DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

PROGRESSIVELY TYPE II CENSORED ORDER STATISTICS, DISTRIBUTIONS, CHARACTERIZATIONS AND CONCOMITANTS

by
TUĞBA YILDIZ

November, 2007
İZMİR

PROGRESSIVELY TYPE II CENSORED ORDER STATISTICS, DISTRIBUTIONS, CHARACTERIZATIONS AND CONCOMITANTS

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Doctor of
Philosophy in Statistics Program

by
TUĞBA YILDIZ

November, 2007
İZMİR

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"PROGRESSIVELY TYPE II CENSORED ORDER STATISTICS, DISTRIBUTIONS, CHARACTERIZATIONS AND CONCOMITANTS"** completed by **TUĞBA YILDIZ** under supervision of **PROF. DR. İSMİHAN BAYRAMOĞLU** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. İsmihan BAYRAMOĞLU

Supervisor

Prof. Dr. Serdar KURT

Thesis Committee Member

Assoc. Prof. Dr. Halil ORUÇ

Thesis Committee Member

Prof. Dr. Aladdin ŞAMİLOV

Examining Committee Member

Assist. Prof. Dr. Esin FİRUZAN

Examining Committee Member

Prof.Dr. Cahit HELVACI
Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I wish to express my sincere thanks and deep appreciation to my advisor, İsmihan BAYRAMOĞLU for his encouragement and guidance throughout the development of my thesis. He not only gave me endless support and attention which is essential accomplishing this work, but also played an important role in my education. It has been an honor and privilege to work with him.

I would like to acknowledge my dissertation committee members, Serdar KURT and Halil ORUÇ for their constructive comments and suggestions. Their invaluable effort and helpful advice improved my thesis.

Finally I would like to express my heartfelt thanks and appreciation to my family who has always been extremely supportive of my endeavors and unconditional in their love. My parents, by word and deed, have emphasized the importance of education. They encouraged me to pursue my own interests, willingly offered any advice and would never ever entertain the possibility that I might not be able to accomplish anything I chose to do.

Last, but not least, I wish to express my gratitude to my husband, Mutlu YILDIZ. I am extremely thankful to him for his tremendous patience and support without which this work would be incomplete. His blessings and love have been a constant source of encouragement and strength.

Tuğba YILDIZ

PROGRESSIVELY TYPE II CENSORED ORDER STATISTICS, DISTRIBUTIONS, CHARACTERIZATIONS AND CONCOMITANTS

ABSTRACT

In this thesis, distributions, characterizations and concomitants of general models of the ordered random variables such as Progressively Type II censored order statistics are studied.

The joint distributions of the nonadjacent Progressively Type II censored order statistics are obtained. A useful identity describing the probability density functions of k nonadjacent Progressively Type II censored order statistics is given. This result is an extension of known results proposed by Balakrishnan, Childs, & Chandrasekar (2002).

It is shown that for a particular Progressive Type II censoring scheme with equal withdrawals the joint distributions can be reduced to the distributions of ordinary order statistics. This result allows obtaining characterization theorems for special families of distributions in terms of conditional expected values of Progressively Type II censored order statistics. Characterization theorems for several important classes of distributions are provided in Chapter 4 of this thesis. Also, some goodness-of-fit test procedures based on characterization theorem are described.

Keywords: Censoring scheme, Order statistics, Progressively Type II censored order statistics, Goodness-of-fit test, Distribution, Characterization, Concomitant.

İLERLEYEN II. TÜR SANSÜRLENMİŞ SIRA İSTATİSTİKLERİ, DAĞILIMLARI, KARAKTERİZASYONLARI VE EŞLENİKLERİ

ÖZ

Bu tezde, ilerleyen II. Tür sansürlenmiş sıra istatistikleri gibi sıralı rassal değişkenlerin genel modellerinin dağılımları, karakterizasyonları ve eşlenikleri çalışıldı.

Ardışık olmayan ilerleyen II. Tür sansürlenmiş sıra istatistiklerinin birleşik dağılımı elde edildi. *k* ardışık olmayan ilerleyen sansürlenmiş sıra istatistiğinin olasılık yogunluk fonksiyonunu tanımlayan önemli bir eşitlik verildi. Bu sonuç, Balakrishnan, Childs, ve Chandrasekar (2002) tarafından elde edilen sonuçların bir uzantısıdır.

Testten çıkışların eşit olduğu, özel bir ilerleyen II. Tür sansürleme şekli için birleşik dağılım fonksiyonunun sıralı istatistiklerin dağılımına indirgenebildiği gösterildi. Bu sonuç, ilerleyen II. Tür sansürlenmiş sıra istatistiklerinin koşullu beklenen değerleri anlamında özel dağılım ailelerinin karakterizasyon teoremlerinin elde edilmesini sağlamıştır. Bu tezin 4. Bölümünde, birçok önemli dağılım sınıfı için karakterizasyon teoremleri verildi. Ayrıca, karakterizasyon teoremlerine dayalı bazı uyum iyiliği testleri tanımlandı.

