramer- Rao inequality in non- regular cases. In Section 3.4, we discuss the role of the identity (3.2.1) in reliability modelling. Finally in Section 3.5, the application of our results in the new area of catastrophe theory is pointed out.

A discrete analogue of the results in the continuous case presented in Chapter 3 is described in Chapter 4 to take the advantage of potential use of the results in analyzing discrete data. Exact expression for lower bounds to the variance is calculated for distributions belonging to the modified power series family, Ord family and Katz family. It is shown that the bounds obtained here contain the Cramer-Rao and Chapman-Robbins inequalities as special cases. Application of the result is illustrated through real data. In conclusion, this chapter arrives at the class of discrete probability distributions that can be used for reliability modelling in terms of characteristic properties represented in terms of reversed failure rate (or failure rate) and right (or left) truncated expectations.

The study on truncated random variables is particularly important in reliability and in survival analysis when the device or system under consideration has lived through a specified time period and the remaining life time is of main concern. In this context the conditional variances play the same role as the variances in usual modelling and analysis of statistical data. The variance residual life pays an important role in identifying life distributions through characterization based on its functional form and in distinguishing appropriate models through criteria for aging and its relationship with failure rate has received considerable attention recently. In Chapter 5 we study the properties of conditional variance and their applications. A new characterization based on variance residual life is established in which it is shown that, the same relationship between the conditional expectations and failure rates used in developing the main results in Chapter 3 has the potential to derive lower bounds on conditional variance as well. Following the methodology by Cacoullos and Papathanasiou (1997) a lower bound to the conditional variance is also established. The bound developed here is compared with Cramer- Rao and Chapman Robbins inequality so that construction of minimum variance unbiased estimators of relevant parametric functions in truncated distributions is made possible.

Recently Arishi (2005) characterized exponential family of distributions by an identity connecting conditional variance and failure rate and deduced some results for binomial and Poisson random variables. This leaves scope for studying the characterization problem addressed by Arishi (2005) in a more general setup. In continuation of the work done in the previous chapters we discuss the properties of conditional variance for a non- negative integer valued random variable in Chapter 6. An identity connecting conditional variance and failure rate is established for the class of distributions satisfying (4.2.1) so that the work by Arishi (2005) on binomial and Poisson random variables are special cases of our findings. Application to unbiased estimation is also discussed.

The thesis is concluded in Section 6.4 with the identification of the problems of future interest that have surfaced during the course of the present study.

Chapter 2

Review of literature*

2.1 Introduction

In the presentation of the research problems considered in our study, it was mentioned that one of the objectives is to unify two important streams of characterizations viz. those based on variance bounds and relation between conditional expectations and failure rates. The basic aim of such a venture is to enrich each stream by exploring new characterization theorems that will enable application of the results to a variety of fields hitherto not covered by the existing results. In doing so, frequent references to the literature on both streams of characterizations appears to be essential and accordingly the present chapter is devoted to a discussion of the important results related to Chernoff-type inequalities as well as those on relation between various reliability concepts. Upper variance bounds derived by Chernoff (1981) to the normal random variable paved the way for other researchers in the earlier part of the work to find the similar bounds for other distributions as well as higher dimensional extensions. However, lower bound to the variances being of more interest to statistical inference, this aspect is pursued in the present study and as such the literature survey takes care of results in this direction more intensively. Other results are only mentioned or references are provided for the sake of completion.

2.2 Variance bounds

By way of the history of the subject, we mention that, although most work on variance bounds cite Chernoff inequality as the starting point and the terminology

* Part of the material in this chapter is reported in Nair and Sudheesh (2007).

'Chernoff- type inequality' is used by various authors to denote them, Houdre and Kagan (1995) noted that the same inequality was obtained much earlier by Nash (1958) in a discussion on continuity of solutions of parabolic and elliptic equations. Another important reference that has precedence over Chernoff (1981) is Brascamp and Lieb (1976) in which they have proved that for a log concave density

$$f(x) = \exp\left[-\phi(x)\right],$$

$$V(c(X)) \le E\left[c'(X)/\phi''(X)\right]^2 \tag{2.2.1}$$

which contains (1.1.1) as a special case when

$$\phi(x) = \frac{1}{2}x^2.$$

We also note that (2.2.1) is a more general result than (1.1.1), but the methods of proof in the two cases are different.

Before taking up the survey of literature we introduce a few notations to bring some uniformity in the presentation of the materials drawn from different sources that vary considerably in description. Let X be a random variable (discrete or continuous as the case may be) with distribution function $F(x)$ and density function $f(x)$, whenever it is assumed to exist, in the range $-\infty \le a < b \le \infty$ in the continuous case or probability mass function $p(x)$ in the range $N = (0,1,2,...)$ when X is discrete. $\mathcal{A}$ is the class of absolutely continuous functions with $c(x) \in \mathcal{A}$ such that the function $c(X)$ has finite variance. Differentiation will be denoted by primes.

Chernoff (1981) makes use of the expansion of $c(x)$ in orthonormalised Hermite polynomials with respect to a normal density

$$c(x) = a_0 + a_1 H_1(X) + a_2 H_2(X) + ... \tag{2.2.2}$$

with probability one, so that we have

$$E(H_i(X)) = 0,\ E(H_i(X) H_j(X)) = \delta_{ij},\ i = 1,2,...$$

where

$$\delta_{ij} = \begin{cases} 0 & i \ne j \\ 1 & i = j \end{cases}.$$

Hence,

$$\frac{dH_i(X)}{dx} = \sqrt{i}\; H_{i-1}(X)$$

and

$$a_i = E\big(c(X)H_i(X)\big).$$

These give,

$$V\big(c(X)\big) = \sum_{i=1}^{\infty} a_i^2$$

and

$$c_i'(X) = \sum_{i=1}^{n} \sqrt{i}\; a_i H_{i-1}(X) + R_n(X)$$

with $H_0 = 1$ and R_n is the remainder. Assuming $E\big(c'(X)\big)^2 < \infty$, one can see that

$$E\big(c'(X)\big)^2 = \sum_{i=1}^{\infty} i a_i^2 \geq V\big(c(X)\big).$$

The proof is indicated here because of its novelty and later reference in the development of a matrix variance inequality. Variance bounds are traditionally established first by developing some covariance identity and then using the Cauchy- Schwarz inequality (e.g. Cramer- Rao Theorem). This was the technique employed by Chen (1982) when he considered independent and identically distributed $N(0,1)$ random variables $X_1, X_2, \ldots, X_k$ and real valued Borel measurable functions $c_1, c_2, \ldots, c_k$ defined on R^k such that

$$c(x_1, x_2, \ldots, x_k) = \int_0^{x_i} c_i(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_k)dt + c_i(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_k)$$

for $i = 1, 2, \ldots, k$ and proved that

$$V\big(c(X_1, X_2, \ldots, X_k)\big) \leq \sum_{i=1}^{k} E\big[c_i(X_1, X_2, \ldots, X_k)\big]^2 \tag{2.2.3}$$

with equality in (2.2.3) holding good if and only if c is linear in $X_1, X_2, \ldots, X_k$. Notice that $c_i(x_1, x_2, \ldots, x_k) = \dfrac{dc}{dx_i}$ and hence for $k = 1$, we have the Chernoff's inequality, thus providing (2.2.3) as multivariate extension of (1.1.1) for independent and identically distributed standard normal variates. A further generalization of (2.2.3) covering non-normal cases also is available in Cacoullos (1982). Defining random variables

$$Y_i = E_i(c) - E_{i-1}(c),\ i = 1,2,\ldots,k. \tag{2.2.4}$$

where

$$E_i(c) = E_i(c(\underline{X})) = \int c(\underline{x}) \prod_{n=i+1}^{k} f_n(x_k)\,dx_k$$

and $\underline{X} = (X_1, X_2, \ldots, X_k)$ are independent random variables with joint density $f(x_1, x_2, \ldots, x_k)$ such that

$$E_0(c(\underline{X})) = E(c(\underline{X}))$$

and

$$E_k(c(\underline{X})) = c(\underline{x}),$$

with the c functions as defined to arrive at (2.2.3), it is easy to see that

$$V(c(\underline{X})) = E\left(\sum_{i=1}^{k} Y_i^2\right). \tag{2.2.5}$$

If the random variable X_i has mean μ_i and $V\left(E_i\left(c(X)\right)\right) < \infty$, Cacoullos (1982) bound is

$$V(c(\underline{X})) \le \sum_{i=1}^{k} E \int_{-\infty}^{\infty} E_i\left[c_i(X_1, \ldots, X_{i-1}, t, X_{i+1}, \ldots, X_k)\right]^2 \int_{-\infty}^{t} (\mu_i - x_i) f_i(x_i)\,dx_i dt \tag{2.2.6}$$

with equality if and only if c is linear in $x_1, x_2, \ldots, x_k$. It is readily seen that when the X_i's are independent and identically distributed as $N(0,1)$, inequality (2.2.3) results. The paper also presents a discrete analogue when $X_1, X_2, \ldots, X_k$ are independent defined on N and c is a real valued function on N^k, stating

$$V(c(\underline{X})) \le \sum_{i=1}^{k} \left\{ \sum_{k=0}^{\infty} E_i\left[\Delta_i c(X_1, \ldots, X_{i-1}, k, X_{i+1}, \ldots, X_k)\right]^2 \sum_{x_i=0}^{k} (\mu_i - x_i)\, p_i(x_i) \right\} \tag{2.2.7}$$

with Δ_i, the forward difference operator

$$\Delta_i c = c_i(x_1, \ldots, x_{i-1}, x_i + 1, x_{i+1}, \ldots, x_k) - c_i(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_k).$$

Cacoullos (1982) also obtained the expression for the lower bound using a version of the Cramer- Rao inequality and further proposed the upper bound

$$V\left(c(X)\right) \le \int_0^\infty \int_0^\infty xf(x)\left[c'(t)\right]^2 dxdt - \int_{-\infty}^0 \int_{-\infty}^t xf(x)\left[c'(t)\right]^2 dxdt \qquad (2.2.8)$$

under the assumption of existence of the density $f(x)$, finiteness of variance of $c(X)$. The discrete analogue of (2.2.8) for X defined on N is

$$V\left(c(\underline{X})\right) \le \sum_{k=0}^{\infty}\left[\Delta c(k)\right]^2 \sum_{x=k+1}^{\infty} xp(x). \qquad (2.2.9)$$

The importance of the above paper (Cacoullos (1982)) is that for the first time the focus of attention was shifted from distributions of X other than the normal including the discrete case. However, as we will see later, these bounds can be improved further.

A series of papers published by Cacoullos and Papathanasiou (1985, 1989, 1992, 1995, 1997), abbreviated as C- P in this review, gave a new dimension to the research on the topic of variance bounds in the form of several theoretical results and applications. For the sake of continuity and uniformity of the subject matter, we slightly deviate from the chronological order and discuss the major results in the above papers. C- P (1985) made use of Legrange's identity

$$\int_{-\infty}^\infty h_1^2 dx \int_{-\infty}^\infty h_2^2 dx - \left(\int_{-\infty}^\infty h_1 h_2 dx\right)^2 = \int_{-\infty}^\infty \int_{-\infty}^\infty \left[h_1(u)h_2(v) - h_1(v)h_2(u)\right]^2 dudv \quad (2.2.10)$$

applied to $V\left(c(X)\right)$ with

$$h_1 = c(x)\sqrt{f(x)} \text{ and } h_2 = \sqrt{f(x)}$$

the Cauchy- Schwarz inequality for $\left(\int_u^v c'(t)dt\right)^2$ and finally Fubini's Theorem to arrive at

$$V\left(c(X)\right) \le \int_{-\infty}^{\infty}[c'(x)]^2\left(\int_{-\infty}^{x}(\mu - t)f(t)dt\right)dx,$$

$$= \int_{-\infty}^{\infty}[c'(x)]^2\left(\int_{x}^{\infty}(t-\mu)f(t)dt\right)dx, \qquad (2.2.11)$$

where $\mu = E(X)$. We may note that the random variable X has zero mean, the bounds (2.2.8) and (2.2.10) remains the same, but for $X \ge 0$, the former exceeds the latter by

$$\int_{\mu}^{\infty}\left[c'(x)\right]^2\left(1-F(x)\right)dx>0.$$

The discrete version of the (2.2.11) is

$$V\left(c(X)\right)\le\sum_{x=0}^{\infty}\left[\Delta c(x)\right]^2\left(\sum_{k=0}^{x}(\mu-k)p(k)\right)=\sum_{x=0}^{\infty}\left[\Delta c(x)\right]^2\left(\sum_{k=x+1}^{\infty}(k-\mu)p(k)\right). \quad (2.2.12)$$

The improved bounds in the case of the normal, exponential, gamma, t, Poisson, binomial and Pascal distributions were also presented in C- P (1985), correcting the results in Cacoullos (1982). C-P (1989) is of special significance in that for the first time a rigorous proof of the lower bound to variance of $c(x)\in\mathcal{A}$ is presented in terms of a special function $w(x)$ defined by

$$\sigma^2 w(x)f(x)=\int_{-\infty}^{x}(\mu-t)f(t)dt \quad (2.2.13)$$

in the continuous case and

$$\sigma^2 w(x)p(x)=\sum_{k=0}^{x}(\mu-k)\,p(k) \quad (2.2.14)$$

in the discrete case, where $\sigma^2=V(X)$. After establishing a pivotal covariance identity

$$Cov\left(X,c(X)\right)=\sigma^2\int_{-\infty}^{\infty}c'(x)w(x)f(x)dx \quad (2.2.15)$$

they use Cauchy-Schwarz inequality to show that

$$\inf_{c\in\mathcal{A}}\frac{V\left(c(X)\right)}{V(X)E^2\left(c'(X)w(X)\right)}=1 \quad (2.2.16)$$

under suitable assumptions on the existence of the expectations involved in (2.2.16). When X is discrete defined on N,

$$V\left(c(X)\right)\ge\sigma^2E^2\left(w(X)\Delta c(X)\right) \quad (2.2.17)$$

with $w(.)$ as in (2.2.14). It is informative to note that the upper bound (2.2.11) is also expressible in terms of the $w(.)$ function so that one can write

$$\sigma^2E^2\left(w(X)c'(X)\right)\le V\left(c(X)\right)\le\sigma^2E\left(\left(c'(X)\right)^2w(X)\right) \quad (2.2.18)$$

and a similar expression for the discrete case replacing $c'(x)$ by $\Delta c(x)$. The lower bounds corresponding to (2.2.6) and (2.2.7) in the multivariable cases involving the w functions are respectively

$$V\left(c(X)\right) \geq \sum_{i=1}^{n} \sigma_i^2 E^2\left(w_i\left(X_i\right) c_i'(X)\right) \tag{2.2.19a}$$

$$V\left(c(X)\right) \geq \sum_{i=1}^{n} \sigma_i^2 E^2\left(w_i\left(X_i\right) \Delta_i c_i\left(X\right)\right) \tag{2.2.19b}$$

for the continuous and discrete cases, where

$$\sigma_i^2 w_i(x_i) f_i(x_i) = \int_{-\infty}^{x_i} (\mu_i - t) f_i(t) dt$$

in (2.2.19a) and

$$\sigma_i^2 w_i(x_i) p_i(x_i) = \sum_{k=0}^{x_i} (\mu_i - k) p_i(k)$$

in (2.2.19b) and the subscript i in all the expressions indicating that they correspond to the random variable X_i, $i = 1, 2, \ldots, n$ in a set of independent random variables $X_1, X_2, \ldots, X_n$.

As a further generalization, C-P (1992) drops the assumption of independence in the above results and consider $X = \left(X_1, X_2, \ldots, X_n\right)$ as a random vector in the n-dimension rectangle $a_i < x_i < b_i$, $-\infty \leq a_i < b_i \leq \infty$, $i = 1, 2, .., n$ with dispersion matrix $\Sigma > 0$. They define functions $w^i\left(x\right)$ by

$$w^i(x) f(x) = \int_{a_i}^{x_i} (\mu^i - q^i\left(u, t_i, v\right)) f(u, t_i, v) dt_i \tag{2.2.20}$$

and

$$q^i\left(x\right) = \sum_{i=1}^{n} \sigma_{ij}^* x_i ,$$

where

$$\Sigma^{-1} = \left(\sigma_{ij}^*\right),\ u_i = \left(x_1, \ldots, x_{i-1}\right),\ v_i = \left(x_{i+1}, \ldots, x_n\right),\ \mu^i = E\left(q^i\left(X\right)\right)$$

to obtain the relationship

$$Cov\left(q^i\left(X\right), c\left(X\right)\right) = E\left(w^i\left(X\right) c_i\left(X\right)\right)$$

with $c(x)$, a real valued function on the range of X satisfying

$$E\left|w^i c_i\right| < \infty,\ E\left|\left(q^i - \mu^i\right)c\right| < \infty$$

and $w^i(x)f(x) \to 0$ monotonically as x approaches any boundary points of the rectangle along the coordinate axes. Then (T denote the transpose of the vector)

$$V(c(X)) \geq E\left(w^1 c_1, \ldots, w^p c_p\right) \Sigma E\left(w^1 c_1, \ldots, w^p c_p\right)^T \tag{2.2.21}$$

with equality holding good if and only if $c(x)$ is linear. We may remark in this connection that the assumption $w^i(x)f(x) \to 0$ is not satisfied by many distributions and as such the question of lower bounds in such cases is still an open problem. Another problem of natural interest when lower bound to variance is prescribed is that of finding the form of probability distributions admitting such bounds. For example, since $w(x)$ appearing in (2.2.13) is unique for each distribution, differentiating (2.2.13) and assuming $\lim_{x \to -\infty} xf(x) = 0$ we get

$$\left(w(x)f'(x) + w'(x)f(x)\right)\sigma^2 = (\mu - x)f(x).$$

Rearranging terms

$$\frac{f'(x)}{f(x)} = \frac{\mu - x - \sigma^2 w'(x)}{\sigma^2 w(x)}, \tag{2.2.22}$$

which provides the distribution of X. On the other hand, in the multivariate case the problem is a little more difficult as the solution of the system of partial differential equations

$$\frac{\partial}{\partial x_i}\left(w^i(x)f(x)\right) = \left(\mu^i - q^i(x)\right)f(x),\ i = 1, 2, \ldots, n \tag{2.2.23}$$

seems much more complicated except in cases where $w^i(x)$ has simple functional forms such as for normal, Dirichlet and certain forms of exponential distributions.

For an absolutely continuous function $h(x)$ of the random variable X defined on an interval (a,b) on the real line where a may be $-\infty$ and b may be $+\infty$, C- P (1997) has the following results

$$Cov(h(X), c(X)) = E(z(X)c'(X))$$

where

$$z(x)f(x)=\int_{a}^{x}(E(h)-h(t))f(t)dt \tag{2.2.24}$$

and

$$V(c(X))\geq\frac{E^{2}(z(X)c'(X))}{E(z(X)h'(X))} \tag{2.2.25}$$

with the equality sign holding good if and only if $c(x)=c_1h(x)+c_2$. There are several particular cases of interest. Obviously when, $h(x)=x$ and $z(x)=\sigma^2w(x)$ equation (2.2.16) derives. Setting

$$h(x)=-f'(x)/f(x),\ z(x)=1$$

and whenever the density vanishes at a or b,

$$V(c(X))\leq cE(c'(X))^2$$

which is the Brascamp and Lieb (1976) bound mentioned earlier. Further, taking $h(X)=X$, X become normally distributed, $c=\sigma^2$ so that we reach at the classical Poincare inequality of Chen and Lou (1987). In the discrete case, corresponding to (2.2.24) and (2.2.25) we have

$$z(x)p(x)=\sum_{k=0}^{x}(E(h)-h(k))p(k) \tag{2.2.26}$$

and

$$V(c(X))\geq\frac{E^{2}(z(X)\Delta c(X))}{E(z(X)\Delta h(X))}. \tag{2.2.27}$$

An interesting feature of the above results is that there exist some nice functional forms for $z(x)$ (or $w(x)$) that characterize continuous and discrete distributions possessing lower variance bounds. Also equations (2.2.24) and (2.2.26) are of fundamental importance in reliability modeling. These aspects are discussed in the next Chapter.

For some other works on Chernoff- type inequalities left out in the chronological sequence in a somewhat different direction we refer to Vitale (1989), Prakasa Rao (1992), Hu Chin Yuan (1986), Johnson (1993) and Prakasa Rao and Sreehari (1997). These are in the form of variance bounds for infinitely divisible distributions, bivariate distributions with specific conditionals such as normal, gamma, binomial, negative binomial and Poisson laws, inequalities for non- linear functions of stochastic integrals, inequalities involving uniform distributions and functions of Pearson variates. Klassen (1985) provides upper bounds to variances of Laplace, logistic, gamma and double Weibull distributions. Some involved extensions covering convolution of probability measures and infinite divisible law in the form of Poincare type inequalities are also considered in Chen (1985). With a shift in the domain of the $c(x)$ function from real values to complex values and assuming the existence of the k^{th} derivatives $c^{(k)}(x)$, Houdre and Kagan (1995) show that

$$\sum_{k=1}^{2n}\frac{(-1)^{k+1}}{k!}E\left|c^{(k)}(X)\right|^2 \le V(c(X)) \le \sum_{k=1}^{2n-1}\frac{(-1)^{k+1}}{k!}E\left|c^{(k)}(X)\right|^2$$

and for $n=1$,

$$E\left|c^{(1)}(X)\right|^2 - \frac{1}{2}E\left|c^{(2)}(X)\right|^2 \le V(c(X)) \le E\left|c^{(1)}(X)\right|^2 .$$

(see also Houdre and Perez (1995)).

Papadatos and Papathanasiou (1996) obtained upper and lower variance bounds of random variable in terms of the quantile density function. Starting from the expression for the upper bound in (2.2.18) (the case of lower bound is similar) with $c = F^{-1}(H)$, where H is the distribution function of the random variable Y, we have identical distribution for X and $c(Y)$ and hence from (2.2.18)

$$V(X) \le V(Y)\int_0^1 w\left(H^{-1}(u)\right)\left[c_Y(u)/c_X(u)\right]^2 du \qquad (2.2.28)$$

where c_Y and c_X becoming quantile density function of Y and X. Equation (2.2.28) generalizes the results of Arnold and Brocket (1988) which states that for a random variable Y in $(0,1)$ and differentiable function $c(.)$ on $(0,1)$

$$V(c(Y)) \le E(Y)\int_0^1 \left(H(u) - H^{(1)}(u)\right)\left(c'(u)\right)^2 du$$

where $H^{(1)}$ is the distribution function corresponding to the first moment distribution of Y, by removing the restriction on the range of Y.

Let X, Y and Z be independent exponential random variable and c, $h \in \mathcal{A}$ such that Ec^2, Eh^2 are finite. Then Bobkov and Houdre (1997) shows that

$$Cov(c(X), h(X)) = E(c'(X+Y)h'(X+Z))$$

Further if c and h are non- decreasing convex function on the real line

$$Cov(c(X), h(X)) \geq kE(c'(X)h'(X)), \ k > 0$$

$$V(c(X)) \geq kE(c'(X))^2$$

and

$$V(X-a)^+ \geq kP(X \geq a),$$

for any real a. For any convex function c, for the standard double exponential distribution with

$$f(x) = 2^{-1} \exp[-|x|],$$

$$E(c'(X))^2 \leq V(c(X)) \leq 4E(c'(X))^2.$$

The paper by Arakelian and Papathansiou (2004) has its main focus on bounds on absolute deviations, but has some implications on variance as well. They provide a sharper bound to $E(|c(X)|)$ than the traditional

$$E(|c(X)|) \leq \left[E(c(X))^2\right]^{\frac{1}{2}}$$

by using Hovenier's (1994) inequality for integrable functions $h_1(x)$, $h_2(x)$, viz.

$$\left(\int_a^b h_1 h_2 dx\right)^2 \leq \left[\left|\int_a^r h_1 h_2 dx\right| + \left(\int_a^r h_1^2 dx\right)^{\frac{1}{2}} \left(\int_r^b h_2^2 dx\right)^{\frac{1}{2}}\right]^2$$

for the choice of

$$h_1 = c(x)\sqrt{f(x)} \text{ and } h_2 = \sqrt{f(x)}$$

to end up with

$$E(|c(X)|) \leq F(r)(E(c(X) \mid X < r)) + \overline{F}(r)\left(E(c^2(X) \mid X \geq r)\right)^{\frac{1}{2}}. \qquad (2.2.29)$$

Writing $U(r)$ for the square of the expression on the right of (2.2.29)

$$U(a) = E\left(\left(c(X)\right)^2\right) \text{ and } U(b) = E^2\left(c(X)\right)$$

so that

$$V\left(c(X)\right) \geq U(a) - U(r),$$

an inequality of some interest. When the point of truncation r becomes the mean μ of X and $c(X) = (X - \mu)^2$,

$$\sigma^2 \leq F(\mu)\left[E\left((X-\mu)^2 \mid X < \mu\right)\right] + \bar{F}(\mu)\left[E\left((X-\mu)^2 \mid X \geq \mu\right)\right]^{\frac{1}{2}}.$$

Similar results for the discrete random variable X derive from the use of the inequality

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leq \left[\left|\sum_{k=1}^{m} x_k y_k\right| + \left(\sum_{k=m+1}^{n} x_k^2 \sum_{k=m+1}^{n} y_k^2\right)^{\frac{1}{2}}\right]^2$$

and we obtain (2.29) with r replaced by m. Borzadaran and Shanbhag (1998b) also provide similar kind of results about the bound on absolute deviation.

The method of deriving variance inequality with the aid of the Hermite polynomials as discussed at the beginning of this section is recalled in Olkin and Shep (2005) to obtain a matrix variance bound. They illustrate the bivariate case that easily extends to the multivariate version by simply increasing the number of components. The expansion (2.2.2) applied to two functions $c(X)$ and $h(X)$ with b_i replacing a_i in (2.2.2) in the expansion for $h(X)$,

$$V\left(c(X)\right) = \sum_{i=1}^{\infty} a_i^2,$$

$$V\left(h(X)\right) = \sum_{i=1}^{\infty} b_i^2$$

and

$$Cov\left|c(X)h(X)\right| = a_i b_i.$$

If $H = (h_{ij})$ and $C = (c_{ij})$ are $k \times k$ matrices, where

$$h_{ij} = E\left(c_i'(X)c_j'(X)\right)$$

and

$$c_{ij} = Cov(c_i(X), c_j(X)),$$

we have

$$C = \begin{pmatrix} \sum_{i=1}^{\infty} a_i^2 & \sum_{i=1}^{\infty} a_i b_i \\ \sum_{i=1}^{\infty} a_i b_i & \sum_{i=1}^{\infty} b_i^2 \end{pmatrix}, \quad H = \begin{pmatrix} \sum_{i=1}^{\infty} i a_i^2 & \sum_{i=1}^{\infty} i a_i b_i \\ \sum_{i=1}^{\infty} i a_i b_i & \sum_{i=1}^{\infty} i b_i^2 \end{pmatrix}$$

so that

$$H - C = (\alpha, \beta)'(\alpha, \beta) \geq 0$$

with $\alpha_i = \sqrt{i-1}\ a_i$ and $\beta_i = \sqrt{i-1}\ b_i$. Some other results on the variance bounds which are in fact characterizations are discussed in the next section for convenience.

2.3 Characterization by variance bound

Characterization problems associated with variance bounds arising from Chernoff- type inequalities can be broadly classified into two categories- results that characterize specific probability distributions such as normal, Poisson etc or certain families and those concerned with a general class of distributions. The work on normal law (Borokov and Utev (1984)), multivariate normal (Prakasa Rao and Sreehari (1986)), Poisson distribution (Prakasa Rao and Sreehari (1987)), uniform distribution (Purkayastha and Bhandari (1990)), Pearson family (Korwar (1991)), Power series and factorial series families (Papathanasiou (1993)) and normal laws with reference to central limit theorem (Cacoullos et. al. (1993, 1994)), belongs to the first category. A further discussion of these results separately is not attempted here in view of the fact that most of them are subsumed in the general theorems that are taken up subsequently.

Srivastava and Sreehari (1987) considered non-negative integer valued random variables with $p(0) > 0$ and $E(X)^2 < \infty$ and proved that

$$\sup_{c(x) \in T} \frac{V(c(X))}{E\left(c(X)(\Delta h(X))^2\right)} = 1 \tag{2.3.1}$$

where T is the class of real valued functions $c(.)$ such that $E\left(c(X)(\Delta h(X))^2\right) < \infty$ and for some $g(x)$ with $E(g(X)) = V(X)$ if and only if g satisfies

$$\sum_{k=0}^{x}(\mu-k)\,p(k)=p(x)g(x). \tag{2.3.2}$$

They also discussed the application of the results for some standard distributions. C- P (1989) characterized (2.2.16) and the discrete version in (2.2.19) provided that in the case of (2.2.16) that expression holds for some $w(x)$ satisfying $E(w(x))=1$ (See also Srivastava and Sreehari (1990) for almost the same results).

A more general result covering distributions that are not purely discrete or continuous, under a uniform framework, is presented in Alharbi and Shanbhag (1996). They considered a non- constant Lebesgue- Steiltjes measure function $h(.)$ in R and a measure v on the Borel field of R determined by it such that $E(h(X))=\mu$ and $E(h(X))^2<\infty$ and Borel- measurable function $w(x)$ such that $w(x)>0$ almost surely and $E(w(X))=V(h(X))$. For $c\in\mathcal{A}$, $E(c(X))^2<\infty$, $0<E\left(w(X)(c'(X))^2\right)<\infty$, and c' is the Radon- Nikodym derivative of c, they proved the following results.

