2.1. Introduction

The normal distribution enjoyed a pivotal role in all kinds of statistical analysis till the end of the nineteenth century and most theoretical developments took place on the assumption that the population is normal or at least approximately so. However, when the interest was focussed on describing natural phenomena by finding statistical distributions that fit the data, it became apparent that the samples from many sources show characteristics that are markedly different from normal. The then prevailing practice of attributing departure from normality to errors in measurement or to imprecise methods of collection, began slowly giving way to the belief that such departures depicted certain inherent features of the population that required alternative models. By the turn of the twentieth century non-normal curves became an accepted fact and efforts were under way to generate systems of curves which include the normal only as a particular case. Of the different approaches initiated to meet this end, Karl Pearson's contribution to describe a system of
distributions by the differential equations

\[ \frac{1}{f(x)} \frac{df(x)}{dx} = \frac{-(x+d)}{b_0 + b_1 x + b_2 x^2} \]  

(2.1)

still stands out as a convenient family that includes many important probability models. Among them the normal, exponential, gamma, beta, Pareto, finite range etc. are used extensively as lifetime distributions and we have reviewed characterizations of these distributions by reliability concepts. Instead of looking at characterizing individual members of the family, in the present chapter we present some results that hold good for the entire system, and then verifying that many of the existing results can be obtained as particular cases of the general theorem.

2.2. Characterization by relation between failure rate and vitality function*

Let \((\Omega, \mathcal{F}, P)\) be a probability space and \(X\) be a random variable thereon such that the range of \(X\) is \(H = (a, b)\) for some real \(a < b\); where \(a\) can be \(-\infty\) and \(b\) \(\infty\).

and $b$ can be $+\infty$. Assume that the distribution function $F(x)$ of $X$ is absolutely continuous with respect to Lebesgue measure and that $f(x)$ is its density. Then the distribution of $X: \mathbb{R} \to H$ belongs to the Pearson system, if the density function is differentiable and satisfies the equation (2.1).

**Theorem 2.1.**

Let $\mu = E(X) < \infty$ and $\lim_{x \to b} (b_0 + b_1 x + b_2 x^2) f(x) \to 0$ if $b = +\infty$ and $\lim_{x \to b} (b_0 + b_1 x + b_2 x^2) f(x) = 0$ if $b < \infty$. A necessary and sufficient condition for the distribution of $X$ to belong to the Pearson family is that for all $x$ in $(a, b)$

$$m(x) = \mu + (a_0 + a_1 x + a_2 x^2) h(x) \quad (2.2)$$

**Proof:**

Suppose that $X$ is a member of the Pearson family.

Then, from the equation (2.1),

$$\int_{x}^{b} (b_0 + b_1 t + b_2 t^2) f'(t) dt = \int_{x}^{b} (t + d) dR(t) \quad (2.3)$$

where, $R(x)$ is the survival function of $X$. 

Conversely, if the relationship (2.5) is satisfied, then

$$\int_{x}^{b} f'(t) dt = -(x + d) R(x) - \int_{x}^{b} R(t) dt$$

$$= -(x + d) R(x) - \int_{x}^{b} R(t) dt$$

$$= -(x + d) R(x) - \int_{x}^{b} R(t) dt$$

$$= -(x + d) R(x) - \int_{x}^{b} R(t) dt$$

$$= -(x + d) R(x) - \int_{x}^{b} R(t) dt$$

$$= -(x + d) R(x) - \int_{x}^{b} R(t) dt$$
Integrating by parts and using the assumptions of the theorem, (2.3) becomes

\[-(b_0 b_1 x + b_2 x^2) f(x) - \int_x^b (b_1 + 2b_2 t) f(t) dt\]

and finally,

\[-(x + d) R(x) - \int_x^b R(t) dt.\]

That is,

\[-f(x)(b_0 b_1 x + b_2 x^2) - (b_1 - d) R(x) = (2b_2 - 1)m(x) R(x).\]  \hspace{1cm} (2.4)

The equation (2.4) can be written as

\[m(x) = \mu + (a_0 + a_1 x + a_2 x^2) h(x),\]

in which,

\[\mu = \frac{b_1 - d}{1 - 2a_2}\]

and

\[a_i = \frac{b_1}{1 - 2b_2} i = 0, 1, 2.\]

Conversely, if the relationship (2.2) is satisfied, then

\[\int_x^b t f(t) dt = (a_0 + a_1 x + a_2 x^2) f(x) + \mu R(x).\]  \hspace{1cm} (2.5)
Differentiating (2.5) with respect to $x$, 

$$-xf(x) = (a_0 + a_1 x + a_2 x^2) f'(x) + (a_1 + 2a_2 x) f(x)$$

(2.6)

and finally,

$$f'(x) = \frac{-(x+d) f(x)}{b_0 + b_1 x + b_2 x^2})$$

(2.7)

which completes the proof.

Deductions.

1. For the gamma distribution, 

$$f(x) = a(\alpha x)^{\beta - 1} e^{-\alpha x} / \Gamma(\beta),$$

direct calculations show that 

$$a_0 = 0, \ a_1 = \alpha^{-1}, \ a_2 = 0 \text{ and } \mu = \beta \alpha^{-1}.$$ 

Accordingly, 

$$m(x) = \mu + \alpha^{-1} x h(x),$$

which is the result of Osaki and Li (1988).
2. When $X$ is normal $N(a, \sigma)$,

$$a_0 = \sigma^2, \ a_1 = 0, \ a_2 = 0 \text{ and } \mu = a$$

so that

$$m(x) = \sigma^2 h(x) + a$$

as observed in Kotz and Shanbhag (1980).

