CHAPTER 1.

NOTATION, TERMINOLOGY AND REVIEW OF THE LITERATURE.

1.0 General Introduction

Characterization problems are located on the borderlines of scientific modelling probability theory and mathematical statistics. This should not lead us to doubt the appropriateness of characterization theorems in the framework of mathematical statistics and its development. Quite to the contrary, in many problems of mathematical statistics we try to transfer the original problem to an equivalent but substantially simpler one by using important properties that certain special distributions possess. The question of how to make full use of the special nature of the parent distribution leads to the study of the characteristic properties of the distribution used in mathematical statistics, and hence to characterization theorems.

As far as the argument that characterization theorems have not offered up to now solutions to real life problems is concerned, one can argue that there are a lot of mathematical ideas, which, when introduced, did not seem to be of any practical use but later became necessary in solving applied problems. However, the main contribution of the characterization problems to statistics was that, since mathematical analysis was required in solving them, they attracted the attention of numerous mathematicians and thus provided the links between mathematics and statistics necessary to justify some statistical arguments and to give proper scientific solutions to practical problems.

Many characterizations of probability distributions have been derived based on various properties, such as independence, order statistics, admissibility and optimality of certain estimation, linearity of regression, etc.

Many of these results can be found in Kagan, Linnik and Rao (1973) as well as in Laha and Lucacs (1964), Kotz (1974) and Lucacs (1956, 1960, 1960-61, 1965). The introduction of the damage model by Rao in (1963) opened the way for a number of characterizations of probability distributions, mostly discrete.

Rao himself, with Rubin (1964), gave a characterization of the Poisson distribution based on the damage model.

In this thesis, we use an extension of the R-R characterization suggested by Shanbhag (1976) to characterize a number of discrete distributions.

Another general form of characterization is subsequently introduced; this enables us to obtain characterizations for finite discrete distributions. Extensions to Bivariate and Multivariate cases of the above result are derived, along with their truncated versions.

Later on in the thesis we apply the damage model theory to certain mixtures of distributions, and make use of the results in order to establish a number of interesting characterizations of distributions and families of distributions.

1.1 Notation and Terminology

Throughout this thesis the following notation will be used.

Capital X, Y and Z will denote random variables (r.v.'s) (these will usually be discrete r.v.'s); P_n stands for the probabilities $P(X=n)$, $n=0,1,2,\ldots$ This notation is extended to the Multivariate case as $P_{\underset{\sim}{n}}$, where $\underset{\sim}{n}=(n_1,\ldots,n_s)$ and $P_{\underset{\sim}{n}} = P(\underset{\sim}{X}=\underset{\sim}{n}) = P(X_1=n_1,\ldots,X_s=n_s)$. In other words $P_{\underset{\sim}{n}}$ denotes the joint probability of $X_1,\ldots,X_s$.

The notation $Y|X$ will be used to denote the conditional r.v. $Y|(X=n)$.

$F(x)$ will represent the distribution function (d.f.) of the r.v. X.

For the probability generating function (p.g.f.) of the r.v. X we will use either $G_X(t)$ or $G(t)$. $G_{\underset{\sim}{X}}(t)$ corresponds to the p.g.f. of the random vector $\underset{\sim}{X}=(X_1,\ldots,X_s)$.

The r-th moment of the r.v. X will be denoted by $\mu_r \equiv \mu_r(x)$, while $\mu_{[r]}(x) \equiv \mu_{[r]}$ will denote the r-th factorial moment of X. As to the generating functions, $M_X(t)$ will represent the moment generating function (m.g.f.) and $M_{[X]}(t)$ the factorial moment generating function (f.m.g.f.).

The distribution

$$F(x) = \int_\theta F_1(x,\theta)dF_2(\theta) ,$$

where θ is the parameter of F_1, will be called "mixed distribution" (alternatively "compound distribution").

The notation $F = F_1 \wedge F_2$ for the resultant distribution will be adopted. We will write $F = F_1 \underset{\theta}{\wedge} F_2$ when we want to specify the parameter of the distribution F_1 over which the mixing is taking place.

We will say that the r.v. X follows a generalized distribution if its p.g.f. is of the form

$$G(t) = G_1(G_2(t))$$

where $G_1(t)$ and $G_2(t)$ are both valid p.g.f.'s.