(i)
$$\sup_{c(x)\in\mathcal{A}}\frac{V(c(X))}{E\left(w(X)(c'(X))^2\right)}=1 \tag{2.3.3}$$

if and only if

$$w(x)dF(x)=\int_{[x,\infty)}(h(t)-\mu)dF(t)dv(x),\ x\in R. \tag{2.3.4}$$

(ii) If $E(w(X)|c'(X)|)<\infty$ and $E(w(X)c'(X))\neq 0$ then

$$\inf_{c(x)\in\mathcal{A}}\frac{V(c(X))V(h(X))}{E^2(w(X)(c'(X)))}=1 \tag{2.3.5}$$

if and only if (2.3.4) holds. It is interesting to note the comparison between (2.3.5) and results in (2.2.24) through (2.2.27) for the essential conditions for the latter results to be valid. Whenever $h(x)=[x]$, $c'(x)$ in (2.3.3) and (2.3.4) becomes $\Delta c(x)$ for each $x=0,\pm1,\pm2,\ldots$ and

$$w(x)p(x)=\sum_{k=x}^{\infty}(k-\mu)\,p(k).$$

Several examples of distributions characterized through the forms of $w(x)$ is given, for example, the Poisson or shifted Poisson has the characteristic property

$$\sup_{c(x)\in\mathcal{A}} \frac{V(c(X))}{V(X)E(\Delta c(X))^2} = 1,$$

improving upon the result of Prakasa Rao and Sreehari (1987). There is also a representation of the distribution function in terms of the $w(x)$ function viz.

$$dF(x) = a\left(w(x)^{-1}\right)\exp\left[-H(x)\right], x\in R \qquad (2.3.6)$$

where

$$H(x) = \int_{\mu}^{x}\frac{(t-\mu)}{w(t)}dt$$

with $\int_{\mu}^{x}$ placed by $-\int_{\mu}^{x}$ if $x<\mu$, a is a normalizing constant and

$$w(x) = \begin{cases} c, & \text{if } x\notin(l,r) \\ \int\limits_{[x,\infty)}\frac{(t-\mu)f(t)}{f(x)}dt, & \text{other wise} \end{cases} \qquad (2.3.7)$$

where l and r are the left and the right extremities of the distribution. Borzadaran and Shanbhag (1998a) in a further refinement of the results mentioned in (2.3.4) through (2.3.6) use a function $a(X)$ of the random variable with finite mean μ^{*} satisfying inf $\left(a(x)-\mu^{*}\right)>0$ as $x\to\infty$ if $r=\infty$ and inf $\left(\mu^{*}-a(x)\right)>0$ as $x\to-\infty$ if $l=-\infty$. Their results are

$$Cov\left(c(X),a(X)\right) = E\left(w(X)c'(X)\right) \qquad (2.3.8)$$

for all $c\in\mathcal{A}$ with $E\left|w(X)c'(X)\right|<\infty$, if and only if

$$w(x)dF(x) = \left\{\int\limits_{[x,\infty)}(a(t)-\mu^{*})dF(t)\right\}dv(x),\ x\in R. \qquad (2.3.9)$$

Further if $a(X)\in\mathcal{A}$ with respect to v and $a(X)$ is square integrable such that $V\left(a(X)\right) = E\left(w(x)a'(X)\right)$ and $E\left(c'(x)w(X)\right)$ is defined and non- zero, then

$$\frac{V(c(X))\,V(a(X))}{E^2(w(X)c'(X))}=1 \tag{2.3.10}$$

if and only if (2.3.9) is true. If $h(.)$ is as in Alharbi and Shanbhag (1996), there results hold with $h(.)=a(.)$. One may note that the results of C- P (1995, 1997), now become special cases of (2.3.8), (2.3.9) and (2.3.10).

Referring to the multivariate case, there are two important characterizations. Retaining the notations in C- P (1992) refered to earlier, they showed that if the inequality (2.2.21) holds for every real valued function $c(x)$ with the equality if c is linear and $w'f$ vanishes at the end points a_i, b_i for $i=1,2,...,n$, then $w'x$ and $f(x)$ are related through (2.2.20) and the $w'x$ characterize the distribution of X. Papadatos and Papathanasiou (2003), choose $h(x)=(h^1(x),...,h^n(x))$ in which X is supported by a convex open subset C_n of R^n with norm $E\|h(X)\|=\sum_{i=1}^{n}E(h^i(X))<\infty$ and define $z(x)=(z^1(x),...,z^n(x)):C_n\to R^n$ by

$$z^i(x)f(x)=\int_{a_i}^{x_i}\left(E(h_i(X))-h^i(u_i,t_i,v_i)\right)f(u_i,t_i,v_i)dt_i\,. \tag{2.3.11}$$

Assuming $z^i(x)f(x)\to 0$ as x_i tends to a_i or b_i, $E\left|h^i(X)-E(h^i(X)c(X))\right|<\infty$, $E(z^i(X)c_i(X))<\infty$ and $c(x)$ is the indefinite integral of its partial derivative

$$c_i(x)=\frac{\partial c(x)}{\partial x_i},$$

then the following results hold.

$$Cov(h^i(X),c(X))=E(z^i(X)c_i(X)),\quad i=1,2,...,n \tag{2.3.12}$$

and if $h:C_n\to R^n$ is an arbitrary function with $E\|h(X)\|<\infty$ with (2.3.12) holding for every bounded function $c:C_n\to R^n$ which are indefinite integrals of their partial derivatives and

$$\int_{a_i}^{b_i}E\left(h^i(X)-E(h^i(X))\right)f(x)dx_i=0$$

then (2.3.11) is true. Equation (2.3.12) provides an extension of the Steins- type identity for continuous multivariate exponential families established by Chou (1988). Papadatos and Papathanasiou (2003) points out the applications of (2.3.12) in inference problems involving normal order statistics.

Before concluding this section, some related problems that have generated interest in other areas also need mention. One of these relates to order statistics. Balakrishnan and Subramonian (1993) shows that the Papathanasiou (1990) bound

$$Cov(X_{1:2}, X_{2:2}) \le \frac{1}{\sigma^2} \tag{2.3.13}$$

concerning order statistics of a sample of size 2 from an absolutely continuous distribution F with variance σ^2 is equivalent to Hartely and David (1954) and Gumbel (1954) bound for $E(X_{2:2})$ and improves (2.3.13) to

$$Cov(X_{1:2}, X_{2:2}) \le \sigma^2 (1+\rho)$$

when X_1 and X_2 are identically distributed but not necessarily independent. Papathanasiou (1990) discuss a case in which $X_1, ..., X_n$ are independent and identically distributed with strictly monotone F and $h(u) = F^{-1}(u)$ to derive

$$V(X_{(i)}) \le \frac{n!}{(i-1)!(n-i)!(n+1)} \int_0^1 u^i (1-u)^{n-i+1} (h(u))^2 du$$

with equality if and only if F is uniform. Another general result due to Papadatos and Papathanasiou (1996) is

$$V(X_{(i)}) \le \frac{n! V(Y_{(i)})}{(i-1)!(n-i)!} \int_0^1 w_i(G^{-1}(u)) \left(\frac{h_Y(u)}{h_X(u)} \right)^2 u^{i-1} (1-u)^{n-i} du$$

where G is the distribution function of another population generated by a random variable Y. Other applications relate to Probability Theory, where infinite divisibility, central limit theorem, stochastic convergence etc are discussed in relation to the variance bounds and we refer to Vitale (1989), C- P (1997), and Cacoullos et. al. (1993, 1994) for details.

2.4 Characterizations by relationships between reliability measures

The first characterization directly connecting the failure rate and conditional expectation appears to have been introduced by Nassar and Mahmoud (1985) when they characterized the mixture of exponential distributions, each with means μ_1 and μ_2 by the relation

$$m_1(x) = x + (\mu_1 + \mu_2) - \mu_1\mu_2 k(x).$$

where $m_1(x) = E[X \mid X > x]$ and $k(x)$ is the failure rate. However the flow of work in this area was spurred by Osaki and Li (1988) when they characterized the gamma distribution with parameter (m, p) by means of the identity

$$m_1(x) = \mu + \frac{x}{m} k(x), \text{ for all } x > 0$$

and negative binomial distribution with parameters r and p by

$$m_1(x) = \mu + \frac{(x+1-r)}{p} k(x+1), \text{ for all } x \geq r-1,$$

where μ stands for $E(X)$. Following Osaki and Li (1988) similar results were developed for other individual distributions as well as for classes of probability distributions and we discuss them in a chronological order. Adatia et. al. (1991) gave a necessary and sufficient condition that a continuous, positive random variable follow a gamma distribution with parameters (α, β) in terms of conditional moments and an expression involving its failure rate in the form

$$m_k(x) = \frac{(\beta+k-1)(\beta+k-2)\ldots(\beta+1)\beta}{\alpha^k} + \sum_{i=1}^{k-1} \frac{(\beta+k-1)(\beta+k-2)\ldots(\beta+i)\beta}{\alpha^{k+1-i}} x^i k(x)$$

$$+ \frac{1}{\alpha} x^k k(x), \tag{2.4.1}$$

where

$$m_k(x) = E\left[X^k \mid X > x\right], \text{ for } k = 1, 2, 3, \ldots.$$

The result was then used to develop a characterization of a mixture of two gamma distributions. Characterizations of beta, binomial, and Poisson distributions were presented by Ahmed (1991). He showed that for the Pearson type I distribution with probability density function

$$f(x)=\frac{(x-a)^{p-1}(b-x)^{q-1}}{B(p,q)(b-a)^{p+q-1}},\ a<x<b,\ a\geq 0,\ b,\ p,\ q>0,$$

$$m_1(x)=\mu\left(1+\frac{(x-a)(b-x)k(x)}{aq+bp}\right),\ \text{for all } a\leq x\leq b, \tag{2.4.2}$$

and in particular the beta distribution is uniquely determined by the relationship

$$m_1(x)=\mu\left(1+\frac{x(1-x)k(x)}{p}\right),$$

obtained by setting $a=0$ and $b=1$ in (2.4.2). Further by setting $q=1$, the result for the power distribution follows as

$$m_1(x)=\mu\left(1+\frac{x(1-x)k(x)}{m}\right),$$

with $\mu=\frac{p}{(p+1)}$. Ahmed (1991) also established that X is distributed as binomial probability mass function

$$p(x)=\binom{n}{x}p^x(1-p)^x,\quad x=0,1,\dots,n$$

if and only if

$$m_1(x)=\mu+(1+x)(1-p)k(x+1)$$

and X is Poisson with probability mass function

$$p(x)=\frac{e^{-\lambda}\lambda^x}{x!},\ x=0,1,\dots$$

if and only if

$$m_1(x)=\mu+(1+x)k(x+1),\ \text{for all integers } x>0.$$

In a more general framework Nair and Sankaran (1991) showed that the distribution of X belongs to the Pearson family specified by

$$\frac{f'(x)}{f(x)}=\frac{-(x+d)}{a_0+a_1x+a_2x^2} \tag{2.4.3}$$

if and only if

$$m_1(x) = \mu + (b_0 + b_1 x + b_2 x^2)k(x) \tag{2.4.4}$$

where

$$c_i = (1-2a_2)^{-1} a_i,\ a_2 \neq \frac{1}{2},\ i = 0,1,2,$$

provided that $\lim_{x \to a} x^2 f(x) = 0$. As a discrete analogue of (2.4.3) they also proved that in the support of the set of integers X has distribution in the Ord family satisfying

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-(x+d)}{a_0 + a_1 x + a_2 x^2} \tag{2.4.5}$$

if and only if

$$m_1(x) = \mu + (c_0 + c_1 x + c_2 x^2)k(x+1), \tag{2.4.6}$$

where

$$c_i = (1-2a_2)^{-1} a_i,\ a_2 \neq \frac{1}{2},\ i = 0,1,2$$

and deduced the formulas of the earlier researchers as a particular cases. Another extension of (2.4.4) for the Pearson family by Glanzel (1991) involved higher order conditional moments resulting in the characteristic property

$$E(X^2 \mid X > x) = P(x)E(X \mid X \geq x) + Q(x),$$

where $P(x)$ and $Q(x)$ are polynomials of degree at most one with real coefficients. Characterization of Koicheva (1993) about gamma distribution also involved higher order moments and failure rate.

Ruiz and Navarro (1994) established that the class of distributions satisfied by the differential equation of the form

$$\frac{f'(x)}{f(x)} = \frac{c - x - q'(x)}{q(x)} \tag{2.4.7}$$

is characterized by the identity

$$m_1(x) = c + q(x)k(x), \tag{2.4.8}$$

where c is a constant, $q(x)$ is function satisfying $\lim_{x \to b} q(x)f(x) = 0$ with end point a and b in the support of X and $\int_a^b q'(x)f(x)dx < \infty$. The results of Shanbhag (1970), Osaki and Li (1988), Ahmed (1991), Nair and Sankaran (1991) are special cases of (2.4.7) and (2.4.8). The discrete analogues of the above results are also given through illustrative examples.

Consul (1995) obtained a general theorem, based on conditional expectation, for the exponential class of distributions. The theorem is then applied to numerous discrete and continuous probability distributions of the exponential class providing specific characterizations for each one of them. He showed that the random variable X, continuous or discrete, has a distribution belonging to the exponential class of the form

$$f(x;\theta)=\exp\left[xQ(\theta)+T(x)+S(\theta)\right], \tag{2.4.9}$$

where $T(x)$ are real valued measurable functions, $Q(\theta)$ and $S(\theta)$ are real functions on $\mathbb{R}$ with $Q(\theta)$ having continuous non-vanishing derivatives in $\mathbb{R}$, if and only if

$$m_1(x)=\mu+\left(Q'(\theta)\right)^{-1}\frac{\partial}{\partial\theta}\left(\log\overline{F}(x)\right) \tag{2.4.10}$$

Note that the equation (2.4.10) can be written as

$$m_1(x)=\mu+\left(f(x)Q'(\theta)\right)^{-1}\frac{\partial\overline{F}(x)}{\partial\theta}k(x),$$

so that the characterization of exponential family is in fact in terms of failure rate and conditional expectations. The characterizations for individual distributions including Lagrangian Poisson, generalized Poisson, Poisson, Borel-Tanner, generalized negative binomial, negative binomial, binomial, Geeta and gamma distributions were deduced from the above results. Of these, characterization for negative binomial and gamma were discussed by Osaki and Li (1988) and for the Poisson case by Ahmed (1991).

Ghitany et. al. (1995) established a characterization result for an absolutely continuous random variable X whose density function is of the form

$$f(x)=\exp\left(-q(x,\theta)\right),\ x\geq 0,\ \theta\in\Theta, \tag{2.4.11}$$

and $q(x,\theta)$ is a real valued function with $q'(x,\theta)\neq 0$ and $q''(x,\theta)$ exist on $(0,\infty)$ for all $\theta\in\Theta$, by the identity

$$E\left[1+\frac{q''(X)}{\left(q'(X)\right)^2}s(X)-\frac{s'(X)}{q'(X)}\,|\,X\geq x\right]=\frac{s(x)}{q'(x)}k(x), \tag{2.4.12}$$

where $s(x,\theta) \neq 0$ is a real valued function with existing $s'(x,\theta)$ on $(0,\infty)$ for all $\theta \in \Theta$. Particularly for gamma distributions

$$E\left[X^{k-1}\left(X-\frac{\lambda+k-1}{\beta}\right)|X \geq x\right]=\frac{x^k}{\beta}k(x) \qquad (2.4.13)$$

and putting $k=1$ it reduces to the results of Osaki and Li (1988). The distribution of X follows Weibull with probability density function given by

$$f(x)=\lambda\beta x^{\beta-1}\exp\left(-\lambda x^{\beta}\right), \quad x \geq 0, \ \lambda,\beta > 0$$

if and only if

$$E\left[X^{\beta}|X \geq x\right]=\frac{1}{\lambda\beta}\left[xk(x)+\beta\right] \qquad (2.4.14)$$

holds for all $x \geq 0$. And for the Gompertz distributions with probability density function given by

$$f(x)=\lambda\exp(\beta x)\exp\left[-\frac{\lambda}{\beta}\left[\exp(\beta x)-1\right]\right], \ x \geq 0, \ \lambda,\beta > 0,$$

the identity (2.4.12) reduces to

$$E\left[\exp(\beta X)|X \geq x\right]=\frac{1}{\lambda}\left[\left[1-\exp(\beta x)\right]k(x)+\lambda+\beta\right].$$

The corresponding results for exponential and Rayleigh distribution derives as special cases of (2.4.14). Navarro et. al. (1998) gave a general method to obtain a distribution function $F(x)$ through the moment of the residual life defined by $m_k{}^*(x)=E\left[(X-x)^k \mid X > x\right]$ and studied the characterization using the relation of the form

$$m_k(x)=c+g(x)k(x), \qquad (2.4.15)$$

where

$$m_k(x)=E\left[X^k \mid X > x\right] \text{ for } k=1,2,3,...$$

and c is a constant. Extending the results given in Adatia et. al. (1991), Koicheva (1993) and Ghitany et. al. (1995) they showed that the random variable X with differentiable density in its support (a,b) is of the form

$$\frac{f'(x)}{f(x)}=\frac{\left(c-g'(x)-x^k\right)}{g(x)} \tag{2.4.16}$$

if and only if

$$m_k(x)=c+g(x)k(x),$$

where c is a constant and $g(x)$ is a real function satisfying $\lim_{x\to b} g(x)f(x)=0$. For $k=1$, we obtained Theorem 3 given in Ruiz and Navarro (1994) for relations of type (2.4.7), as well as particular characterizations given in Kotz and Shanbhag (1980), Osaki and Li (1988), Ahmed (1991) and Nair and Sankaran (1991) for some usual distributions. In a similar direction, Sankaran and Nair (2000) obtained a further extension of their (Nair and Sankaran (1991)) and Glanzel (1991) characterizations for Pearson family. They gave a necessary and sufficient condition for the distribution of X belongs to the Pearson system specified by (2.2.3) given by

$$m_k(x)=\left(a_{0,r}+a_{1,r}x+a_{2,r}x^2\right)x^{r-1}k(x)+d_r m_k(x)+a_{0,r}(r-1)m_{k-2}(x)$$

provided that $\lim_{x\to a} x^r f(x)=0$, where

$$d_r=\frac{b_i r-d}{1-(1+r)b_2},\ b_2\neq\frac{1}{r+1} i=0,1,2$$

and

$$a_{i,r}=b_i\left(1-(1+r)b_2\right)^{-1},\ i=0,1,2.$$

Nair et. al. (1999) obtained a relation of the form

$$m_1(x)=x+\left(\frac{1}{p_1}+\frac{1}{p_2}\right)-\frac{1}{p_1p_2}k(x+1),\ 0<p_1,\ p_2<1$$

to characterize the mixture of geometric law with probability mass function

$$p(x)=\alpha p_1(1-p_1)^x+(1-\alpha)p_2(1-p_2)^x,\ 0<\alpha<1.$$

They also proved similar results for Waring distribution, and extended this approach in several directions which include the higher order moment and factorial moments. Abraham and Nair (2001) established an identity connecting the failure rate and mean residual life to characterize a class of continuous distributions containing finite mixtures of exponential, Lomax and beta distributions. The identity

$$m_1(x) = x + (1+ax)(\mu_1 + \mu_2 + a\mu_1\mu_2) - \mu_1\mu_2(1+ax)^2 k(x)$$

is satisfied for all x for a random variable X with density

$$f(x) = pf_1(x) + (1-p) f_2(x), \quad 0 < p < 1$$

if and only if for $i = 1,2,$ the component densities are

$$f_i(x) = \lambda_i \exp(-\lambda_i x), \quad \lambda_i > 0, x > 0, \text{ for } a = 0$$

$$f_i(x) = \alpha_i \beta^{\alpha_i} (x+\beta)^{-(\alpha_i+1)}, \quad \alpha_i,\ \beta > 0, x > 0, \text{ for } a > 0$$

$$f_i(x) = \frac{C_i}{R}\left(1 - \frac{x}{R}\right)^{C_i - 1}, \quad C_i, R > 0, 0 < x < R, \text{ for } a < 0.$$

Of these, the results in the exponential case appeared in Nassar and Mahmoud (1985) is contained in the above formulation. A further generalization aimed at including more distributions than in (2.4.3) has appeared in Sankaran et. al. (2003) who extended the Pearson family by replacing the linear functions in the numerator on the right of (2.4.3) with a quadratic function $b_0 + b_1 x + b_2 x^2$ to claim the characteristic property

$$E\left[\left(b_2 X^2 + (b_1 + 2a_2)X + b_0 + a_1\right) \mid X > x\right] + \left(a_0 + a_1 x + a_2 x^2\right) k(x) = 0,$$

provided that $\lim_{x \to b} x^r f(x) = 0$, $r = 0,1,2$ so that for a choice of

$$h(x) = px^2 + qx + r$$

with

$$p = b_2,\ q = b_1 + 2a_2 \text{ and } r = b_0 + a_1 + \mu,$$

the above equation can be written

$$m(x) = E(h(X) \mid X > x) = \mu + g(x) k(x), \tag{2.4.17}$$

where

$$g(x) = -\left(a_0 + a_1 x + a_2 x^2\right).$$

Sindhu (2003) gave the discrete analogue of the (2.4.17) which states that the random variable X belongs to the family of distributions satisfying

$$\frac{p(x+1) - p(x)}{p(x)} = \frac{-\left(b_0 + b_1 x + b_2 x^2\right)}{a_0 + a_1 x + a_2 x^2} \tag{2.4.18}$$

if and only if

$$m_1(x) = \mu + g(x)k(x+1),$$

where

$$g(x) = \left(a_0 + a_1 x + a_2 x^2\right)$$

and

$$h(x) = Px^2 + Qx + R$$

with

$$P = b_2,\ Q = b_1 + 2a_2 \text{ and } R = a_1 - a_2 + \mu.$$

More recently, Gupta and Bradley (2003) characterized the class of distributions satisfying

$$\frac{f'(x)}{f(x)} = \frac{(\mu - x)}{g(x)} - \frac{g'(x)}{g(x)}, \tag{2.4.19}$$

where μ is a constant and g satisfies the first- order linear differential equation,

$$g'(x) + \frac{f'(x)}{f(x)} g(x) = \mu - x,$$

with the identity given by

$$m_1(x) = \mu + g(x)k(x). \tag{2.4.20}$$

The characterizations based on (2.4.19) and (2.4.20) were discussed in Ruiz and Navarro (1994) but Gupta and Bradley (2003) also emphasized the ageing behaviour concerning the family (2.4.19).

Nair et. al. (2005) developed a characterization for continuous probability distributions using the relationship between the reversed failure rate and conditional expectations. Related results using the failure rate and the conditional expectations are also given, which generalize the results of Ruiz and Navarro (1994). A continuous random variable X in the support of (a,b) has a probability density function of the form

$$f(x) = \frac{c}{g(x)} \exp\left[-\int_a^x \frac{h(t) - \mu}{\mu g(t)} dt\right], \tag{2.4.21}$$

where $g(x)$ is a real function and c is a constant which makes $f(x)$ a density, if and only if

$$r(x) = \mu(1 - g(x)\lambda(x)), \qquad (2.4.22)$$

provided $\lim_{x\to\infty} g(x) f(x) = 0$, where $h(x)$ is function such that $E(h^2(X)) < \infty$ and $E(h(X)) = \mu$. The discrete analogue of the above work has been introduced by Gupta et. al. (2005) and applied to Ord family and modified power series family for characterizing them. Let X be discrete random variable defined on N with $\mu = E(h(X)) < \infty$, where $h(.)$ is a real function, then for a real function $g(x)$ with $\lim_{x\to\infty} g(x) p(x) = 0$, the condition

$$r(x) = \mu(1 - g(x)\lambda(x)) \qquad (2.4.23)$$

is satisfied for all x in N if and only if the probability mass function $p(x)$ of X satisfies the difference equation

$$\frac{p(x+1)}{p(x)} = \frac{\mu g(x)}{h(x+1) - \mu(1 - g(x+1))}. \qquad (2.4.24)$$

The results of Nair et. al. (2005) and Gupta et. al. (2005) subsumes all the available results in literature discussed in this context for specified values of $h(x)$ and $g(x)$. And further they enlighten the role of reversed failure rate and the right truncated expectation of $h(X)$ in characterizing life distributions, which was not considered by the earlier authors.

The research problem considered in Chapter 1 will be addressed for possible solutions in the subsequent chapters taking advantage of the above results as background material.

Chapter 3

Characterization of Continuous Distributions by Variance Bound and Its Implication to Reliability Modelling*

3.1 Introduction

The significance of the lower variance bound in the inference procedure and in characterization problems were discussed in the previous chapter. Taking this into account, the present chapter attempts to establish lower bounds to the variance of functions of random variables that characterize a wide class of probability distributions, by extending some results from literature. In the process, as mentioned in the research problem, a link is established between these characterizations and those based on relationship between conditional expectations and failure (reversed failure) rates developed during the last decade. This approach has useful implication to reliability modelling and minimum variance unbiased estimation, through Cramer-Rao type inequalities. Our work subsumes the existing results and unifies two topics in characterization theory that were independently carried out. Application of the results to the catastrophe theory is also pointed out.

3.2 Characterizations

The background material required for the work in this chapter consists of a random variable X supported by the interval (a,b), $-\infty \le a < b \le \infty$ with absolutely continuous distribution function F and density function f. We denote by $\mathcal{A}$, the class

---

* Part of the work done in this chapter was published in Nair and Sudheesh (2006).

of absolutely continuous distribution functions and by $\mathfrak{B}$, the set of absolutely continuous functions of the random variable X.

Let $h(X)$ be a Borel- measurable function of the random variable X such that $E(h^2(X)) < \infty$. We define the conditional expectations

$$m(x) = E\left[h(X) \mid X > x\right]$$

and

$$r(x) = E\left[h(X) \mid X \le x\right]$$

so that when $h(x) = x$, they become closely related to the mean residual life function

$$E(X - x \mid X > x) = m_1(x) - x$$

and the reversed mean residual life function

$$E(x - X \mid X \le x) = x - r_1(x)$$

where

$$m_1(x) = E(X \mid X > x) \text{ and } r_1(x) = E(X \mid X \le x).$$

Since $m_1(x)$, $r_1(x)$, the failure rate $k(x)$ and the reversed failure rate $\lambda(x)$ uniquely determine the distribution of X, their functional forms are extensively used in modelling lifetime data. However, for many distributions the expression for these functions are not in simple closed forms and this has prompted reliability analysts to develop relationship between $k(x)$ and $m(x)$ or $\lambda(x)$ and $r(x)$ that characterize different lifetime models. The major results in this context were reviewed in Chapter 2. We first establish a characterization theorem that brings the equivalence of these characterizations with those based on variance bounds in a more general framework, and show that many of the earlier results in both steams of characterizations can be unified into a single format.

Theorem 3.2.1:

For a random variable X with support (a,b), $-\infty \le a < b \le \infty$ with distribution function $F(x) \in \mathcal{A}$ and functions $c(.)$ and $g(.)$ belonging to $\mathfrak{B}$ such that $c'(.)$ exist on (a,b) with $\mathrm{E}(\mathrm{c}'(\mathrm{X}).\mathrm{g}(\mathrm{X})) < \infty$, the following statements are equivalent.

(i) $\int_a^x (\mu - h(t)) f(t)dt = \sigma g(x) f(x)$ (3.2.1)

(ii) $V(c(X)) \geq E^2(c'(X)g(X))$ (3.2.2)

with $\sigma = E(h'(X)g(X)) < \infty$

(iii) $r(x) = \mu - \sigma\lambda(x)g(x)$ or $m(x) = \mu + \sigma k(x)g(x)$ (3.2.3)

(iv) $\frac{f'(x)}{f(x)} = -\frac{g'(x)}{g(x)} + \frac{\mu - h(x)}{\sigma g(x)}$ (3.2.4)

provided that both g and f are differentiable almost everywhere in (a,b).

Here $h(x)$ stands for a Borel measurable function satisfying $E(h^2(X)) < \infty$ with mean μ and variance σ^2.