Keywords: Sansürleme şekli, Sıra istatistikleri, İlerleyen II. Tür sansürlenmiş sıra istatistikleri, Uyum iyiliği testi, Dağılım, Karakterizasyon, Eşlenik.

CONTENTS

Page

CHAPTER ONE
INTRODUCTION

One of the important topics of mathematical statistics is the theory of ordered random variables which finds wide applications in many fields of statistical science and inference. The theory of ordered random variables starting with order statistics and developing to the record values, progressively censored order statistics and generalized order statistics are described in several monographs including David (1981) and Arnold, Balakrishnan, & Nagaraja (1992).

There are several models of ordered random variables with different interpretations and interesting applications in many fields, for example, in survival analysis, reliability theory, financial economics, etc.

In this thesis, properties of ordered random variables in a model of Progressively Type II censored order statistics are considered. One of the interesting modifications of order statistics is the concept of Progressively Type II censored order statistics which is very useful in reliability and lifetime studies.

“Progressive Type II censoring is a useful and more general scheme in which a specific fraction of individuals at risk may be removed from the study (i.e., censored) at each of several ordered failure times” (Fernandez, 2004).

This scheme of censoring appears to be of great importance in planning duration experiments in reliability studies. In many industrial experiments involving lifetimes of machines or units, experiments have to be terminated early or the number of experiments must be limited due to a variety of circumstances (e.g. when expensive items must be destroyed, when experiments are time-consuming and expensive, etc.). In addition, some life tests require removals of functioning test specimens to collect degradation-related information to failure time data. The samples that arise from such experiments are called censored. The planning of experiments with the aim of reducing both the number of failures and the total duration of the experiment leads naturally to the well known Type I and Type II right censoring schemes. "Type-I and Type-II censoring schemes do not allow removal of units at points other than the terminal point of experiment" (Balakrishnan, Kannan, Lin, & Ng, 2003). Many references and historical notes on this subject are presented by Balakrishnan & Aggarwala (2000). They developed parametric methods to analyze Progressively Type II right censored data.

Considerable amount of research has been carried out on progressive censoring methodology. The progressive censoring scheme, as will be introduced in Chapter 2, is constructed with the aim of moderating the loss of information by reducing the number of failures with respect to the full sample approach. Point estimation for the scale and location parameters of the Extreme value distribution by linear functions of order statistics from Progressively Type II censored samples is investigated by Thomas & Wilson (1972). Montanari & Cacciari (1988) reported results of progressively censored data aging tests on XLPE-insulated cable models under combined thermal-electrical stresses. By using progressively censored data, Viveros & Balakrishnan (1994) proposed interval estimation of parameters of life. Balasooriya, Saw & Gadag (2000) and Balasooriya & Balakrishnan (2000) studied progressively censored reliability sampling plans for Weibull and Lognormal distributions. Ng, Chan, & Balakrishnan (2002) discussed the estimation of parameters from progressively censored data using the EM algorithm, while Ng et al. (2004) used this approach to determine optimal progressive censoring schemes and reliability sampling plans for the Weibull distribution. Balakrishnan, Cramer, Kamps & Schenk (2001) and Balakrishnan & Lin (2002) discussed inferential procedures for exponential distribution, while Balakrishnan, Kannan et al. (2003) discussed for Normal distribution based on Progressively Type II censored samples.

The first description of the theory of Progressively Type II right censored order statistics can be found in Balakrishnan & Sandhu (1995, 1996). There are also numerous works published in statistical literature discussing inference problems for a wide range of distributions under this sampling scheme. For instance, see Cohen (1963, 1991). Details of theory, methods and applications of progressive censoring can be found in Balakrishnan et al. (2000).

"In recent years characterization problems have become of increasing interest" (Navarro, Ruiz, & Zoroa, 1998), and considerable attention has been paid to problem of characterizing the distribution function of a random variable based on conditional expectations. Also, characterizations based on different class of distributions have been introduced in several studies. In Chapter 4, some of these studies are given and by using the conditional expectations of order statistics and Progressively Type II censored order statistics, some general class of distributions are characterized.