2.3. **Discrete Case**

The literature on reliability analysis is heavily biased towards continuous models and the use of discrete distributions in this context has not been properly investigated or exploited. Frequently, in reliability analysis, we need to know the probability that a specific number of events will occur or to calculate the average number of events that are taking place. For example, suppose that $p$ is the probability that light bulb will fail during the first 100 hours of service. Then on string of 25 lights, what is the probability that there will be 'n' failures during this 100 hours period. To answer this question, we have to consider a discrete model representing by probability mass function. Xekalaki (1983) points out that the discrete models are more appropriate in a variety of
applied problems due to the limitations in measuring equipments and to the fact that many continuous life length distributions can be very well approximated by the corresponding discrete counterparts. Gupta (1985) has given an example of discrete random variables that occur naturally, such as the case with the time to failure in fatigue studies measured in terms of the number of cycles to failure. Moreover, in the type I censoring, the number of failed units up to a certain time period can be represented by a discrete distribution and this may be used to study the failure process of the system. These considerations have opened up a spurt in characterization of discrete models using reliability concepts. For details we refer to Xekalaki (1983), Gupta (1984), Nair and Hitha (1989), and Hitha and Nair (1989).

Let $X$ be a random variable in the support of the set of non-negative integers with the probability mass function $f(x)$ and the survival function $R(x) = P(X \geq x)$. Then the failure rate $h(x)$ of $X$ is defined (Kalbfleisch and Prentice (1980)) as

$$h(x) = \frac{f(x)}{R(x)}, \quad x = 0,1,2,\ldots$$  \hspace{1cm} (2.8)
Further, \( R(x) \) can be written in terms of \( h(x) \) as

\[
R(x) = \prod_{t=0}^{x-1} (1-h(t))
\]

(2.9)

As in the continuous case, here also the failure rate \( h(x) \) uniquely determines the survival function \( R(x) \) or the distribution of \( X \).

The discrete analogue of MRL function \( r(x) \) of \( X \) given in equation (1.4) is defined in the same fashion as

\[
r(x) = E[X-x|X>x]
\]

(2.10)

and the vitality function \( m(x) \) of \( X \) is

\[
m(x) = x + r(x)
\]

(2.11)

Further, \( h(x) \), \( r(x) \) and \( m(x) \) are related to one another by the following identities (Hitha and Nair, 1989)

\[
m(x) = \mu + (\sigma^2 + \alpha \mu + \alpha^2 \mu^2) \ h(x+1)
\]

(2.14)

where, \( \mu = E(X) \), is satisfied for all non-negative integer values of \( x \).
and

\[ 1 - h(x+1) = \frac{r(x)-1}{r(x+1)}. \quad (2.13) \]

With the aid of the above definitions we establish the following characterization theorem for the family of discrete distributions described by the difference equation

\[ f(x+1) - f(x) = \frac{-(x+d) f(x)}{b_0 + b_1 x + b_2 x^2} \quad (2.14) \]

which is the extension of the Pearson system to the discrete case given in Ord (1967). This theorem includes the result of Osaki and Li (1988) concerning the negative binomial distribution as a particular case.

**Theorem 2.2.**

Let \( X \) be a random variable with support as the set of non-negative integers with finite mean. Then the distribution of \( X \) belongs to the family mentioned in (2.14) if and only if the relationship

\[ m(x) = \mu + (a_0 + a_1 x + a_2 x^2) h(x+1) \quad (2.15) \]

where, \( \mu = E(X) \), is satisfied for all non-negative integer values of \( X \).
Proof:

We first prove the necessary part of the theorem. When (2.15) holds

\[
\sum_{t=T}^{n} tf(t) = (a_0 + a_1 x + a_2 x^2) f(x+1) + \mu R(x+1),
\]

(2.16)

where \( n > x+1 \), can be finite or infinite. The last equation reduces to

\[
\sum_{t=T}^{n} R(t+1) = (a_0 + a_1 x + a_2 x^2) f(x+1) + \mu R(x+1).
\]

(2.17)

Now, changing the variable \( x \) to \( (x-1) \), in equation (2.17) and subtracting (2.17) from it, we get

\[
(f(x+1)-f(x))(a_0 + a_1 x + a_2 x^2) = ((1+2a_2)x+a_1-a_2-\mu)f(x).
\]

(2.18)

On simplification, the equation (2.18) leads to (2.14) with the constants \( b_0, b_1, b_2 \) and \( d \) as

\[
b_i = \frac{a_i}{1+2a_2} \quad i = 0,1,2
\]
and
\[ d = \frac{a_1 - a_2 - \mu}{1 + 2a_2}. \]

The converse part of the theorem follows by retracing the above steps.

Corollary 2.1.

Taking \( a_0 = (1-r)p^{-1} \), \( a_1 = p^{-1} \), \( a_2 = 0 \) and \( \mu = rp^{-1} \) we find that

\[
m(x) = \frac{r}{p} + \frac{[(1-r)x]}{p} h(x)
\]

characterizes the negative binomial distribution with probability mass function

\[
f(x) = \binom{x-1}{r-1} p^r (1-p)^{n-r} \quad x \geq r
\]

as proved in Osaki and Li (1988).

The result of the last two theorems are operational in a practical situation once we know the value of \( d, a_0, a_1 \) and \( a_2 \) for various members of the respective families. The \( a_i \)'s are related to the \( b_i \)'s in the systems and expressions for the latter in terms of the moments of the
distribution are well known (see Johnson and Kotz (1969) and Ord (1967)). For easy reference, however, we present in Tables 1 and 2 the values of $b_1$ for some popular models that are members of the two families.