Proof:

(i) $\Rightarrow$ (ii)

The covariance between the functions $h(X)$ and $c(X)$ can be written as

$$Cov(h(X),c(X)) = \int_a^b (h(x)-\mu)(c(x)-c(\mu))f(x)dx$$

$$= \int_a^b (h(x)-\mu)\left(\int_\mu^x c'(t)dt\right)f(x)dx$$

$$= \int_a^b c'(x)\left(\int_a^x (\mu - h(t))f(t)dt\right)dx$$

by Fubini's Theorem. Using (3.2.1)

$$Cov(h(X),c(X)) = \sigma\int_a^b c'(x)g(x)f(x)dx$$

$$= \sigma E(c'(X)g(X)) \qquad (3.2.5)$$

Applying Cauchy- Schwartz inequality

$$Cov^2(h(X),c(X)) \leq V(h(X)).V(c(X))$$

in (3.2.5), we get (3.2.2).To prove the second part of (ii) we note that

$$E\left(h'(X)g(X)\right)=\int_a^b h'(x)g(x)f(x)dx$$

$$=\int_a^b h'(x)\left(\frac{1}{\sigma}\int_a^x\left(\mu-h(t)\right)f(t)dt\right)dx\,.$$

Integrating by parts

$$=\frac{1}{\sigma}\left[\left\{h(x)\int_a^x(\mu-h(t))f(t)dt\right\}_a^b\right.$$

$$\left.-\int_a^b\left(\mu-h(x)\right)h(x)f(x)dx\right].$$

The first term vanishes since $E\left(h'(X)g(X)\right)<\infty$ and the second term is

$$-\frac{1}{\sigma}\int_a^b\left(\mu-h(x)\right)h(x)f(x)dx=-\frac{1}{\sigma}\left[\mu E\left(h(X)\right)-E\left(h^2(X)\right)\right]$$

$$=\frac{1}{\sigma}\left[E\left(h^2(X)\right)-\mu^2\right]$$

$$=\frac{1}{\sigma}\sigma^2=\sigma\,.$$

(ii) $\Rightarrow$ (iii)

For an arbitrary real θ choose $c(x)=h(x)+\theta p(x)$ such that $c(x)\in\mathfrak{B}$. Then

$$V\left(c(X)\right)=\sigma^2+\theta^2V\left(p(X)\right)+2\theta Cov\left(h(X),p(X)\right)\qquad(3.2.6)$$

and

$$E^2\left(c'(X)g(X)\right)=\left\{E\left[h'(X)g(X)+\theta p'(X)g(X)\right]\right\}^2$$

$$=E^2\left(h'(X)g(X)\right)+\theta^2E^2\left(p'(X)g(X)\right)$$

$$+2\theta E\left(h'(X)g(X)\right).E\left(p'(X)g(X)\right)$$

$$=\sigma^2+\theta^2E^2\left(p'(X)g(X)\right)+2\theta\sigma E\left(p'(X)g(X)\right).$$

(3.2.7)

When (3.2.2) is true, substituting (3.2.6) and (3.2.7) in (3.2.2) and rearranging terms,

$$\theta^2\left[V(p(X))-E^2\left(p'(X)g(X)\right)\right]+2\theta\left[Cov(h(X),p(X))-\sigma E\left(p'(X)g(X)\right)\right]\geq 0.$$

(3.2.8)

Since the first term is always positive, for (3.2.8) to be non-negative for all real θ one must have

$$2\theta\left[Cov(h(X),p(X))-\sigma E\left(p'(X)g(X)\right)\right]=0$$

$$Cov\left(h(X),p(X)\right)=\sigma E\left(p'(X)g(X)\right)$$

or

$$\int_a^b p'(x)\left(\int_a^x(\mu-h(t))f(t)dt\right)dx=\sigma\int_a^b p'(x)g(x)f(x)dx. \qquad (3.2.9)$$

Setting $p(x)=\cos tx$

$$-\int_a^b t\sin tx\left(\int_a^x(\mu-h(t))f(t)dt\right)dx=-\sigma\int_a^b t\sin tx.g(x)f(x)dx.$$

Similarly when $p(x)=\sin tx$

$$\int_a^b t\cos tx\left(\int_a^x(\mu-h(t))f(t)dt\right)dx=\sigma\int_a^b t\cos tx.g(x)f(x)dx.$$

Hence from the last two equations

$$t\int_a^b\left(\cos tx+i\sin tx\right)\left(\int_a^x(\mu-h(t))f(t)dt\right)dx=t\sigma\int_a^b\left(\cos tx+i\sin tx\right)g(x)f(x)dx$$

or

$$\int_a^b e^{itx}\left(\int_a^x(\mu-h(t))f(t)dt\right)dx=\sigma\int_a^b e^{itx}g(x)f(x)dx,$$

which is the same as

$$\int_a^b e^{itx}\left(\int_a^x\frac{(\mu-h(t))f(t)dt}{f(x)}\right)f(x)dx=\sigma\int_a^b e^{itx}g(x)f(x)dx.$$

By the uniqueness of Fourier transforms

$$\int_a^x\frac{(\mu-h(t))f(t)dt}{f(x)}=\sigma g(x)$$

$$\frac{\mu}{f(x)}\int_a^x f(t)dt-\frac{1}{f(x)}\int_a^x h(t)f(t)dt=\sigma g(x)$$

$$\mu \frac{F(x)}{f(x)} - E\left(h(X) \mid X \le x\right) \frac{F(x)}{f(x)} = \sigma g(x)$$

$$\mu - E\left(h(X) \mid X \le x\right) = \sigma g(x) \lambda(x)$$

$$r(x) = \mu - \sigma g(x) \lambda(x).$$

Further, since $E\left(h(X)\right) = \mu$ we can write

$$\int_a^x h(t) f(t) dt + \int_x^b h(t) f(t) dt = \mu\left[F(x) + 1 - F(x)\right]$$

or

$$F(x) r(x) + \left[1 - F(x)\right] m(x) = \mu\left[F(x) + \left(1 - F(x)\right)\right],$$

$$F(x)\left(r(x) - \mu\right) = \left(1 - F(x)\right)\left(\mu - m(x)\right)$$

$$\frac{f(x)}{\lambda(x)}\left(r(x) - \mu\right) = \frac{f(x)}{k(x)}\left(\mu - m(x)\right)$$

which leads to the identity

$$\frac{m(x) - \mu}{k(x)} = \frac{\mu - r(x)}{\lambda(x)},$$

this establishes the equivalence of the two expressions in (3.2.3) and the implication (ii)$\Rightarrow$ (iii).

(iii) $\Rightarrow$ (iv)

We rewrite the second equality in (3.2.3) as

$$\int_x^b h(t) f(t) dt = \mu\left[1 - F(x)\right] + \sigma f(x) g(x).$$

Differentiating and assuming $\lim_{x \to b} h(x) f(x) = 0$

$$h(x) f(x) = \mu f(x) - \sigma\left(f(x) g'(x) + f'(x) g(x)\right) \tag{3.2.10}$$

from which (3.2.4) results.

Finally, to establish the equivalences stated in the Theorem we complete the chain of implications by showing that (iv) $\Rightarrow$ (i). Assuming (3.2.4) we can reach at (3.3.10). Now from (3.3.10),

$$\sigma \frac{d(f(x)g(x))}{dx} = (\mu - h(x)) f(x),$$

and the integration with respect to x from a to x gives the identity (3.2.1) so that the proof of the theorem is completed.

Remark 3.2.1: The value of $g(x)$ is unique for a particular choice of $h(x)$. But we can have different forms for $g(x)$ for the same distribution, when $h(x)$ is different. For example, in the case of Maxwell distribution with

$$f(x) = 4\sqrt{\frac{\lambda^3}{\pi}} x^2 e^{-\lambda x^2}, \; x > 0, \; \lambda > 0$$

$$g(x) = -\frac{x}{\sigma} \text{ when } h(x) = -2\lambda x^2 + 3 + \mu$$

and

$$g(x) = \frac{1}{2\lambda}\left(1 + \frac{1}{\lambda x^2}\right) \text{ when } h(x) = x.$$

Remark 3.2.2: The results in Cacoullos and Papathanasiou (1989) and Srivastava and Sreehari (1990) are special cases of Theorem 3.2.1 when $h(x) = x$. To realize characterization of distributions by the lower bound attained by a particular absolutely continuous function in $\mathfrak{B}$, it is necessary and sufficient that (i), (iii) or (iv) is satisfied. This is equivalent to finding a unique $g(x)$ that satisfies one of these identities. The earlier papers look at $h(x) = x$ and the $g(x)$ values corresponding to $E(X \mid X > x)$, that limits characterizations to a class of distributions for which

$$E(X \mid X > x) = \mu + \sigma k(x) g(x).$$

Our generalization enables characterization through the variance of a wider class of functions through a simple calculation of $g(x)$ implied by (3.2.4), as illustrated through the following theorems. Since for modelling and inference, families of distributions are more desirable and results for individual distributions can be easily deduced, we look at the exponential family, Pearson family and the generalized Pearson family.

Theorem 3.2.2:

The distribution of X belongs to the Pearson family specified by (2.4.3) if and only if

$$\inf_{c(x)\in \mathfrak{B}} \frac{V\left(c\left(X\right)\right)}{E^2\left(b_0+b_1X+b_2X^2\right)}=1 \tag{3.2.11}$$

where

$$b_i = a_i\sigma^{-1}(1-2a_2)^{-1}\ ,\ a_2 \neq \frac{1}{2},\ i=0,1,2,$$

provided that $\lim_{x\to b} x^2 f(x) = 0$.

Proof:

From (2.4.3) we have

$$\left(a_0+a_1x+a_2x^2\right)f'(x) = -(x+d)f(x)$$

$$\left(a_0+a_1x+a_2x^2\right)f'(x)+\left(a_1+2a_2x\right)f(x) = -(x+d)f(x)+\left(a_1+2a_2x\right)f(x),$$

that is

$$\frac{d}{dx}\left[\left(a_0+a_1x+a_2x^2\right)f(x)\right] = (a_1-d)f(x)+\left(2a_2-1\right)xf(x). \tag{3.2.12}$$

Integrating with respect to x from t to b, we get,

$$-\left(a_0+a_1t+a_2t^2\right)f(t) = (a_1-d)R(t)+\left(2a_2-1\right)R(t)E\left(X\mid X>t\right),$$

dividing through out by $R(t)\left(2a_2-1\right)$ we get

$$E\left(X\mid X>t\right) = \frac{(a_1-d)}{\left(1-2a_2\right)}+\frac{\left(a_0+a_1t+a_2t^2\right)k(t)}{\left(1-2a_2\right)}$$

or

$$E\left(X\mid X>x\right) = \frac{(a_1-d)}{\left(1-2a_2\right)}+\frac{\left(a_0+a_1x+a_2x^2\right)k(x)}{\left(1-2a_2\right)}. \tag{3.2.13}$$

Assuming $\lim_{x\to a} x^2 f(x) = 0$ and integrating both side of (3.2.12) with respect to x over the whole range we get

$$(a_1-d)+\left(2a_2-1\right)E\left(X\right)=0$$

or

$$E(X) = \frac{(a_1 - d)}{(1-2a_2)}.$$

Now comparing (3.2.13) with second inequality in (3.2.3)

$$E(X \mid X > x) = \mu + \sigma g(x) k(x),$$

where

$$d = a_1 - (1-2a_2)\mu$$

and

$$g(x) = (b_0 + b_1 x + b_2 x^2),$$

with b_i values as sated in the theorem. Substituting for $g(x)$ in Theorem 3.2.1 we arrive at (3.2.11)

The following examples provide the values of $g(x)$ for the different members of the Pearson family

Example 3.2.1: Consider the gamma distribution with probability density function

$$f(x) = \frac{m^p}{\Gamma p} e^{-mx} x^{p-1}, x > 0,$$

then

$$\frac{f'(x)}{f(x)} = \frac{-(mx + 1 - p)}{x} = \frac{-(x + (1-p)/m)}{x/m},$$

so that $a_0 = a_2 = 0$ and $a_1 = m^{-1}$, hence $g(x) = (\sigma m)^{-1} x$. Substituting this value of $g(x)$ in (3.2.3) we get the result of Osaki and Li (1988) stated in Section 2.4 and that of Adatia et. al. (1991) in equation (2.4.1) when $k = 1$. The lower bound for the gamma distribution derived in Srivastava and Sreehari (1990) is a special case of (3.2.11).

Example 3.2.2: For the normal random variable with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty,$$

we have

$$\log f(x) = \log k - \frac{(x-\mu)^2}{2\sigma^2},$$

differentiating

$$\frac{f'(x)}{f(x)} = \frac{-(x-\mu)}{\sigma^2},$$

so that $a_1 = a_2 = 0$ and $a_0 = \sigma^2$, hence $g(x) = \sigma$. The relationship between failure rate and $E(X \mid x > x)$ stated in Kotz and Shanbhag (1980) is a special case of (3.2.3) when $g(x) = \sigma$. Further, the corresponding lower bound to the variance given in Cacoullos and Papathanasiou (1989) also drawn from (3.2.11).

Example 3.2.3: Consider the beta distribution with probability density function

$$f(x) = \frac{1}{B(p,q)} x^{p-1} (1-x)^{q-1}, \ x > 0$$

then

$$\frac{f'(x)}{f(x)} = \frac{(p-1)(1-x)-(q-1)x}{x(1-x)} = \frac{-(x(p+q-2)-p+1)}{x(1-x)}$$

so that $a_0 = 0$, $a_1 = \frac{-1}{(p+q-2)}$ and $a_2 = \frac{-1}{(p+q-2)}$, hence $\sigma g(x) = (p+q)^{-1} x(1-x)$.

Now the equation (3.2.3) becomes a generalization of the result for the beta distribution proved in Ahmed (1991) and the lower bound in Srivastava and Sreehari (1990).

Example 3.2.4: Consider the Pearson type II distribution with probability density function

$$f(x) = k(1 - x^2 / a^2), -a < x < a.$$

Then

$$\frac{f'(x)}{f(x)} = \frac{-k2x}{a^2 k(1-x^2/a^2)} = \frac{-2x}{(a^2 - x^2)}$$

so that $a_0 = \frac{a^2}{2}$, $a_1 = 0$ and $a_2 = \frac{-1}{2}$, hence $\sigma g(x) = 4^{-1}(a^2 - x^2)$.

Example 3.2.5: The Pearson type IV distribution with probability density function

$$f(x) = k(1+x^2/a^2)\exp\left(-c\tan^{-1}x/a\right), -\infty < x < \infty.$$

or

$$\log f(x) = \log k + \log\left(1+x^2/a^2\right) - c\tan^{-1}x,$$

has

$$\frac{f'(x)}{f(x)} = \frac{2x - a^2c}{a^2 + x^2}$$

so that $a_0 = \dfrac{-a^2}{2}$, $a_1 = 0$ and $a_2 = \dfrac{-1}{2}$, hence $\sigma g(x) = \dfrac{-1}{4}(a^2 + x^2)$.

Example 3.2.6: Consider the Pearson type VI distribution with probability density function

$$f(x) = k(x-a)^b x^{-c},\ x > a.$$

Then

$$\frac{f'(x)}{f(x)} = \frac{k(x-a)^{b-1}x^{-c}\left(b - cx^{-1}(x-a)\right)}{k(x-a)^b x^{-c}} = \frac{-(b+c)x + ac}{x(x-a)}$$

so that $a_0 = 0$, $a_1 = \dfrac{-a}{(b+c)}$ and $a_2 = \dfrac{-1}{(b+c)}$, hence $\sigma g(x) = -(b+c+2)^{-1}x(x+a)$.

Example 3.2.7: For the Pearson type VII distribution with probability density function

$$f(x) = k(1+x^2/a^2)^{-m},\ -\infty < x < \infty,$$

$$\frac{f'(x)}{f(x)} = \frac{-km(1+x^2/a^2)^{-(m+1)}2a^{-2}x}{k(1+x^2/a^2)^{-m}} = \frac{-2mx}{a^2 + x^2}$$

so that $a_0 = 0$, $a_1 = \dfrac{a^2}{2m}$ and $a_2 = \dfrac{1}{2m}$, hence $\sigma g(x) = 2^{-1}(m-1)^{-1}x\left(a^2 + x\right)$.

Example 3.2.8: Consider the Pearson type VIII distribution with probability density function

$$f(x) = k(1+x/a)^{-m}, -a < x < 0.$$

then

$$\frac{f'(x)}{f(x)} = -\frac{km(1+x/a)^{-(m+1)}}{ka(1+x/a)^{-m}} = -\frac{mx}{x(a+x)}$$

so that $a_0 = 0$, $a_1 = \frac{a}{m}$ and $a_2 = \frac{1}{m}$, hence $\sigma g(x) = (m-2)^{-1} x(a+x)$.

Example 3.2.9: Consider the Pearson type IX distribution with probability density function

$$f(x) = k(1+x/a)^m, -a < x < 0.$$

then

$$\frac{f'(x)}{f(x)} = \frac{km(1+x/a)^{m-1}}{ka(1+x/a)^m} = \frac{mx}{x(a+x)}$$

so that $a_0 = 0$, $a_1 = \frac{-a}{m}$ and $a_2 = \frac{-1}{m}$, hence $\sigma g(x) = -(m+2)^{-1} x(a+x)$.

Example 3.2.10: The inverted gamma distribution (Pearson type V) with probability density function

$$f(x) = k.x^{-p} e^{-q/x},\ x > 0$$

has

$$\frac{f'(x)}{f(x)} = \frac{kx^{-p} e^{-q/x}\left(-px^{-1} + qx^{-2}\right)}{kx^{-p} e^{-q/x}} = \frac{-px+q}{x^2}$$

so that $a_0 = a_1 = 0$, and $a_2 = \frac{-1}{p}$, hence $\sigma g(x) = -(p+2)^{-1} x^2$.

The values of $g(x)$ for various distributions covering under Theorem 3.2.2 is summarized in Table 3.2.1 for easy reference.

In an effort to improve the richness in members of the Pearson family and there by extend the domain of application, Sindhu (2003) has replaced the linear term in the differential equation of (2.4.3) by a quadratic, obtaining

$$\frac{d \log f(x)}{dx} = \frac{\left(b_0 + b_1 x + b_2 x^2\right)}{a_0 + a_1 x + a_2 x^2}.$$

Table 3.2.1

Values of $g(x)$ for members of Pearson family

Distribution	$f(x)$	$\sigma g(x)$
*Beta	$x^{p-1}(1-x)^{q-1}, 0<x<1$	$(p+q)^{-1}x(1-x)$
Type II	$k(1-x^2/a^2), -a<x<a$	$4^{-1}(a^2-x^2)$
**Gamma	$ke^{-mx}x^{p-1}, x>0$	$m^{-1}x$
Type IV	$k(1+x^2/a^2)\exp\left(-c\tan^{-1}x/a\right)$ $-\infty<x<\infty$	$-4^{-1}(a^2+x^2)$
Inverted Gamma	$kx^{-p}e^{-q/x}, x>0$	$-(p+2)^{-1}x^2$
Type VI	$k(x-a)^b x^{-c}, x>a$	$-(b+c+2)^{-1}x(a+x)$
Type VII	$k(1+x^2/a^2)^{-m}, -\infty<x<\infty$	$2^{-1}(m-1)^{-1}x(a^2+x)$
Type VIII	$k(1+x/a)^{-m}, -a<x<0$	$(m-2)^{-1}x(a+x)$
Type IX	$k(1+x/a)^{m}, -a<x<0$	$-(m+2)^{-1}x(a+x)$
Normal	$k\exp\left[\frac{-(x-\mu)}{2\sigma^2}\right], -\infty<x<\infty$	σ^2

* Power distribution when $p=1$ or $q=1$ and uniform when $p=q=1$.

** Exponential distribution when $p=1$.

Besides containing all the members of the Pearson family (corresponding to $b_2 = 0$), this extended Pearson system consists of many new members like the inverse Gaussian, random walk, Maxwell and Rayleigh models. It is also shown in the study by Sindhu (2003) that the new system is quite useful in reliability modelling and analysis by virtue of several characterizations in terms of reliability concepts and also through several interesting ageing properties the members of the family posses. Our next theorem extends the result of Theorem 3.2.2 to the extended family.

Theorem 3.2.3:

The density function $f(x)$ satisfies the differential equation

$$\frac{f'(x)}{f(x)} = \frac{\left(b_0 + b_1 x + b_2 x^2\right)}{a_0 + a_1 x + a_2 x^2}, \quad b_2 \neq 0 \tag{3.2.14}$$

if and only if

$$\inf_{c(x) \in \mathfrak{B}} \frac{\sigma^2 V\left(c(X)\right)}{E^2\left[c'(X)\left(a_0 + a_1 X + a_2 X^2\right)\right]} = 1$$

for a choice of

$$h(x) = px^2 + qx + r$$

with

$$p = b_2,\ q = b_1 + 2a_2 \text{ and } r = b_0 + a_1 + \mu,$$

provided that $\lim_{x \to b} x^2 f(x) = 0$.

Proof:

From (3.2.14)

$$f'(x)\left(a_0 + a_1 x + a_2 x^2\right) + \left(a_1 + 2a_2 x\right) f(x) = \left(a_1 + b_0 + 2a_2 x + b_1 x + b_2 x^2\right) f(x)$$

$$\frac{d}{dx}\left[f(x)\left(a_0 + a_1 x + a_2 x^2\right)\right] = \left(a_1 + b_0 + 2a_2 x + b_1 x + b_2 x^2\right) f(x)$$

Integrating with respect to x from t to b, we get

$$-f(t)\left(a_0 + a_1 t + a_2 t^2\right) = \left(a_1 + b_0\right) R(t) + \left(2a_2 + b_1\right) R(t) E\left(X \mid X > t\right) + R(t) b_2 E\left(X^2 \mid X > t\right).$$

Dividing through out by $R(t)$ we get

$$E\left(\left(b_2X^2+(b_1+2a_2)X+b_0+a_1\right)\mid X>x\right)+\left(a_0+a_1x+a_2x^2\right)k(x)=0$$

or

$$E\left(\left(b_2X^2+(b_1+2a_2)X+b_0+a_1+\mu\right)\mid X>x\right)=\mu-\left(a_0+a_1x+a_2x^2\right)k(x).$$

Taking $h(x)$ as stated in the theorem, the above equation can be written as

$$E\left(h(X)\mid X>x\right)=\mu-\left(a_0+a_1x+a_2x^2\right)k(x).$$

Comparing with second inequality in (2.3) we get

$$g(x)=-\sigma^{-1}\left(a_0+a_1x+a_2x^2\right).$$

The rest of the proof of the theorem follows from Theorem 3.2.1.

Example 3.2.11: Consider the Rayleigh distribution with probability density function

$$f(x)=2\lambda xe^{-\lambda x^2},\ x>0,\ \lambda>0,$$

then

$$\frac{f'(x)}{f(x)}=\frac{2\lambda e^{-\lambda x^2}\left(1-2\lambda x^2\right)}{2\lambda xe^{-\lambda x^2}}=\frac{\left(1-2\lambda x^2\right)}{x}$$

so that $a_0=a_1=0$, and $a_2=1$, $b_0=2$, $b_1=0$ and $b_2=-2\lambda$, hence $\sigma g(x)=-x$ with $h(x)=-2\lambda x^2+2+\mu$.

Example 3.2.12: For the inverse gaussian distribution with probability density function

$$f(x)=\sqrt{\frac{\lambda}{2\pi x^3}}\exp\left[\frac{-\lambda(x-\alpha)^2}{2x\alpha^2}\right],\ x>0,\ \lambda>0,$$

we have

$$\log f(x)=\log k+\log x^{-\frac{3}{2}}-\frac{\lambda(x-\alpha)^2}{2x\alpha^2},$$

differentiating

$$\frac{f'(x)}{f(x)}=-\frac{3}{2x}-\frac{\lambda(x-\alpha)(x+\alpha)}{2x^2\alpha^2}=\frac{-\lambda x^2-3\alpha^2x+\lambda\alpha^2}{2x^2\alpha^2}$$

so that $a_0=a_1=0$, and $a_2=2\alpha^2$, $b_0=\lambda\alpha^2$, $b_1=-3\alpha^2$ and $b_2=-\lambda$, hence $\sigma g(x)=-2\alpha^2x^2$ with $h(x)=-\lambda x^2+\alpha^2x+\lambda\alpha^2+\mu$.

Example 3.2.13: Consider the Maxwell distribution with probability density function

$$f(x)=4\sqrt{\frac{\lambda^3}{\pi}}x^2e^{-\lambda x^2},\ x>0,\ \lambda>0.$$

then

$$\frac{f'(x)}{f(x)}=\frac{4\sqrt{\frac{\lambda^3}{\pi}}xe^{-\lambda x^2}\left(2-2\lambda x^2\right)}{4\sqrt{\frac{\lambda^3}{\pi}}x^2e^{-\lambda x^2}}=\frac{\left(2-x^2\right)}{x}$$

so that $a_0=a_1=0$, $a_2=1$, $b_0=1$, $b_1=0$ and $b_2=-2\lambda$, hence $\sigma g(x)=-x$ with $h(x)=-2\lambda x^2+3+\mu$.

Example 3.2.14: For the random walk distribution with probability density function

$$f(x)=\sqrt{\frac{\lambda}{2\pi x}}\exp\left[\frac{-\lambda(x\alpha-1)^2}{2x\alpha^2}\right],\ x>0,\ \alpha,\ \lambda>0,$$

we have

$$\log f(x)=\log k+\log x^{-\frac{1}{2}}-\frac{\lambda(x\alpha-1)^2}{2\alpha^2 x},$$

differentiating

$$\frac{f'(x)}{f(x)}=\frac{1}{2x}-\frac{\lambda(x\alpha-1)(x\alpha+1)}{2\alpha^2 x}=\frac{-\lambda\alpha^2x^2-\alpha^2x+\lambda}{2\alpha^2 x}$$

so that $a_0=a_1=0$, $a_2=2\alpha^2$, $b_0=\lambda$, $b_1=-\alpha^2$ and $b_2=-\lambda\alpha^2$, hence $\sigma g(x)=-2\alpha^2x^2$ with $h(x)=-\lambda\alpha^2x^2+3\alpha^2x+\lambda+\mu$.

Table 3.2.2 shows some of the distributions satisfying (3.2.14) and the corresponding $g(x)$ values that enable to obtain individual bounds for these distributions.

The one- parameter exponential family is extensively used in classical as well as in Bayesian inference in view of the pleasant properties it posses in developing optimal estimators and tests. The completeness property of the family is widely exploited in unbiased estimation problem. In the case of a likelihood which belongs to the exponential family there exists a conjugate prior, which is often in the exponential makes easy

Table 3.2.2

Distributions in the extended Pearson family with $h(x)$ and $g(x)$ values

Distributions	$f(x)$	$h(x)$	$-\sigma g(x)$
Inverse Gaussian	$\sqrt{\frac{\lambda}{2\pi x^3}}\exp\left[\frac{-\lambda(x-\alpha)^2}{2x\alpha^2}\right]$ $x>0, \lambda>0$	$-\lambda x^2+\alpha^2 x+\lambda\alpha^2+\mu$	$2\alpha^2 x^2$
Rayleigh	$2\lambda x e^{-\lambda x^2}, x>0, \lambda>0$	$-2\lambda x^2+2+\mu$	x
Maxwell	$4\sqrt{\frac{\lambda^3}{\pi}}x^2 e^{-\lambda x^2}$ $x>0, \lambda>0$	$-2\lambda x^2+3+\mu$	x
Random Walk	$\sqrt{\frac{\lambda}{2\pi x}}\exp\left[\frac{-\lambda(x\alpha-1)^2}{2x\alpha^2}\right]$ $x>0, \alpha, \lambda>0$	$-\lambda\alpha^2 x^2+3\alpha^2 x+\lambda+\mu$	$2\alpha^2 x^2$

analysis of the model in Bayesian setup. A probability density function $f(x,\theta)$ is said to belong to the one- parameter exponential family of distributions defined over a measure space $(\Omega, \mathfrak{B}, \nu)$, if it can be written in the form

$$f(x,\theta) = \exp\left(T(x)Q(\theta) + \boldsymbol{P}(x) + S(\theta)\right) \tag{3.2.15}$$

where $P(x)$ and $T(x)$ are real valued measurable functions over Ω and $Q(\theta)$ and $S(\theta)$ are real functions on $\mathbb{R}$ with $Q(\theta)$ having continuous non-vanishing derivatives in $\mathbb{R}$. If the probability density function of a random variable X belongs to an exponential family expressed in the above form, then

$$E(T(X)) = \mu = -\frac{S'(\theta)}{Q'(\theta)}$$

and

$$V(T(X)) = \frac{Q''(\theta)S'(\theta) - Q'(\theta)S''(\theta)}{\left(Q'(\theta)\right)^3} = -\frac{1}{Q'(\theta)}\frac{\partial \mu}{\partial \theta}.$$

Now we look at the application of the Theorem 3.2.1 to exponential family of distributions.

Theorem 3.2.4: The distribution of X belongs to the exponential family specified by (3.2.15) if and only if

$$\inf_{c(x) \in \mathfrak{B}} \frac{V(c(X))}{E^2\left[g(X)c'(X)\right]} = 1 \tag{3.2.16}$$

with

$$g(x) = \frac{1}{\sigma f(x) Q'(\theta)} \frac{\partial \bar{F}(x)}{\partial \theta}$$

where $\bar{F}(x) = 1 - F(x)$.

Proof:

For the family (3.2.15) we have,

$$\log f(x;\theta) = P(x)Q(\theta) + T(x) + S(\theta).$$

Differentiating with respect to θ,

$$\frac{f_\theta'(x;\theta)}{f(x;\theta)} = P(x)Q'(\theta) + S'(\theta).$$

Integrating with respect to x form t to b

$$\frac{d\bar{F}(t)}{d\theta} = Q'(\theta)\int_t^b P(x)f(x;\theta)dx + S'(\theta)\bar{F}(t),$$

$$\frac{d\bar{F}(t)}{d\theta} = Q'(\theta)\bar{F}(t)E(P(X)X > t) + S'(\theta)\bar{F}(t).$$

Rearranging,

$$E(P(X) \mid X > x) = -\frac{S'(\theta)}{Q'(\theta)} + \frac{1}{Q'(\theta)f(x)}\frac{d\bar{F}(x)}{d\theta}k(x).$$

Now taking $P(X) = h(X)$ for applying Theorem 3.2.1, so that the above equation can be written as

$$E(h(X) \mid X > x) = \mu + \frac{1}{Q'(\theta)f(x)}\frac{d\bar{F}(x)}{d\theta}k(x), \qquad (3.2.17)$$

where

$$\mu = E(P(X)) = -\frac{S'(\theta)}{Q'(\theta)}.$$

Comparing with the second identity in (3.2.3), we have the value of $g(x)$ as

$$g(x) = \frac{1}{\sigma f(x)Q'(\theta)}\frac{\partial\bar{F}(x)}{\partial\theta}. \qquad (3.2.18)$$

Substituting (3.2.18) in (3.2.2) the conclusion (3.2.16) of the Theorem follows.

So far we have specialized Theorem 3.2.1 to some families of distributions that include many of the continuous distributions used in reliability studies. However there are some important distributions like the Weibull, Burr that are not members of these families, and at the same time quite useful in lifetime models. Our general framework in the Theorem 3.2.1 permits us to include them also with appropriate choice of $h(x)$ in each case and this is illustrated by the following example.

Example 3.2.15: For the Weibull distribution with probability density function

$$f(x)=\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^{\alpha}},\quad x>0,\ \alpha,\ \beta>0,$$

we have

$$\frac{f'(x)}{f(x)}=\frac{(\alpha-1)}{x}-\frac{\alpha}{\beta^{\alpha}}x^{\alpha-1}$$

or

$$xf'(x)=(\alpha-1)f(x)-\frac{\alpha}{\beta^{\alpha}}x^{\alpha}f(x).$$

Integrating with respect to x from t to ∞ and assuming $\lim_{x\to\infty}xf(x)=0$, we get

$$-tf(t)-R(t)=(\alpha-1)R(t)-\frac{\alpha}{\beta^{\alpha}}R(t)E\left[X^{\alpha}\mid X>t\right].$$

Multiplying through out by $\beta^{\alpha}\alpha^{-1}R^{-1}(t)$ and rearranging

$$E\left[X^{\alpha}\mid X>t\right]=\beta^{\alpha}+\frac{\beta^{\alpha}}{\alpha}tk(t)$$

or

$$E\left[X^{\alpha}\mid X>x\right]=\beta^{\alpha}+\frac{\beta^{\alpha}}{\alpha}xk(x)$$

which gives

$$E\left[X^{\alpha}+\mu-\beta^{\alpha}\mid X>x\right]=\mu+\frac{\beta^{\alpha}}{\alpha}xk(x).$$

Hence the Weibull distribution has

$$h(x)=x^{\alpha}+\mu-\beta^{\alpha}\ \text{ and }\ g(x)=(\alpha\sigma)^{-1}x\beta^{\alpha}.$$

3.3 Application to unbiased estimation

The data arising from some stochastic phenomena is often modelled by a distribution function that involves some unknown parameter θ. The aim of the analysis is to describe the essential characteristics of the phenomenon that generated the data, which is generally expressed in terms of θ. This raises the problem of specifying a plausible value for θ that is consistent with the observations. Such a value is generally obtained by considering a suitable function of the observations called estimators. In general infinite number of functions can be proposed as estimators of θ, with the property that some estimators far better than the others for certain values of θ. The property unbiasedness is used to restrict the class of estimators in to a smaller one. An estimate T is said to be unbiased for θ if, for every $\theta \in \Theta$, $E(T) = \theta$. This condition ensures that in a long run the estimated value will be correct "on the average". For detailed discussions on unbiasedness one can refer to Chapter 2 of Lehman and Casella (1998). There may be more than one unbiased estimators for a given parametric function and that leads to finding a best estimator from the class of unbiased estimators. An unbiased estimator with uniformly minimum variance (UMVUE) is considered to be the best estimator. This has led to the derivation of expressions for the lower bound of the variance of an unbiased estimator. The following theorems provide two important results in this connection.