The Laplace Stieltjes transform of the function $F(t)$, $0 < t < \infty$ will be denoted as

$$\text{L.S.}\ \{F(t),s\} = \int_0^\infty e^{-st}\,dF(t).$$

The notation $\{a_n\} * \{b_n\}$ will be used to denote the convolution of the sequences $\{a_n\}$ and $\{b_n\}$.

The following notations will be adopted for the Gamma and Beta functions respectively

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x}\,dx \qquad a > 0$$

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \qquad a,b > 0.$$

The incomplete Beta function will be defined as

$$I_p(a,b) = \frac{1}{B(a,b)} \int_0^p x^{a-1}(1-x)^{b-1}\,dx$$

and we will make use of the following property (e.g. see Uppuluri and Blot (1970)).

$$\sum_{x=0}^{k-1} \binom{r+x-1}{x} p^r q^x = I_p(r,k) \ ; \qquad 0 < p < 1,\ p+q = 1.$$

A summary of the properties of the Incomplete Beta Function can be found in Abramowitz and Stegun (1965) and Erdéyli (1953).

The ascending and descending factorials will be denoted by

$$a_{(n)} = a(a+1)\ldots(a+n-1)$$

and

$$a^{(n)} = a(a-1)\ldots(a-n+1)$$

respectively with

$$a_{(0)} = a^{(0)} = 1,$$

$$a_{(n)} = (-1)^n(-a)^{(n)} \text{ and } a_{(n+m)} = a_{(n)}(a+n)_{(m)} = a_{(m)}(a+m)_{(n)}.$$

For the generalized Hypergeometric function we will use the notation

$${}_pF_q(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;z) = \sum_{x=0}^{\infty} \frac{(a_1)_{(x)}\cdots(a_p)_{(x)}}{(b_1)_{(x)}\cdots(b_p)_{(x)}} \frac{z^x}{x!}.$$

The following integral representation of the Hypergeometric function (the Generalized Hypergeometric for p=2, q=1) will be used

$${}_2F_1(a,b;c;z) = \frac{1}{B(b,c-b)} \int_0^1 x^{b-1}(1-x)^{c-b-1}(1-zx)^{-a}\,dx.$$

(for b,c-b > 0).

In addition, we make use of the integral representation of the Confluent Hypergeometric function (the generalized Hypergeometric with p=q=1)

$${}_1F_1(b;c;z) = \frac{1}{B(b,c-b)} \int_0^1 x^{b-1}(1-x)^{c-b-1}e^{zx}\,dx.$$

(b,c-b > 0).

Another form of the Hypergeometric function we use, is the Confluent Hypergeometric function of two variables (a Bivariate generalization of the Confluent Hypergeometric Function).

This is defined as

$$\Phi_1[a,b;c;x,y] = \sum_{m\geq 0} \sum_{n\geq 0} \frac{a_{(m+n)} b_{(m)}}{c_{(m+n)} m!n!} x^m y^n$$

with integral representation

$$\Phi_1[a,b;c;x,y] = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 u^{a-1}(1-u)^{c-a-1}(1-ux)^{-b} e^{uy} du.$$

Note 1

Definitions and properties of the various forms of the Hypergeometric functions mentioned, can be found in Appel and Kampé de Feriet (1926), Slater (1966) and Erdélyi (1953).

Note 2

There are many distributions whose p.g.f.'s can be expressed in terms of Hypergeometric functions. A list of these distributions is given by Dacey (1972), while conditions for which the generalized Hypergeometric series becomes a valid p.g.f. have been studied by Kemp (1968).

The thesis is composed of nine chapters. Each chapter is divided into sections. The formulae, theorems, lemmata and corollaries are numbered treblewise; the first number refers to the chapter, the second refers to the section of the chapter and the third indicates the successive item within the section. Numbers in parentheses are used for formulae, e.g. (4.3.2) refers to the second formula of the third section in Chapter 4. Theorem 3.5.1, means the first theorem of Section 5 in Chapter 3. Notes and remarks are single numbered since they invariably follow a theorem or a corollary. Whenever we refer to a note or a remark we specify the theorem or corollary to which the note or remark is related. We will say, for example, Note 1, theorem 3.1.1. When we wish to use theorems or corollaries that already exist in the literature, these will appear with the name of the author following the number of the theorem. Hence, when we write Theorem 3.1.1 (Shanbhag 1976), we will be quoting a theorem introduced by Shanbhag in 1976. All other theorems, corollaries and lemmata, which are mentioned just by their number, appear for the first time in this thesis in the form in which they are presented.