This thesis is organized as follows. The definition and the distribution properties of some ordered random variables such as order statistics, Progressively Type II censored order statistics and generalized order statistics are given in Chapter 2. New results on the joint probability density function of nonadjacent Progressively Type II right censored order statistics are introduced in Chapter 3. In Chapter 4, some important results for Progressively Type II censored order statistics for the case $\mathbf{R} = (R, R, \ldots, R)$ where $\mathbf{R}$ is the censoring scheme are discussed. New contributions to the characterizations of the general class of distributions through the conditional expectations of order statistics and Progressively Type II censored order statistics are obtained in Chapter 4 as well. Goodness-of-fit tests are described for order statistics and Progressively Type II censored order statistics through the conditional expectations of these ordered random variables in Chapter 5. Concluding remarks are given in the final chapter, 6.

CHAPTER TWO

MODELS OF ORDERED RANDOM VARIABLES

One of the models of ordered random variables is order statistics and its distribution functions are given in Section 2.1. In Section 2.2, some descriptions and distribution functions of Progressively Type II censored order statistics are represented. Distribution functions and some special cases of generalized order statistics are introduced in Section 2.3. Finally, in Section 2.4, concomitants of order statistics and Progressively Type II censored order statistics are given.

2.1 Order Statistics

If the random variables $X_1, X_2,..., X_n$ are arranged in ascending order of magnitude and then $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ will denote ordered random variables where $X_{i:n}$ corresponds to the *i*th order statistic $i = 1, 2,..., n$. While the random variable X_i is assumed to be independent and identically distributed (i.i.d.), the $X_{i:n}$ are dependent because of the inequality relations among them.

Suppose that $X_1, X_2,..., X_n$ are random variables each with distribution function $F(x)$. Let $F_r(x),\ r = 1,..., n$ denote the distribution function of the *r*th order statistic $X_{r:n}$. Then the distribution function of the largest order statistics $X_{n:n} = \max\{X_1,..., X_n\}$ is given by

$$F_n(x) = P\{X_{n:n} \leq x\} = P\{all\ X_i \leq x\} = F^n(x). \tag{2.1}$$

Likewise, the distribution function of the $X_{1:n} = \min\{X_1,..., X_n\}$ is

$$F_1(x) = P\{X_{1:n} \leq x\} = 1 - P\{all\ X_i > x\} = 1 - [1 - F(x)]^n. \tag{2.2}$$

The marginal distribution function $F_r(x)$ of rth order statistic $X_{r:n}$, $1 \le r \le n$ is given by

$$F_r(x) = \sum_{j=r}^{n} \binom{n}{j} (F(x))^j (1-F(x))^{n-j} . \tag{2.3}$$

From the relation between binomial series and incomplete beta function, $F_{r:n}(x)$ can be written as

$$F_r(x) = I_{F(x)}(r, n-r+1) \tag{2.4}$$

where $I_x(a,b) = (B(a,b))^{-1} \int_0^x t^{a-1}(1-t)^{b-1} dt$ and $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$ $= \Gamma(a)\Gamma(b)/\Gamma(a+b)$ (David, 1981).

The joint probability density function of $X_{r:n}$ and $X_{s:n}$ ($1 \le r < s \le n$, $x \le y$) is denoted by $f_{r,s}(x, y)$ and can be expressed as

$$f_{r,s}(x,y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} (F(x))^{r-1} f(x) \times (F(y)-F(x))^{s-r-1}$$

$$\times f(y)(1-F(y))^{n-s} . \tag{2.5}$$

From (2.5), the joint probability density function of $X_{r_1},...,X_{r_k}$ $(1 \le r_1 < \ldots < r_k \le n)$ and $k < n$ can be written easily. Then the probability density function is given by

$$f_{r_1,...,r_k}(x_1,...,x_k) = \frac{n!}{(r_1-1)!(r_2-r_1-1)!...(n-r_k)!} (F(x))^{r_1-1} (F(x_2)-F(x_1))^{r_2-r_1-1}$$

$$\times (F(x_3)-F(x_2))^{r_3-r_2-1} ... (1-F(x_k))^{n-r_k} f(x_1) f(x_2) ... f(x_k). \tag{2.6}$$

As it is said before, each variable in the vector of $X_{1:n}$, $X_{2:n}...,X_{n:n}$ are independent although X_1, X_2,..., X_n are i.i.d. random variables.

In discrete case, the joint distribution function of $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$ is given by

$$F_{r,s}(x,y) = \sum_{i=r}^{n} \sum_{j=\max(0,s-i)}^{n-i} \frac{n!}{i!\,j!(n-i-j)!}$$

$$\times (F(x))^i (F(y)-F(x))^j (1-F(y))^{n-i-j} \ , \ x < y. \tag{2.7}$$

If $x \geq y$ then $F_{r,s}(x,y) = F_s(y) = \sum_{i=s}^{n} C_n^i (F(x))^i (1-F(x))^{n-i}$ (David, 1981).