Table 2.1.
Values of $a_0, a_1, a_2, \mu$ for some continuous distributions

<table>
<thead>
<tr>
<th>Model</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma $\frac{\alpha^\beta x^{\alpha-1} e^{\beta x}}{\Gamma\beta}$ (exponential for $\beta=1$)</td>
<td>0</td>
<td>$\alpha^{-1}$</td>
<td>0</td>
<td>$\beta \alpha^{-1}$</td>
</tr>
<tr>
<td>Beta $\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}$</td>
<td>0</td>
<td>$(p+q)^{-1}$</td>
<td>$-(p+q)^{-1}$</td>
<td>$p(p+q)^{-1}$</td>
</tr>
<tr>
<td>Lomax $\alpha c^{c(x+\alpha)-(c+1)}$</td>
<td>0</td>
<td>$\alpha(c-1)^{-1}$</td>
<td>$(c-1)^{-1}$</td>
<td>$\alpha(c-1)^{-1}$</td>
</tr>
<tr>
<td>Pareto $ak^a x^{-a-1}, x \geq k$</td>
<td>0</td>
<td>$-k(a-1)^{-1}$</td>
<td>$(a-1)^{-1}$</td>
<td>$ak(a-1)^{-1}$</td>
</tr>
<tr>
<td>Finite range $\frac{\alpha}{R}(1- \frac{x}{R}); 0 &lt; x &lt; R$</td>
<td>0</td>
<td>$R(d+1)^{-1}$</td>
<td>$-(d+1)^{-1}$</td>
<td>$R(d+1)^{-1}$</td>
</tr>
<tr>
<td>Normal $\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$</td>
<td>$\sigma^2$</td>
<td>0</td>
<td>0</td>
<td>$\mu$</td>
</tr>
</tbody>
</table>
| Student's $t$ $\frac{1}{\sqrt{\nu} B(\frac{\nu}{2}, \frac{\nu}{2})}$ | $(1+\frac{x}{\nu})^{-\frac{(v+1)}{2}}$ | $v(v-1)^{-1}$ | 0     | $(v-1)^{-1}$ | 0
Table 2.2.
Values of $a_0, a_1, a_2$ and $\mu$ for discrete distributions

<table>
<thead>
<tr>
<th>Model</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson $e^{-\frac{p}{x!}}$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$p$</td>
</tr>
<tr>
<td>Binomial $(\begin{pmatrix} n \ x \end{pmatrix})p^xq^{n-x}$</td>
<td>$q$</td>
<td>$q$</td>
<td>0</td>
<td>np</td>
</tr>
<tr>
<td>Negative $(x-1)p^r(x-r)^{(l-r)p-1}$</td>
<td>$p-1$</td>
<td>0</td>
<td>np-1</td>
<td></td>
</tr>
<tr>
<td>Binomial $(r-1)p^rq^{n-x}$</td>
<td>$(N-D-n+1)N^{-1}$</td>
<td>$(N-D-n+2)N^{-1}$</td>
<td>$N-1$</td>
<td>$nDN^{-1}$</td>
</tr>
<tr>
<td>Hyper geometric $N^{-1}$</td>
<td>$N^{-1}$</td>
<td>$nDN^{-1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waring $(a-b)(b)^{x}_{(a)}x+1$</td>
<td>$(a+1)(a-b-1)^{-1}$</td>
<td>$(a+2)(a-b-1)^{-1}$</td>
<td>$(a-b-1)^{-1}$</td>
<td>$b(a-b-1)^{-1}$</td>
</tr>
<tr>
<td>Negative hyper geometric</td>
<td>$N^{-1}$</td>
<td>$nDN^{-1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta Pascal $N^{-1}$</td>
<td>$N^{-1}$</td>
<td>$nDN^{-1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{(k+x-1)A+B-1}{(A+k)}$</td>
<td>$A+B+k$</td>
<td>$A+B+k+1$</td>
<td>$\frac{1}{A-1}$</td>
<td>$\frac{AB}{A-1}$</td>
</tr>
</tbody>
</table>
It appears to be in order to observe that the above theorems are of importance in reliability modeling in consideration of the following aspects.

1. Many distributions belonging to the two families such as gamma, beta, normal, hypergeometric and binomial do not have simple closed form expressions either for the failure rate or for the MRL function to extract a useful identity connecting the two. The theorems provides such expressions.

2. Extra flexibility is imparted in the choice of the model, as one can use the system as the basis of the model and then select a particular member depending on the values of the constants $d, a_0, a_1, a_2$ dictated by physical considerations or empirical evidence.

3. Not only many models that are extensively used in reliability analysis belong to the systems, but as an inherent property of the system, their truncated versions also are members. This helps when the data is truncated from the left or right.
2.4 Length biased models

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and \(X: \Omega \to \mathbb{R}\) be a random variable where \(\mathbb{R} = (a, b)\) is the subset of the real line with \(a > 0\) and \(b > a\) can be finite or infinite. The distribution function \(F(x)\) is assumed to be absolutely continuous with respect to Lebesgue measure with probability density function \(f(x)\) and \(w(x)\) is a non-negative function of \(X\) such that \(\mu = Ew(X) < \infty\). The random variable \(Y\) with probability density function

\[
g(x) = \frac{w(x)f(x)}{\mu}, \quad x > 0
\]

is said to have a weighted distribution associated with \(X\). When the counting measure is employed instead of the Lebesgue measure, the same equation (2.19) holds for the weighted distribution in the discrete case for \(x = 1, 2, \ldots\), where \(g(x)\) and \(f(x)\) obviously are interpreted as the respective probability mass functions.

While, different weight functions such as \(x^\alpha (\alpha > 0)\), \(e^{\alpha x}\) etc in the continuous case and again \(x^\alpha\), \(1 - (1 - \alpha)x\) (\(0 < \alpha < 1\)), \(x+1\), \(t^x\) etc. when \(X\) is discrete have been used by various researchers, the simplest and most extensively studied form appears to be \(x\). In this
situation (2.19) specialises to
\[ g(x) = \frac{xf(x)}{\mu}, \quad x > 0 \] (2.20)
and \( \mu = E(X) \). The version (2.20), often called the length biased distribution corresponding to \( X \), will be the focus of attention in the rest of this chapter.