.2 Probability Distribution

The probability distributions which will be used in this study are the following.

1.2.1 Univariate Probability Distributions.

The Poisson Distribution ($P_n(\lambda)$)

$$P_n = P(X=n) = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n=0,1,2,\ldots,\lambda > 0. \qquad (1.2.1)$$

The Binomial Distribution B(n,N,p)

$$P(X=n) = \binom{N}{n} p^n q^{N-n} \qquad n=0,1,2,\ldots,N=0,1,2,\ldots \qquad (1.2.2)$$

$$0 < p < 1 \quad p+q = 1$$

where

$$P_n = 0 \text{ for } n > N.$$

The Hypergeometric Distribution H(N,m,n)

$$P(X=r) = \frac{\binom{m}{r}\binom{N-m}{n-r}}{\binom{N}{n}} \qquad r,m,n,N > 0 \qquad (1.2.3)$$

$$r \leqslant n, \; m \leqslant N.$$

The Beta Distribution $B(\alpha,\beta)$

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad 0 < x < 1,\ \alpha > 0,\ \beta > 0 \qquad (1.2.4)$$

where $B(\alpha,\beta)$ is the Beta function defined in the previous section.

The Exponential Distribution

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \qquad 0 < x < \infty,\ \theta > 0 \qquad (1.2.5)$$

The Geometric Distribution

$$P(X=n) = p\, q^{n-1} \qquad n=1,2,\ldots,\ 0 < p < 1,\ p+q = 1. \qquad (1.2.6)$$

The Gamma Distribution

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \qquad x > 0,\ \alpha > 0,\ \beta > 0. \qquad (1.2.7)$$

For a=1 the Gamma distribution becomes exponential with parameter β.

The Negative Binomial

$$P(X=n) = \binom{-N}{n} p^{N}(-q)^{n} \qquad n=0,1,\ldots,N > 0,\ 0 < p < 1,\ p+q = 1 \qquad (1.2.8)$$

The Negative Binomial can also be derived as the mixture on λ of the Poisson distribution with parameter λ, where λ has the Gamma distribution with parameters p/1-p and N.

The Negative Hypergeometric N.H. (n,m,p)

$$P(X=r) = \frac{\binom{-m}{r}\binom{-\rho}{n-r}}{\binom{-m-\rho}{n}} \qquad r=0,1,\ldots,n \quad m>0,\ \rho>0. \tag{1.2.9}$$

The Negative Hypergeometric is also obtained as the mixture on p of the Binomial B(r,n,p) if p has the Beta distribution with parameters m and ρ.

Gurland Distribution (see Gurland (1958))

The p.g.f. of the Gurland distribution is given by

$$G(t) = {}_1F_1\{\alpha;\alpha+\beta;\lambda(t-1)\} \qquad \alpha,\beta,\lambda > 0\ . \tag{1.2.10}$$

This distribution was first examined by Gurland, (1958) who derived it as a result of compounding the Binomial (r,n,p) with a Poisson (λ) on the parameter n and then compounding the resulting distribution by a Beta (α,β) on the parameter p.

1.2.2 Truncated Univariate Distributions

The Truncated Poisson

The Poisson distribution, truncated at the point k-1; k=1,2,... is defined as

$$P_n = P(X=n) = \frac{\frac{\lambda^n}{n!}}{\sum_{n=k}^{\infty}\frac{\lambda^n}{n!}} \qquad n=k,k+1,\ldots;\lambda>0,\ k=1,2,\ldots \tag{1.2.11}$$

The Truncated Binomial

$$P(X=n) = \frac{\binom{N}{n} p^n q^{N-n}}{\sum_{n=k}^{\infty} \binom{N}{n} p^n q^{N-n}} \qquad n=k,k+1,\ldots;N=k,k+1\ldots. \quad 0 < p < 1,\ q = 1-p. \qquad (1.2.12)$$

The Truncated Hypergeometric

$$P(X=r) = \frac{\binom{m}{r}\binom{N-m}{n-r}}{\sum_{r=k}^{n} \binom{m}{r}\binom{N-m}{n-r}} \qquad r=k,k+1,\ldots;\ r \leq n \quad n \geq k,\ m \geq k,\ N \geq k. \qquad (1.2.13)$$