In continuous case, the joint distribution function of $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$ can be obtained by integration of (2.5) as well as by a direct argument valid also in the discrete case. For $x < y$

$$F_{r,s}(x,y) = \sum_{j=s}^{n} \sum_{i=r}^{j} \tfrac{n!}{i!(j-i)!(n-j)!} (F(x))^i (F(y)-F(x))^{j-i} (1-F(y))^{n-j} \ . \tag{2.8}$$

Also for $x \geq y$, the inequality $X_{s:n} \leq y$ implies $X_{r:n} \leq x$, so that

$$F_{r,s}(x,y) = F_s(y). \tag{2.9}$$

The joint probability density function of all n order statistics is

$$f_{1,2,\ldots,n}(x_1,\ldots,x_n) = n!\prod_{j=1}^{n} f(x_j) \tag{2.10}$$

where $x_1 \le x_2 \le \ldots \le x_n$ (David, 1981).

The following theorem shows the equality of probability density function of order statistics $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$ and the probability density function of $Y_{p:p+q-1}$.

Theorem 2.1 Let $X_1, X_2, \ldots, X_n$ be a sample with distribution function $F(x)$ and probability density function $f(x)$. Let $X_{1:n}$, $X_{2:n}, \ldots$, $X_{n:n}$ be the order statistics from this sample with continuous distribution. From

$$(X_{j:n} \mid X_{j-p:n} = x, X_{j+q:n} = y) \equiv Y_{p:p+q-1} \tag{2.11}$$

the conditional probability density function of order statistics $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$ can be represented as

$$f(x_{j:n} \mid x_{j-p:n} = x, x_{j+q:n} = y) = \frac{(p+q-1)!}{(p-1)!(q-1)!}\frac{f(t)}{F(y)-F(x)}$$

$$\times\left(\frac{F(t)-F(x)}{F(y)-F(x)}\right)^{p-1}\left(1-\frac{F(t)-F(x)}{F(y)-F(x)}\right)^{q-1} \tag{2.12}$$

$$f(y) = \frac{f(t)}{F(y)-F(x)}, \qquad x \le t \le y.$$

Proof. Let r, k and s be real numbers such that $r \le k \le s$. The conditional probability density function of order statistics $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$ is

$$f_{r,k,s:n}(x_k \mid x_r = x, x_s = y) = \frac{(s-r-1)!}{(k-r-1)!(s-k-1)!} f(x_k)$$

$$\times \frac{[F(x_k) - F(x_r)]^{k-r-1}[F(x_s) - F(x_k)]^{s-k-1}}{\left[F(x_s) - F(x_r)\right]^{s-r-1}}. \qquad (2.13)$$

Then, if $x_k = t, x_r = x$ and $x_s = y$ in equation (2.13), the following equation is obtained

$$f_{r,k,s:n}(x_k = t \mid x_r = x, x_s = y) = \frac{(s-r-1)!}{(s-k-1)!(k-r-1)!} f(t)$$

$$\times \frac{[F(t) - F(x)]^{k-r-1}[F(y) - F(t)]^{s-k-1}}{\left[F(y) - F(x)\right]^{s-r-1}}. \qquad (2.14)$$

From $\left(\dfrac{F(t) - F(x)}{F(y) - F(x)}\right)' = \dfrac{f(t)}{F(y) - F(x)}$ and $f(t) = F'(t)$, the equation (2.14) can be written as

$$f_{r,k,s:n}(x_k = t \mid x_r = x, x_s = y) = \frac{(s-r-1)!}{(k-r-1)!(s-k-1)!} \frac{f(t)}{F(y) - F(x)}$$

$$\times \left[\frac{F(t) - F(x)}{F(y) - F(x)}\right]^{k-r-1} \left[1 - \frac{F(t) - F(x)}{F(y) - F(x)}\right]^{s-k-1}. \qquad (2.15)$$

If $j = k$, $j - p = r$ and $j + q = s$, then from the equivalence of equation (2.12) and (2.15), the theorem is proved.

2.1.1 Order Statistics for Multivariate Observations

"In applications there are situations when we need to order multivariate random variables (or random vectors). For example, consider a model in reliability analysis when the system has n independent components each consisting of p arbitrarily dependent elements connected by using a parallel structure" (Bairamov, 2006). The life length of the ith component of the system is denoted by $\mathbf{X}_i = (X_i^1, X_i^2, ..., X_i^p)$, $i = 1, 2, ..., n$, where X_i^j denotes the life length of the jth element of the ith component. Then the first failure in the system occurs at time

$$\min\{\max(X_1^1, X_1^2, ..., X_1^p), \max(X_2^1, X_2^2, ..., X_2^p), ..., \max(X_n^1, X_n^2, ..., X_n^p)\}$$