Theorem 3.3.1 (Cramer (1946) – Rao (1945) Inequality)

Let $\Theta \subseteq R$ be an open interval and suppose that the family $\{x : f_\theta(x) > 0\}$ has a common support which does not depends on θ and satisfies the regularity conditions.

i) For any statistic h with $E(h(X)) < \infty$ for all θ, then

$$\frac{\partial}{\partial\theta}\int h(x) f_\theta(x)dx = \int h(x)\frac{\partial}{\partial\theta} f_\theta(x)dx.$$

ii) For all θ, the derivative $\frac{\partial}{\partial\theta}\log f_\theta(x)dx$ exist and finite.

If $T(x)$ be an unbiased estimator of $g(\theta)$, then

$$V(T(X)) \geq \frac{(g'(\theta))^2}{E\left(\frac{\partial}{\partial\theta}\log f_\theta(X)\right)^2}.$$

The equality holds if and only if $T(x)$ is a UMVUE of $g(\theta)$.

Theorem 3.3.2 (Chapman Robbins (1951) Inequality)

Let $\Theta \subseteq R$ and $\{x : f_\theta(x) > 0\}$ be the class of probability density (mass) functions. Let $T(x)$ be an unbiased estimator of $\psi(\theta)$ with $E_\theta(T)^2 < \infty$ for all $\theta \in \Theta$. Assume that $\varphi \neq \theta \in \Theta$ such that $f_\theta(x)$ and $f_\varphi(x)$ are different satisfying $\{f_\theta(x) > 0\} \supset \{f_\varphi(x) > 0\}$, then

$$V(T(X)) \geq \sup_{\{\varphi\{f_\theta(x)>0\}\supset\{f_\varphi(x)>0\}\}} \frac{[\psi(\varphi) - \psi(\theta)]^2}{V(f_\varphi(X)/f_\theta(X))},$$

for all $\theta \in \Theta$.

The one parameter exponential family plays an important role in statistical inference, being a class of distributions which admits many desirable properties that lead to optimum inference procedures. Our main theorem (3.2.1) speaks of lower bound to the variance of a function of the random variable and this result as shown to hold for the one parameter exponential family. When the random variable in the Theorem 3.2.1 is replaced by a statistic, the results is transformed to the variance of a statistic whose bounds are of special interest in unbiased estimation. Since there are many types of lower bounds under different conditions, available in the theory of unbiased estimation, it is of interest to compare them with the bound we have obtained in Theorem 3.2.1. First we look at the Cramer- Rao bound.

3.3.1 Comparison with the Cramer- Rao lower bound

In the light of the relationship between conditional expectations and failure rates derived for the one parameter exponential family and the consequent lower bound for the variances it is informative to look and this bound from the estimation point of view. In the search for UMVUE, one criteria that is often used to check whether the estimator chosen satisfy the lower bound prescribed by the Cramer- Rao inequality. The exponential family is one in which the regularity conditions of the Cramer- Rao Theorem is satisfied and offer cases in which the lower bound is attained. Therefore, we continue our discussion with the one parameter exponential family.

Under the conditions on the density (3.2.15) Theorem 3.2.1 apply to $f(x;\theta)$, we can write by taking $h(X)=p(X)$,

$$m(x)=\mu+\sigma k(x)g(x), \tag{3.3.1}$$

where $g(x)$ is given by (3.2.18) and

$$\sigma^2=V\left(p\left(X\right)\right)=\frac{d\mu}{d\theta}/Q'(\theta),\ \mu=-\frac{S'(\theta)}{Q'(\theta)}.$$

When (3.3.1) holds, it is equivalent to

$$\int_x^b h(t)f(t)dt=\mu\overline{F}(x)+\sigma f\left(x\right)g\left(x\right).$$

On substituting (3.2.18) for $g(x)$ in the last equation we have

$$\int_x^b h(t)f(t)dt=\mu\overline{F}(x)+\left[Q'(\theta)\right]^{-1}\frac{\partial\overline{F}\left(x\right)}{\partial\theta}$$

or

$$\int_x^b h(t)f(t)dt=\mu\int_x^b f(t)dt+\left[Q'(\theta)\right]^{-1}\frac{\partial}{\partial\theta}\int_x^b f(t)dt.$$

Differentiating and assuming $\lim_{x\to b}h\left(x\right)f(x)=0$ gives

$$h\left(x\right)f(x)=\mu f\left(x\right)+\left[Q'(\theta)\right]^{-1}\frac{\partial f_\theta\left(x\right)}{\partial\theta}.$$

Dividing through out by $f(x)$ and rearranging

$$\frac{\partial\log f\left(x\right)}{\partial\theta}=Q'(\theta)\left(h\left(x\right)-\mu\right).$$

Thus the Cramer- Rao lower bound becomes

$$\frac{\left(\frac{d\mu}{d\theta}\right)^2}{E\left(\frac{\partial\log f\left(x\right)}{\partial\theta}\right)^2}=\frac{\left(\frac{d\mu}{d\theta}\right)^2}{\left[Q'(\theta)\right]^2\sigma^2}$$

$$=\left(\frac{d\mu}{d\theta}\right)/Q'(\theta). \tag{3.3.2}$$

Also, assuming $\lim_{x\to a}h\left(x\right)f(x)=0$ and using the value of $g(x)$ given in (3.2.18)

$$\sigma E\left(g(X)p'(X)\right)=\int_a^b \frac{h'(x)}{Q'(\theta)}\frac{\partial \overline{F}(x)}{\partial \theta}dx$$

$$=\int_a^b \frac{h'(x)}{Q'(\theta)}\left[\int_x^\infty \frac{\partial f(t)}{\partial t}dt\right]dx.$$

Integrating by parts gives

$$\sigma E\left(g(X)p'(X)\right)=\left.\int_x^\infty \frac{\partial f(t)}{\partial t}dt.h(x)\right]_a^b+\frac{1}{Q'(\theta)}\int_a^b \frac{\partial f(t)}{\partial \theta}h(x)dx$$

$$=\frac{1}{Q'(\theta)}\int_a^b \frac{\partial f(x)}{\partial \theta}h(x)dx$$

$$=\frac{1}{Q'(\theta)}\frac{\partial}{\partial \theta}\left[\int_a^b h(x)f(x)dx\right]$$

$$=\left(\frac{d\mu}{d\theta}\right)/Q'(\theta). \tag{3.3.3}$$

Substituting (3.3.3) in the expression (3.2.2)

$$V\left(h(X)\right)\geq \sigma^2\left(\frac{d\mu}{d\theta}\right)^2/\left(Q'(\theta)\right)^2=\left(\frac{d\mu}{d\theta}\right)/Q'(\theta),$$

so that the Cramer- Rao bound is attained at the same value where the equality in (3.2.2) is holding. Thus in cases where regularity conditions of Cramer- Rao inequality are satisfied the two bounds are equal.

Now we examine the non-regular cases with the aid of an example. When X is uniform over $(0,\theta)$ and $\Theta=(0,\infty)$, it is well known that the estimator $T=\frac{n+1}{n}X_{(n)}$ is unbiased for θ, where $X_{(n)}$ is the largest order statistics in a sample of size n from $U\left(0,\theta\right)$. The distribution of T is given by

$$f(t)=\frac{n^{n+1}}{(n+1)^n}\frac{t^{n-1}}{\theta^n},\ 0<t<\frac{n+1}{n}\theta.$$

Thus

$$F\left(t\right)=\left(\frac{n}{n+1}\right)^n\left(\frac{t}{\theta}\right)^n$$

giving

$$\lambda(t)=\frac{f(t)}{F(t)}=\frac{n^{(n+1)}t^{n-1}}{(n+1)^n\theta^n}\frac{(n+1)^n\theta^n}{n^nt^n}=\frac{n}{t},$$

$$r(t)=\frac{1}{F(t)}\int_0^t tf(t)dt=\frac{(n+1)^n\theta^n}{n^nt^n}\int_0^t t\frac{n^{(n+1)}t^{n-1}}{(n+1)^n\theta^n}dt$$

$$=\frac{n}{t^n}\int_0^t t^ndt=\frac{n}{t^n}\frac{t^{(n+1)}}{(n+1)}=\frac{n}{n+1}t,$$

$$V(T)=E(T^2)-E^2(T)=\frac{(n+1)^2\theta^2}{n(n+2)}-\theta^2=\frac{\theta^2}{n(n+2)},$$

so that from (3.2.3)

$$g(t)=\sqrt{\frac{n+2}{n}}\frac{t}{\theta}\left(\theta-\frac{n}{n+1}t\right).$$

Taking T as the random variable and $c(T)=T$ we see that the equality in (3.2.2) is attained. The corresponding value is

$$V(T)=E(g(T))=\frac{\theta^2}{n(n+2)}$$

which is infact the actual variance of T. Thus there are situations in which our bound is sharper (infact is the minimum variance) than the Cramer- Rao bound.

3.3.2 Comparison with the Chapman- Robbins inequality

Here we compare (3.2.2) with Chapman- Robbins inequality. Assuming $E(h(X))=\mu(\theta)$, $\theta\in\Theta\subset R$ and $\varphi\in\Theta$ such that $f_\theta(x)$ and $f_\varphi(x)$ are different satisfying $\{f_\theta(x)>0\}\supset\{f_\varphi(x)>0\}$, we set

$$c(x)=\left(\frac{f_\varphi(x)}{f_\theta(x)}-1\right)$$

so that (on using (3.2.1))

$$E[c'(X)g(X)]=\sigma^{-1}\int_a^b c'(x)\int_a^x(\mu-h(t))f_\theta(t)dtdx$$

$$=\sigma^{-1}\int_a^b c(x)(h(x)-\mu)f_\theta(x)dx$$

$$=\sigma^{-1}\int_a^b\left(\frac{f_\varphi(x)-f_\theta(x)}{f_\theta(x)}\right)(h(x)-\mu)f_\theta(x)dx$$

$$=\sigma^{-1}\int_a^b(f_\varphi(x)-f_\theta(x))(h(x)-\mu)dx$$

$$=\sigma^{-1}\left[\mu(\varphi)-\mu(\theta)\right]. \tag{3.3.4}$$

Substituting (3.3.4) in (3.2.2)

$$V(c(X))\geq\left[\mu(\varphi)-\mu(\theta)\right]^2/V(h(X))$$

or

$$V(h(X))\geq\left[\mu(\varphi)-\mu(\theta)\right]^2\left[V(f_\varphi(X)/f_\theta(X))\right]^{-1},$$

which is the Chapman-Robbins inequality.

The different conditions under which the bounds discussed here hold form an interesting aspect to be considered in actual minimum variance unbiased estimation problems. Our conclusion is that when the regularity condition of Cramer-Rao inequality are met, the bounds obtained there is identical with (3.2.2) and in addition in non-regular cases latter may provide better bounds than the former. The generalization attempted here paves way for UMVUE's for many parametric functions than in the existing results. The Chapman- Robbin's inequality is obtained as a particular cases of (3.2.2), when the conditions for the former inequality are met in our general framework discussed for the latter. In the sense we have discussed it appears that the Chernoff- type inequality is a more general result and could be employed to extract minimum variance unbiased estimators. Though relationships between failure rate and conditional expectations are necessary and sufficient conditions, it is only the $g(x)$ value that matters in the inference situations. The method discussed here provides an alternate methodology to arrive at UMVUE's.

3.4 Application to reliability modelling

We now examine some implication of the above theorems in reliability modelling. In the following X is restricted to be a non- negative random variable representing the lifetime so that $F(0-)=0$ satisfying the other conditions of Theorem 3.2.1. The expression for $g(x)$ satisfying (3.2.3) is unique for a given distribution and one of the expressions in (3.2.3) depending on whether data is left or right truncated, can be used to find $g(x)$ based on the observations. Then equation (3.2.4) enables to identify the appropriate model followed by the data. With suitable choice of $h(x)$ for a desired μ, one can also obtain estimate having minimum variance. In other words, the Theorem 3.2.1 enables model identification as well as choice of estimator for a given parametric function. In respect of the families mentioned in Theorem 3.2.2 and 3.2.3 the form of the distribution depends on the nature of the roots of the equation $a_0 + a_1 x + a_2 x^2 = 0$. Since $g(x) = 0$ has the very same roots as $a_0 + a_1 x + a_2 x^2 = 0$, the distribution can be identified from $g(x)$ itself, without going through the integration involved in (3.2.4). Further the solutions derived from $\frac{f'(x)}{f(x)}$ and $\frac{g'(x)}{g(x)}$ end up with the same form of distributions, except for a possible change in the parameters. The identity (3.2.4) is more helpful while finding the monotoncity of probability distributions using the well known results of Glaser (1980).

3.5 Statistical catastrophe

In a broad sense catastrophe theory describes how small changes in the control parameters (viewed as independent variables that affect system) can have sudden discontinuous effects on the dependent variables. The theory is extensively used in ecology, biology, physics, chemistry, economics and psychology.

Stochastic catastrophe models are governed by the differential equation

$$\frac{dx}{dt} = -\frac{\partial u}{\partial x} + \sqrt{g(x)}\frac{dw(t)}{dt} \tag{3.5.1}$$

where x is a real valued state variable, $w(t)$ is the standard Weiner process, and $g(x)$ that controls the random input is called the infinitesimal variance function. The dynamics of $X(t)$ are controlled by the function $U(x)$ and the statistical catastrophe theory studies the behaviour of (3.5.1) at theoretical points for which

$$\frac{\partial u}{\partial x} = 0 \text{ and } \frac{\partial^2 u}{\partial x^2} = 0.$$

If $f(u, t, x_0)$ is the probability density function of the random variable X at t, given an initial position x_0, it satisfies the differential equation

$$\frac{\partial f}{\partial t} = -\frac{\partial (gf)}{\partial u} + \frac{1}{2}\frac{\partial^2 (gf)}{\partial u^2}. \tag{3.5.2}$$

As $t \to \infty$, f converges to stationary for f^* such that $\frac{\partial f^*}{\partial t} = 0$, and takes the form

$$f^*(x) = K \exp\left(-\int^x \frac{2m(x) - g'(s)}{g(s)} ds \right). \tag{3.5.3}$$

Identifying $2m(x)$ with $-\frac{\partial u}{\partial x}$, the function $f^*(x)$ will be the theoretical probability density function for use in statistical catastrophe theory. We can write (3.5.3) as

$$\frac{d \log f^*}{dx} = -\frac{1}{g}\left(\frac{\partial U}{\partial x} + \frac{\partial g}{\partial x} \right). \tag{3.5.4}$$

If we denote

$$v(x) = \frac{\partial U}{\partial x} + \frac{\partial g}{\partial x},$$

the equation (3.5.4) becomes

$$\frac{d \log f^*}{dx} = -\frac{v(x)}{g(x)} \tag{3.5.5}$$

or $f^*(x)$ has the form

$$f^*(x) = K \exp\left(-\int^x \frac{g(t)}{v(t)} dt \right),$$

in which $v(x)$ is called a shape function (See Cobb (1981) for details). It is observed in Cobb (1981) that apart from a constant, the values $1, x$ and $1-x$ provide reasonable catastrophe models (normal, gamma and beta) corresponding to a linear $v(x)$ in equation (3.5.5). It is evident that his $g(x)$ functions coincide with $g(x)$ in Table 3.2.1 and the form of the density in Theorem 3.2.1 and the characterization thereof, provides a statistical basis for generating catastrophe models. Moreover Theorem 3.2.1 and Table 3.2.1 throw up more plausible models. Catastrophes densities correspond to $v(x)$ that are quadratic or higher. Table 3.2.2 presents $g(x)$ that conforms to quadratic $v(x)$. In general, it is easy to convert (3.5.5) into the form of (3.2.3) using the methods described earlier for any functional form of $v(x)$ and $g(x)$ to derive general form of density $f^*(x)$, that is appropriate for a given data situation.

An important aspect of statistical catastrophe theory is the estimation of the model parameters. Our discussions can also help to find UMVUE's of desired parametric functions.

Chapter 4

Characterization of Discrete Distributions by Variance Bound*

4.1 Introduction

The literature on reliability theory mainly deals with non- negative absolutely continuous random variables. However, quiet often we come across with situations where the product life can be described through non- negative integer valued random variable. For example (i) a device can be monitored only once per time period and the observation is taken as the number of time periods successfully completed prior to the failure of a device and (ii) a piece of equipment may operate in cycles and we measure the number of cycles completed prior to failure. Reliability theory, therefore, needs to be developed for discrete description of life maintaining similarity with the continuous counterpart. Salvia and Bollinger (1982) and Xekalaki (1983) pointed out situations where the product life is discrete in nature and gave characterization results concerning the geometric, Waring and negative hyper geometric distributions in terms of failure rates. The topic of characterization based on discrete reliability concepts is discussed subsequently in many papers by Nair (1983), Hitha and Nair (1989) and Nair and Hitha (1989). For surveys of discrete concepts and distributions one can refer to Pudget and Spurrier (1985), Ebrahimi (1986), Guess and Perk (1988) and Bracquemond and Gaudoin (2002). Recently, Kemp (2004) exhaustively studied the ageing behaviour of discrete life distributions and gave some new results in this direction.

In this chapter, as a continuation of the work done in Chapter 3, a general theorem that links characterizations of discrete life distributions based on relationship between failure rate and conditional expectations with those in terms of Chernoff-type inequalities

* Part of the work done in this chapter is to appear in Nair and Sudheesh (2007).

is proposed. Exact expression for lower bounds to the variance is calculated for distribution belonging to the modified power series family, Ord family, Katz family and mixture geometric models. It is shown that the bounds obtained here contain the Cramer-Rao and Chapman-Robbins inequalities as special cases. An application of the results to real data is also provided.

4.2 Main result

Several papers in literature address the problem of characterizing probability distributions through Chernoff-type inequalities satisfying specific conditions when the domain of application is measured in continuous scale and the results in discrete setup is established as analogous results of continuous counterpart. Alharbi and Shanbhag (1996) address the same problem with measure theoretical framework and pointed out the application of the results in characterizing life distributions through a result similar to Cox representation of the survival function in terms of failure rate and suggest cases of some continuous distribution as illustrations. In the present chapter we establish a general characterization theorem that combines the results available in the two approaches described in Sections 3 and 4 of Chapter 2.

The concepts and definitions required for the work in the subsequent sections consist of a class of discrete probability distributions supported by the set N of non-negative integers, the set $\mathfrak{C}$ of real valued functions $c(.)$, of a random variable X defined on N having finite variance along with

$$m(x) = E\left(h(X) \mid X > x\right)$$

$$r(x) = E\left(h(X) \mid X \le x\right)$$

for a function $h(X) \in \mathfrak{C}$ such that

$$E\left(h^2(X)\right) < \infty,\ E\left(h(X)\right) = \mu \text{ and } V\left(h(X)\right) = \sigma^2.$$

In the above formulation $p(x)$, $F(x)$ and $R(x) = P(X \ge x)$ denote respectively the probability mass function, distribution function and survival function of X so that

$$k(x) = \frac{p(x)}{R(x)}$$

and

$$\lambda(x) = \frac{p(x)}{F(x)}$$

are the failure rate and reversed failure rate of X respectively. Note that in discrete setup the failure rates and the reversed failure rates are probabilities which is not the case in the continuous domain.

Next we present a general result that meets the objective and also subsumes most of the existing results that were taken up in the two streams of characterizations mentioned above.

Theorem 4.2.1:

Let X be a discrete random variable supported on N or a subset thereof and $g(.)$, $c(.)$, $h(.)$ be functions in $\mathfrak{C}$ such that $E\left(c^2(X)\right) < \infty$, $E\left(g(X)\Delta c(X)\right) < \infty$ and $E\left(h^2(X)\right) < \infty$. Then for every $c(x) \in \mathfrak{C}$ and some $g(x)$ and $h(x)$, the following statements are equivalent.

(i) $$\frac{p(x+1)}{p(x)} = \frac{\sigma g(x)}{\sigma g(x+1) - \mu + h(x+1)}, \quad x = 0,1,2,... \tag{4.2.1}$$

with $g(0) = \left(\mu - h(0)\right)/\sigma$ and $p(0)$ is evaluated from $\sum_0^x p(x) = 1$.

(ii) $$r(x) = \mu - \sigma\lambda(x)g(x) \tag{4.2.2}$$

(iii) $$m(x) = \mu + \frac{\sigma k(x) g(x)}{1 - k(x)} \tag{4.2.3}$$

(iv) $$V\left(c(X)\right) \geq E^2\left(g(X)\Delta c(X)\right), \tag{4.2.4}$$

provided

$$E\left(g(X)\Delta h(X)\right) = \sigma. \tag{4.2.5}$$

Here μ and σ^2 denote respectively the mean and variance of $h(X)$ and $\Delta c(x) = c(x+1) - c(x)$.

Proof:

Assuming (4.2.1) we have

$$\sigma g(x+1)p(x+1) - \mu p(x+1) + h(x+1)p(x+1) = \sigma g(x)p(x)$$

or

$$(\mu - h(x))\,p(x) = \sigma p(x)g(x) - \sigma p(x-1)g(x-1).$$

Summation from 1 to x and the use of the values of $g(0)$ from (i) leads to

$$\sigma p(x)g(x) = \sum_{0}^{x}(\mu - h(y))\,p(y) \tag{4.2.6}$$

or

$$\sigma p(x)g(x) = \mu F(x) - F(x)r(x).$$

Dividing by $F(x)$, we reach (4.2.2). Retracing the steps we get (4.2.1). Thus (i) $\Leftrightarrow$ (ii).

Now, from

$$r(x)F(x) + m(x)(1 - F(x)) = \mu$$

one can solve for $F(x)$ and $R(x)$, and then use the definition of failure rate and reversed failure rate to reach the identity

$$\frac{\mu - r(x)}{\lambda(x)} = \frac{(m(x)-\mu)(1-k(x))}{k(x)}$$

which proves the equivalence of (ii) and (iii). From the results of Cacoullos and Papathanasiou (1997) stated at (2.2.24) and (2.2.25), we take $z(x) = \sigma g(x)$ and obtain

$$V(c(X)) \geq \frac{\sigma E^2(g(X)\Delta c(X))}{E(g(X)\Delta h(X))}$$

if and only if (ii) or equivalently (iii) is satisfied. Further

$$\begin{aligned}
E(g(X)\Delta h(X)) &= \sum_{0}^{\infty}(h(x+1) - h(x))g(x)p(x) \\
&= \sigma^{-1}\sum_{0}^{\infty}(h(x+1) - h(x))\left(\sum_{0}^{x}(\mu - h(y)p(y))\right) \\
&= \sigma^{-1}\sum_{0}^{\infty}h(x)(h(x) - \mu)p(x) \\
&= \sigma^{-1}V(h(X)) = \sigma.
\end{aligned}$$

This proves that (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv). Since (iv) $\Rightarrow$ (ii) $\Rightarrow$ (i), the chain of implications in the Theorem is complete.

Remark 4.2.1: The equality in (4.2.4) holds if and only if $c(x)$ is linear in $h(x)$.

Remark 4.2.2: The results of Cacoullos and Papathanasiou (1989, 1997) and Ruiz and Navarro (1994) are special cases of Theorem 4.2.1

Remark 4.2.3: $E(g(X)) = \sigma^{-1}Cov(X, h(X))$.

This follows from

$$E(g(X)) = \sigma^{-1}\sum_{x=0}^{\infty}\sum_{y=0}^{x}(\mu - h(y))p(y)$$

$$= \sigma^{-1}\sum_{x=0}^{\infty}\sum_{y=x+1}^{\infty}(h(y) - \mu)p(y)$$

$$= \sigma^{-1}\sum_{x=0}^{\infty}x(h(x) - \mu)p(x).$$

Remark 4.2.4: For a given $h(x)$ the value of $g(x)$ characterizes the distribution of X. Thus for $h(x) = x$, the random variable X has the Poisson distribution in the class of discrete probability distributions supported by the set N of non- negative integers if and only if $g(x) \equiv \lambda^{1/2}$ for all x.

We now consider some illustrations of the above results to probability modelling. Since the modified power series family, the Ord family and the Katz family which together covers most of the discrete life distributions in common use we find the characterization results for the same families. And then provide some examples of individual distribution to check the validity of the results.

The power series distributions is specified by the probability mass function

$$p(x) = \frac{a(x)\theta^x}{A(\theta)}, \quad x = 1, 2, \ldots, \quad \theta > 0 \tag{4.2.7}$$

where $a(x) \geq 0$, $\sum_{x=0}^{\infty} a(x)\theta^x = A(\theta)$, contains several discrete life distributions used in practice. Patil (1961, 1962) allowed the set of values that the variable can take to be any non- empty enumerable set S of non- negative integers and called the resulting distribution as generalized power series distributions. Gupta (1974) replace θ^x in (4.2.7) by $(u(\theta))^x$ to obtain the modified power series distributions. This distributions is also called linear exponential and it admits several desirable properties helpful in inference problems.

Theorem 4.2.2:

The distribution of X follows the modified power series distributions specified by

$$p(x) = \frac{a(x)(u(\theta))^x}{A(\theta)}, \tag{4.2.8}$$

where $X \in N$, $a(x) \geq 0$, $u(\theta)$ and $A(\theta)$ are positive, finite and differentiable if and only if

$$\inf_{c(x) \in \mathfrak{C}} \frac{V(c(X))}{E^2(g(X)\Delta c(X))} = 1 \tag{4.2.9}$$

with

$$g(x) = -\frac{u(\theta)}{u'(\theta)} \frac{1}{\sigma p(x)} \frac{\partial F(x)}{\partial \theta}.$$

Proof:

From (4.2.8)

$$A(\theta)F(x) = \sum_{y=0}^{x} a(y)(u(\theta))^y,$$

differentiation with respect to θ yield

$$A'(\theta)F(x) + A(\theta)F'(x) = \frac{u'(\theta)}{u(\theta)} \sum_{y=0}^{x} ya(y)(u(\theta))^y \tag{4.2.10}$$

or

$$A'(\theta) + A(\theta)(\log F(x))' = u'(\theta)(u(\theta))^{-1} r(x). \tag{4.2.11}$$

Using the expressions for $E(X)$ (see Johnson et. al. (1992))

$$\mu = \frac{A'(\theta)}{A(\theta)} \frac{u(\theta)}{u'(\theta)}, \tag{4.2.11a}$$

the equation (4.2.11) reduces to

$$r(x) = \mu\left(1 + (\log A(\theta))'(\log F)'\right)$$

or

$$r(x) = \mu\left(1 + \frac{A'(\theta)}{A(\theta)} \frac{1}{F(x)} \frac{\partial F(x)}{\partial \theta}\right),$$

when $h(x) = x$ and comparing with (4.2.2),

$$g(x) = -\frac{\mu}{\sigma} \frac{A(\theta)}{A'(\theta)} \frac{1}{p(x)} \frac{\partial F(x)}{\partial \theta}$$

$$= -\frac{u(\theta)}{u'(\theta)} \frac{1}{\sigma p(x)} \frac{\partial F(x)}{\partial \theta}. \tag{4.2.12}$$

Then the proof follows from Theorem 2.1 with $g(x)$ values as stated in (4.2.12).

Remark 4.2.5: Differentiating (4.2.10) again with respect to θ , by similar way as above we can arrive at

$$r(x) = (\log F)'\left(\log F'\right)' + 2(\log F)'(\log A(\theta))' + \left(\log A'(\theta)\right)'(\log A(\theta))'$$

when

$$h(x) = \left(\log u'(\theta)\right)'^{2} x(x-1) + \left(\log u'(\theta)\right)'(\log u(\theta))' x - \left(\log A'(\theta)\right)'(\log A(\theta))'$$

and

$$g(x) = \frac{1}{\sigma p(x)}\left(\frac{\partial^2 F(x)}{\partial \theta^2} + 2\frac{A'(\theta)}{A(\theta)} \frac{\partial F(x)}{\partial \theta}\right), \tag{4.2.13}$$

since in that case $E(h(X)) = \mu = 0$ and this shows that we can have different vales of $g(x)$ for different choice of $h(x)$.

Remark 4.2.6: When $u(\theta)=\theta$ in the above discussion, the results for the sub-class of generalized power series distributions can be obtained.

The results given in the Theorem 4.2.1 suggest that the characterizations in terms of relationship between failure rate and mean residual lives are not independent of those between the corresponding reversed concepts. But the two sets of results are useful in their own right depending on whether the data is left or right truncated. It may be noted that, though (4.2.12) appears to be complicated, it ends up with a simple forms for various members and can be verified from the following example.

Example 4.2.1: Consider the Poisson distributions with probability mass function

$$p(x)=\frac{e^{-\lambda}\lambda^{x}}{x!},\ x=0,1,2,...$$

then

$$\frac{dF(x)}{d\theta}=\frac{d}{d\theta}\left[\sum_{0}^{x}\frac{e^{-\lambda}\lambda^{x}}{x!}\right]$$

$$=\sum_{0}^{x}\frac{e^{-\lambda}\lambda^{x}}{x!}\left(\frac{x}{\lambda}-1\right)=-\frac{e^{-\lambda}\lambda^{x}}{x!}.$$

Hence

$$g(x)=-\frac{u(\theta)}{u'(\theta)}\frac{1}{\sigma p(x)}\frac{\partial F(x)}{\partial\theta}=\lambda^{1/2}$$

and

$$\inf_{c(x)\in\mathfrak{C}}\frac{V(c(X))}{\lambda.E^{2}(\Delta c(X))}=1.$$

Substituting the value of $g(x)$ in the identity (4.2.3) we have the result of Ahmed (1991) about the Poisson distribution.

The Ord family of distributions is considered as a discrete analogue of Pearson system where $f'(x)$ in the equation (2.4.5) is replaced by $\Delta f(x)=f(x+1)-f(x)$ and the probability mass function satisfies the difference equation (2.4.13). Most of the discrete distributions even if do not belong to the modified power series family are the members of the Ord family.