Beta Truncated at t, $0 < t \leq 1$

$$f(x) = \frac{\alpha x^{\alpha-1}(1-x)^{\beta-1}}{t^{\alpha}\, {}_2F_1(\alpha,1-\beta;\alpha+1;t)} \qquad 0 < x < t,\ \alpha > 0,\ \beta > 0. \qquad (1.2.14)$$

Right Truncated Exponential

$$f(x) = \frac{\frac{1}{\theta} e^{-\frac{x}{\theta}}}{1 - e^{-\frac{1}{\theta}}} \qquad 0 < x < 1,\ \theta > 0. \qquad (1.2.15)$$

Right Truncated Gamma

$$f(x) = \frac{\alpha x^{\alpha-1} e^{-\frac{x}{\beta}}}{{}_1F_1(\alpha;\alpha+1;-\frac{1}{\beta})} \qquad 0 < x < 1,\ \alpha > 0,\ \beta > 0. \qquad (1.2.16)$$

The Truncated Negative Binomial

$$P(X=n) = \frac{\binom{-N}{n} p^N (-q)^n}{\sum_{n=k}^{\infty} \binom{-N}{n} p^N (-q)^n} \qquad n=k,k+1,\ldots;\ N=k,k+1,\ldots \tag{1.2.17}$$

The Truncated Negative Hypergeometric

$$P(X=r) = \frac{\binom{-m}{r}\binom{-\rho}{n-r}}{\sum_{r=k}^{\infty} \binom{-m}{r}\binom{-\rho}{n-r}} \qquad r=k,k+1,\ldots;\ r \leqslant n. \tag{1.2.18}$$

The Convolution of a Poisson (λ) with a Truncated Poisson (μ)

This is also a truncated distribution and is defined as follows

$$P(X=n) = \frac{e^{-\mu} \sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r}}{n! \sum_{n=k}^{\infty} \frac{\lambda^n}{n!}} \qquad n=k,k+1,\ldots;\ \lambda > 0,\ \mu > 0. \tag{1.2.19}$$

Samaniego (1976) examines a particular case of the above model (k=1) and studies some estimation problems.

The Convolution of a Binomial with a Truncated Binomial

The distribution which is the Convolution of a Binomial (N-m,p) and a Truncated Binomial (m,p) has p.d.f.

$$P(X=n) = \frac{p^n q^{N-n} \sum_{r=k}^{n} \binom{m}{r}\binom{N-m}{n-r}}{I_p(k,m-k+1)} \qquad n,N=k,k+1,\ldots;\ k=1,2,\ldots \tag{1.2.20}$$

The Convolution of a Negative Binomial with a Truncated Negative Binomial

Consider a Negative Binomial (p,m) and a T.N.B. (p,ρ). Then the p.d.f. of their convolution is given by

$$P(X=n) = \frac{\sum_{r=k}^{n} \binom{-m}{r}\binom{-\rho}{n-r} (-q)^{n} (1-q)^{\rho}}{\sum_{r=0}^{\infty} \binom{-m}{r} (-q)^{r}} \qquad \begin{array}{l} n=k,k+1,\ldots \\ m,\rho > 0 \\ 0 < q < 1. \end{array} \qquad (1.2.21)$$

1.2.3 Bivariate Distributions

In this thesis we will make use of a particular class of Bivariate distributions, which we will call double distributions. These are obtained when one considers the product of two independent and identically distributed r.v.'s.

The truncated versions of those distributions can be derived by considering each of the r.v.'s X_i i=1,2 truncated at a point k_i -1, i=1,2 ; $k_i > 0$.

In particular we will use the following.

The Double Poisson Distribution

$$P(X_1=n_1, X_2=n_2) = e^{-\lambda_1-\lambda_2} \frac{\lambda_1^{n_1}}{n_1!} \frac{\lambda_2^{n_2}}{n_2!} \qquad \begin{array}{l} n_i=1,2,\ldots \\ \lambda_i > 0 \\ i=1,2. \end{array} \qquad (1.2.22)$$

The Double Binomial

$$P(X_1=n_1, X_2=n_2) = \binom{N_1}{n_1} p_1^{n_1} q_1^{N_1-n_1} \binom{N_2}{n_2} p_2^{n_2} q_2^{N_2-n_2} \qquad (1.2.23)$$

The Double Poisson and Binomial have been studied by Talwalker (1970) and Shanbhag (1974).