and the corresponding vector of the length is denoted by $\mathbf{X}^{(1)}$. To know the time of the first failure in such a system, it is needed to arrange random vectors $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ by the magnitude of function $\mathbf{N}(\mathbf{X}_i) = \max(X_i^1, X_i^2, ..., X_i^p)$, $i = 1, 2, ..., n$ (or $\mathbf{N}(\mathbf{X}_i) = \min(X_i^1, X_i^2, ..., X_i^p)$, etc.). The distributions of the norm-ordered statistics, i.e. the random vectors ordered with respect to the norm function $\mathbf{N}(\mathbf{x}) = \|\mathbf{x}\|$ and statistics $\mathbf{x} = (x_1, x_2, ..., x_p)$ in a linear normed space, are introduced in Bairamov & Gebizlioglu (1997).

The conditionally ordered and Progressive Type II right censored conditionally ordered statistics for multivariate observations and their distributional properties are introduced by Bairamov (2006). Next the concept of conditionally ordered random vectors will be investigated in view of conditionally **N**-ordered statistics.

Let $(\Omega, \Im, P)$ be a probability space, where Ω is a non-empty set of points, $\omega, \Im$ is a σ-field of subsets of Ω and P is a probability measure defined on $\{\Omega, \Im\}$. The real

Euclidean space IR^p is considered. Let $\Re^p$, $p \geq 1$ be the Borel σ -algebra of subsets of IR^p. Let $\mathbf{X}(\omega)$, $\omega \in \Omega$ be the random variable mapping Ω into IR^p, so $\mathbf{X}^{-1}(B) \in \Im$ for any $B \in \Re^p$. If ω is fixed, then $\mathbf{X} = \mathbf{X}(\omega)$ is a point of IR^p. If $\mathbf{N}(\mathbf{x})$ is a measurable function with respect to the Borel σ -algebra $\Re^p$, then $\mathbf{N}(\mathbf{x})(\omega)$ is a random variable. It is assumed that $\mathbf{N}(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, ..., x_p)$ is a continuous function of its arguments satisfying $\mathbf{N}(\mathbf{x}) \geq \mathbf{0}$, for all $\mathbf{x} \in IR^p$ and $\mathbf{N}(\mathbf{x}) = \mathbf{0}$ if and only if $\mathbf{x} = 0$, where $\mathbf{0} = (0, 0, ..., 0)$.

Suppose that $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n \in \mathbf{S} \subseteq IR^p$ are i.i.d. random variables ($p \geq 1$ random vectors) with p-variate distribution function $F(\mathbf{x})$ and probability density function $f(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, ..., x_p)$ and $\mathbf{S}$ is the support of $\mathbf{X}$. It is clear that $\mathbf{N}(\mathbf{X}_1), \mathbf{N}(\mathbf{X}_2), ..., \mathbf{N}(\mathbf{X}_n)$ are i.i.d. random variables with probability distribution $P\{\mathbf{N}(\mathbf{X}_i) \leq x\}$, $x \in IR$. If F is assumed to be continuous, the probability of any two or more of these random vectors assuming equal magnitudes is zero. Therefore, there exists a unique ordered arrangement within the random vectors $\mathbf{N}(\mathbf{X}_i)$, $i = 1, 2, ..., n$. It is said that $\mathbf{X}_1$ is conditionally less than $\mathbf{X}_2$ if $\mathbf{N}(\mathbf{X}_1) \leq \mathbf{N}(\mathbf{X}_2)$ and written $\mathbf{X}_1 \prec \mathbf{X}_2$. $\mathbf{X}^{(1)}$ denotes the smallest of the set $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$; $\mathbf{X}^{(2)}$ denotes the second smallest, etc.; and $\mathbf{X}^{(n)}$ denotes the largest in the sense of conditional ordering with respect to a function $\mathbf{N}(.)$. $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, ..., \mathbf{X}^{(n)}$ is called the conditionally $\mathbf{N}$-ordered statistics. $h(\mathbf{x})$ is called the structural function and plays an important role in the probability density function of random vectors and is denoted by $h(\mathbf{x}) = P\{\mathbf{N}(\mathbf{X}) \leq \mathbf{N}(\mathbf{x})\}$, $\mathbf{x} = (x_1, x_2, ..., x_p) \in \mathbf{S} \subseteq IR^p$.