Although Rao (1965) introduced distributions of type (2.19) and cited practical examples where \( w(x) = x \) or \( x^\alpha \) are appropriate, the form (2.20) has found a place much earlier in the discussions given in Cox (1962) relating to renewal theory. Instead of the usual practice in random sampling of selecting units from the population, with probability of selection of each unit the same, regardless of the values of \( x \) it carries, Cox (1962) perceived the idea that from a population of failure times distributed according to \( f(x) \), the selection of any unit in the population is proportional to its length (or size), the random variable \( Y \) which is the failure time of the component whose life falls in the sample, has probability density function (2.20). This explains the terminology length biased distribution to such a model. It seems however that the same idea has originally been conceived much before as evidenced from Daniels (1942) who discusses length biased sampling in the
analysis of the distribution of fibre lengths in wool. Practical problems where length biased models arise in a natural way include the analysis of a family size and in the study of albinism (Rao, 1965), family data in human heredity (Neel and Schull, 1966), aerial survey and visibility bias (Cook and Martin, 1974), forest disease and line transcend sampling (Patil and Rao, 1977), renewal theory (Cox, 1962), cell cycle analysis and pulse labelling (Takahashi, 1966) and efficacy of family screening for disease (Zelen, 1974). An exhaustive account of the research in this area is available in Patil and Rao (1977) and in Gupta and Kirmani (1990).

There are many situations when an investigator collects observations from real world phenomena and such data may not reproduce the original distribution believed to be true. Since the characteristics of the original distribution are the object of inference, one has to look into the structural relationship existing between the original model and the one that is realised in practice. This is especially important when length biased sampling is resorted to in drawing observations from the population. The paper by Gupta and Kirmani (1990) address this question and they develop several relationships that are of
relevance to reliability analysis concerning the random variables X and Y. If G(y), k(y) and s(y) represent respectively the survival function, failure rate and MRL of Y, they show that

\[ G(x) = \frac{m(x)}{\mu} R(x) \]  
\[ k(x) = \frac{x}{m(x)} h(x) \]  
\[ s(x) = \frac{r(x)}{m(x)} \int_{x}^{\infty} \frac{t+r(t)}{r(t)} \exp[- \int_{x}^{t} \frac{du}{r(u)}} dt. \]

The above identities along with some characterization theorems cited in Gupta and Kirmani (1990) show how length biased sampling affects the original distribution and how the corresponding reliability characteristics change under such a scheme of sampling. While comparing the distribution under length biased sampling with the parent model, it will be of some definite advantage if the original distribution keeps the same form under length biased sampling also, except possibly for a change in the parameters. In the next section, we prove a general theorem in this direction by identifying those distributions of X belonging to the Pearson family that retain the same form of the distribution of Y. Since the
parameters of the distribution can often be interpreted in terms of the population characteristics this would mean that, a theorem of this kind provides a tool to ascertain the changes in such characteristics as a result of length biased sampling.

2.4.1. A Closure Property

When the distribution of $Y$ is of the same form as that of $X$, we say that the distribution of $X$ is closed with respect to length biased sampling. Retaining the notations of Section 2.2, we investigate the conditions under which the family in equation (2.1) induce the closure property. A major distinction made in the present section from the previous discussion of the Pearson family is that, the discussion is confined now only to distributions of non-negative random variables.

**Theorem 2.3.**

Among the members of (2.1) with $b_2 \neq 1$, in the support of the real line having the subset $(a,b)$, $a \geq 0$ and $b > a$ can be finite or infinite, $X$ and $Y$ have the same type of distribution if and only if $b_0 = 0$ and the probability density function of $Y$ satisfies
where, 
\[ c_i = \frac{b_i}{1-b_2}, \quad i = 1, 2. \]
and 
\[ d_1 = \frac{d-b_1}{1-b_2}. \]

Proof.

From (2.20), we have

\[ \frac{1}{g(x)} \frac{dg(x)}{dx} = \frac{1}{f(x)} \frac{df(x)}{dx} + \frac{1}{x}. \] (2.25)

First, suppose that \( X \) belongs to the family (2.1) and that \( X \) and \( Y \) have the same type of distributions. Since \( Y \) also must belong to the Pearson family, the equation (2.25) leads to

\[ \frac{1}{(1+c_2) + c_1 x + c_2 x^2} = \frac{x+d}{\frac{b_0}{x} + \frac{b_1 x + b_2 x^2}{x}} \]

This identity is satisfied if and only if the following conditions hold.
The three different possibilities arising out of condition (2.26) are

(i) \( b_0 \neq 0 \) and \( c_0 = 0 \)

(ii) \( b_0 = 0 \) and \( c_0 \neq 0 \)

(iii) \( b_0 = 0 \) and \( c_0 = 0 \).

When \( b_0 \neq 0 \) and \( c_0 = 0 \), from (2.28)

\[
c_1 = -d_1 = b_0 / d(1-b_2),
\]

while,

\[
b_0(l-c_2)+dd_1 = 0,
\]

from (2.29) along with (2.27), leave the equation,

\[
b_0-b_1d + b_2d^2 = 0.
\]

Thus \( (x+d) \) is a factor of \( (b_0+b_1x+b_2x^2) \) and therefore from (2.1)
Thus the distribution of $X$ will be of the form

$$f(x) = c \alpha^c (x+a)^{-(c+1)}$$

(2.32)

with

$$c = \frac{1-b_2}{b_2^2}$$

(2.34)

For $c > 0$, $\alpha > 0$ (2.32) represents the Pareto II model. From (2.24) we find in this case that,

$$g(x) = c(c-1)\alpha^{c-1} x(x+a)^{-(c+1)}$$

which does not have the same form of density as $X$.

In the second case, when $b_0 = 0$ and $c_0 \neq 0$, we have a situation parallel to case (i) and the densities of $X$ and $Y$ do not have identical forms. Thus the first two cases lead to inadmissible solutions. This leaves us to examine the third possibility $c_0 = 0$ and $b_0 = 0$. 