The Double Hypergeometric

$$P(X_1=r_1, X_2=r_2) = \frac{\binom{m_1}{r_1}\binom{N_1-m_1}{n_1-r_1}}{\binom{N_1}{n_1}} \frac{\binom{m_2}{r_2}\binom{N_2-m_2}{n_2-r_2}}{\binom{N_2}{n_2}} \qquad \begin{array}{l} r_i, m_i, n_i, N_i > 0 \\ r_i \leqslant n_i,\ m_i \leqslant N_i \\ i=1,2. \end{array} \tag{1.2.24}$$

The Double Negative Binomial

$$P(X_1=n_1, X_2=n_2) = \binom{-N_1}{n_1}\binom{-N_2}{n_2} p_1^{N_1} p_2^{N_2} (-q_1)^{n_1} (-q_2)^{n_2} \qquad \begin{array}{l} n_i, N_i = 0,1,\ldots \\ 0 < p_i < 1 \\ i=1,2. \end{array} \tag{1.2.25}$$

The Double Hypergeometric and Double Negative Binomial have been examined by Ratnaparkhi (1975).

Note: The bivariate extensions of the distributions defined in Section 1.2.2 can be obtained in the same way.

1.2.4 Multivariate Distributions

The Multinomial Distribution

$$P_{\underset{\sim}{n}} = P(X_1=n_1, \ldots, X_s=n_s) = \frac{n!}{n_0!n_1!\ldots n_s!} p_0^{n_0} p_1^{n_1} \ldots p_s^{n_s}$$

$$0 < p_i < 1, \quad \sum_{i=1}^{s} n_i = n-n_0, \quad \sum_{i=1}^{s} p_i = 1-p_0, \quad n_i \geqslant 0. \tag{1.2.26}$$

The Multivariate Hypergeometric Distributions

$$P(X_1=n_1,\ldots,X_s=n_s) = \frac{\binom{N_0}{n_0}\binom{N_1}{n_1}\cdots\binom{N_s}{n_s}}{\binom{N}{n}} \tag{1.2.27}$$

$$N_i \geq 0, \quad \sum_{i=1}^{s} n_i = n-n_0, \quad \sum_{i=1}^{s} N_i = N-N_0.$$

The Negative Multinomial Distribution

$$P(X_1=n_1,\ldots,X_s=n_s) = \frac{\Gamma(m+n_1+\ldots+n_s)}{\Gamma(m)} \; p_0^m \prod_{i=1}^{s} \frac{p_i^{n_i}}{n_i!} \tag{1.2.28}$$

$$n_i=0,1,\ldots;\ 0<p_i<1,\ \sum_{i=1}^{s} p_i<1,\ i=1,2,\ldots,s;\ p_0=1-\sum_{i=1}^{s} p_i$$

The Multivariate Inverse Hypergeometric Distribution

$$P(X_1=n_1,\ldots,X_s=n_s) = \frac{B(m+n_1+\ldots+n_s,\rho+(N_1-n_1)+\ldots+(N_s-n_s))}{B(m,\rho)} \prod_{i=1}^{s} \binom{N_i}{n_i} \tag{1.2.29}$$

$$n_i=0,1,\ldots,N_i,\ m>0,\ \rho>0,\ i=1,2,\ldots,s.$$

Note 1

The Class of Multiple distributions will also be used, in particular the Multiple Poisson, Binomial, Negative Binomial, Hypergeometric and Negative Hypergeometric, along with their truncated versions. They are the straightforward extensions of the Double distributions, i.e. the product of s independent and identically distributed r.v.'s.

Note 2

Detailed study of the distributions which were mentioned in this section, as well as a list of references, can be found in Johnson and Kotz (1969), and Patil and Joshi (1968).

Also, a full account on the Poisson distribution appears in Haight (1967), while Kemp and Kemp (1956) examine the Hypergeometric distribution.

.3 Literature Review

The problem of characterizing discrete statistical distributions using properties of the conditional distribution of one random variable given another has been considered by many mathematicians and statisticians from a number of different angles.

Let X and Y be two random variables taking non-negative integer-values, such that X is greater than, or equal to Y. Let Z denote the difference X-Y. Then the various results which have been obtained, using the conditional distributions of Y on X, can be broadly divided into the following major classes.

1.3.1 The Damage Model and its Applications

Rao (1963) in a pioneer paper gives the following physical interpretation to the r.v.'s X, Y and Z, for the damage model.