The probability density function of $\mathbf{X}^{(r)}$, $1 \leq r \leq n$ is given by Bairamov (2006)

$$f_r(\mathbf{x}) = \frac{n!}{(r-1)!(n-r)!} [h(\mathbf{x})]^{r-1} [1 - h(\mathbf{x})]^{n-r} f(\mathbf{x}). \tag{2.16}$$

Let $1 \le r < s \le n$ and $\mathbf{X}^{(r)} \prec \mathbf{X}^{(s)}$. Consider $B_1, B_2 \in \Re^p$, such that $\mathbf{N}(\mathbf{X}_1) \le \mathbf{N}(\mathbf{X}_2)$ for any $\mathbf{x}_1 \in B_1$, $\mathbf{x}_2 \in B_2$, $\mathbf{x}_k = (x_k^1, x_k^2, ..., x_k^p)$, $k = 1,2$. Then, the joint probability density function of $\mathbf{X}^{(r)}$ and $\mathbf{X}^{(s)}$ is given in Bairamov (2006)

$$f_{r,s}(\mathbf{x}_1, \mathbf{x}_2) = \begin{cases} \dfrac{n!}{(r-1)!(s-r-1)!(n-s)!}[h(\mathbf{x}_1)]^{r-1}[h(\mathbf{x}_2) - h(\mathbf{x}_1)]^{s-r-1} \\ \times[1 - h(\mathbf{x}_2)]^{n-s} f(\mathbf{x}_1) f(\mathbf{x}_2) \quad \text{if } \mathbf{N}(\mathbf{x}_1) \le \mathbf{N}(\mathbf{x}_2) \\ 0 \qquad\qquad\qquad\qquad \text{otherwise.} \end{cases} \tag{2.17}$$

The joint probability density function of random vectors $\mathbf{X}^{(r_1)}, \mathbf{X}^{(r_2)}, ..., \mathbf{X}^{(r_k)}$, $1 \le r_1 < r_2 < \cdots < r_k \le n$ is

$$f_{r_1, r_2, ..., r_k}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k) = \frac{n!}{(r_1 - 1)!(r_2 - r_1 - 1)! \cdots (n - r_k)!}[h(\mathbf{x}_1)]^{r_1 - 1}$$

$$\times[(h(\mathbf{x}_2) - h(\mathbf{x}_1)]^{r_2 - r_1 - 1} \cdots [(h(\mathbf{x}_k) - h(\mathbf{x}_{k-1})]^{r_k - r_{k-1} - 1}$$

$$\times[1 - h(\mathbf{x}_k)]^{n - r_k} f(\mathbf{x}_1) f(\mathbf{x}_2) \cdots f(\mathbf{x}_k), \text{ if } \mathbf{N}(\mathbf{x}_1) \le \mathbf{N}(\mathbf{x}_2) \tag{2.18}$$

and $f_{r_1, r_2, ..., r_k}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k) = 0$; otherwise where $\mathbf{x}_i = (x_i^1, x_i^2, ..., x_i^p)$, $i = 1, 2, ..., n$ (Bairamov, 2006).

2.2 Progressively Type II Censored Order Statistics

In many life testing, reliability studies and industrial experiments, experiments have to be terminated early or the experimenter might not always obtain complete information on failure times for all experimental units. For example, individuals in a clinical trial may drop out of the study or number of experiments must be limited. There are, however, many situations in which the removal of units prior to failure is pre-planned. One of the main reasons for this is to save time and cost associated with testing. Data obtained from such experiments are called censored data.

2.2.1 Censoring

It is considered a life testing experiment where n items are kept under observation until failure. These items could be some systems, components in reliability study experiments or they could be patients put under certain drug or clinical conditions. Often the event of interest; death of patient, failure of component, recovery of patient has not occurred by the time of recording so that all that is known is that the lifetime for that subject is at least some value and may well be greater than this value. Such censoring cannot be ignored since they carry important information about the effectiveness of the treatment. This introduces a complication in the statistical description and analysis of the data. The most common censoring schemes are Type I and Type II censoring. Other types of censoring are as the following:

- Type I (Time) censoring
- Type II (Failure) censoring
- Random censoring
- Right censoring
- Left censoring
- Simple censoring
- Progressive censoring

2.2.1.1 Type I (Time) Censoring

Suppose the experiment is terminated at a predetermined time t, then only the failure times of the items that failed prior to this time are recorded. Any failures that occur after t are not observed. The data obtained this way denotes a Type-I censored sample. Termination point t of the experiment is fixed and known at the start of the study and the number of failures observed is then random.

2.2.1.2 Type II (Failure) Censoring

If the experimenter decides to terminate the experiment at the rth failure, then a Type II censored sample would be obtained. For example, experimenter assumes n patients in study at the start and experiment terminates after r deaths (do not specify the end of the experiment initially, carry on until r out of n patients are dead, i.e. record $t_{(1)}$, $t_{(2)},\ldots,t_{(r)}$ up to a time $t_{(r)}$ when r have died). In this scenario, only the smallest lifetimes are observed. This type of censoring is less common in medical studies but is widely used in electronic component testing and reliability studies. The number of failures is not random in Type II censoring but is fixed in advance and the endpoint is random. Type II censoring is one part of the censoring schemes which is subject to this thesis.