These values when inserted into equations (2.22) and (2.23) result in

$$\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{-1}{b_2(x+a)}$$

(2.31)

where,

$$\alpha = \frac{b_1-b_2 d}{b_2}$$

Thus after using (2.27) in (2.31), the resulting values are

$$f(x) = c \alpha^c (x+a)^{-(c+1)}$$

(2.32)
These values when inserted to equations (2.29) and (2.30) result in

\[ b_1d_1 = (d-b_1)c_1 \]

and

\[ b_1+b_2d_1 = c_1(1-b_2)+(d-b_1)c_2 \]

Solving the last two equations for \( c_1 \) and \( d_1 \) after using (2.27) in (2.31), the resulting values are

\[ d_1 = \frac{d-b_1}{1-b_2} \] \hspace{1cm} (2.33)

and

\[ c_i = \frac{b_i}{1-b_2} \hspace{0.2cm}, \hspace{0.2cm} i = 1,2. \] \hspace{1cm} (2.34)

The various models generated by the above two equations depend on the nature of the roots of the equations

\[ c_1x + c_2x^2 = 0 \] and \( b_1x + b_2x^2 = 0 \). The roots of the first equation are 0 and \(-c_1/c_2\), while that of the second are 0 and \(-b_1/b_2\). However, since \( c_1/c_2 = b_1/b_2 \) from (2.34), the two roots of these equations have identical nature and therefore \( X \) and \( Y \) have the same type of distribution with possibly different parameters. The change in parameters are governed by equation (2.33) and (2.34).
Conversely, suppose that $b_0 = 0$ in (2.1). Then from (2.24) and (2.25)
\[
\frac{1}{f(x)} \frac{df(x)}{dx} = -\frac{x+d_1}{c_1x+c_2x^2} - \frac{1}{x}
\]
\[
= -\frac{(1+c_2)x+d_1+c_1}{c_1x+c_2x^2}
\]
\[
= -\frac{x+d}{b_1x+b_2x^2}
\]
where, $d = \frac{d_1+c_1}{1+c_2}$ and $b_i = \frac{c_i}{1+c_2}$, $i = 1, 2$.

This completes the proof.

The idea of form-invariant length biased distributions have been discussed earlier in an investigation by Patil and Ord who shows that a necessary and sufficient condition for a distribution to be closed with respect to formation of length biased distribution is that it must belong to the log exponential family. The result in Theorem 2.3 does not provide any new model that is not presented in the log exponential family. The major
difference between the two investigation is that our study is confined to the Pearson family and utilises a different approach with application pointed towards reliability analysis. The result will be used in the next section to characterize some important distributions belonging to the sub-class of the Pearson family defined in the differential equation (2.35). This sub-class, to be denoted by \( C \), contains the beta distribution of the first kind and second kind and their translations (by the transformation \( Z = \alpha X \)), the gamma, the inverted gamma and the Pareto type I. The exponential, Pareto type II and finite range models discussed earlier have no form-invariant structure on their own, but when regarded as special cases of the above mentioned families, the same property can be attributed to them.

2.4.2. Characterization of form-invariant length biased distributions by reliability concepts.

In the light of Theorem 2.1 reflecting the relationship between vitality function and failure rate for the Pearson family, it is possible to achieve several characterizations of the class of models in \( C \).
Theorem 2.4.

If \( \lim_{x \to b} (b_1x + b_2x^2) f(x) = 0 \), the probability density function \( f(x) \) belongs to \( C \) if and only if

\[
k(x) = \frac{xh(x)}{\mu + (a_1x + a_2x^2)h(x)}
\]

where,

\[
a_i = \frac{b_i}{1 - 2b_i}, \quad i = 1, 2.
\]

Proof:

When \( X \) belongs to \( C \), from the Theorem 2.1, \( m(x) \) and \( h(x) \) can be related as

\[
m(x) = \mu + (a_1x + a_2x^2)h(x).
\]  

(2.37)

From (2.22) and (2.37) we recover (2.36). The only if part follows from the equations (2.22) and (2.37) and the Theorem 2.1.

Theorem 2.5.

Under the conditions of Theorem 2.4, \( f(x) \) belongs to \( C \) if and only if

\[
v(x) - m_2 = \frac{(m(x) - \mu)(1 - 2b_2)x}{m(x)(1 - 3b_2)}
\]  

(2.38)
where,

\[ v(x) = E(Y|Y>x) \]  and  \[ m_2 = E(Y). \]

Proof:

Suppose that \( f(x) \) belongs to \( C \). Then from (2.37)

\[ m(x) - \mu = (a_1x + a_2x^2)h(x). \]

(2.39)

And similarly,

\[ v(x) - m_2 = (q_1x + q_2x^2)k(x) \]

(2.40)

where,

\[ q_i = b_i/(1 - 2b_2) \]

\[ = a_i/(1 - 3a_2), \quad i = 1, 2. \]

Eliminating \( k(x) \) and \( h(x) \) using the equations (2.39) and (2.40) we obtain (2.38). The only if part results from retracing the above steps.

There is an elegant relationship that characterize the Pareto I law among the class of all absolutely continuous distributions with non-negative support. In the following, we denote that MRL function of \( X \) by \( r(x) \) and that of \( Y \) by \( s(x) \).
Theorem 2.6.

For a continuous non-negative random variable with \( E(X^2) < \infty \), \( s(x) = kr(x) \) for \( k > 1 \) if and only if \( X \) has probability density function

\[
f(x) = ak^a x^{-(a+1)} \quad x \geq k > 0.
\]  

(2.41)

Proof:

We find from Gupta and Kirmani (1990) that

\[
k(x) = xh(x)/(x+r(x))
\]

and hence

\[
\frac{1+s'(x)}{s(x)} = \frac{x(1+r'(x))}{r(x)[r(x)+x]} \quad (2.42)
\]

where, the primes denote differentiation.

Substituting \( s(x) = kr(x) \) in the equation (2.42), we get,

\[
r(x) + kr'(x) r(x) = x(k-1).
\]

Accordingly \( r(x) \) must be linear and the only solution is \( r(x) = x/(a-1) \) with \( a = (2k-1)/(k-1) \). In the Pareto case, \( r(x) = x/a-1 \) and \( s(x) = x/a-2 \), and therefore, the condition of Theorem 2.6 is verified.
Corollary:

For a continuous non-negative random variable with $E(X) < \infty$, $k(x) r(x) = 1$ if and only if $X$ has probability density function (2.41).