Let X denote an observation produced by some natural process (e.g. number of eggs, number of accidents, etc.). This observation may be partially destroyed, or may be only partially ascertained. In such circumstances the original distribution (i.e. the distribution of X) will be distorted. Rao then points out that if the model underlying the partial destruction of original observations (or the survival distribution) is known we can derive the distribution appropriate to the observed values knowing the original distribution.

The notation used for the damage model varies from author to author. According to the notation introduced by Rao (1963), X represents the original r.v. and Y the resulting r.v. The distribution of X is called the "original distribution" and the distribution of Y the "resulting distribution". The r.v. Z = X-Y represents the damage part. The distribution of $Y|X$ is called the "survival distribution".

Kagan, Linnik and Rao (1973) use the term "ruin process" instead of survival distribution. Patil has switched the role of the letters; he uses Z instead of X, X instead of Y and Y instead of Z.

In the sequel, Rao's notation is adopted.

Rao sets up the problem in the following way. Let P_n be the probability that the original observation is n, where n=0,1,... and let the chance that there are r survivals from the original n be s(r,n). Then, the probability of observing r is

$$P(Y=r) = \sum_{n=r}^{\infty} P_n s(r,n).$$

He then examines in depth the following two particular cases.

Firstly, he assumes s(r,n) to be binomial; secondly he assumes that either all or none survive out of n with probability π_n and $(1-\pi_n)$. The second case corresponds to the situation where the investigator does not record observations which are partially damaged.

In the case where the original distribution is Poisson, Binomial or Negative Binomial and the survival distribution is Binomial, Rao proves that the resulting distribution is of the same form as the original distribution, but with the original parameter λ multiplied by p, where p is the binomial parameter; he shows that λ and p get confounded, i.e. cannot be separately estimated.

He also shows that in this case the probability distribution of the damaged observations alone, the distribution of the undamaged observations alone and the distribution of the observations when the classification as damaged or undamaged is not known, are all the same. Moreover he found that this result is true even when the original distribution is truncated at an arbitrary point.

Revankar, Hartley and Pagano (1974) use the same model in a slightly different form to study the distribution of under-reported income.

Talwalker (1975) applies the damage model to medical and toxicological problems.

Patil and Rao (1976) consider the distribution of the "undamaged observation" of the damage model as a special case of the weighting model. This model gives rise to distributions which have been modified by the method of ascertainment as a result of sampling with unequal probabilities of observation. The concept was also introduced by Rao (1963).

The form of the resulting weighted distribution can be deduced from the original distribution provided of course that the model for the sampling chance (weight) is known.

1.3.2 The R-R Characterization and its Variants

Using damage model theory, Rao and Rubin (1964) obtained the following characterization of the Poisson distribution.

Suppose that the survival distribution $s(r,n)$ is Binomial with parameters n and p, where p is fixed. Then the condition

$$P(Y=r) = P(Y=r|X=Y) = P(Y=r|X>Y)$$

is necessary and sufficient for the distribution of X to be Poisson. This is known as the Rao-Rubin condition. (R-R condition.) A modified version of the R-R condition was used by them to characterize the truncated Poisson distribution.

This was the first characterization based on the damage model theory. The mathematical importance of it lies in the fact that independence of Y and the event $X=Y$ is sufficient to determine the distribution of the r.v.

X as Poisson when Y|X is Binomial.

A number of interesting problems have been generated by the introduction of the R-R condition.

The first is to examine whether the R-R condition characterizes the Binomial distribution as the only form of the distribution of Y|X when X is Poisson. Srivastava and Srivastava (1970) gave a positive answer to the problem, but subject to the additional assumption that the parameter λ of the Poisson distribution is variable.

The second is to examine whether the variant of the R-R condition, namely

$$P(Y=r|Y\geqslant k) \;\; = \;\; P(Y=r|X=Y)$$

determines uniquely the distribution of Y|X when X is truncated Poisson. Srivastava and Singh (1975) conjectured that a "modified" Binomial provides the only solution. However, their conjecture is not valid as we will show later on in Chapter 2.

The third is to find pairs of distributions other than the Poisson, Binomial for which the R-R property is characteristic. Many papers of that kind have appeared recently. However, most of these require some additional conditions in order to obtain the required result. Patil and Ratnaparkhi (1975) used a result of Patil and Seshadri (1964) and assumed that the r-th order derivative of $G_x(t)$ exists; this enabled them to derive a characterization for the Negative Binomial (with Y|X ~ Negative Hypergeometric). They also proved that the R-R condition holds when Y|X is Hypergeometric and X is Binomial. They left unsolved the problem as to whether if Y|X is again Hypergeometric and the R-R condition holds then X must be Binomial. (The answer to this question will be given in Chapter 5.)