Theorem 4.2.3:

The distribution of X follows the Ord family of distributions specified by (2.4.5) if an only if

$$\underset{c(x)\in\mathfrak{C}}{inf} \frac{V(c(X))}{E^2\left(\left(b_0+b_1X+b_2X^2\right)\Delta c(X)\right)}=1 \tag{4.2.14}$$

where

$$b_0=\frac{\mu}{\sigma}+\frac{a_0-a_1+a_2}{(1-2a_2)\sigma},\ b_1=\frac{a_1-1}{(1-2a_2)\sigma},\ b_2=\frac{a_2}{(1-2a_2)\sigma} \text{ and } a_2\neq\frac{1}{2}.$$

Proof:

In analogy with (4.2.3) the equation (2.4.5) can be written as

$$\frac{p(x+1)}{p(x)}=\frac{(a_0-d)+(a_1-1)x+a_2x^2}{a_0+a_1x+a_2x^2}$$

$$=\frac{\sigma g(x)}{\sigma g(x+1)-\mu+h(x+1)}. \tag{4.2.15}$$

Obviously, $g(x)$ must be a quadratic function of x of the form $g(x)=b_0+b_1x+b_2x^2$ for $h(x)=x$. Substituting in to the above equation and identifying the coefficient of x we get the values of b_i, $i=1,2,3$ as stated in the Theorem. Then applying Theorem 4.2.1 we have the result (4.2.14).

Remark 4.2.7: The results stated in Theorem 4.2.3 can be compared to that of Korwar (1991) where he discussed the characterization of Ord family by variance bound and hence is a special case of Theorem 4.2.1. Further, the results of Nair and Sankaran (1991) about the same family is also a special case of Theorem 4.2.1 where they characterize (2.4.5) by the identity (2.4.6).

Example 4.2.2: Consider the binomial distributions with probability mass function

$$p(x)=\binom{n}{x}p^x(1-p)^{n-x},\ x=1,2,\ldots n,\ 0<p<1,$$

then

$$\frac{p(x+1)-p(x)}{p(x)}=\frac{(n-x)p}{(x+1)(1-p)}-1$$

$$=\frac{(n-x)p-(x+1)(1-p)}{(x+1)(1-p)}$$

$$=\frac{(n-x)p-(x+1)(1-p)}{(x+1)(1-p)}$$

$$=\frac{(n-x)p-(x+1)(1-p)}{(x+1)(1-p)},$$

so that

$$a_0=a_1=(1-p),\ a_2=0$$

and

$$\sigma g(x)=p(n-x),$$

hence

$$m_1(x)=\mu+(1+x)(1-p)k(x+1),$$

$$\inf_{c(x)\in\mathfrak{C}}\ \frac{V(c(X))}{E^2(p(n-x)\Delta c(X))}=1.$$

The characterization of binomial distribution by the relation between failure rate and conditional expectation obtained by Ahmed (1991) is identical with the above result.

Pearson (1895) use the difference equation

$$\frac{p(x)-p(x-1)}{p(x-1)}=\frac{(a-x)}{b_0+b_1x+b_2x(x-1)}$$

as the starting point for obtaining the differential equation defining the Pearson system of continuous distributions. Katz's (1946) restriction $b_2=0,\ b_0=b_1$ gives rise to the Katz family of distributions and the following theorem is the application of the Theorem 4.2.1 to the same family.

Theorem 4.2.4:

The distribution of X follows the Katz family specified by

$$\frac{p(x+1)}{p(x)}=\frac{\alpha+\beta x}{1+x} \tag{4.2.16}$$

if and only if

$$\underset{c(x)\in\mathfrak{C}}{Inf}\ \frac{\sigma^2(1-\beta)V(c(X))}{E^2((\alpha+\beta X)\Delta c(X))}=1.$$

Proof:

When $h(x)=x$, comparing (4.2.16) with (4.2.3) we have

$$\frac{\sigma g(x)}{\sigma g(x+1)-\mu+(1+x)}=\frac{\alpha+\beta x}{1+x}. \tag{4.2.17}$$

Clearly $g(x)$ is linear function of x, substitute $g(x)=a_0+a_1x$ in equation (4.2.17) we have

$$\frac{\sigma(a_0+a_1x)}{\sigma(a_0+a_1(1+x))-\mu+(1+x)}=\frac{\alpha+\beta x}{1+x}.$$

Comparing the coefficient gives

$$a_0=\frac{\alpha}{\sigma(1-\beta)} \text{ and } a_1=\frac{\beta}{\sigma(1-\beta)}.$$

Now the proof follows from Theorem 4.2.1.

Example 4.2.3: For the negative binomial distribution with probability mass function

$$f(x)=\binom{k+x-1}{k-1}p^k(1-p)^x,\ x\in N$$

we have

$$\frac{p(x+1)}{p(x)}=\frac{(1-p)(k+x)}{(1+x)}.$$

Comparing with (4.2.16) gives

$$\alpha=k(1-p) \text{ and } \beta=(1-p).$$

Hence the value of $g(x)$ is given by

$$\sigma g(x)=\frac{(1-p)}{p}(k+x)$$

and

$$\underset{c(x)\in\mathfrak{C}}{inf} \frac{\sigma^2(1-\beta)V(c(X))}{E^2((\alpha+\beta X)\Delta c(X))}=1.$$

The result of Osaki and Li (1988) about the negative binomial distribution can be had from the identity (4.2.3) by means of $g(x)$ value given above.

We now show that Theorem 4.2.1 can be applied to some finite mixture of discrete distributions as well. The mixture of geometric laws

$p(x)=\alpha p_1 q_1^x+(1-\alpha)p_2 q_2^x$, $0<p_i<1$, $q_i=(1-p_i)$, $i=1,2; 0\le\alpha\le 1$; $x\in N$

is characterized by (Nair et. al (1999))

$$E(X-x\mid X>x)=\frac{p_1+p_2}{p_1p_2}-\frac{1}{p_1p_2}k(x+1)$$

so that by taking $h(X)=X-x$ and comparing with (4.2.3)

$$g(x)=\left[\sigma p_1p_2 f(x)\right]^{-1}\left[(p_1+p_2-\mu p_1p_2)R(x+1)-f(x+1)\right].$$

Table 4.2.1 shows the values of $g(x)$ for some well known discrete distributions so that the characterizations in terms of (4.2.2), (4.2.3) and (4.2.4) can be easily deduced from the Theorem 4.2.1.

4.3 Unbiased estimation

The present section is a discussion of the implications of Theorem 4.2.1 to unbiased estimation and a comparison of inequality (4.2.4) with the Cramer-Rao and Chapman-Robbins lower bounds to the variance of an unbiased estimator. First we take $h(x)=x$ and note that $h(x)\in\mathfrak{C}$. The lower bound in (4.2.4) is attained when $c(x)=h(x)$, in which case

$$V(c(X))=\sigma^2. \tag{4.3.1}$$

A necessary and sufficient condition for this is

$$r(x)=\mu-\sigma\lambda(x)g(x),$$

which is equivalent to

Table 4.2.1

Values of $g(x)$ for some discrete distribution.

Distribution	$p(x)$	$\sigma g(x)$
Uniform	$(n+1)^{-1}$	$\mu n^{-1}(n-x)(x+1)$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$	$p(n-x)$
Poisson	$(x!) e^{-\lambda} \lambda^x$	λ
Negative binomial	$\binom{k+x-1}{k-1} p^x (1-p)^{k-x}$	$\frac{(1-p)}{p}(x+k)$
Hypergeometric	$\binom{a}{x}\binom{b}{n-x} / \binom{a+b}{n}$	$\mu(na)^{-1}\left[x^2+(n-a)x\right]+\mu$
Yule	$P(P!)x!/(x+P+1)!$	$\mu(x^2+2x+1)$
Waring	$(c-a)(a)_x/(c)_{x+1}$	$\mu a^{-1}\left(x^2+(1+a)x\right)+\mu$
Discrete student's t	$\alpha\left[\prod_{j=1}^{k}\left\{(j+x+a)^2+b^2\right\}\right]^{-1}$	$-(4k)^{-1}\mu\left[2akx^2-(1+a)(2a+k)x\right]+$ $1-(4k)^{-1}\mu(2a+k)(k^2+2k+a^2-a)$
Lagrangian Poisson	$(x!)^{-1}(1+\alpha x)^{x-1}\theta^x e^{-\theta(1+\alpha x)}$	$-\frac{\mu}{f}\frac{\partial F}{\partial \theta}$

$$\sum_0^x h(t)p(t) = \mu\sum_0^x p(t) + \mu\frac{A(\theta)}{A'(\theta)}\frac{\partial}{\partial\theta}\sum_0^x p(t) \tag{4.3.2}$$

on using (4.2.12). From (4.3.2),

$$h(x)p(x) = \mu p(x) + \mu\frac{A(\theta)}{A'(\theta)}\frac{\partial p(x)}{\partial\theta},$$

dividing through out by $p(x)$ and using the relation (4.2.11a) we have

$$\frac{\partial \log p(x)}{\partial\theta} = \frac{u'(\theta)}{u(\theta)}(h(x)-\mu).$$

Now the Cramer-Rao lower bound for unbiasedly estimating μ using $h(x)$ is

$$V(c(X)) = \frac{[\mu'(\theta)]^2}{E\left(\frac{\partial\log p}{\partial\theta}\right)^2}$$

$$= \frac{u(\theta)}{u'(\theta)}\frac{\partial\mu}{\partial\theta} = \sigma^2. \tag{4.3.3}$$

Hence the two bounds in (4.3.1) and (4.3.3) are equal. Notice that modified power series distribution is linear exponential and hence include cases in which the Cramer-Rao lower bound is attained, under regularity conditions.

Second popular lower bound to the variance of an unbiased estimator is provided by the Chapman-Robbins inequality. When $E(h(X)) = \mu(\theta)$ where $\theta \in \Theta \subset R$ and $\varphi \in \Theta$ such that $p_\theta(x)$ and $p_\varphi(x)$ are different, satisfying $\{p_\theta(x) > 0\} \supset \{p_\varphi(x) > 0\}$, we can set

$$c(x) = \left(\frac{p_\varphi(x)}{p_\theta(x)} - 1\right)$$

in (4.2.4). Now consider

$$E(g(X)\Delta c(X)) = \sigma^{-1}\sum_{x=0}^{\infty}\left(\sum_{y=0}^{x}(\mu - h(y))p(y)\right)\left(\frac{p_\varphi(x+1)}{p_\theta(x+1)} - \frac{p_\varphi(x)}{p_\theta(x)}\right) \tag{4.3.4}$$

$$= \sigma^{-1} \left\{ \sum_{x=1}^{\infty} \left(\sum_{y=0}^{x-1} (\mu - h(y)) p_\theta(y) \frac{p_\varphi(x)}{p_\theta(x)} \right) - \right.$$

$$\left. \sum_{x=0}^{\infty} \left(\sum_{y=0}^{x} (\mu - h(y)) p_\theta(y) \frac{p_\varphi(x)}{p_\theta(x)} \right) \right\}$$

$$= -\sigma^{-1} \sum_{0}^{\infty} (\mu - h(x)) p_\varphi(x)$$

$$= \sigma^{-1} \left[\mu(\varphi) - \mu(\theta) \right].$$

Inequality (4.2.4) then reduces to

$$V(c(X)) \geq \left[\mu(\varphi) - \mu(\theta) \right]^2 / V(h(X))$$

or

$$V(h(X)) \geq \frac{\left[\mu(\varphi) - \mu(\theta) \right]^2}{V\left(p_\varphi(X) / p_\theta(X)\right)},$$

which is the Chapman- Robbins inequality. It is well known that this bound does not require the regularity conditions of the Cramer-Rao inequality, is valid when Θ is discrete and provides bounds sharper than the latter. The last statement is also true for the Chernoff-type inequality (4.2.4) derived in Theorem 4.2.1, which is more general. Moreover, (4.2.4) provides a more general alternative methodology to extract UMVUE's when $h(x)$ is taken as a statistic that is unbiased for μ.

4.4 Illustration

Although the above deliberations were essentially directed towards modelling and inference of lifetime data, the methodology is applicable for identification of distribution and estimation of parameters in other contexts as well. We illustrate the procedure for the data on the count of alpha particles giving rise to Poisson distribution reported in Mould (2005)

X	0	1	2	3	4	5	6	7	8	9	10	11	12
Frequency:	57	203	383	525	532	408	273	139	45	27	10	4	2

The failure rate, mean residual life (which has no physical interpretation in the present data) and $g(x)$ are plotted in figures 4.4.1, 4.4.2 and 4.4.3. Values of $g(x)$ corresponding to x from 0 through 12 are respectively 2.00, 2.06, 2.07, 1.96, 1.87, 1.85, 1.65, 1.61, 2.81, 2.01, 2.23 and 2.12. Except for a small aberration around the value 2.81 (Caused due to the observed frequency 45 at $x=8$ which in the Poisson fit gives a clear distant theoretical value of 60. It is also seen from the graph of the failure rate that at this point the failure rate is decreasing which is not so for the Poisson model) $g(x)$ remains constant about it average value 2.02 showing that Poisson model adequately describes the data. If we consider the random variable as X, the sum of n independent and identically distributed observations following Poisson distribution so that $E(X)=n\lambda,\ n=2608$, from the results in Section 3, the UMVUE for λ is 3.87. This is also very close to the unbiased estimate of λ obtained from $E(g(X))$.

If we look at the unbiased estimation of the probability mass function

$$p(r)=\frac{e^{-\lambda}\lambda^{r}}{r!}, \qquad r=0,1,2,... \tag{4.4.1}$$

then for a real function $h(x)$ we have

$$E(h(X))=\frac{e^{-\lambda}\lambda^{r}}{r!}$$

$$\sum_{0}^{\infty}h(x)e^{-n\lambda}(n\lambda)^{x}/x!=p(r)=\frac{e^{-\lambda}\lambda^{r}}{r!}, \tag{4.4.2}$$

which gives

$$\sum_{0}^{\infty}\frac{h(x)(n\lambda)^{x}}{x!}=\sum_{0}^{\infty}\frac{\left[(n-1)\lambda\right]^{x}\lambda^{r}}{x!r!}.$$

Comparing the coefficient of λ^{x+r} we have

$$\frac{h(x+r)n^{x+r}}{(x+r)!}=\frac{(n-1)^{x}}{x!r!},$$

which on simplification gives

$$h(x)=\binom{x}{r}\left(\frac{1}{n}\right)^{r}\left(1-\frac{1}{n}\right)^{x-r} \tag{4.4.3}$$

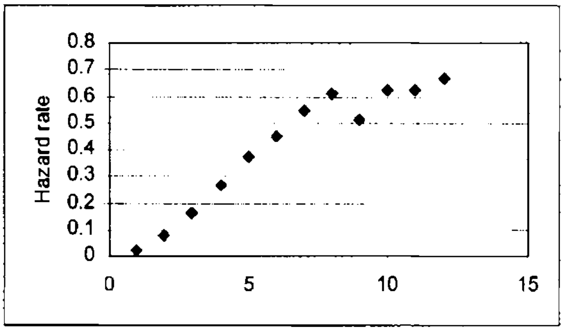

Fig.4.4.1: Hazard rate of X

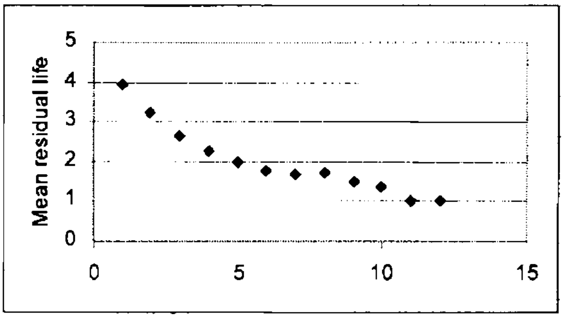

Fig.4.4.2: Mean residual life of X

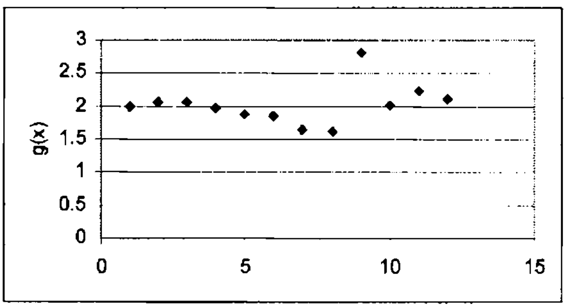

Fig.4.43: The value of $g(x)$

Taking in $c(x)=h(x)$ in Theorem 4.2. 1, equation (4.2. 2) can be written as

$$\sum_{y=0}^{x} h(y)f(y)=\sum_{y=0}^{x} f(r)f(y)-\sigma g(x)f(x),$$

and the value of $g(x)$ is given by

$$\sigma g(x)=\frac{x!}{(n\lambda)^{x}}\sum_{y=0}^{x}\{f(r)-h(y)\}(n\lambda)^{y}/y!$$

so that $h(x)$ in (4.4.3) is unbiased for $f(r)$ and attain the minimum variance bound (4.2.4).

In conclusion, in this chapter we arrives at the class of discrete probability distributions typified by (4.2.1) that can be used for reliability modelling in terms of characteristic properties represented in (4.2.2) and (4.2.3) in terms of failure rate (or reversed failure rate) and right (or left) truncated expectations. The link established with Chernoff-type inequalities further enables to assist in the unbiased estimation of parametric functions with properties that subsumes some of the well-known results in the classical theory of estimation.

Chapter 5

Characterization of Continuous Distributions by Properties of Conditional Variance

5.1 Introduction

In the preceding chapters we have discussed the properties of variance of a function of a random variable and its implications in reliability modeling, with the main focus on the mean residual life (or reversed mean residual life) and the failure rate (or reversed failure rate). Other than the mean, a second characteristic of the residual life that plays a similar role in identifying life distributions and distinguishing them is the variance residual life (VRL) denoted by $v(x)$ and is defined as

$$v(x)=V(X-x\mid X>x)=V(X>x)=E\left(X^2\mid X>x\right)-E^2\left(X\mid X>x\right),$$

first discussed by Launer (1984) while classifying life distributions based on the monotonic behaviour of VRL. Mukherjee and Roy (1986) used relations of the form

$$v(x)=c\left(m_1(x)-x\right) \text{ and } h(x)\left(m_1(x)-x\right)=c,$$

c a real constant, to characterize the exponential, Pearson type XI and finite range distributions. As a discrete analogue of the results of Mukherjee and Roy (1986) Hitha and Nair (1989) showed that the relations

$$v(x)=c\left(m_1(x)-x\right)\left(m_1(x)-x-1\right) \text{ and } h(x)\left(m_1(x)-x\right)=c$$

uniquely determine geometric, negative hyper geometric and Waring distributions for $c=1$, $c<1$ and $c>1$. Gupta et. al. (1987) have studied conditions under which the variance residual life function is monotone and showed that the DMRL (decreasing mean residual life) class is contained in the decreasing variance residual life (DVRL) class.

Gupta and Kirmani (1987) investigated the connection between MRL and VRL for the equilibrium distributions. Gupta (1987) studied the monotonic behavior of VRL in terms of the residual coefficient of variation defined by $s(x)=\sqrt{v(x)}/m(x)$ while the square of the residual coefficient of variation is the interest of Gupta and Kirmani (1998). Bounds on the residual moments and residual variance are obtained by Gupta and Kirmani (1990) and some examples were furnished as illustrations. Gupta and Kirmani (2000) showed that $v(x)$ and $s(x)$ characterizes the life distributions in univariate as well as in bivariate case and proved that the constancy of $v(x)$ characterizes the univariate exponential distribution. They also studied the monotonic behaviour of $s(x)$ in terms of convex (concave) nature of mean residual life function. Gupta and Kirmani (2004) established that the ratio of failure rate and mean residual life characterized the distributions and result is then used to show that the second residual moment characterize the distributions. They also discuss the application of the results to non- homogeneous Poisson process. Defining two classes, decreasing variance residual life (V_D) and increasing variance residual life (V_I), Stoyanov and Al- sadi (2004) discuss some ageing pattern of life distributions and then studied the properties of coherent system based on these classes. Kundu and Gupta (2003) gave two simple characterizations theorem on proportional (reversed) hazard model based on conditional variance. They proved that for any real number t such that $F_X(t)>0,\ \lambda_X(t)=\alpha\lambda_Y(t)$ with $\alpha>0$ if and only if

$$V\left(-\ln F_X(Y)\,|\,Y<t\right)=\frac{1}{\alpha^2},$$

and for any real number t such that $\bar{F}_X(t)>0,\ k_X(t)=\alpha k_Y(t)$ with $\alpha>0$ if and only if

$$V\left(-\ln \bar{F}_X(Y)\,|\,Y>t\right)=\frac{1}{\alpha^2}.$$

Recently, Arishi (2005) characterized exponential family of distribution specified by (3.2.13) by the identity

$$V(X\,|\,X>x)=\frac{Q''S'-Q'S''}{(Q')^3}-\frac{Q''}{(Q')^3}\frac{\partial\log\bar{F}}{\partial\theta}+\frac{1}{(Q')^2}\frac{\partial^2\log\bar{F}}{\partial\theta^2} \qquad (5.1.1)$$

As an illustrative example he proved that the random variable X follows binomial distributions if and only if

$$V(X \mid X > x) = np(1-p) + (x+1)(1-p)(x+1-p(n+1))k(x+1) - (x+1)^2 k^2(x+1). \tag{5.1.2}$$

The Poisson random variable with mean μ is characterized by the identity

$$V(X \mid X > x) = \lambda + (x+1)(x+1-\lambda)k(x+1) - (x+1)^2 k^2(x+1). \tag{5.1.3}$$

A close examination of the properties of VRL and characterizations through its relationship with MRL so far studied in literature leaves scope for generalization of these results by forging an identity connecting VRL and MRL for a class of probability models.

In this chapter we study the properties of the left and right truncated variance of a function of a non- negative random variable, that characterize a class of continuous distributions. These properties include characterizations by relationships the conditional variance has with truncated expectations and/or the failure rate as well as lower bound to the conditional variance. Various results in literature become special cases of our formula and consequently they produce characteristic properties of families of distributions as well as individual models. It is shown that the characteristic properties are linked to those based on relationship between conditional means and failure rates, discussed in the earlier chapters. The lower bound developed here compares favourably with that given by Cramer- Rao inequality so that the bound developed here can be utilized to find the UMVUE in the case of truncated random variable.

5.2 Covariance identity and related characterization

For the sake of continuity, through out this chapter we retain the notations described in Chapter 3. For a real function $h(X)$ of X, from Theorem 3.2.1 we have seen that the density function $f(x)$ satisfies the differential equation (3.2.4) if and only if (3.2.3) is satisfied. Utilizing conditions in (3.2.3) in deriving the necessary relationship between $V(h(X) \mid X > x)$ and $E(h(X) \mid X > x)$ will lead to a characterization of $f(x)$ in terms of (3.2.4). The advantages of using this approach are (i) a link can be established between various characterizations of distributions by relationship between conditional expectations and failure rates so far discussed in literature and the present work (ii) A lower bound to the conditional variances can be obtained in the sense of Cacoullos and Papathanasiou (1997). The study of the lifetime of a system after the elapse of a prescribed time is important in reliability theory and in this case, the conditional variances play the same role as the usual variances in other statistical models. (iii) While the results implied in (i) helps the identification of the appropriate distribution for the data through the characteristic property, (ii) provides construction of minimum variance unbiased estimators of relevant parametric functions through judicious choice of $h(X)$, where X is a statistic.

Next we derive an expression for the covariance between $c(x)$ and $h(x)$ conditioning on the random variable X in terms of $m(x)$ and present it as the following theorem.

Theorem 5.2.1:

For every $c(x)$ in $\mathfrak{B}$ and all $x>0$ satisfying $E\left(g(X)|c'(X)|\right)<\infty$,

$$Cov\left(c(X),h(X)\mid X>x\right)=\sigma E\left(c'(X)g(X)\mid X>x\right)+\left(\mu-m(x)\right)\left(a(x)-c(x)\right) \tag{5.2.1}$$

where

$$a(x)=E\left(c(X)\mid X>x\right),$$

if and only if

$$m(x)=\mu+\sigma k(x)g(x). \tag{5.2.2}$$

Proof:

To prove the if part, we consider

$$\mathrm{E}\left[c(\mathrm{X})\left(h(\mathrm{X})-\mu\right)\mid X>x\right]=\left[\bar{\mathrm{F}}(\mathrm{x})\right]^{-1}\left\{\int_x^\infty\left(c(t)-c(x)\right)\left(h(t)-\mu\right)f(t)\,dt\right.$$

$$\left.+c(x)\int_x^\infty\left(h(t)-\mu\right)f(t)\,dt\right\}$$

$$=\left[\bar{\mathrm{F}}(\mathrm{x})\right]^{-1}\int_x^\infty\left(\int_x^t c'(u)\,du\right)\left(h(t)-\mu\right)f(t)\,dt+c(x)E\left[\left(h(X)-\mu\right)\mid X>x\right] \tag{5.2.3}$$

By changing the order of integration in the first term on the right of (5.2.3) and then on using (3.2.1), it becomes.

$$[\bar{F}(x)]^{-1}\int_x^\infty c'(u)\left(\int_u^\infty (h(t)-\mu)f(t)\,dt\right)du = [\bar{F}(x)]^{-1}\int_x^\infty c'(u)\left(\int_0^u (\mu-h(t))f(t)\,dt\right)du$$

$$= [\bar{F}(x)]^{-1}\int_x^\infty c'(u)\sigma g(u)f(u)\,du$$

$$= \sigma E\left(c'(X)g(X) \mid X>x\right).$$

Thus (5.2.3) reduces to

$$E\left[c(X)(h(X)-\mu) \mid X>x\right] = \sigma E\left[c'(X)g(X) \mid X>x\right] + c(x)(m(x)-\mu) \quad (5.2.4)$$

or

$$E\left[c(X)h(X) \mid X>x\right] = \sigma E\left[c'(X)g(X) \mid X>x\right] + c(x)(m(x)-\mu) + \mu a(x). \quad (5.2.5)$$

By definition

$$Cov\left(c(X),h(X) \mid X>x\right) = E\left(c(X)h(X) \mid X>x\right) - E\left(c(X) \mid X>x\right).E\left(h(X) \mid X>x\right). \quad (5.2.6)$$

Substituting (5.2.5) in (5.2.6) the identity (5.2.1) follows.

Conversely assuming (5.2.1), we have (5.2.4) and hence

$$E\left[(c(X)-c(x))(h(X)-\mu) \mid X>x\right] = \sigma E\left[c'(X)g(X) \mid X>x\right].$$

This is equivalent to

$$[\bar{F}(x)]^{-1}\int_x^\infty c'(u)\left(\int_u^\infty (h(t)-\mu)f(t)\,dt\right)du = \sigma[\bar{F}(x)]^{-1}\int_x^\infty c'(u)g(u)f(u)\,du. \quad (5.2.7)$$

For any (a,b) contained in $(0,\infty)$, choose absolutely continuous $c(x)$ such that

$$c'(x) = \begin{cases} 1, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases}$$

so that (5.2.7) simplifies to

$$\int_a^b\left(\int_u^\infty (h(t)-\mu)f(t)\,dt\right)du = \sigma\int_a^b g(u)f(u)\,du$$

or

$$\sigma g(b)f(b) = \int_b^\infty (h(t)-\mu)f(t)\,dt, \quad \text{for all } b.$$

Since $E(h(X)) = \mu$ the above equation can be written as

$$\sigma g(b) f(b) = \int_a^b (h(t) - \mu) f(t) dt, \quad \text{for all } b$$

which is the same as (3.2.1) then from Theorem 3.2.1 we can arrive at (5.2.2) and this completes the proof of the theorem.

Corollary 5.2.1: When $h(X) \in \mathfrak{B}$,

$$V(h(X) \mid X > x) = \sigma E[h'(X) g(X) \mid X > x] + (\mu - m(x))(m(x) - h(x)) \quad (5.2.8)$$

if and only (5.2.2) holds.

Proof follows by taking $c(X) = h(X)$ in (5.2.1).

Using (5.2.2) we can write

$$(m(x) - h(x)) = (m(x) - \mu) - (h(x) - \mu) = \sigma g(x) k(x) - (h(x) - \mu),$$

hence from Corollary 5.2.1 we can deduce the following result.

Corollary 5.2.2: When $h(X) \in \mathfrak{B}$,

$$V(h(X) \mid X > x) = \sigma E[h'(X) g(X) \mid X > x] + (h(x) - \mu) \sigma g(x) k(x) - \sigma^2 g^2(x) k^2(x) \quad (5.2.9)$$

if and only (5.2.2) holds.

Remark 5.2.1: We see from Corollary 5.2.1 that the conditions (5.2.8) and (5.2.2) are equivalent and the latter characterizes (see Theorem 3.2.1) the distributional form in (3.2.4). Now the equation (3.2.4) can be written as

$$\frac{d \log f(x)}{dx} + \frac{d \log g(x)}{dx} = \frac{(\mu - h(x))}{\sigma g(x)}.$$

Integrating with respect to x from 0 to x and assuming $\lim_{x \to 0} g(x) f(x) = 0$

$$\log f(x) + \log g(x) = \int_0^x \frac{(\mu - h(t))}{\sigma g(t)} dt,$$

which can be written as

$$f(x)=\left[g(x)\right]^{-1}\exp\left[\int_0^x \frac{(\mu-h(t))}{\sigma g(t)}dt\right]. \tag{5.2.10}$$

Hence for a given functional form of $g(x)$ that satisfies (5.2.2), the relationship (5.2.8) uniquely determines $f(x)$ as (5.2.10). Characterizations of probability distributions by means of relationships between conditional means and failure rates therefore extend to characteristic properties by relationships between conditional variances and conditional expectations. A further feature of Theorem 5.2.1 is that many existing characterizations proved separately for families and distributions by different methods can be brought under a single framework. These facts are illustrated in the following remarks.

Remark 5.2.2: When $h(X)=X$, from (5.2.8) we have characterization of distributions through relationships between VRL and the mean residual life $m_0(x)=m(x)-x$.

In the above results $g(x)$ for a particular choice of $h(X)$ is unique and therefore, identifies the distribution.

As pointed out in the introduction next we discuss the implication of the results with existing characterization theorems.