2.5. Discrete Length Biased Distribution

Analogous to the continuous case, the length biased distribution of a discrete random variable $X$ with the set of non-negative integers as the support is defined as (Gupta, 1979),

$$g(x) = \frac{x f(x)}{\mu} \quad x = 1, 2, ... \quad (2.43)$$

where, $\mu = E(X) < \infty$. Clearly, the above random variable $Y$ will have no zero in its support. Applying a displacement of $Y$ to the left, by taking $Z = Y - 1$, $Z$ would be realized by length biased sampling on $X$ with the above displacement and the support becomes the set of non-negative integers (see Patil and Ord, 1976). The resulting probability mass function of $Z$ is

$$p(x) = g(x+1) \text{ for } x = 0, 1, 2, ...$$

With the notations of Section 2.3, we investigate the conditions under which the family in equation (2.14) induce the closure property.
Theorem 2.7.

Among the members of the family (2.14) with $b_2 \neq 1$, in the support of non-negative integers, $X$ and $Z$ have the same type of distribution if and only if either

1) $b_0 = d$ and the distribution of $Z$ satisfies

$$
\frac{g(x+1)-g(x)}{g(x)} = -\frac{x+d_1}{c_0+c_1x+c_2x^2},
$$

(2.44)

with

$$
d_1 = \frac{1+d-b_1}{1-b_2}, \quad c_1 = \frac{b_i}{1-b_2}, \quad i=0,1,2.;
$$

or

2) $b_1 = b_0 + b_2$ and the distribution of $Z$ satisfies

$$
\frac{g(x+1)-g(x)}{g(x)} = -\frac{(x+d_1)}{x(c_1+c_2x)}
$$

(2.45)

with

$$
c_1 = \frac{b_0}{1-b_2}, \quad c_2 = \frac{b_2}{1-b_2} \quad \text{and} \quad d_1 = \frac{d-b_0}{1-b_2}.
$$

Proof:

From the definition,

$$
\frac{g(x+1)-g(x)}{g(x)} = \frac{(x+1)f(x+1) - xf(x)}{xf(x)}
$$

(2.46)
First, the necessary part will be proved. For this, suppose that $Z$ satisfies
\[ \frac{g(x+1)-g(x)}{g(x)} = -\frac{x+d_1}{c_0+c_1x+c_2x^2}. \] (2.47)

From (2.14) and (2.47), using (2.46), we have
\[ -\frac{(x+d_1)}{c_0+c_1x+c_2x^2} = \frac{(b_0-d)+(b_1-d-1)x+(b_2-1)x^2}{(b_0+b_1x+b_2x^2)x}. \]

That is,
\[ -(x^2+dx)(b_0+b_1x+b_2x^2) = [(b_2-1)x^2+(b_1-1-d)x+(b_0-d)](c_0+c_1x+c_2x^2). \]

Equating coefficients on either side, the following conditions hold.

\[ b_2 = c_2(l-b_2). \] (2.48)
\[ c_2(b_1-1-d)+c_1(b_2-1) = -b_2d_1-b_1. \] (2.49)
\[ c_2(b_0-d)+c_1(b_1-1-d)+c_0(b_2-1) = -b_0b_1d_1. \] (2.50)
\[ c_1(b_0-d)+c_0(b_1-1-d) = -b_0d_1. \] (2.51)
\[ c_0(b_0-d) = 0. \] (2.52)
The three different possibilities arising out of the last equation are

i) \( b_0 = d \) and \( c_0 \neq 0 \),

ii) \( b_0 \neq d \) and \( c_0 = 0 \),

and

iii) \( b_0 = d \) and \( c_0 = 0 \).

When \( b_0 = d \) and \( c_0 \neq 0 \),

\[
\frac{f(x+1)-f(x)}{f(x)} = -\frac{x+d}{d+b_1x+b_2x^2}
\]

and

\[
\frac{g(x+1)-g(x)}{g(x)} = \frac{(b_2-1)x+b_1-1-d}{b_0+b_1x+b_2x^2}
\]

\[
= -\frac{x+d_1}{c_0+c_1x+c_2x^2}
\]

where,

\[
c_i = b_i/(1-b_2), \quad i = 0,1,2.
\]

and

\[
d_1 = (1+d-b_2)/(1-b_2).
\]

The roots of \( c_0 + c_1x + c_2x^2 = 0 \) are,

\[
-\frac{c_1 \pm \sqrt{c_1^2-4c_0c_2}}{2c_2} = -\frac{b_1 \pm \sqrt{b_1^2-4b_0b_2}}{2b_2},
\]

equal to the
roots of \( b_0 + b_1 x + b_2 x^2 = 0 \) by virtue of the relationship mentioned above between \( c_i \) and \( b_i \) and the two equations therefore produce roots of identical nature. Hence \( X \) and \( Z \) have form-invariant distributions.

In the second case, substituting \( c_0 = 0 \) in equations (2.48), (2.49) and (2.51), the following equalities result.

\[
c_1 = \frac{b_0 (b_2 d + b_2 - b_1)}{(b_2 d - b_0)(1 - b_2)}
\]

\[
c_2 = \frac{b_2}{(1 - b_2)}
\]

\[
d_1 = \frac{(b_2 d + b_2 - b_1)(d - b_0)}{(b_2 d - b_0)(1 - b_2)}
\]

Substituting these values in the equation (2.48) we have,

\[
(b_1 - b_0 - b_2)(b_2 d^2 - b_1 d - b_1 d + b_0) = 0 .
\] (2.53)

Thus, whenever \( b_1 - b_0 - b_2 \neq 0 \), \( (x + d) \) is a factor of \( (b_0 + b_1 x + b_2 x^2) \) and then

\[
\frac{f(x+1) - f(x)}{f(x)} = -\frac{1}{b_2 x + b_1 - b_2 d}
\]
and
\[ \frac{f(x+1)}{f(x)} = \frac{x+a}{x+b+1} \]

where,
\[ a = \frac{(b_1-b_2d-1)}{b_2} \]

and
\[ b = \frac{(b_1-b_2d-b_2)}{b_2} \].