Consul (1974, 1975) studied the characterization of the Lagrangian-Poisson distribution as the original distribution with the Quasi-Binomial as the survival distribution in the damage model. (For definition and properties of the Lagrangian-Poisson and Quasi-Binomial distributions see Consul and Jain (1973).)

The fourth is to find pairs of distributions other than the truncated Poisson, Binomial for which the property

$$P(Y=r|Y \geq k) \quad = \quad P(Y=r|X=Y)$$

is characteristic. No work seems to have been published in this field.

The fifth is to obtain characterizations for families of distributions, using the R-R property. Shanbhag (1976) used a technique existing in the renewal theory to characterize the form of the distribution of X using the R-R condition, in the case where the distribution of $X|Y$ satisfies a given condition. This result provides many other characterizations as particular cases.

The sixth is to study the R-R property under the assumption that $Y|X$ follows a truncated distribution. Again here, there do not seem to be any results available in the literature.

The seventh is to extend the results mentioned previously to the Bivariate and Multivariate cases. Talwalker (1970) extended Rao and Rubin's result to the Bivariate and Multivariate case. Patil and Ratnaparkhi (1975) and Ratnaparkhi (1975) have obtained the Bivariate versions of their result which has already been mentioned; so has Shanbhag (1976). Clearly there are many gaps in the literature so far as the Bivariate amd Multivariate cases are concerned. Except for the Double and Multiple Poisson, and Negative Binomial, no other bivariate or multivariate distribution has been characterized through the R-R condition. Truncated bivariate and multivariate distributions have been totally ignored, both when truncation concerns the distribution of X, and also where truncation concerns the distribution of $Y|X$.

The eighth is to derive similar characterizations based on variants of the R-R condition. This has been done by Ratnaparkhi (1975), Kumar and Consul (1976), and Srivastava and Singh (1975). Talwalker (1975) has used the condition

$$P(Y=r|X \text{ damaged when prob. of survival is } p)$$
$$= P(Y=r|X \text{ undamaged when prob. of survival is } p')$$
$$\text{with } p' = \frac{p}{1+dqa} \quad 0 < p < 1;\ q = 1-p,\ 0 < a < 1,\ d = -1,0,+1$$

to characterize the Negative Binomial, Poisson and Binomial distributions when $Y|X$ is Binomial. She argues that the above condition reduces to the R-R condition when d=0. However, this is not the case, because in the course of her proof she treats p, p' as variables, while in the situation examined by Rao and Rubin p is fixed.

Many authors have tried to improve or simplify some of the results mentioned above. Aczél (1972) and Van der Vaart (1972) attempted to give simpler methods for deriving the result of Rao and Rubin and Talwalker (1970). However, they did this by allowing p (the parameter of the Binomial survival in the Univariate case) and p_1, p_2 (the parameter of the double Binomial survival in the bivariate case) to vary over (0,1), a condition which is of course very stringent. This was pointed out by Shanbhag (1974), who in fact used an elementary method to arrive at the R-R characterization and also to improve Talwalker's result by relaxing some of the conditions. Ord (1975) derived a version of Talwalker's bivariate extension (again by assuming p variable) as a special case of a characterization of his Bivariate dependent Poisson distribution. A characterization of the Bivariate Hermite was derived as a special case.

The conditional expectation has been used, instead of the R-R condition, by a number of research workers. Shanbhag and Clark (1972) established that if X has a power series distribution with parameter λ, and $s(r,n)$ has mean np and variance $np(1-p)$ with p and $s(n,n)$ independent of λ, then

$$E(Y) = E(Y|X=Y) \text{ and } \operatorname{Var} Y = \operatorname{Var}(Y|X=Y)$$

iff P_n is Poisson and $s(n,n) = p^n$. (For the definition of a power series distribution the reader is referred to Patil (1962).)

Patil and Ratnaparkhi (1975) made use of the conditional expectation of $Y|X$ to characterize the Poisson, Binomial, Negative Binomial, Factorial and Hypergeometric distributions. The same authors in their (1977) paper considered a bivariate observation with the second component subjected to damage. Under this assumption they characterized the Poisson, Binomial and Negative Binomial distributions within the framework of the damage model using the invariance of linearity of regression of the first component on the second.