2.2.1.3 Random Censoring

Suppose each lifetime is censored by an independent time and whether the observation is the life length or the corresponding censoring time is known. Then, random censoring sample is obtained, which is very common in clinical trials. In such experiments, patients enter into the study at random time points, when the experiment itself is terminated at a prespecified time.

2.2.1.4 Right Censoring

In survival analysis, most common type of censoring is right censoring, and occurs when we only know that the survival time exceeds a certain value. Consider a five year study of mortality from cancer. Survival times will be right censored for patients: who are alive at the end of the five year period, drop out or become lost to follow up during the study or die from some other cause during the study. Right censoring is the other part of the censoring schemes which is subject to this thesis.

2.2.1.5 Left Censoring

When the event of interest has already occurred by the time the individual enters the study, left censoring is present. Hence only the time to event that is less than a certain value is known. For example, in a study of times of occurrence of developmental milestones in children, some of the children will already have achieved the milestone prior to entry to the study. In other words, left censoring might occur if subjects are only observed at fixed appointments, and only then it is discovered that death occurred sometime before then, so survival time is less than the period of observation. For example, when the lifetime observed is the time to recurrence of a tumor observable only during surgery.

Censoring times can also be categorized as “simple” or “progressive”.

2.2.1.6 Simple Censoring

When observation of each individual cases at a common, predetermined time, simple Type-I censoring occurs. This situation occurs, for example, in industrial life testing where a batch of identical items are put on test simultaneously and observed for a fixed time interval.

2.2.1.7 Progressive Censoring

When there are potentially different fixed censoring times for each individual, Progressive Type-I censoring occurs. Of the n items put on test, suppose n_1 surviving items at time t_1, n_2 surviving items at time t_2 and n_i surviving items at time t_i are removed where $t_1 < t_2 <\ldots< t_i$ are pre-specified times. The experiment will be terminated at time t_i if n_i exceeds the number of surviving items remaining at that time.

"*In the classical type-II right censoring scheme only the first m failures are observed for a sample of size n whereas for a progressively censored sample, the loss of information of n – m durations (as for the type-II plan) is organized sequentially as it is described subsequently*" (Bordes, 2004).

The scheme of Progressive Type II censoring is important in life-testing experiments. It allows the experimenter to remove units from a life test at various stages during the experiment. The withdrawal of units may be seen as a model describing drop outs of units due to failures, which have causes other than the specific one under study. Progressive censoring schemes are also considered in clinical trials. Here, the drop out of patients may be caused by migration, lack of interest or by ethical decisions. These reasons may be regarded as random withdrawals during the carrying out of the study. For a detailed discussion of progressive censoring can be found in Sen (1986). A saving of costs and of time may be the consequence of such a sampling scheme (Cohen, 1963).

Consider the following Progressive Type II censoring scheme. Assume that the experiment consists of life testing of n units. The first failure is observed. Immediately following the first observed failure, R_1 surviving items are removed from the test at random. Similarly, following the second observed failure, R_2 surviving items are randomly removed (censored) from the test. This process continues until, immediately following the mth observed failure, all the remaining $R_m = n - R_1 - \ldots - R_m - 1 - m$ items are removed from the experiment. In here, $\sum_{i=1}^{m} R_i + m = n$ and $\mathbf{R} = (R_1, R_2,\ldots, R_m)$ are called censoring scheme (Balakrishnan et al., 2000). The resulting m ordered failure times, $X_{1:m:n}^{\mathbf{R}}, X_{2:m:n}^{\mathbf{R}},\ldots, X_{m:m:n}^{\mathbf{R}}$ are called a Progressively Type II right censored order statistics.

Since $X_{1:m:n}^{\mathbf{R}}, X_{2:m:n}^{\mathbf{R}},\ldots, X_{m:m:n}^{\mathbf{R}}$ are the lifetimes of the completely observed items to fail, and that $\mathbf{R} = (R_1, R_2,\ldots, R_m)$ represents the numbers of items withdrawn at these failure times, a Progressive Type II right censoring scheme is specified by integer numbers n, m and $R_1, R_2,\ldots, R_{m\text{-}1}$ with the constraints $n - m - R_1 -\ldots- R_m - 1 \geq 0$ and $n \geq m \geq 1$. A schematic illustration is depicted in Figure 2.1.