Remark 5.2.3: Dallas (1981) proved that an absolutely continuous random variable distributed as

$$F(x)=1-\exp(-h(x)/k),\ \ a\le X<c,\ k>0 \tag{5.2.11}$$

where $h(x)$ is strictly increasing from $[a,c)$ to $(0,\infty)$, twice differentiable if and only if

$$V(h(X)\,|\,X>x)=k^2, \text{ for all } x.$$

Using the expressions

$$\begin{aligned} k(x)&=\frac{f(x)}{\bar{F}(x)}\\ &=\frac{h'(x)}{k}\exp\left(-\frac{h(x)}{k}\right)/\exp\left(-\frac{h(x)}{k}\right)\\ &=\frac{h'(x)}{k}, \end{aligned}$$

$$m(x)=\frac{1}{\bar{F}(x)}\int_x^{\infty} h(x)f(x)dx$$

$$=\frac{1}{\bar{F}(x)}\int_x^{\infty} h(x)\frac{h'(x)}{k}\exp\left(-\frac{h(x)}{k}\right)dx$$

$$=\frac{-1}{\bar{F}(x)}\left[h(x)\exp\left(-\frac{h(x)}{k}\right)\right]_x^{\infty}+\int_x^{\infty} h'(x)\exp\left(-\frac{h(x)}{k}\right)dx\Bigg]$$

$$=\frac{1}{\bar{F}(x)}\left[h(x)\exp\left(-\frac{h(x)}{k}\right)+k\exp\left(-\frac{h(x)}{k}\right)\right]$$

$$=h(x)+k$$

and

$$\mu=E(h(X))=k,$$

from (3.2.3) one could find the value of $g(x)$ as

$$g(x)=\frac{m(x)-\mu}{\sigma k(x)}$$

$$=\frac{k(h(x)+k-k)}{\sigma h'(x)}$$

$$=\frac{kh(x)}{\sigma h'(x)},$$

so that the substitution of these in (5.2.8) gives

$$V(h(X)\,|\,X>x)=km(x)+(\mu-m(x))(m(x)-h(x)), \quad (5.2.11a)$$

The result by Dallas (1981) now follows from (5.2.11a). Assuming (5.2.11) and substituting the values of $m(x)$, μ and $h(x)$ on the right side of (5.2.11a) gives the value k^2. Conversely if the conditional variance is k^2, from (5.2.11a) we get $\mu=k$ by letting x tend to zero. Further (5.2.11a) simplifies to

$$(m(x)-k)(m(x)-k-h(x))=0,$$

giving solutions $m(x)=k$ or $k+h(x)$. The first solution is inadmissible and second leads to (5.2.11). This completes the proof. The characterization results in Nagaraja

(1975) for the exponential distribution by the constancy of the conditional variance is further special case when $h(X)=X$.

Remark 2.5: The random variable X belongs to the exponential family specified by (3.2.15) if and only if

$$V(P(X)\mid X>x)=[Q'(\theta)]^{-1}\frac{\partial m(x)}{\partial\theta}. \tag{5.2.12}$$

To prove the if part we observe from Theorem 3.2.4 that for the exponential family (3.2.15)

$$m(x)=\mu+[Q'(\theta)]^{-1}\frac{\partial\log\bar{F}(x)}{\partial\theta}$$

when $h(X)=P(X)$, so that

$$\sigma g(x)=[f(x)Q'(\theta)]^{-1}\frac{\partial\bar{F}(x)}{\partial\theta}.$$

Now, assuming the differentiation under the integral sign, consider

$$\begin{aligned}\sigma E[P'(X)g(X)\mid X>x]&=[\bar{F}(x)Q'(\theta)]^{-1}\frac{\partial}{\partial\theta}\int_x^\infty P'(t)\bar{F}(t)dt\\&=[\bar{F}(x)Q'(\theta)]^{-1}\frac{\partial}{\partial\theta}\left[-\bar{F}(x)P(x)+\int_x^\infty P(t)f(t)dt\right]\\&=[\bar{F}(x)Q'(\theta)]^{-1}\frac{\partial}{\partial\theta}[(\bar{F}(x)m(x)-\bar{F}(x)P(x))]\\&=[\bar{F}(x)Q'(\theta)]^{-1}\left[\bar{F}(x)\frac{\partial m(x)}{\partial\theta}+(m(x)-P(x))\frac{\partial\bar{F}(x)}{\partial\theta}\right].\end{aligned} \tag{5.2.13}$$

Now

$$\begin{aligned}[\bar{F}(x)Q'(\theta)]^{-1}\left[(m(x)-P(x))\frac{\partial\bar{F}(x)}{\partial\theta}\right]&=(m(x)-P(x))\sigma[\sigma f(x)Q'(\theta)]^{-1}\frac{\partial\bar{F}(x)}{\partial\theta}\frac{f(x)}{\bar{F}(x)}\\&=(m(x)-P(x))\sigma g(x)k(x)\\&=(m(x)-P(x))(m(x)-\mu),\end{aligned}$$

on using (5.2.2). Substituting in (5.2.13) gives

$$\sigma E[P'(X)g(X)\mid X>x]=[Q'(\theta)]^{-1}\frac{\partial m(x)}{\partial\theta}+(m(x)-\mu)(m(x)-P(x)).$$

Thus from (5.2.8) we have (5.2.12). Conversely, if we assume (5.2.12), from the resulting equation by retracing the above steps we reach at

$$\sigma \int_x^\infty P'(t) g(t) f(t) dt = \left[Q'(\theta).\bar{F}(x)\right]^{-1} \frac{\partial}{\partial\theta} \int_x^\infty P'(t) \bar{F}(t) dt.$$

Differentiating with respect to x we find $\sigma g(x)$ as stated earlier. Using $g(x)$ in (5.2.11) we have

$$f(x) = [\sigma Q'(\theta)] \left(\frac{\partial \bar{F}(x)}{\partial\theta}\right)^{-1} \exp\left[\int_0^x \left(\frac{\partial \bar{F}(t)}{\partial\theta}\right)^{-1} (\mu - P(t)) Q'(\theta) f(t) dt\right]$$

or

$$\log[\sigma Q'(\theta)]^{-1} \frac{\partial \bar{F}(x)}{\partial\theta} = \int_0^x \left(\frac{\partial \bar{F}(t)}{\partial\theta}\right)^{-1} (\mu - P(t)) Q'(\theta) f(t) dt.$$

Differentiating with respect to x

$$\left(\frac{\partial \bar{F}(x)}{\partial\theta}\right)^{-1} \frac{\partial \bar{F}(x)}{\partial\theta} = \left(\frac{\partial \bar{F}(x)}{\partial\theta}\right)^{-1} (\mu - P(x)) Q'(\theta) f(x)$$

or

$$\frac{\partial \log f(x)}{\partial\theta} = (\mu - P(x)) Q'(\theta).$$

Integrating with respect to θ, we have the exponential form and our assertion is established.

Since for the family (3.2.13)

$$m(x) = -\frac{S'(\theta)}{Q'(\theta)} + \frac{1}{Q'(\theta)} \frac{\partial \log \bar{F}(x)}{\partial\theta},$$

differentiation with respect to θ gives

$$\frac{\partial m(x)}{\partial\theta} = \frac{\partial}{\partial\theta}\left[-\frac{S'(\theta)}{Q'(\theta)}\right] + \frac{\partial}{\partial\theta}\left[\frac{1}{Q'(\theta)} \frac{\partial \log \bar{F}(x)}{\partial\theta}\right]$$

$$= \frac{Q''S' - Q'S''}{(Q')^2} - \frac{Q''}{(Q')^2} \frac{\partial \log \bar{F}(x)}{\partial\theta} + \frac{1}{(Q')} \frac{\partial^2 \log \bar{F}(x)}{\partial\theta^2}.$$

Thus (5.2.12) takes the form

$$V\left(P(X)\mid X>x\right)=\frac{Q''S'-Q'S''}{(Q')^3}-\frac{Q''}{(Q')^3}\frac{\partial\log\bar{F}(x)}{\partial\theta}+\frac{1}{(Q')^2}\frac{\partial^2\log\bar{F}(x)}{\partial\theta^2},$$

which is the expression obtained by El- Arishy (2005) when $P(X)=X$, using a completely different approach and was stated in (5.1.1).

Remark 5.2.5: We note that when $h(X)=X^r$, from (5.2.2),

$$E\left(X^r\mid X>x\right)=E\left(X^r\right)+\sigma k(x)g(x)$$

or

$$E\left(X^r\mid X>x\right)=E\left(X^r\right)+k(x)g^*(x)$$

where $g^*(x)=\sigma g(x)$ and by Theorem 3.2.1 the corresponding form for $f(x)$ is obtained as

$$\frac{f'(x)}{f(x)}=\frac{\left(c-g^{*\prime}(x)-x^k\right)}{g^*(x)},$$

which is the result in Theorem 5.4.1 of Navarro et. al. (1998) and was reported in (2.4.15) and (2.4.16). For the gamma distribution with probability density function

$$f(x)=\frac{m^p}{\Gamma p}e^{-mx}x^{p-1},\ x>0$$

$$\log f(x)=\log\frac{m^p}{\Gamma p}+(p-1)\log x-mx,$$

differentiating with respect to x gives

$$\frac{f'(x)}{f(x)}=\frac{(p-1)}{x}-m.$$

Multiplying throughout by x^r and rearranging

$$x^r f'(x)=(p-1)x^{r-1}f(x)-mx^r f(x).$$

Assuming $\lim_{x\to\infty}x^r f(x)=0$ and integrating with respect to x from x to ∞

$$-x^r f(x)-\int_x^\infty rx^{r-1}f(x)=(p-1)\int_x^\infty x^{r-1}f(x)-m\int_x^\infty x^r f(x).$$

Rearranging and multiplying throughout by $\left(m.R(x)\right)^{-1}$ we get

$$E\left(X^r - \frac{r+p-1}{m}X^{r-1} \mid X > x\right) = \frac{x^r}{m}k(x). \tag{5.2.14}$$

Taking

$$h(x) = x^r - \frac{r+p-1}{m}x^{r-1}$$

and comparing with (5.2.2) we find the value of $g(x)$ as

$$\sigma g(x) = \frac{x^r}{m} \text{ and } \mu = 0.$$

Applying the last formula (5.2.14) recursively for r, we can arrive at the identity (2.4.1) so that we find the characterization in Theorem 1 in Adatia et. al. (1991) (see also Koicheva (1993)). Now

$$V(X^r \mid X > x) = E(X^{2r} \mid X > x) - E^2(X^r \mid X > x). \tag{5.2.15}$$

Proceeding in similar way as to obtain (5.2.14), we have

$$E\left(X^{2r} - \frac{2r+p-1}{m}X^{2r-1} \mid X > x\right) = \frac{x^{2r}}{m}k(x)$$

or

$$E(X^{2r} \mid X > x) = E\left(\frac{2r+p-1}{m}X^{2r-1} \mid X > x\right) + \frac{x^{2r}}{m}k(x).$$

Substituting the last identity in (5.2.15)

$$V(X^r \mid X > x) = \frac{2r+p-1}{m}E(X^{2r-1} \mid X > x) - E^2(X^r \mid X > x) + \frac{x^{2r}}{m}k(x),$$

and in particular

$$V(X \mid X > x) = \frac{p+1}{m}m_1(x) - m_1^2(x) + \frac{x^2}{m}k(x), \tag{5.2.16}$$

where

$$m_k(x) = E(X^k \mid X > x),\ k = 1,2,...,$$

characterize the gamma distribution. When $r = 1$ in (5.2.14) we have the Osaki and Li (1988) characterization and the corresponding conditional variance is in (5.2.16). Further for $r = 1$ and $g(x) = \sigma$ in (5.2.16) the result of Kotz and Shanbhag (1980) for the normal distribution is obtained by relaxing the support X to be $(-\infty, \infty)$. Since we are considering only non- negative random variables, the truncated normal law with probability density function

$$f(x)=\frac{2}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\quad x>0 \tag{5.2.17}$$

has

$$\log f(x)=\log\frac{2}{\sqrt{2\pi}\sigma}-\frac{(x-\mu)^2}{2\sigma^2},$$

differentiating with respect to x gives

$$\frac{f'(x)}{f(x)}=-\frac{(x-\mu)}{\sigma^2}$$

or

$$\sigma^2 f'(x)=(x-\mu)f(x).$$

Assuming that the density function vanishes at the end point and integrating with respect to x from x to ∞ we obtain

$$-\sigma^2 f(x)-\int_x^\infty \sigma^2 f(x)\,dx=-(x-\mu)\int_x^\infty f(x)dx.$$

Hence

$$\sigma^2 f(x)+\sigma^2 R(x)=(m(x)-\mu)R(x)$$

or

$$m(x)=(\mu+\sigma^2)+\sigma^2 k(x).$$

Since $E(X)=\mu+\sigma^2$, comparing with (5.2.2) we have

$$g(x)=\sigma,$$

so that (5.2.17) can be characterized by the conditional variance relationship

$$V(X\mid X>x)=\sigma^2+(\mu-m_1(x))(m_1(x)-x).$$

Remark 5.2.6: Ghitany et. al. (1995) have shown that for a real valued function $s(x)\neq 0$ the identity (2.4.12) holds for all $x\geq 0$ if and only if the random variable X belongs to the class of absolutely continuous distributions with density (2.4.11). Taking

$$h(X)=\left(1+\frac{q''(X)p(X)}{(q'(X))^2}-\frac{s'(X)}{q'(X)}\right)$$

we have

$$\sigma g(x)=s(x)/q'(x)$$

and hence the conditional variance characterization is valid according to (5.2.8).

Remark 5.2.7: Our results hold for mixtures of distributions as well. Nassar and Mahmoud (1985) derives

$$m_1(x)=x+\left(\alpha_1^{-1}+\alpha_2^{-1}\right)-\left(\alpha_1\alpha_2\right)^{-1}k(x) \tag{5.2.18}$$

as a necessary and sufficient condition for the distribution of X to be the mixture of exponentials

$$f(x)=\lambda\alpha_1\exp(-\alpha_1 x)+(1-\lambda)\alpha_2\exp(-\alpha_2 x). \tag{5.2.19}$$

The $g(x)$ value worked out from (5.2.18) leads to

$$V(X\mid X>x)=\alpha_1^{-2}+\alpha_2^{-2}-\left(\alpha_1\alpha_2\right)^{-2}h^2(x)$$

as a characteristic property of (5.2.19). The Lomax and beta mixtures of Abraham and Nair (2001) and gamma mixture of Adatia et. al. (1991) also admits similar identities by adopting this procedure.

Remark 5.2.8: For the Pearson family of distributions specified by (2.4.3) from Theorem 3.2.2 we have seen that

$$\sigma g(x)=b_0+b_1x+b_2x^2$$

with

$$b_i=\left(1-2a_2\right)^{-1}a_i,\ i=0,1,2,\ \ a_2\neq\frac{1}{2}.$$

Consider

$$\sigma E\left(h'(X)g(X)\mid X>x\right)+\left(\mu-m(x)\right)\left(m(x)-h(x)\right)$$

$$=E\left(b_0+b_1X+b_2X^2\mid X>x\right)+\mu m_1(x)-\mu x-m_1^2(x)+xm_1(x)$$

$$=\left[b_0+\left(b_1+\mu+x\right)m_1(x)-\mu x\right]+b_2m_2(x)-b_2m_1^2(x)-\left(1-b_2\right)m_1^2(x)$$

$$= \left[b_0 + (b_1 + \mu + x) m_1(x) - \mu x\right] + b_2 V(X \mid X > x) - (1 - b_2) m_1^2(x).$$

Accordingly we have from (5.2.8)

$$V(X \mid X > x) = (1 - b_2)^{-1} \left[b_0 + (b_1 + \mu + x) m_1(x) - \mu x\right] - m_1^2(x). \quad (5.2.20)$$

Conversely, when (5.2.20) holds, we can work backwards to obtain

$$\sigma E(g(X) \mid X > x) = E\left((b_0 + b_1 X + b_2 X^2) \mid X > x\right)$$

or

$$\sigma \int_x^\infty g(t) f(t)\, dt = \int_x^\infty (b_0 + b_1 t + b_2 t^2) f(t)\, dt.$$

Differentiating, we get

$$\sigma g(x) = b_0 + b_1 x + b_2 x^2,$$

and by Corollary 5.2.1

$$m(x) = \mu + (b_0 + b_1 x + b_2 x^2) k(x),$$

so that X belongs to the Pearson family by (2.4.3) and (2.4.4) (see Sankaran and Nair (1991)). A special case is the beta distribution with probability density function

$$f(x) = \frac{1}{B(p,q)} x^{p-1} (1-x)^{q-1}$$

considered in Ahmed (1991), for which (see Table 3.1)

$$\sigma g(x) = (p+q)^{-1} x(1-x).$$

Using (5.2.8)

$$V(X \mid X > x) = (p+q)^{-1} \left[m_1(x) - m_2(x) + (p+q)(\mu - m_1(x))(m_1(x) - x)\right]$$

$$= (p+q)^{-1} \left[m_1(x) - m_1^2(x) + m_1^2(x) - m_2(x) + (p+q)(\mu - m_1(x))(m_1(x) - x)\right]$$

$$= (p+q)^{-1} \left[m_1(x) - m_1^2(x) - V(X \mid X > x)(p+q)(\mu - m_1(x))(m_1(x) - x)\right].$$

Hence

$$V(X \mid X > x) = (p+q+1)^{-1} \left[m_1(x) - m_1^2(x) + (p+q)(\mu - m_1(x))(m_1(x) - x)\right].$$

Further interesting special cases of the last result in reliability modelling are the power distribution $(q = 1)$, finite range distribution $(p = 1)$ and uniform distribution

$(p=1,\ q=1)$. Results for some members of the family have already been covered and hence we have as additional illustration, the inverted gamma

$$f(x)=cx^{-p}\exp(-q/x),\quad x>0$$

with (from table 3.2.1)

$$\sigma g(x)=(p-2)x^2$$

and

$$\begin{aligned}V(X\mid X>x)&=\left[(p-2)m_2(x)-(p-2)m_1^2(x)+(p-2)m_1^2(x)+(\mu-m_1(x))(m_1(x)-x)\right]\\&=\left[(p-2)V(X\mid X>x)+(p-2)m_1^2(x)+(\mu-m_1(x))(m_1(x)-x)\right]\end{aligned}$$

or

$$V(X\mid X>x)=(3-p)^{-1}\left[(p-2)m_1^2(x)+(\mu-m_1(x))(m_1(x)-x)\right].$$

The Pareto model

$$f(x)=\frac{a}{k}\left(\frac{x}{k}\right)^{-(a+1)},\quad x>k$$

has

$$\log f(x)=\log ak^a-(a+1)\log x.$$

Differentiating with respect to x

$$\frac{f'(x)}{f(x)}=\frac{-(a+1)}{x}=\frac{-(a+1)x}{x^2}.$$

Hence the given probability density function belongs to the family specified by (2.4.3) and the value of $g(x)$ is given by

$$\sigma g(x)=\frac{x^2}{(a-1)}.$$

Consider

$$\sigma E(h'(X)g(X)\mid X>x)=\frac{1}{(a-1)}E(X^2\mid X>x)$$

$$=\frac{1}{(a-1)}E(X^2\mid X>x)-\frac{1}{(a-1)}m_1^2(x)+\frac{1}{(a-1)}m_1^2(x)$$

$$= \frac{1}{(a-1)} V(X \mid X > x) + \frac{1}{(a-1)} m_1^2(x).$$

Hence from 5.2.8

$$V(X \mid X > x) = \frac{1}{(a-1)}\left(V(X \mid X > x) + m_1^2(x)\right) + \mu\left(m_1(x) - x\right) + \left(x - m_1(x)\right) m_1(x)$$

$$= (a-2)^{-1}\left[m_1^2(x) + (a-1)\mu\left(m_1(x) - x\right) + \left(x - m_1(x)\right) m_1(x)\right].$$

Remark 5.2.9: The generalization of the Pearson family given by (3.2.12) has (see Sankaran et. al. (2003))

$$\sigma g(x) = \left(c_0 + c_1 x + c_2 x^2\right)$$

where c_0, c_1 and c_2 are the solution of

$$-pa_2 = \mu c_2 b_2, \; -\left(pa_1 + qa_2\right) = \mu\left(c_2 b_1 + c_1 b_2 + 2c_2 a_2\right) \text{ and } \left(\mu - r\right) a_0 = \mu\left(b_0 c_0 + a_0 c_1\right)$$

for the choice $h(x) = px^2 + qx + r$. We have already noticed that the family specified by (3.2.12), in addition to generating all members of (2.4.3), has provision for new members like the inverse Gaussian, Maxwell, random walk and Rayleigh distributions. The inverse Gaussian distribution

$$f(x) = \sqrt{\frac{\beta}{2\pi x^3}} \exp\left[-\frac{\beta(x-\theta)^2}{2\theta^2 x}\right], \quad x > 0$$

verifies

$$\sigma g(x) = 2\theta^2 x^2$$

for a choice of

$$h(x) = \beta x^2 - \theta^2 x - \beta\theta^2$$

and hence

$$V(X \mid X > x) = \theta^2 \beta^{-1}\left(\beta + m_1(x) + 2x^2 k(x)\right) - m_1^2(x).$$

For the Rayleigh distribution

$$f(x) = 2\alpha x \exp\left(-\alpha x^2\right), \quad x > 0$$

we observe that

$$\sigma g(x) = x$$

for choice of

$$h(x) = 2\alpha x^2 - 2$$

and

$$V(X \mid X > x) = (2\alpha)^{-1}(xk(x) + 2) - m_1^2(x).$$

For the Random walk distribution with probability density function

$$f(x) = \sqrt{\frac{\lambda}{2\pi x}} \exp\left[\frac{-\lambda(x\alpha - 1)^2}{2x\alpha^2}\right], \quad x > 0,\ \lambda,\ \alpha > 0,$$

we have

$$\sigma g(x) = 2\alpha^2 x^2$$

for a choice of

$$h(x) = \lambda\alpha^2 x^2 - 3\alpha^2 x - \lambda.$$

And

$$V(X \mid X > x) = \left(3\lambda m_1(x) + 2\lambda x^2 k(x) + \alpha^{-2}\right) - m_1^2(x).$$

Consider the Maxwell distribution with probability density function

$$f(x) = 4\sqrt{\frac{\lambda^3}{\pi}} x^2 e^{-\lambda x^2}, x > 0, \lambda > 0.$$

then

$$\sigma g(x) = x$$

with

$$h(x) = 2\lambda x^2 - 3.$$

Hence

$$V(X \mid X > x) = (2\lambda)^{-1}(xk(x) + 3) - m_1^2(x).$$

These examples throw up sufficient illustration of the utility of our results in establishing characterizations by relationship the conditional variance has with conditional means that include mean residual life as well. Involving failure rates in these equations is easily accomplished by means of (5.2.2) or by Corollary 5.2.2 in all cases, though this is not explicitly stated.

5.3 Lower bound to the conditional variance

The covariance identity in Section 2 provides scope for introducing bounds on variance using the traditional inequality between the covariance and the variances of the functions involved. We follow the methodology used to characterize distributions by such bounds in Cacoullos and Papathanasiou (1997). When the function $c(X)$ appearing in the bounds are chosen as estimators of the desired parametric functions, the sampling distribution of X will determine $g(x)$ and the bounds have the potential to identify through a different approach, the minimum variance unbiased estimators obtained by the well known classical theorems.

Theorem 3.1:

Under the condition of the Theorem 5.2.1, for every $c(x)$ in $\mathfrak{B}$,

$$V(c(X)|X>x) \geq [V(h(X)|X>x)]^{-1}[\sigma E[c'(X)g(X)|X>x] + (\mu - m(x))(a(x)-c(x))]^2 \quad (5.3.1)$$

if (5.2.2) holds, with equality whenever $c(X)$ is a linear function of $h(X)$. Conversely, if the conditional variance is (5.3.1), then (5.2.2) is true provided $V(h(X)|X>x)$ is as in (5.2.8).

Proof:

Since

$$Cov^2(c(X),h(X)|X>x) \leq V(c(X)|X>x)V(h(X)|X>x),$$

using (5.2.1) we have (5.3.1). To prove the second part of the Theorem assume

$$c(x) = h(x) + \theta l(x)$$

for some arbitrary real θ. Then

$$a(x) = m(x) + \theta b(x) \text{ and } b(x) = E(l(X)|X>x)$$

so that (5.3.1) becomes

$$V(h(X)|X>x)[V(h(X)|X>x) + \theta^2 V(l(X)|X>x) + 2\theta Cov(l(X),h(X)|X>x)]$$
$$\geq [\sigma E[h'(X)g(X)|X>x] + \sigma\theta E[l'(X)g(X)|X>x] + \{m(x)+\theta b(x) - h(x) - \theta l(x)\}(\mu - m(x))]^2.$$

$$=\left[\sigma\theta E\left[l'(X)g(X)\mid X>x\right]+\theta\left(b(x)-l(x)\right)\left(\mu-m(x)\right)+V\left(h(X)\mid X>x\right)\right]^2,$$
(5.3.2)

using (5.2.8) and the inequality (5.3.2) simplifies to

$$\theta^2\left[V\left(h(X)\mid X>x\right)V\left(l(X)\mid X>x\right)-\left\{\sigma E\left[l'(X)g(X)\mid X>x\right]\right.\right.$$

$$\left.\left.+\left(\mu-m(x)\right)\left(b(x)-l(x)\right)\right\}^2\right]+2\theta V\left(h(X)\mid X>x\right)\left[Cov\left(l(X),h(X)\mid X>x\right)\right.$$

$$\left.-\left\{\sigma E\left[l'(X)g(X)\mid X>x\right]+\left(\mu-m(x)\right)\left(b(x)-l(x)\right)\right\}\right]\geq 0. \qquad (5.3.3)$$

For (5.3.3) to be true for all θ, the coefficient of θ must vanish and therefore,

$$Cov\left(l(X),h(X)\mid X>x\right)=\sigma E\left[l'(X)g(X)\mid X>x\right]+\left(\mu-m(x)\right)\left(b(x)-l(x)\right).$$

Using the expression for the covariance the above identity reduces to

$$E\left(l(X)\left(h(X)-\mu\right)\mid X>x\right)=\sigma E\left(l'(X)g(X)\mid X>x\right)+l(x)E\left(\left(h(X)-\mu\right)\mid X>x\right)$$

or

$$\int_x^\infty \left(l(t)-l(x)\right)\left(h(t)-\mu\right)f(t)dt=\sigma\int_x^\infty l'(t)g(t)f(t)dt$$

$$\int_x^\infty \left(\int_x^t l'(u)du\right)\left(h(t)-\mu\right)f(t)dt=\sigma\int_x^\infty l'(t)g(t)f(t)dt.$$

The rest of the proof is as in Theorem 5.2.1 to arrive at (5.2.2).

From the point of estimation, it is worth examining how our inequality compares with the lower bound of the variance in the Cramer- Rao theorem. Let X belong to the regular exponential family of distributions specified by (3.2.13). Then the random variable $(X\mid X>t)$ has density

$$f_1(x;\theta)=\exp\left[P(x)Q(\theta)+T(x)+S(\theta)-\log\bar{F}(t)\right], \quad x>t$$

which is also of the exponential form so that the Cramer- Rao lower bound is attained for the unbiased estimation of

$$E\left(P(X)\mid X>t\right)=m(t).$$

This bound is

$$V(T)=\frac{\left[m'(t)\right]^2}{E\left(\frac{\partial \log f_1(t)}{\partial \theta}\right)^2}.$$

From the discussion on Section 3.3.1 we have

$$\frac{\partial \log f_1}{\partial \theta}=Q'(\theta)\left(P(X)-\mu\right),$$

hence

$$V(T)=\frac{\left[m'(t)\right]^2}{\left[Q'(\theta)\right]^2 V\left(p(X)\mid X>t\right)}=\frac{m'(t)}{Q'(\theta)},$$

which is also the bound obtained in (5.3.1). Thus in regular cases, the two inequalities produce the same result.

To see the position in non- regular case, note that Theorem 5.3.1 does not require the regularity conditions of the Cramer – Rao Theorem and accordingly it is applicable in non- regular cases. Note that in the uniform distribution over $(0,\theta)$, the random variable $(X\mid X>x)$ has the density $f^*(x)$ given by

$$f^*(x)=\frac{f(x)}{R(x)}=\frac{1}{\theta}\frac{\theta}{(\theta-x)}=\frac{1}{(\theta-x)}.$$

Taking $\left(h(X)\mid X>x\right)=2X-x$,

$$E(2X-x)=\frac{1}{(\theta-x)}\int_x^\theta (2t-x)\,dt$$

$$=\frac{1}{(\theta-x)}\left[t^2-xt\right]_x^\theta$$

$$=\frac{1}{(\theta-x)}\left(\theta^2-\theta x\right)=\theta$$

and $\left(h(X)\mid X>x\right)=2X-x$ is unbiased for θ. The Cramer- Rao inequality provides the lower bound

$$V\left(h(X)\mid X>x\right)=\frac{(\theta')^2}{E\left(\frac{\partial \log f^*(x)}{\partial \theta}\right)^2}. \qquad (5.3.4)$$

Now

$$\log f^{*}(x) = -\log(\theta - x),$$

differentiating with respect to θ gives

$$\frac{\partial \log f^{*}(x)}{\partial \theta} = -\frac{1}{(\theta - x)}.$$

Substituting in (5.3.4) we obtain the lower bound provided by Cramer- Rao inequality as

$$V(h(X) \mid X > x) = (\theta - x)^2.$$

Again

$$V(h(X) \mid X > x) = E(4X^2) - E^2(2X), \tag{5.3.5}$$

and

$$E^2(h(X) \mid X > x) = (\theta + x)^2$$

$$E(h^2(X) \mid X > x) = \frac{1}{(\theta - x)} \int_x^\theta 4x^2 dx$$

$$= \frac{1}{(\theta - x)} \left[\frac{4}{3} x^3 \right]_x^\theta$$

$$= \frac{1}{(\theta - x)} \left[\frac{4}{3}\theta^3 - \frac{4}{3} x^3 \right]$$

$$= \frac{1}{(\theta - x)} \left[\frac{4}{3}(\theta - x)(\theta^2 + x^2 + \theta x) \right]$$

$$= \frac{4}{3}(\theta^2 + x^2 + \theta x).$$

Hence from (5.3.5)

$$V(h(X) \mid X > x) = \frac{4}{3}(\theta^2 + x^2 + \theta x) - (\theta + x)^2$$

$$= \frac{1}{3}\left[4(\theta^2 + x^2 + \theta x) - 3(\theta + x)^2 \right]$$

$$= \frac{(\theta - x)^2}{3}. \tag{5.3.6}$$

For the random variable $c(X)=2X$, now we consider the bound obtained by (5.3.1). From the identity (5.2.4) we have

$$\sigma E\left(c'(X)g(X)\mid X>x\right)+\left(\mu-m(x)\right)c(x)=\mathrm{E}\left[\mathrm{c}(\mathrm{X})\left(h(\mathrm{X})-\mu\right)\mid X>x\right]$$

$$=E\left[2X\left(2X-x-\theta\right)\right]$$

$$=\frac{1}{(\theta-x)}\int_x^\theta\left(4t^2-2tx-2t\theta\right)dt$$

$$=\frac{1}{(\theta-x)}\left(\frac{4}{3}t^3-t^2x-t^2\theta\right)\Bigg]_x^\theta$$

$$=\frac{1}{(\theta-x)}\left(\frac{4}{3}\theta^3-\theta^2x-\theta^3-\frac{4}{3}x^3+x^3+\theta x^2\right)$$

$$=\frac{1}{(\theta-x)}\left(\frac{1}{3}\left(\theta^3-x^3\right)-\theta x(\theta-x)\right)$$

$$=\frac{1}{3(\theta-x)}\left((\theta-x)\left(\theta^2+x^2+\theta x\right)-3\theta x(\theta-x)\right)$$

$$=\frac{(\theta-x)^2}{3}.$$

Hence the random variable $c(X)=2X$ attain the bound in (5.3.1) which is in fact the actual variance of $h(X)$ given in (5.3.6), so that (5.3.1) improves upon the Cramer- Rao bound.