Korwar (1975) derived a characterization for a class of distributions assuming that the distribution of $Y|X$ is Binomial and the regression of X on Y is linear.

1.3.3 Conditionality Characterizations

Let us consider integer-valued non-negative r.v.'s Y and Z, where X=Y+Z. Then one can obtain characterizations relating the distributions of X and Y when the distribution of $Y|X$ is given. Another problem is to study the relation between the distribution of Y and the distribution of $Y|X$ when the distribution of X is known. One can also examine the form of the distributions of Y and Z for a given form of the distribution of $Y|X$.

Many papers have appeared in the literature dealing with these problems. The main assumption in all of them is that Y and Z are independent. Patil and Seshadri (1964) proved that if the distribution of $Y|X$ satisfies a certain condition then the distributions of Y and Z are uniquely determined. Menon (1966), clarified certain points of Patil and Seshadri's work.

This work was motivated by a very interesting result derived by Moran (1952); he showed that if

$$P(Y=r|X=n) = \binom{n}{r} p_n^r q_n^{n-r},$$

then p_n is independent of n, and Y, Z have Poisson distributions. A smilar result was given by Chatterji (1963).

Using Menon's theorem, Kemp (1974b) derived a general characterization for generalized Hypergeometric probability distributions, which contains the result of Patil and Seshadri, Kemp and Kemp (1975) and Moran (for $p_n=p$) as special cases. (For definition and properties of the g.h.p. distributions the reader is referred to Kemp (1968, 1974a).

Some of Patil and Seshadri's results were extended to the multivariate case by Janardan (1974). Janardan (1975) proved that if $Y|X$ is mixed Quasi-Binomial distribution (M.Q.B.D.), then Y and Z follow generalized Poisson distributions (G.P.D.) (G.P.D. is called Lagrangian-Poisson by Consul.) Janardan also showed that if $Y|X$ is mixed Quasi-Hypergeometric then Y, Z have generalized Negative Binomial distributions. Govindarajulu and Leslie (1970) examined the same models in more detail.

Consul (1974, 1975) has also studied the above models.

Most of these results require independence of Y and Z; this condition is very stringent.

Svensson (1969) avoided the assumption of independence. He proved that for a given r.v. X there exists a two-dimensional random variable (X,Y) with the property that the p.g.f. of the distribution of Y|X is of the form $(q+pt)^n$ iff

$$G_X(t) = G_Y\left(\frac{t-q}{p}\right).$$

Seshadri and Patil (1963) obtained the form of the distribution of Z in various cases when the distributions of Y and Y|X are given. Haight (1972) discussed in detail some of the results that have appeared in the literature in connection with Svensson's theorem.

Considering the distribution of the r.v. Y to be the result of mixing the distribution of X with the distribution of Y|X, Skibinsky (1970) proved that, with the assumption that Y|X is Hypergeometric, X is Binomial iff Y is Binomial. A physical interpretation of this result has been given by Mood (1943) and by Hald (1960). Nevill and Kemp (1975) and Janardan (1973) have extended this characterization to the multivariate case. Krishnaji (1970) derived a characterization for the Yule distribution truncated at zero using the fact that when the survival distribution is uniform, the resulting distribution truncated at zero coincides with the original distribution.

It is interesting that conditionality models (and hence damage models) with the distribution of X or the distribution of Y|X having a mixed form have not been considered in detail. Only Krishnaji (1974) studied the independence of Y and Z for mixed original or survival distributions.

It is clear that the conditionality model is more general than the damage model. This is so because the conditionality model is a mathematical model whereas the damage model corresponds to a particular application of this. The assumption of independence of Y and Z, existing in most cases of the former, is dropped in the latter.

The introduction of the damage model has extended considerably the existing conditionality results. The reason for this is that the R-R property, introduced in the damage model theory, is clearly a weaker assumption than independence. Moreover, the R-R property has given rise to many interesting characterizations.

Note: Other papers not directly of use in this thesis which deal with certain of these topics and which may be of some interest to the reader include Aczél (1975a), Medhi (1975), Samaniego (1976), Svensson (1975), Teicher (1954, 1961), Gani and Shanbhag (1974), Ottaviani (1957), Chatfield and Theobald (1973), Gurland (1957, 1958), Gupta (1974), Moran (1951), Daboni (1959), Lamberti (1959), Zippin (1956).