Solving the above equation,
\[ f(x) = k(a)^x/(b)^{x+1} \]

with \( k \) as the normalising constant, and

\[ (a)^x = a(a+1)...(a+x-1). \]

When \( b > a \) or \( 0 < b_2 < 1 \), the distribution of \( X \) is Waring with
\[ f(x) = \frac{(b-a)(a)^x}{(b)^{x+1}}, \quad x = 0, 1, 2, ... \]  \hspace{1cm} (2.54)

but, the distribution of \( Y \) is
\[ g(x) = \frac{(b-a)(b-a-1) x(a)^x}{a(b)^{x+1}}, \quad x = 1, 2, ... \]  \hspace{1cm} (2.58)

and that of \( Z \) is
Thus $X$ and $Z$ are not form-invariant. The remaining case arising out of equation (2.53) is, when $b_1 = b_0 + b_2$,

$$d_1 = \frac{d-b_0}{1-b_2},$$

$$c_2 = \frac{b_2}{1-b_2},$$

and

$$c_1 = \frac{b_0}{1-b_2}.$$}

In this case,

$$\frac{f(x+1)-f(x)}{f(x)} = -\frac{(x+d)}{(x+1)(b_0+b_2x)}$$

(2.56)

while,

$$\frac{g(x+1)-g(x)}{g(x)} = -\frac{(x+d_1)}{x(c_1+c_2x)}.$$ (2.57)

The distribution of $Z$ satisfies

$$\frac{p(x+1)-p(x)}{p(x)} = -\frac{(x+d_1+1)}{(x+1)(c_1+c_2+c_2x)}.$$ (2.58)

The roots of $(x+1)(b_0+b_2x)=0$ are $x=-1$ and $x=-b_0/b_2$

and those of $(x+1)(c_1+c_2+c_2x)=0$ are $x=-1$ and
\[ x = \frac{-c_1 + c_2}{c_2} = -\frac{b_1}{b_2}. \]

Hence, when \( b_1 = b_0 \), the distribution of \( X \) is closed with respect to the formation of the distribution of \( Z \).

On the other hand if \( b_0 = d \) and \( c_0 = 0 \) as in case (iii), we have again three different cases, namely

(a) \( b_0 = 0 \) and \( d_1 \neq 0 \)

(b) \( b_0 \neq 0 \) and \( d_1 = 0 \)

and (c) \( b_0 = 0 \) and \( d_1 = 0 \).

Considering case (a), we must have

\[ \frac{f(x+1) - f(x)}{f(x)} = -\frac{1}{b_1 + b_2 x} \]

and

\[ f(x) = k_1 \frac{(m)_x}{(n)_x+1} \]

with \( k_1 \) as the normalising constant, \( m = \frac{b_1 - 1}{b_2} \) and

\[ n = \frac{b_1 - b_2}{b_2}. \]

Notice that if \( n > m \) or \( 0 < b_2 < 1 \), the distribution of \( X \) will be Waring with probability mass function

\[ f(x) = \frac{(n-m)(m)_x}{(n)_x+1} x, \quad x = 0, 1, 2, \ldots. \]
The corresponding length biased distribution, specified by

\[ g(x) = \frac{(n-m)(n-m-1)}{m} \frac{x(a)_x}{(b)_{x+1}}, \quad x = 1, 2, \ldots \]

or

\[ p(x) = \frac{(n-m)(n-m-1)}{m} \frac{(x+1)(a)_{x+1}}{(b)_{x+2}}, \quad x = 0, 1, 2, \ldots \]

is not form-invariant.

When \( b_0 \neq 0 \) and \( d_1 = 0 \), we have a parallel result that the distribution of \( Z \) is Waring, but that of \( X \) has form (2.55).

Finally, the values \( b_0 = 0 \) and \( d_1 = 0 \) provide us

\[ c_1 (b_1 - 1) = 0. \]

This implies either \( c_1 = 0 \) or \( b_1 = 1 \).

If \( c_1 = 0 \), from (2.48) and (2.49), we should have \( b_1 = b_2 \) or \( b_1 = b_0 + b_2 \); the cases already discussed leading to form-invariance.

When \( b_1 = 1 \), from (2.48) and (2.49)

\[ c_2 = \frac{b_2}{1 - b_2} \]

This completes the proof.
and 
\[ c_1 = \frac{1}{1-b_2}. \]

In this case,
\[ \frac{f(x+1)-f(x)}{f(x)} = -\frac{1}{b_2x+1} \]

and the distribution of \( X \) has the form (2.54).

Accordingly, the distribution of \( Z \) is not identical with that of \( X \).

Conversely, when \( b_0=d \) and \( g(x) \) satisfies (2.44).

Then,
\[ \frac{f(x+1)-f(x)}{f(x)} = -\frac{x+d}{b_2x^2+b_1x+b_0} \]

Thus \( X \) and \( Z \) have form-invariant and \( g(x) \) satisfies

where,

\[ b_i = \frac{c_i}{1+c_2}, \quad i = 0,1,2. \]

When \( b_2=b_1+b_0 \) and \( g(x) \) satisfies (2.45),
\[ \frac{f(x+1)-f(x)}{f(x)} = -\frac{x+d}{(x+1)(b_0+b_2x)} \]

where,

\[ b_0 = \frac{c_1}{1+c_2} \]
\[ b_2 = \frac{c_2}{1+c_2} \]
\[ d = \frac{(d_1+c_1)}{(1+c_2)}. \]

This completes the proof.
We give two examples to illustrate the results of the theorem.

1. For the negative binomial distribution with probability mass function

\[ f(x) = \binom{x-1}{r-1} p^r q^{x-r} \quad x=r, r+1, \ldots, \]

The survival function \( G(x) \), corresponding to (2.43) is given by

\[ G(x) = \frac{m(x-1)}{m(x)} R(x) \]

(2.59)

and therefore, \( b_0 = d = \frac{1-r}{1-q} \), \( b_1 = \frac{1}{1-q} \) and \( b_2 = 0 \).