Now we compare our bound with variance provided by Chapman- Robbins inequality. Let $E_\theta\left(h(X)\mid X>x\right)=m_\theta(x)$, where $\theta,\ \varphi\in\Theta\subset R$ such that $f_\theta(x)$ and $f_\varphi(x)$ are different, satisfying $\{f_\theta(x)>0\}\supset\{f_\varphi(x)>0\}$, choosing $c(x)=\left(\frac{f_\varphi(x)}{f_\theta(x)}-1\right)$

Consider

$$\sigma E\left[c'(X)g(X)\mid X>x\right]=\frac{1}{\bar{F}(x)}\int_x^\infty c'(u)g(u)f_\theta(u)du$$

$$=\frac{1}{\bar{F}(x)}\int_x^\infty c'(u)\left(\int_0^x\left(\mu-h(t)\right)f_\theta(t)dt\right)du$$

$$=\frac{1}{\bar{F}(x)}\int_x^{\infty}c'(u)\left(\int_u^{\infty}(h(t)-\mu)f_{\theta}(t)dt\right)du.$$

Interchanging the order of integration by using Fubini's Theorem

$$\sigma E\left[c'(X)g(X)\mid X>x\right]=\frac{1}{\bar{F}(x)}\int_x^{\infty}(h(t)-\mu)\left(\int_x^{t}c'(u)du\right)f_{\theta}(t)dt$$

$$=\frac{1}{\bar{F}(x)}\int_x^{\infty}c(t)(h(t)-\mu)f_{\theta}(t)dt-c(x)(m(x)-\mu)$$

$$=\frac{1}{\bar{F}(x)}\int_x^{\infty}\left(\frac{f_{\varphi}(x)}{f_{\theta}(x)}-1\right)(h(t)-\mu)f_{\theta}(t)dt-c(x)(m(x)-\mu)$$

$$=\frac{1}{\bar{F}(x)}\int_x^{\infty}(f_{\varphi}(x)-f_{\theta}(x))(h(t)-\mu)dt-c(x)(m(x)-\mu)$$

$$=\left[m_{\varphi}(x)-m_{\theta}(x)\right]-c(x)(m(x)-\mu).$$

Now the substitution in (5.3.1) gives

$$V(c(X)\mid X>x)\geq\left[m_{\varphi}(x)-m_{\theta}(x)\right]^2/V(h(X)\mid X>x)$$

or

$$V(h(X)\mid X>x)\geq\frac{\left[m_{\varphi}(x)-m_{\theta}(x)\right]^2}{V(p_{\varphi}(X)/p_{\theta}(X)\mid X>x)},$$

which is the Chapman-Robbins inequality.

5.4 Reversed variance residual life

Like the reversed mean residual life, we can define the variance of reversed residual life (RVRL) by

$$V(X\mid X\leq x)=E\left[(x-X)^2\mid X\leq x\right]-r_0^2(x).$$

. As before, defining

$$r(x)=E(h(X)\mid X\leq x),$$

covariance identity and variance bounds can be derived for $V(h(X)\mid X\leq x)$ in terms of $r(x)$ and the reversed failure rate $\lambda(x)$ and we present it as following theorem.

Theorem 5.4.1:

For every $c(x) \in \mathfrak{B}$ and all $t > 0$ and $h(x)$ satisfying $E\left(h^2(X)\right) < \infty$,

$$Cov\left(c(X), h(X) \mid X \le x\right) = \sigma E\left(c'(X) g(X) \mid X \le x\right) + \left(\mu - r(x)\right)\left(b(x) - c(x)\right) \quad (5.4.4)$$

where

$$b(x) = E\left(c(X) \mid X \le x\right),$$

if and only if

$$\sigma g(x) \lambda(x) = \mu - r(x), \quad x > 0. \quad (5.4.5)$$

Proof:

To prove the if part, we consider

$$\mathrm{E}\left[c(X)\left(h(X) - \mu\right) \mid X \le x\right] = \left[F(x)\right]^{-1}\left\{\int_0^x \left(c(t) - c(x)\right)\left(h(t) - \mu\right) f(t)\, dt\right.$$

$$\left. + c(x) \int_0^x \left(h(t) - \mu\right) f(t)\, dt\right\}$$

$$= \left[F(x)\right]^{-1} \int_0^x \left(\int_x^t c'(u)\, du\right)\left(h(t) - \mu\right) f(t)\, dt + c(x) E\left[\left(h(X) - \mu\right) \mid X \le x\right]$$

$$= \left[F(x)\right]^{-1} \int_0^x \left(\int_t^x c'(u)\, du\right)\left(\mu - h(t)\right) f(t)\, dt + c(x) E\left[\left(h(X) - \mu\right) \mid X \le x\right]$$

$$= \left[F(x)\right]^{-1} \int_0^x c'(u) \left(\int_0^u \left(\mu - h(t)\right) f(t)\, dt\right) du + c(x) E\left[\left(h(X) - \mu\right) \mid X \le x\right]$$

$$= \left[F(x)\right]^{-1} \int_0^x c'(u) \sigma g(u) f(u)\, du + c(x) E\left[\left(h(X) - \mu\right) \mid X \le x\right],$$

on using (3.2.1).Hence

$$\mathrm{E}\left[c(X)\left(h(X) - \mu\right) \mid X \le x\right] = = \sigma E\left(c'(X) g(X) \mid X \le x\right) + c(x)\left(r(x) - \mu\right) \quad (5.4.6)$$

or

$$\mathrm{E}\left[c(X) h(X) \mid X \le x\right] = \sigma E\left[c'(X) g(X) \mid X \le x\right] + c(x)\left(r(x) - \mu\right) + \mu b(x). \quad (5.4.7)$$

Now by definition

$$Cov\left(c(X), h(X) \mid X \le x\right) = E\left(c(X) h(X) \mid X \le x\right) - E\left(c(X) \mid X \le x\right).E\left(h(X) \mid X \le x\right). \quad (5.4.8)$$

Substituting (5.4.7) in (5.4.8) the identity (5.4.4) follows.

Conversely assuming (5.4.4), we have (5.4.6) and hence

$$\mathrm{E}\left[\left(c(X)-c(x)\right)\left(h(X)-\mu\right)\mid X\le x\right]=\sigma E\left[c'(X)g(X)\mid X\le x\right].$$

This is equivalent to

$$\left[F(x)\right]^{-1}\int_0^x\left(\int_x^t c'(u)\,du\right)\left(h(t)-\mu\right)f(t)\,dt=\sigma\left[F(x)\right]^{-1}\int_0^x c'(u)g(u)f(u)\,du$$

or

$$\left[F(x)\right]^{-1}\int_0^x\left(\int_t^x c'(u)\,du\right)\left(\mu-h(t)\right)f(t)\,dt=\sigma\left[F(x)\right]^{-1}\int_0^x c'(u)g(u)f(u)\,du$$

and changing the order of integration on the right hand side of the last equation we have

$$\left[F(x)\right]^{-1}\int_0^x c'(u)\left(\int_0^u\left(\mu-h(t)\right)f(t)\,dt\right)du=\sigma\left[F(x)\right]^{-1}\int_0^x c'(u)g(u)f(u)\,du. \quad (5.4.9)$$

For any (a,b) contained in $(0,\infty)$, choose absolutely continuous $c(x)$ such that

$$c'(x)=\begin{cases}1, & x\in(a,b)\\ 0, & x\notin(a,b)\end{cases}$$

so that (5.4.9) simplifies to

$$\int_a^b\left(\int_0^u\left(\mu-h(t)\right)f(t)\,dt\right)du=\sigma\int_a^b g(u)f(u)\,du$$

or

$$\sigma g(b)f(b)=\int_0^b\left(\mu-h(t)\right)f(t)\,dt,\quad \text{for all } b.$$

which is the same as (3.2.1) then from Theorem 3.2.1 we can arrive at (5.4.5) and this completes the proof of the theorem.

Corollary 5.4.1: When $h(X)\in\mathfrak{B}$,

$$V\left(h(X)\mid X\le x\right)=\sigma E\left[h'(X)g(X)\mid X\le x\right]+\left(\mu-r(x)\right)\left(r(x)-h(x)\right) \quad (5.4.10)$$

if and only if (5.4.5) holds.

Proof follows by taking $c(X)=h(X)$ in (5.4.4).

Using (5.4.5) we can write

$$\left(r(x)-h(x)\right)=\left(r(x)-\mu\right)-\left(h(x)-\mu\right)=-\sigma g(x)\lambda(x)-\left(h(x)-\mu\right),$$

hence from Corollary 5.4.1 we can deduce the following result.

Corollary 5.4.2: When $h(X) \in \mathfrak{B}$,

$$V(h(X) \mid X \le x) = \sigma E\left[h'(X) g(X) \mid X \le x\right] - \left(h(x) - \mu\right)\sigma g(x)\lambda(x) - \sigma^2 g^2(x)\lambda^2(x)$$

if and only if (5.4.5) holds. (5.4.11)

Remark 5.4.1: RVRL characterizing $f(x)$ is a special case of (5.2.10) or (5.2.11) when $h(X) = X$.

Remark 5.4.2: Note that the $g(x)$ function appearing in Theorem 5.2.1 and Theorem 5.4.1 are the same and further the relationship (5.4.5) is equivalent to (5.2.2). Thus the distribution characterized by (5.2.1) and (5.4.4) have the same form (3.2.4).

Theorem 5.4.2:

Under the condition on $c(X)$ in Theorem 5.3.1, if (5.4.2) is satisfied, then

$$V(c(X) \mid X \le x) \ge \left[V(h(X) \mid X \le x)\right]^{-1}$$

$$\left[\sigma E\left(h'(X) g(X) \mid X \le x\right) + \left(r(x) - \mu\right)\left(b(x) - c(x)\right)\right]^2 \quad (5.4.12)$$

with equality whenever $c(x)$ is linear function of $h(x)$. Conversely if $V(h(X) \mid X \le x)$ is as in (5.4.10) and (5.4.12) holds then the equation (5.4.5) is satisfied for all $x > 0$.

Proof:

Since

$$Cov^2\left(c(X), h(X) \mid X \le x\right) \le V\left(c(X) \mid X \le x\right) V\left(h(X) \mid X \le x\right),$$

using (5.4.4) we have (5.4.12). To prove the second part of the Theorem assume

$$c(x) = h(x) + \theta l(x)$$

for some arbitrary real θ. Then

$$b(x) = r(x) + \theta e(x) \text{ and } e(x) = E\left(l(X) \mid X \le x\right)$$

so that (5.4.12) becomes

$$V(h(X) \mid X \le x)\left[V(h(X) \mid X \le x) + \theta^2 V(l(X) \mid X \le x) + 2\theta Cov\left(l(X), h(X) \mid X \le x\right)\right]$$

$$\ge \left[\sigma E\left[h'(X) g(X) \mid X \le x\right] + \sigma\theta E\left[l'(X) g(X) \mid X \le x\right]\right.$$

$$\left. + \left\{r(x) + \theta e(x) - h(x) - \theta l(x)\right\}\left(\mu - r(x)\right)\right]^2. \quad (5.4.13)$$

Using (5.4.10) inequality (5.4.13) simplifies to

$$\theta^2\Big[V\big(h(X)\,|\,X\le x\big)V\big(l(X)\,|\,X\le x\big)-\big\{\sigma E\big[l'(X)g(X)\,|\,X\le x\big]$$

$$+\big(\mu-r(x)\big)\big(b(x)-l(x)\big)\big\}^2+2\theta V\big(h(X)\,|\,X\le x\big)\Big[Cov\big(l(X),h(X)\,|\,X\le x\big)$$

$$-\big\{\sigma E\big[l'(X)g(X)\,|\,X\le x\big]+\big(\mu-r(x)\big)\big(e(x)-l(x)\big)\big\}\Big]\ge 0. \qquad (5.4.14)$$

For (5.4.14) to be true for all θ, the coefficient of θ must vanish and therefore,

$$Cov\big(l(X),h(X)\,|\,X\le x\big)=\sigma E\big[l'(X)g(X)\,|\,X\le x\big]+\big(\mu-r(x)\big)\big(e(x)-l(x)\big).$$

On using (5.4.4) the above identity reduces to

$$E\big(l(X)\big(h(X)-\mu\big)\,|\,X\le x\big)=\sigma E\big(l'(X)g(X)\,|\,X\le x\big)+l(x)E\big(\big(h(X)-\mu\big)\,|\,X\le x\big).$$

This is equivalent to

$$\int_0^x\big(l(t)-l(x)\big)\big(h(t)-\mu\big)f(t)\,dt=\sigma\int_0^x l'(t)g(t)f(t)\,dt$$

or

$$\int_0^x\Big(\int_x^t l'(u)\,du\Big)\big(h(t)-\mu\big)f(t)\,dt=\sigma\int_0^x l'(t)g(t)f(t)\,dt$$

or

$$\int_0^x\Big(\int_t^x l'(u)\,du\Big)\big(\mu-h(t)\big)f(t)\,dt=\sigma\int_0^x l'(t)g(t)f(t)\,dt.$$

The rest of the proof is as in Theorem 5.4.1 to arrive at (5.4.5).

Chapter 6

Characterization of Discrete Life Distributions by Properties of Conditional Variance

6.1 Introduction

In continuation of the work done in the previous chapter here we discuss the properties of conditional variance for a non- negative integer valued random variable. First we establish a relation between VRL and conditional expectations or failure rate to characterize a general class of discrete life distributions and the results is then applied to families of distributions. It is shown that the recent work by Arishi (2005) on binomial and Poisson random variables are special cases of our findings. Also we discuss the link between the characterizations of distributions by relationship between conditional expectations and failure rates so far discussed in literature and the present work. A lower bound to the conditional variances is obtained in the sense of Cacoullos and Papathanasiou (1997) and compared with the minimum variance unbiased estimators obtained by the well known classical theorems on inference, when the function $c(X)$ appearing in the bounds are chosen as estimators of the desired parametric functions.

6.2 Main results

Let $h(X)$ be real valued non- constant function of a random variable X defined on N such that $E\left(h^2(X)\right)<\infty$. To obtain the variance expression for $h(X)$ first we establish a covariance identity for a real valued function $c(X)$ and $h(X)$ in terms of $m(x)$ and $g(x)$, where $g(x)$ is a positive real function on N satisfying the condition

(4.2.6). From Theorem 4.2.1 we observed that the condition (4.2.6) is equivalent to (4.2.3) and the corresponding probability mass function satisfies the difference equation (4.2.1). Taking these in to account next we present a general theorem as a discrete analogue of Theorem 5.2.1, using the notations in chapter IV.

Theorem 6.2.1:

Let X be a discrete random variable supported on N or a subset thereof and $g(.)$, $c(.)$, $h(.)$ be functions in $\mathfrak{C}$ such that $E\left(c^2(X)\right)<\infty$, $E\left(g(X)\Delta c(X)\right)<\infty$, $E\left(\left|c(X)h(X)\right|\right)<\infty$ and $E\left(h^2(X)\right)<\infty$. Then the probability mass function satisfies the equation (4.2.1) with $g(0)$ and $p(0)$ as stated in the Theorem 4.2.1, if and only if for every $c(x)\in\mathfrak{C}$ the covariance expression satisfies

$$Cov\left(c(X),h(X)\mid X>x\right)=\sigma E\left(\Delta c(X).g(X)\mid X>x\right)+\left(\mu-m(x)\right)\left(a(x)-c(x+1)\right), \tag{6.2.1}$$

where $a(x)=E\left(c(X)\mid X>x\right)$.

Proof:

To prove the if part, we consider

$$\mathrm{E}\left[\mathrm{c}(\mathrm{X})\left(h(\mathrm{X})-\mu\right)\mid X>x\right]=\left[R(x+1)\right]^{-1}\sum_{k=x+1}^{\infty}c(k)\left(h(k)-\mu\right)p(k)$$

$$=\left[R(x+1)\right]^{-1}\Bigg[c(x+1)\left[\sum_{y=x+1}^{\infty}\left(h(k)-\mu\right)p(k)-\sum_{y=x+2}^{\infty}\left(h(k)-\mu\right)p(k)\right]$$

$$+c(x+2)\left[\sum_{y=x+2}^{\infty}\left(h(k)-\mu\right)p(k)-\sum_{y=x+3}^{\infty}\left(h(k)-\mu\right)p(k)\right]+\ldots\Bigg]$$

$$=\left[R(x+1)\right]^{-1}\Bigg[\left(c(x+2)-c(x+1)\right)\left[\sum_{y=x+2}^{\infty}\left(h(k)-\mu\right)p(k)\right]$$

$$+\left(c(x+3)-c(x+2)\right)\left[\sum_{y=x+2}^{\infty}\left(h(k)-\mu\right)p(k)\right]+\ldots\Bigg]$$

$$+c(x+1)\left(m(x)-\mu\right)$$

$$=\left[R(x+1)\right]^{-1}\left[\sum_{y=x+1}^{\infty}\Delta c(y)\sum_{k=y+1}^{\infty}(h(k)-\mu)p(k)\right]+c(x+1)(m(x)-\mu)$$

$$=\left[R(x+1)\right]^{-1}\left[\sum_{y=x+1}^{\infty}\Delta c(y)\sum_{k=0}^{y}(\mu-h(k))p(k)\right]+c(x+1)(m(x)-\mu),\quad (6.2.2)$$

since

$$E(h(X)-\mu)=\sum_{k=0}^{\infty}(h(k)-\mu)p(k)=0.$$

Now, assuming (4.2.1) from Theorem 4.2.1 we have (4.2.6) and substituting in (6.2.2) we obtain

$$\mathrm{E}\left[c(X)(h(X)-\mu)\mid X>x\right]=\left[R(x+1)\right]^{-1}\left[\sigma\sum_{y=x+1}^{\infty}\Delta c(y).g(y)p(y)\right]+c(x+1)(m(x)-\mu)$$

$$=\sigma E\left[\Delta c(X).g(X)\mid X>x\right]+c(x+1)(m(x)-\mu)\quad (6.2.3)$$

or

$$\mathrm{E}\left[c(X)h(X)\mid X>x\right]=\sigma E\left[\Delta c(X).g(X)\mid X>x\right]+c(x+1)(m(x)-\mu)+\mu a(x).$$

By definition

$$Cov(c(X),h(X)\mid X>x)=E(c(X)h(X)\mid X>x)-E(c(X)\mid X>x).E(h(X)\mid X>x)$$

$$=\sigma E\left[\Delta c(X).g(X)\mid X>x\right]+c(x+1)(m(x)-\mu)+\mu a(x)-m(x)a(x)$$

$$=\sigma E\left[\Delta c(X).g(X)\mid X>x\right]+(\mu-m(x))(a(x)-c(x+1)),$$

which is the expression given in (6.2.1).

Conversely assuming (6.2.1) and retracing the above steps we can arrive at (6.2.3). Now the left hand side of (6.2.2) and (6.2.3) are same so the right hand sides are equal, hence we have

$$\left[R(x+1)\right]^{-1}\left[\sum_{y=x+1}^{\infty}\Delta c(y)\sum_{k=0}^{y}(\mu-h(k))p(k)\right]=\left[R(x+1)\right]^{-1}\left[\sigma\sum_{y=x+1}^{\infty}\Delta c(y).g(y)p(y)\right].$$

Let $c(x)=\theta s^{x}$ for some real θ and $0\le s\le 1$, then the above equation can be written as

$$\sum_{y=x+1}^{\infty}s^{y}\sum_{k=0}^{y}(\mu-h(k))p(k)=\sigma\sum_{y=x+1}^{\infty}s^{y}g(y)p(y)$$

or

$$\sum_{y=x+1}^{\infty} s^y \frac{1}{p(y)}\left(\sum_{k=0}^{y}(\mu-h(k))\,p(k)\right)p(y)=\sigma\sum_{y=x+1}^{\infty} s^y g(y)p(y). \tag{6.2.4}$$

Using the uniqueness of probability generating function, from (6.2.4) we can write

$$\frac{1}{p(y)}\sum_{k=0}^{y}(\mu-h(k))\,p(k)=\sigma g(y), \text{ for every } y$$

or

$$\sigma g(x)\,p(x)=\sum_{k=0}^{x}(\mu-h(k))\,p(k).$$

Multiplying by $\left[F(x)\right]^{-1}$ and rearranging we get

$$r(x)=\mu-\sigma g(x)\lambda(x),$$

this is the same as (4.2.2), hence rest of the proof follows from Theorem 4.2.1.

Corollary 6.2.1: When $h(X)\in\mathfrak{C}$,

$$V(h(X)\mid X>x)=\sigma E\left[\Delta h(X).g(X)\mid X>x\right]+(\mu-m(x))(m(x)-h(x+1)) \tag{6.2.5}$$

if and only if (4.2.1) holds.

The proof follows from (6.2.1) by taking $c(X)=h(X)$.

Corollary 6.2.2: When $h(X)\in\mathfrak{C}$,

$$V(h(X)\mid X>x)=\sigma E\left[\Delta h(X).g(X)\mid X>x\right]+(h(x+1)-\mu)\sigma g(x)k(x)(1-k(x))^{-1}$$

$$-\sigma^2 g^2(x)k^2(x)(1-k(x))^{-2} \tag{6.2.6}$$

if and only if (4.2.1) holds.

Proof:

Using (4.2.3) we can write

$$(\mu-m(x))(m(x)-h(x+1))=(h(x+1)-\mu)\sigma g(x)\frac{k(x)}{(1-k(x))}-\sigma^2 g^2(x)\frac{k^2(x)}{(1-k(x))^2},$$

hence from Corollary 6.2.1 the result follows.

Remark 6.2.1: VRL characterizing $p(x)$ is a special case of (6.2.5) or (6.2.6) when $h(X)=X$.

Next we look at the characterization results in terms of (6.2.5) or (6.2.6) for the modified power series family, the Ord family and the Katz family of distributions which together cover most of the discrete probability distributions used in reliability modelling. And then give some examples of individual distributions to check the validity of the results.

Theorem 6.2.2:

The distribution of X belongs to MPSD with probability mass function (4.2.8) then the conditional variance satisfies

$$V(X \mid X>x)=\frac{u(\theta)}{u'(\theta)}\frac{\partial m(x)}{\partial \theta}. \tag{6.2.7}$$

Proof:

From the proof of Theorem 4.2.1 we observed that for the family specified by (4.2.8) the value of $g(x)$ is given in (4.2.12). Now consider

$$\sigma E\left[\Delta h(X) g(X) \mid X>x\right]=E\left[\Delta h(X).\frac{-u(\theta)}{u^{'}(\theta)p(X)}\frac{\partial F(X)}{\partial \theta} \mid X>x\right],$$

$$=-\frac{u(\theta)}{u^{'}(\theta)R(x+1)}\sum_{k=x+1}^{\infty}\Delta h(k).\frac{\partial F(k)}{\partial \theta},$$

$$=-\frac{u(\theta)}{u^{'}(\theta)R(x+1)}\frac{\partial}{\partial \theta}\left[\sum_{k=x+1}^{\infty}\Delta h(k).F(k)\right]. \tag{6.2.8}$$

Again

$$-\sum_{k=x+1}^{\infty}\Delta h(k)F(k)=\sum_{k=x+1}^{\infty}h(k)F(k)-\sum_{k=x+1}^{\infty}h(k+1)F(k),$$

$$=h(x+1)F(x+1)+\sum_{k=x+2}^{\infty}h(k)p(k),$$

$$= h(x+1)F(x) + \sum_{k=x+1}^{\infty} h(k)p(k),$$

$$= m(x)R(x+1) + h(x+1)F(x).$$

Hence

$$-\frac{\partial}{\partial\theta}\left[\sum_{k=x+1}^{\infty} \Delta h(k).F(k)\right] = \frac{\partial}{\partial\theta}\left[m(x)R(x+1) + h(x+1)F(x)\right],$$

$$= R(x+1)\frac{\partial m(x)}{\partial\theta} + m(x)\frac{\partial R(x+1)}{\partial\theta} + h(x+1)\frac{\partial F(x)}{\partial\theta},$$

$$= R(x+1)\frac{\partial m(x)}{\partial\theta} - m(x)\frac{\partial F(x)}{\partial\theta} + h(x+1)\frac{\partial F(x)}{\partial\theta},$$

$$= R(x+1)\frac{\partial m(x)}{\partial\theta} + \big(h(x+1) - m(x)\big)\frac{\partial F(x)}{\partial\theta}. \qquad (6.2.9)$$

Substituting (6.2.9) in (6.2.8) gives

$$\sigma E\left[\Delta h(X).g(X) \mid X > x\right] = \frac{u(\theta)}{u'(\theta)R(x+1)}\left[R(x+1)\frac{\partial m(x)}{\partial\theta} + \big(h(x+1) - m(x)\big)\frac{\partial F(x)}{\partial\theta}\right]$$

$$= \frac{u(\theta)}{u'(\theta)}\frac{\partial m(x)}{\partial\theta} - \frac{1}{R(x+1)}\frac{\partial F(x)}{\partial\theta}\big(m(x) - h(x+1)\big),$$

$$= \frac{u(\theta)}{u'(\theta)}\frac{\partial m(x)}{\partial\theta} + \big(m(x) - \mu\big)\big(m(x) - h(x+1)\big), \qquad (6.2.10)$$

so that the identity (6.2.5) reduces to (6.2.7).

Remark 6.2.2: When $u(\theta) = \theta$ in the above equations, the results for the sub-class of generalized power series distributions can be obtained.

Theorem 6.2.3:

The distribution of X belongs to Ord family of distributions specified by the difference equation (2.4.5) if and only if

$$V(X \mid X > x) = E\left[\left(b_0 + b_1 X + b_2 X^2\right) \mid X > x\right] + \big(\mu - m_1(x)\big)\big(m_1(x) - x - 1\big) \qquad (6.2.11)$$

where

$$b_0 = \mu + \frac{a_0 - a_1 + a_2}{(1-2a_2)},\ b_1 = \frac{a_1 - 1}{(1-2a_2)},\ b_2 = \frac{a_2}{(1-2a_2)} \text{ and } a_2 \neq \frac{1}{2}.$$

Proof:

Taking $h(X) = X$, from Theorem 4.2.3 we have obtained the values of $g(x)$ for the family specified by (2.4.5) as

$$\sigma g(x) = b_0 + b_1 x + b_2 x^2$$

with b_i's as stated in the Theorem. Hence the proof follows from Corollary 6.2.1. To prove the converse supposes that (6.2.11) holds. Using the identity (6.2.5) we have

$$\sigma E\left[g(X) \mid X > x\right] = E\left[\left(b_0 + b_1 X + b_2 X^2\right) \mid X > x\right], \text{ for all } x$$

or

$$\sigma \sum_{k=x+1}^{\infty} g(k) f(k) = \sum_{k=x+1}^{\infty} \left(b_0 + b_1 k + b_2 k^2\right) f(k). \tag{6.2.12}$$

Changing from $x+1$ to x in (6.2.12) and then subtracting (6.2.12)

$$\sigma g(x) f(x) = \left(b_0 + b_1 x + b_2 x^2\right) f(x)$$

or

$$\sigma g(x) = \left(b_0 + b_1 x + b_2 x^2\right).$$

Hence by Corollary 6.2.1, (4.2.3) holds with the above $g(x)$ value so that the distribution of X belongs to the Ord family by Nair and Sankaran (1991).

Theorem 6.2.4:

The distribution of X belongs to Katz family of distributions specified by (4.2.16) if and only if

$$V(X \mid X > x) = (1-\beta)^{-1}\left(\alpha + \beta m_1(x)\right) + \left(\mu - m_1(x)\right)\left(m_1(x) - x - 1\right). \tag{6.2.13}$$

Proof:

We see from Theorem 4.2.4 that for the family specified by (4.2.16) the value of $g(x)$ is given by

$$\sigma g(x) = (1-\beta)^{-1}(\alpha+\beta x).$$

Substituting in the identity (6.2.5) we have the result stated in the theorem. Conversely suppose that (6.2.13) holds, using (6.2.5) we obtained

$$\sigma E\left[g(X) \mid X > x\right] = E\left[\left((1-\beta)^{-1}(\alpha+\beta m_1(x))\right) \mid X > x\right]$$

or

$$\sigma \sum_{k=x+1}^{\infty} g(k) f(k) = \sum_{k=x+1}^{\infty} \left(b_0 + b_1 k + b_2 k^2\right) f(k). \qquad (6.2.14)$$

Changing from $x+1$ to x in (6.2.14) and then subtracting (6.2.14)

$$\sigma g(x) f(x) = (1-\beta)^{-1}(\alpha+\beta x) f(x)$$

or

$$\sigma g(x) = (1-\beta)^{-1}(\alpha+\beta x).$$

Hence by Corollary 6.2.1, equation (4.2.3) holds with the above $g(x)$ and therefore X has Katz family of distributions by results of Theorem 4.2.4.

Example 6.2.1: For the binomial random variable with probability mass function

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

the value of $g(x)$ is given by (see Example 4.2.2)

$$\sigma g(x) = p(n-x),$$

so that

$$m(x) - \mu = (1-k(x))^{-1} k(x)(n-x)p$$

$$= \left(\frac{R(x)-p(x)}{R(x)}\right)^{-1} \frac{p(x)}{R(x)}(n-x)p$$

$$= (n-x)p\frac{p(x)}{p(x+1)}k(x+1)$$

$$= (x+1)(1-p)k(x+1) \qquad (6.2.15)$$

and the variance expression using (6.2.5) is given by

$$V(X \mid X > x) = np(1-p) + (x+1)(1-p)(x+1-p(n+1))k(x+1) - (x+1)^2 k^2(x+1)$$
(6.2.16)

Example 6.2.2: For the Poisson distribution with probability mass function

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,2,...,$$

from Example 4.2.2 we found the value of $g(x)$ as

$$\sigma g(x) = \lambda,$$

hence by using (4.2.3)

$$\begin{aligned} m(x) &= \lambda + \lambda(1-k(x))^{-1} k(x) \\ &= \lambda + \lambda\left(\frac{R(x)-p(x)}{R(x)}\right)^{-1} \frac{p(x)}{R(x)} \\ &= \lambda + \lambda \frac{p(x)}{p(x+1)} k(x+1) \\ &= \lambda + (x+1)k(x+1), \end{aligned} \tag{6.2.17}$$

substituting in equation (6.2.5) gives

$$V(X \mid X > x) = \lambda + (x+1)(x+1-\lambda)k(x+1) - (x+1)^2 k^2(x+1). \tag{6.2.18}$$

Remark 6.2.3: The characterizations by relationship between failure rate and conditional variance for the binomial and the Poisson random variables discussed in Example (6.2.1) and (6.2.2) are similar to those by respective identities (5.1.2) and (5.1.3). Hence the results given by Arishi (2005) for the binomial and the Poisson random variables are special cases of Theorem 6.2.3. Also the identities (6.2.15) and (6.2.17) provide the same characterization results of Ahmed (1991) concerning binomial and Poisson random variables.

Example 6.2.3: For the negative binomial distribution with probability mass function

$$p(x) = \binom{k+x-1}{k-1} p^k (1-p)^x, \quad x \in N, \tag{6.2.19}$$

we see from example (4.2.3) that the value of $g(x)$ is given by

$$\sigma g(x)=\frac{(1-p)}{p}(x+k).$$

Now

$$m(x)=\mu+\sigma g(x)(1-k(x))^{-1}k(x)$$

$$=\mu+\frac{(1-p)}{p}(x+k)\left(\frac{R(x)-p(x)}{R(x)}\right)^{-1}\frac{p(x)}{R(x)}$$

$$=\mu+\frac{(1-p)}{p}(x+k)\frac{p(x)}{p(x+1)}k(x+1)$$

$$=\mu+\frac{(1-p)}{p}(x+k)\frac{(x+1)}{(1-p)(x+k)}k(x+1)$$

$$=\mu+p^{-1}.(x+1)k(x+1)$$

Hence the conditional variance identity is valid and written as

$$V(X\mid X>x)=\frac{\mu}{p}+\frac{(x+1)}{p}k(x+1)+\frac{(x+1-\mu)}{p}k(x+1)-\frac{(x+1)^2}{p^2}k^2(x+1).$$

A special case of (6.2.19) is the geometric distribution with probability mass function

$$p(x)=p(1-p)^x,\ \ x=0,1,2,...$$

and

$$\sigma g(x)=\mu(x+1),$$

hence we have

$$m(x)=\mu+\mu(x+1)\frac{p(x)}{p(x+1)}k(x+1)$$

$$=\mu+(x+1)$$

and the variance expression using Corollary 6.2.1 is given by

$$V(X\mid X>x)=\mu^2+\mu(x+2)-\mu(x+1)$$

$$=\mu^2+\mu$$

$$=q/p^2=V(X).$$

It is interesting to note from the above that the conditional as well as unconditional variances are the same for the geometric distribution, a fact observed earlier as a characteristic property.

6.3 Lower bound to the conditional variance

As discussed in the previous chapter, here also we explore the scope for introducing bounds on variance using the methodology in Cacoullos and Papathanasiou (1997). The following theorem gives the analogues result of Theorem 5.3.1.

Theorem 6.3.1:

Under the condition of the Theorem 6.2.1, for every $c(x)$ in $\mathfrak{C}$,

$$V\left(c(X)\mid X>x\right)\geq\left[V\left(h(X)\mid X>x\right)\right]^{-1}\left[\sigma E\left[\Delta c(X)g(X)\mid X>x\right]\right.$$

$$\left.+\left(\mu-m(x)\right)\left(a(x)-c(x+1)\right)\right]^{2} \qquad (6.3.1)$$

if (4.2.3) holds, with equality whenever $c(X)$ is a linear function of $h(X)$. Conversely, if the conditional variance is (6.3.1), then (4.2.3) is true provided $V\left(h(X)\mid X>x\right)$ is as in (6.2.5).

Proof:

Since

$$Cov^{2}\left(c(X),h(X)\mid X>x\right)\leq V\left(c(X)\mid X>x\right)V\left(h(X)\mid X>x\right),$$

using (6.2.1) we have (6.3.1). To prove the second part of the Theorem assume

$$c(x)=h(x)+\theta l(x)$$

for some arbitrary real θ. Then

$$a(x)=m(x)+\theta b(x)\ , \qquad b(x)=E\left(l(X)\mid X>x\right)$$

so that (6.3.1) becomes

$$V\left(h(X)\mid X>x\right)\left[V\left(h(X)\mid X>x\right)+\theta^{2}V\left(l(X)\mid X>x\right)+2\theta Cov\left(l(X),h(X)\mid X>x\right)\right]$$

$$\geq\left[\sigma E\left[\Delta h(X).g(X)\mid X>x\right]+\sigma\theta E\left[\Delta l(X).g(X)\mid X>x\right]\right.$$

$$\left.+\left\{r(x)+\theta b(x)-h(x+1)-\theta l(x+1)\right\}\left(\mu-m(x)\right)\right]^{2}$$

$$=\left[\sigma\theta E\left[\Delta l(X).g(X)\mid X>x\right]+\theta\left(b(x)-l(x)\right)\left(\mu-m(x)\right)+V\left(h(X)\mid X>x\right)\right]^2. \tag{6.3.2}$$

Using (6.2.5), inequality (6.3.2) simplifies to

$$\theta^2\Big[V\left(h(X)\mid X>x\right)V\left(l(X)\mid X>x\right)-\big\{\sigma E\left[\Delta l(X).g(X)\mid X>x\right]$$
$$+\left(\mu-m(x)\right)\left(b(x)-l(x+1)\right)\big\}^2\Big]+2\theta V\left(h(X)\mid X>x\right)\Big[Cov\left(l(X),h(X)\mid X>x\right)$$
$$-\big\{\sigma E\left[\Delta l(X).g(X)\mid X>x\right]+\left(\mu-m(x)\right)\left(b(x)-l(x+1)\right)\big\}\Big]\geq 0. \tag{6.3.3}$$

For (6.3.3) to be true for all θ, the coefficient of θ must vanish and therefore,

$$Cov\left(l(X),h(X)\mid X>x\right)=\sigma E\left[\Delta l(X).g(X)\mid X>x\right]+\left(\mu-m(x)\right)\left(b(x)-l(x+1)\right),$$

which on simplification reduces to

$$E\left(l(X)\left(h(X)-\mu\right)\mid X>x\right)=\sigma E\left(\Delta l(X).g(X)\mid X>x\right)+l(x)E\left(\left(h(X)-\mu\right)\mid X>x\right),$$

which is similar to the expression (6.2.3), hence rest of the proof follows from Theorem 6.2.1 and 4.2.1.

As in the continuous case now we are examining how our inequality compares with the lower bound of the variance in the Cramer- Rao theorem and Chapman- Robbins inequality.

Let X belong to the family specified by (4.2.8) which is also known as discrete exponential family, then the random variable $(X\mid X>t)$ has density

$$p_1(x)=a(x)\left[u(\theta)\right]^x/R(t+1)A(\theta),\quad X>t.$$

For this density the Cramer- Rao lower bound is attained for the unbiased estimation of of

$$m(t)=E\left(h(X)\mid X>t\right)$$

and this bound is

$$V(t)=\frac{\left[m'(t)\right]^2}{E\left(\frac{\partial\log f_1(x)}{\partial\theta}\right)^2}.$$

For the attainment of (6.3.1) a necessary and sufficient condition is (4.2.3), subsequently from the discussion of Section 4.3 we have the expression

$$\frac{\partial \log p(x)}{\partial \theta} = \frac{u'(\theta)}{u(\theta)}\left(h(x)-\mu\right),$$

so that

$$E\left(\frac{\partial \log p(x)}{\partial \theta} \mid X > t\right)^2 = \left(\frac{u'(\theta)}{u(\theta)}\right)^2 E\left(\left(h(X)-\mu\right)^2 \mid X > t\right).$$

$$= \left(\frac{u'(\theta)}{u(\theta)}\right)^2 V\left(h(X) \mid X > t\right)$$

Now the Cramer-Rao lower bound for unbiasedly estimating $m(t)$ using $h(x)$ is

$$V(t) = \frac{\left[m'(t)\right]^2}{\left(\frac{u'(\theta)}{u(\theta)}\right)^2 V\left(h(X) \mid X > t\right)}$$

$$= \frac{\left[m'(t)\right]^2}{\left(\frac{u'(\theta)}{u(\theta)}\right)^2 \frac{u(\theta)}{u'(\theta)} m'(t)}$$

$$= \frac{u(\theta)}{u'(\theta)} m'(t),$$

which is also the bound obtained by (6.3.1) and is given in (6.2.7).

Now we compare our bound with variance provided by Chapman- Robbins inequality. Let $E_\theta\left(h(X) \mid X > x\right) = m_\theta(x)$, where $\theta,\ \varphi \in \Theta \subset R$ such that $p_\theta(x)$ and $p_\varphi(x)$ are different, satisfying $\{p_\theta(x) > 0\} \supset \{p_\varphi(x) > 0\}$, choosing

$$c(x) = \left(\frac{p_\varphi(x)}{p_\theta(x)} - 1\right)$$

and using the identity (6.2.4)

$$\sigma E\left[\Delta c(X).g(X) \mid X > x\right] + c(x+1)\left(m(x)-\mu\right) = \mathrm{E}\left[\mathrm{c}(\mathrm{X})\left(h(\mathrm{X})-\mu\right) \mid X > x\right]$$

$$= E\left[\left(\frac{p_\varphi(X)}{p_\theta(X)} - 1\right)(h(X) - \mu) \mid X > x\right]$$

$$= E\left[\left(\frac{p_\varphi(X)}{p_\theta(X)} - 1\right)h(X) \mid X > x\right]$$

$$= \left[m_\varphi(x) - m_\theta(x)\right].$$

Now, the inequality (6.3.1) reduces to

$$V(c(X) \mid X > x) \geq \left[m_\varphi(x) - m_\theta(x)\right]^2 / V(h(X) \mid X > x)$$

or

$$V(h(X) \mid X > x) \geq \frac{\left[m_\varphi(x) - m_\theta(x)\right]^2}{V(p_\varphi(X)/p_\theta(X) \mid X > x)},$$

which is the Chapman-Robbins inequality.

In this chapter, some characterizations based on VRL is obtained. A lower bound to the variance proposed for the left truncated random variable can be utilized to find the UMVUE for the desired parametric functions.

6.4 Some open problems

The present thesis is concentrated on variance bounds in the discrete and continuous cases with respect to functions of a single random variable. As a natural extension, the possibility of similar results for the multivariate case is an open problem. Since definition of failure rate and mean residual life in multivariate case exist in more than one way, each case has to be dealt with separately. An important aspect of modelling using multivariate distributions is the dependency structure in the data. Hence covariance between residual lives also can play a significant role in the process. These and other concepts for the study of bounds for multi dimensional random variables will be taken up in a future work.

References

1. Abraham, B and Nair, N.U (2001). On characterizing mixtures of some life distributions, *Statistical Papers,* 42, 387- 393.
2. Adatia, A.A, Law, G and Wang, Q (1991). Characterization of a mixture of gamma distributions via. Conditional moments, *Communications in Statistics-Theory and Methods,* 20, 1937- 1949.
3. Ahmed, A.N (1991). Characterization of beta, binomial and Poisson distribution, *IEEE Transactions on Reliability,* 40, 290-293.
4. Alharbi, A.A and Shanbhag, D.N (1996). General characterization theorems based on versions of the Chernoff inequality and Cox representation, *Journal of Statistical Planning and Inference*, 55, 139-150.
5. Arkalian, V and Papathanasiou, V (2004). On bounding the absolute mean value, *Statistics and Probability Letters,* 69, 447- 450.
6. Arnold, B.C and Brocket, P.L (1988). Variance bounds using a theorem of Polya, *Statistics and Probability Letters,* 62, 321- 326.
7. Balakrishnan, N and Subramonian, K (1993). Equivalence of Hartely- David-Gumbel- Papathanasiou bound and some further remarks, *Statistics and Probability Letters,* 16, 39-41.
8. Block H.W, Savits, T.H and Singh H (1998). The reversed hazard rate function, *Probability in Engineering and Informational Sciences*, 12, 69- 90.
9. Bobkov, S.G and Houdre, C (1997). Converse Poincare- type inequalities for convex functions, *Statistics and Probability Letters,* 34, 37- 42.
10. Borokov, A.A and Utev, S.A (1984). On an inequality and a related characterization of the normal distribution, *Theory of Probability and its Applications,* 28, 219-228.
11. Borzadaran, G.R.M and Shanbhag, D.N (1998a). Further results based on Chernoff- Type inequalities, *Statistics and Probability Letters*, 39, 109–117.

12. Borzadaran, G.R.M and Shanbhag, D.N (1998b). General characterization theorems via the mean absolute deviation, *Journal of Statistical Planning and Inference*, *74,* 205- 214.

13. Bracquemond, C and Gaudoin, O (2003). A survey on discrete lifetime distributions, *International Journal on Reliability, Quality, and Safety Engineering*, 10, 69- 98.

14. Brascamp, H.J and Lieb, E.H (1976). On extension of Brunn- Minkoswki and Prekop- Liender theorems including inequalities for log concave functions and with an application to the diffusion equation, *Journal of Functional Analysis,* 22, 366-389.

15. Cacoullos, T (1982). On upper and lower bounds for the variance of a function of a random variable, *Annals of Probability*, 10, 709-809.

16. Cacoullos, T and Papathanasiou, V (1985). On upper bounds on the variance of random variables, *Statistics and Probability Letters,* 3, 175-184.

17. Cacoullos, T and Papathanasiou, V (1989). Characterization of distribution by variance bounds, *Statistics and Probability Letters,* 7, 351-356.

18. Cacoullos, T and Papathanasiou, V (1992). Lower variance bounds and a new proof of the central limit theorem, *Journal of Multivariate Analysis*, 43, 173–184.

19. Cacoullos, T and Papathanasiou, V (1995). A generalization of covariance identity and related characterizations, *Mathematical Methods of Statistics*, 4, 106-113.

20. Cacoullos, T and Papathanasiou, V (1997). Characterizations of distributions by generalizations of variance bounds and simple proofs of the C. L. T, *Journal of Statistical Planning and Inference*, 63, 157– 171.

21. Cacoullos, T, Papathanassiou, V and Utev, S (1993).Another Characterization of the normal law and a proof of central limit theorem connected with it, *Theory of Probability and its Application,* 37, 581- 582.

22. Cacoullos, T, Papathanassiou, V and Utev, S (1994). Variational inequalities with example and application to central limit theorem, *Annals of Probability,* 22, 1607-1618.

23. Chapman, D.G and Robbins, H (1951). Minimum variance estimation without regularity assumptions, *Annals of Mathematical Statistics,* 22, 581- 586.

24. Chen, L (1982). An inequality for multivariate normal distributions, *Journal of Multivariate Analysis*, 12, 306- 315.

25. Chen, L (1985). Poincaré- type inequalities via stochastic integrals, *Probability Theory and Related Fields,* 69, 251- 257.

26. Chen, L (1988). The central limit theorem and Poincare- type inequalities, *Annals of Probability,* 16, 300-304.

27. Chen, L and Lou, J.N (1987). Characterization of probability distribution by Poincare inequality under normal probability measure, *Annales de Institute Henri Poincaré,* 23, 91- 110.

28. Chernoff, H (1981). A note on an inequality involving the normal distribution, *Annals of Probability*, 9, 533-535

29. Chou, J.P (1988). An inequality for multi dimensional continuous exponential families and its applications, *Journal of Multivariate Analysis*, 24, 129–142.

30. Cobb, L (1981). The multimodal exponential families of statistical catastrophe theory. Statistical Distributions in Scientific Work, 4, 67-90, Ed. Taille, C; Patel G. P and Baldessari, B.A, D. Reidel Publishing Company, Holland.

31. Consul, P.C (1995). Some characterization of exponential class of distributions, *IEEE Transaction on Reliability*, 44, 403-407.

32. Cramer, H (1946). A contribution to the theory of statistical estimation, *skand. Aktuarietidskr,* 29, 85- 94.

33. Dallas, A.C (1981). A characterization using conditional variance, *Metrika*, 28, 151- 153.

34. Ebrahimi, N (1986). Classes of discrete decreasing and increasing mean residual life distributions, *IEEE Transaction on Reliability*, 35, 403-405.

35. El- Arishi, S (2005). A conditional variance characterization of some discrete probability distributions, *Statistical Papers*, 46, 31- 45.

36. Ghitany, M.E, El-Saidi, M.A and Khalil, Z (1995). Characterization of a general class of life testing models, *Journal of Applied Probability*, 32, 548-553.

37. Glanzel, W (1991). Characterization through some conditional moments of Pearson-type distributions and discrete analogues, *Sankhya- B*, 53, 17- 24.

38. Glaser, R.E (1980). Bath- tub and related failure rate characterizations, *Journal of American Statistical Association*, 75, 667- 672.

39. Guess, F.M and Park, D.H (1988). Modelling discrete bathtub and upside down bathtub mean residual life functions, *IEEE Transactions on Reliability*, 37, 545-549.

40. Gumbel, E.J (1954). *Statistical theory of extreme values and some practical applications, a series of lectures*. National Bureau of Standards, Applied Mathematics Series, 33. US Govt. Printing Office, Washington. USA.

41. Gupta, R.C (1974). Characterization of distributions by property of discrete order statistics. *Communications in Statistics–Theory and Methods*, 3, 287-289.

42. Gupta, P.L (1985). Some characterizations of distributions by truncated moments, *Statistics*, 16, 465- 473.

43. Gupta, R.C (1987). On the monotonic properties of residual variance and their applications in reliability, *Journal of Statistical Planning and Inference*, 16, 329–335.

44. Gupta, R.C and Bradley, D.M (2003). Representing mean residual life in terms of the failure rate, *Mathematical and Computer Modeling*, 1, 1-10.

45. Gupta, R.C and Kirmani, S.N.U.A (1987). On order relations between reliability measures, *Communications in Statistics–Theory and Methods*, 16, 149–156.

46. Gupta, R.C and Kirmani, S.N.U.A (1990). The role of weighted distributions on stochastic modeling, *Communications in Statistics–Theory and Methods*, 19, 3147-3162.

47. Gupta, R.C and Kirmani, S.N.U.A (1998). Closure and monotonicity properties of non homogeneous Poisson processes and record values, *Probability in Engineering and Informational Sciences,* 2, 475–484.

48. Gupta, R.C and Kirmani, S.N.U.A (2000). Residual coefficient of variation and some characterization results, *Journal of Statistical Planning and Inference*, 91, 23–31.

49. Gupta, R.C and Kirmani, S.N.U.A (2004). Some characterization of distributions by functions of failure rates and mean residual life, *Communications in Statistics–Theory and Methods*, 33, 3115-3131.

50. Gupta, R.C, Kirmani, S.N.U.A and Launer, R.L (1987). On life distributions having monotone residual variance, *Probability in Engineering and Informational Sciences*, 1, 299- 307.

51. Gupta, R.D and Nanda, A.K (2001). Some results on reversed hazard process ordering, *Communications in Statistics– Theory and Methods*, 30, 2447-2457.

52. Gupta, R.P Nair, N.U and Asha, G (2005). Characterizing discrete life distributions by relation between reversed failure rates and conditional expectations, Far East Journal of Theoretical Statistics, 20, 113- 122.

53. Hartley, H.O. and David, H.A. (1954). Universal Bounds for Mean Range and Extreme Observation, *Annals of Mathematical Statistics,* 25, 85-99.

54. Hitha, N and Nair, N.U (1989). Characterization of some models by properties of residual life, *Calcutta Statistical Association Bulletin*, 38, 219-225.

55. Houdre, C and Kagan, A (1995). Variance inequalities for functions of Gaussian variables, *Journal of Theoretical Probability,* 8, 23- 30.

56. Houdre, C and Perez- Abreu, V (1995). Covariance identities and inequalities for functions in Weiner and Poisson spaces, *Annals of Probability,* 23, 400- 419.

57. Hovenier, J.W (1994). Sharpening Cauchy's Inequality, *Journal of Mathematical Analysis and Applications,* 186, 156-160

58. Hu Chin- Yuan (1986). An inequality involving rectangular distribution, *Bulletin of the Institute of Mathematics Academia Sinica,* 14, 21- 23.

59. Johnson, N.L, Kotz, S and Kemp, A.W (1992). *Univariate discrete distributions*, John Wiley, New York.

60. Johnson, R.W (1993). A note on variance bounds for a function of the Pearson variate, *Statistical Decision,* 11, 273- 278.

61. Katz, L (1946). On the class of functions defined by the difference equation $(x+1)f(x+1)=(a+bx)f(x)$, *Annals of Mathematical Statistics,* 17, 501.

62. Kemp, A.W (2004). Classes of discrete life distributions, *Communications in Statistics –Theory and Methods,* 33, 3069-3093

63. Klassen, C.A.J (1985). On an inequality of Chernoff, *Annals of Probability,* 13, 966- 974.

64. Koicheva, M. A (1993). Characterization of the Gamma distribution in terms of conditional moments, *Applied Mathematics,* 38, 19-22.

65. Korwar, R.M (1991). On characterization of distributions by mean absolute deviation and variance bounds, *Annals of Institute of Statistical Mathematics*, 43, 287-295.

66. Kotz, S and Shanbhag, D.N (1980). Some new approaches to probability distributions, *Advances in Applied Probability,* 12, 903- 921.

67. Kundu, D and Gupta, R.D (2003). Characterization of the proportional hazard class, *Communications in Statistics –Theory and Methods,* 32, 3095-3102.

68. Lehman, E.L and Casella, G (1998). *Theory of point estimation,* Second Edition, Springer, New York.

69. Launer, R.L (1984). Inequalities for NBUE and NWUE distributions, *Operation Research*, 32, 660- 667.

70. Mould, R.F (2005). *Introductory medical statistics*, Overseas Press India Private limited, New Delhi.

71. Mukherjee, S.P and Roy, D (1986). Some characterizations of the exponential and related life distributions, *Calcutta Statistical Association Bulletin,* 36, 189- 197.

72. Nagaraja, H.N (1975). Characterization of some distributions by conditional moments, *Journal of Indian Society for Probability and Statistics,* 13, 57- 61.

73. Nair, N.U (1983). A measure of memory of some discrete distributions, *Journal of Indian Society for Probability and Statistics*, 21,141-147.

74. Nair, N.U and Asha, G (2004). Characterization using failure rate and reversed failure rate, *Journal of Indian Society for Probability and Statistics*, 8, 45-57.

75. Nair, N.U, Geetha, K.G and Priya, P (1999). Modelling life length data using mixture distributions, *Journal of Japanese Statistical Society*, 29, 65- 73.

76. Nair, N.U and Sankaran, P.G (1991). Characterization of Pearson family of distribution, *IEEE Transactions on Reliability*, 40, 75- 77.

77. Nair, N.U and Hitha, N (1989). Characterization of discrete models by distribution based on their partial sums, *Statistics and Probability Letters,* 8, 335-337.

78. Nair, N.U, Sankaran, P.G and Asha, G (2005). Characterization of distribution using reliability concepts, *Journal of Applied Statistical Science,* 14, 237- 241.

79. Nair, N.U and Sudheesh. K. K (2006). Characterization of continuous distributions by variance bound and its implications to reliability modeling and catastrophe theory, *Communication in Statistics- Theory and Methods*, 35, 1985-1995.

80. Nair, N.U and Sudheesh. K. K (2007). Some results on lower variance bound useful in reliability and estimation, *Annals of the Institute of Statistical Mathematics* (to appear).

81. Nair, N.U and Sudheesh. K. K (2007). Some results on variance bound, *Journal of Indian Society for Probability and Statistics* (to appear)

82. Nanda, A.K and Sengupta, D (2005). Discrete life distribution with decreasing hazard, *Sankhya- A*, 55, 164-168.

83. Nanda, A.K and Gupta, R.D (2001), Some Properties of Reversed Hazard Rate Function, *Statistical Methods*, 3, 108- 124.

84. Nanda, A.K, Singh, H, Misra, N and Paul P (2003), Reliability properties of reversed residual life, *Communications in Statistics- Theory and Methods*, 32, 2031- 2042.

85. Nash, J (1958). Continuity of solutions of parabolic and elliptic equations, *American Journal of Mathematics*, 80, 931- 954.

86. Nassar, M.M and Mahmoud, M.R (1985).On characterizations of a mixture of exponential distributions, *IEEE Transaction on Reliability*, 34, 484- 488.

87. Navarro, J, Franco, M and Ruizz, J.M (1998). Characterization through moments of residual life and conditional spacing, *Sankhya- A*, 60, 36- 48.

88. Navarro, J and Ruizz J.M (2004). Characterizations from relationships between failure rate functions and conditional moments, *Communications in Statistics- Theory and Methods*, 33, 3159 – 3171.

89. Olkin, I and Shep, L (2005). A matrix variance inequality, *Journal of Statistical Planning and Inference*, 130, 351–358.

90. Osaki, S and Li, X (1988). Characterization of gamma and negative binomial distributions, *IEEE Transactions on Reliability*, 37, 379-382.

91. Papadotos, N and Papathanasiou, V (1996). A generalization of variance bounds, *Statistics and Probability Letters*, 28, 191- 194.

92. Papadotos, N and Papathanasiou, V (2003). Multivariate covariance identity with application to order statistics, *Sankhya- A*, 65, 307- 316.

93. Papathanasiou, V (1990). Some characterization of distributions based on order statistics, *Statistics and Probability Letters*, 9, 145- 147.

94. Papathanasiou, V (1993). Characterization of power series family and factorial series distributions, *Sankhya- A*, 55, 164-168.

95. Patil G. P (1961). Asymptotic bias and variance of ratio estimates in generalized power series distributions and certain applications, *Sankhya- A*, 23, 269-280.

96. Patil G. P (1962). Certain properties of the generalized power series distribution. *Annals of the Institute of Statistical Mathematics*, 14, 179-182.

97. Pearson, K (1895). Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material, *Philosophical Transactions of the Royal Society of London A,* 186: 343–414.

98. Prakasa Rao, B.L.S (1979). Characterizations of distribution through some identities, *Journal of Applied Probability,* 16, 903-909.

99. Prakasa Roa, B.L.S (1993). On some inequalities of Chernoff- type, *Theory of Probability and its Applications,* 37, 392- 398.

100. Prakasa Rao, B.L.S and Sreehari, M (1986). Another characterization of multivariate normal distribution, *Statistics and Probability Letters*, 4, 208- 210.

101. Prakasa Rao, B.L.S and Sreehari, M (1987). On a characterization of Poisson distribution through inequalities of Chernoff type, *Australian Journal of Statistics*, 29, 38-43.

102. Prakasa Rao, B.L.S and Sreehari, M (1997). A Chernoff- type inequality and variance bound, *Journal of Statistical Planning and Inference*, 63, 325- 335.

103. Pudget, W.J and Spurrier, J.D (1985), On discrete failure models, *IEEE Transactions on Reliability*, 34, 253- 256.

104. Purkayastha, S and Bhandari, S.K (1990). Characterization of uniform distribution by inequalities of Chernoff type, *Sankhya-A,* 52, 376-382.

105. Rao, C.R (1945). Information and accuracy attainable in the estimation of statistical parameters, *Bulletin of Calcutta Mathematical Society,* 37, 81- 91.

106. Ruizz, J.M and Navarro, J (1994). Characterization of distribution by relationship between failure rate and mean residual life, *IEEE Transactions on Reliability*, 43, 640-644.

107. Salvia, A.A and Bollinger, R.C (1982). On discrete hazard functions, *IEEE Transactions on Reliability*, 31, 458- 459.

108. Sankaran, P.G and Nair, N.U (2000). On some reliability aspect of the Pearson family of distributions, *Statistical Papers*, 41, 109-117.

109. Sankaran, P.G, Nair, N.U and Sindhu, T.K (2003). A generalized Pearson system useful in reliability analysis, *Statistical Papers*, 4, 125-130.

110. Shanbagh, D.N (1970). Characterization of exponential and geometric distributions, *Journal of American Statistical Association*, 65, 1256- 1259.

111. Sindhu, T.K (2003). An extended Pearson system useful in reliability analysis, Unpublished Ph. D thesis, Cochin University of Science and Technology.

112. Srivastava, D and Sreehari, M (1987). Characterization of family of distribution via Chernoff- type inequality, *Statistics and Probability Letters*, 5, 293-294.

113. Srivastava, D and Sreehari, M (1990). Characterization results on variance bounds, *Sankhya- A*, 52, 297-302.

114. Stoyanov, J and Al- Sadi, M. H. M. (2004), Properties of class of life distributions based on the conditional variance, *Journal of Applied Probability*, 41, 953- 960.

115. Vitale, R (1989). A differential version of the Effron- Stein inequality: Bounding variance of a function of an infinitely divisible random variable, *Statistics and Probability Letters*, 7, 105-112.

116. Xekalaki, E (1983). Hazard function and life distributions in discrete time, *Communications in Statistics-Theory and Methods*, 12, 2503- 2509.