Thus \( X \) and \( Z \) have form-invariant and \( g(x) \) satisfies

\[ \frac{g(x+1)-g(x)}{g(x)} = -\frac{x+(1-r)/(1-q)}{(x/1-q)+(1-r)/(1-q)} \]

(2.60)

2. When \( X \) is Poisson,

\[ f(x+1)-f(x) \quad \frac{f(x+1)}{f(x)} = -\frac{x+1-\lambda}{x+1} \]

Since \( b_1 = b_0 + b_2 \), with \( b_2 = 0 \), the distribution of \( X \) is closed with respect to the length biased distribution and probability mass function is given by

\[ g(x+1) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \ldots \]
Besides the above two models some others belonging to the family that satisfy the closure property are the binomial, the hypergeometric, the binomial beta and the beta Pascal.

2.5.1. Characterizations of discrete life distributions

The survival function $G(x)$, corresponding to (2.43) is given by

$$G(x) = \frac{m(x-1) R(x)}{\mu}$$

(2.59)

where, $m(x)$ is the vitality function of $X$. Then the survival function $P(x)$ of $Z$ is specified by

$$P(x) = \frac{m(x) R(x+1)}{\mu}$$

(2.60)

Denoting by $u(x)$ and $e(x)$ the failure rate and the MRL function of $Z$, we have,

$$u(x) = \frac{(x+1) h(x+1)}{m(x)}$$

(2.61)

and

$$e(x-1) = \frac{\sum_{t=0}^{x} R(t+1)m(t)}{R(x+1)m(x)}$$

(2.62)
In analogy with the continuous case, the above identities connecting reliability characteristics of \( X \) and \( Z \) can be employed in the characterization of the distribution of \( X \).

Theorem 2.8.

The probability mass function \( f(x) \) satisfies (2.14) if and only if

\[
u(x) = \frac{(x+1)h(x+1)}{\mu+(a_0+a_1x+a_2x^2)h(x+1)}.
\]

(2.63)

Proof:

When \( f(x) \) satisfies (2.14), from Theorem 2.2

\[
m(x) = \mu+(a_0+a_1x+a_2x^2)h(x+1),
\]

(2.66)

Substituting the above identity in the equation (2.61) we have (2.63). The only if part follows from (2.61) and (2.63) along with Theorem 2.2.

Theorem 2.9.

The probability mass function \( f(x) \) satisfies (2.14) if and only if

\[
v(x)-m_2 = \frac{(x+1)(m(x)-\mu)(1-2b_2)}{m(x)(1-3b_2^2)}.
\]

(2.64)
where,
\[ v(x) = E[Y|Y > x] \]
and \[ m_2 = E(Y). \]

Proof:
When \( f(x) \) belongs to the family (2.14)
\[ m(x)-\mu = (a_0 + a_1x + a_2x^2) \cdot h(x+1). \tag{2.65} \]
Similarly,
\[ v(x)-m_2 = (q_0 + q_1x + q_2x^2) \cdot u(x), \tag{2.66} \]
where,
\[ q_i = b_i/1-3b_2, \quad i = 0,1,2,... \]
Dividing (2.65) by (2.66) and simplifying the resulting identity, using (2.61), we have (2.64). The converse part of the theorem follows by retracing the above steps.

As a point of departure from the family specified by Theorem 2.7, some characterization will be presented associated with distributions which are not its members. The geometric, Waring and negative hypergeometric distributions are shown to be unique models from the
class of all discrete distributions with non-negative integers as support satisfying certain simple properties.

Theorem 2.10.

Let \( X \) be a non-negative integer valued random variable. Then a necessary and sufficient condition that

\[
\frac{P(x-1)}{R(x)} = 1 + dx
\]  

is that the distribution of \( X \) is

(i) geometric, \( G(p) \) with probability mass function

\[
f(x) = pq^x, \quad x = 0, 1, 2, \ldots
\]  

when \( \mu_d - 1 = 0 \)

(ii) Waring, \( W(a, b) \) specified by

\[
f(x) = \frac{(b-a)(a)_x}{(b)_x} \left( 1 + \frac{k}{n+x} \right), \quad x = 0, 1, 2, \ldots
\]  

when \( \mu_d - 1 > 0 \)

(iii) negative hypergeometric distribution, \( H(k, n) \) with

\[
f(x) = \binom{-1}{x}(-k)_{n-x} / \binom{-1-k}{n} \quad x = 0, 1, 2, \ldots
\]  

when \( \mu_d - 1 < 0 \).
Proof:

When (2.67) holds, from the equation (2.60)

\[ m(x-1) = \mu + \mu dx \]

and consequently

\[ r(x-1) = (\mu d - 1)x + \mu - l. \]  \hspace{1cm} (2.71)

The MRL function of \( X \) of the form \( a_1 + b_1 x \) characterizes the distributions (2.68), (2.69) and (2.70) (see Nair and Hitha (1989)).

Conversely, when \( X \) is \( G(p) \) \((W(a,b); H(k,n))\)

\[ \frac{P(x-1)}{R(x)} = (1 + \frac{b-a}{a} x; 1 + \frac{k}{n+k} x) \]

This completes the proof.

Theorem 2.11.

For the random variable in Theorem 2.10, the relationship

\[ u(x) = \frac{(x+1)h(x+1)}{d(l+cx)} \] \hspace{1cm} (2.72)

holds if and only if \( X \) is \( G(p) \) \((W(a,b) \) or \( H(k,n))\) according as \( \text{dc} - \text{l} = 0 \) (\( \text{dc} - \text{l} > 0 \) or \( \text{dc} - \text{l} < 0 \)).
Proof:

When (2.72) holds, from the equation (2.61),

\[ m(x) = d + dc x , \]

which leads to

\[ r(x) = d + (dc - 1)x . \]

The remaining part of the proof will follow from Nair and Hitha (1989).

The results in sections 2.4 and 2.5 are reported in Sankaran and Nair (1992d).