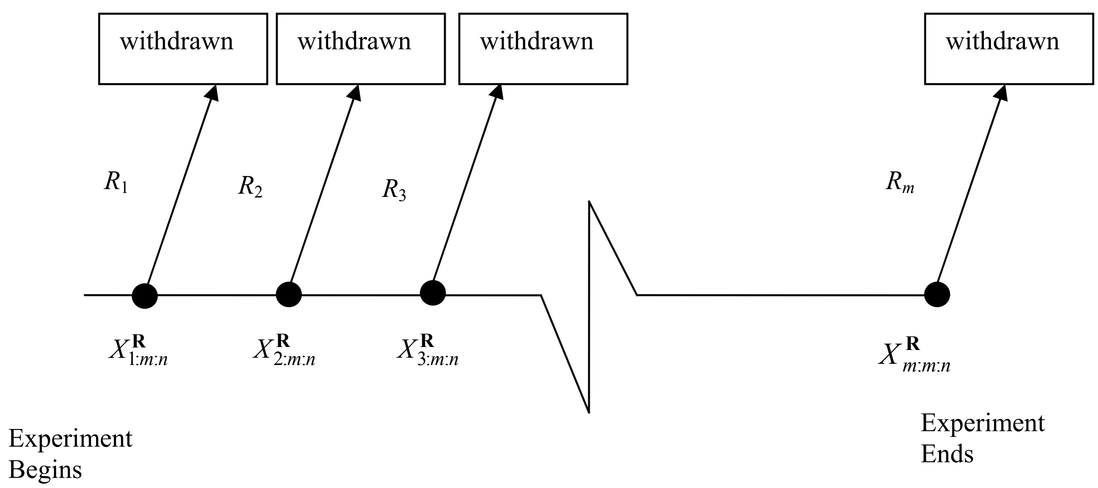

Figure 2.1 Schematic representation of a Progressively Type II right censored life test.

If $R_1 = R_2 = \ldots = R_{m\text{-}1} = 0$, then the usual Type II right censored sample is obtained. Thus the first m order statistics $X_{1:n},\ldots,X_{m:n}$ are observed and the $n - m$ remaining times are right censored by $X_{m:n}$. Moreover if $m = n$, the usual order statistics is obtained.

2.2.2 Distribution Functions of Progressively Type II Right Censored Order Statistics

Let $f(x)$ and $F(x)$ denote the probability density function and the distribution function of an absolutely continuous random variable X, and let $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ be censoring scheme. The probability density function of $(X^{\mathbf{R}}_{1:m:n}, X^{\mathbf{R}}_{2:m:n}, \ldots, X^{\mathbf{R}}_{m:m:n})$ is given in Balakrishnan et al. (2000)

$$f^{\mathbf{R}}_{1,\ldots,m:m:n}(x_1,\ldots,x_m) = c\prod_{j=1}^{m} f(x_j)(1-F(x_j))^{R_j},\ -\infty < x_1 < \ldots < x_m < \infty \qquad (2.19)$$

where $c = n\,(n - R_1 - 1)\ldots(n - R_1 - \ldots - R_{m-1} - m + 1)$. Using the joint probability density function of $(X^{\mathbf{R}}_{1:m:n}, X^{\mathbf{R}}_{2:m:n}, \ldots, X^{\mathbf{R}}_{m:m:n})$ the probability density function of $(X^{\mathbf{R}}_{1:m:n}, X^{\mathbf{R}}_{2:m:n}, \ldots, X^{\mathbf{R}}_{s:m:n})$ is obtained as (Balakrishnan et al., 2000)

$$f^{\mathbf{R}}_{1,\ldots,s:m:n}(x_1,\ldots,x_s) = c'\prod_{j=1}^{s-1} f(x_j)(1-F(x_j))^{R_j}$$

$$\times f(x_s)(1-F(x_s))^{R_s^*}, -\infty < x_1 < \ldots < x_m < \infty, \qquad (2.20)$$

where $c' = n\,(n - R_1 - 1)\ldots(n - R_1 - \ldots - R_{s-1} - s + 1)$ and $R_s^* = n - s - R_1 - \ldots - R_{s-1}$.

The marginal probability density function of $X^{\mathbf{R}}_{s:m:n}$ $(1 \le s \le m \le n)$ is given by (Balakrishnan, Childs et al., 2002)

$$f^{\mathbf{R}}_{s:m:n}(x_s) = c'\sum_{i=0}^{s-1} c_{i,s-1}(R_1+1,\ldots,R_{s-1}+1)$$

$$\times f(x_s)(1-F(x_s))^{R_i''-1}, -\infty < x_s < \infty \qquad (2.21)$$

where $R_i'' = (R_s^* + 1) + \sum_{j=s-i}^{s-1}(R_j + 1)$, $R_s^* = n - s - R_1 - \ldots - R_{s\text{-}1}$ and $\sum_{i=1}^{m} R_i + m = n$.

They obtained the probability density function of $X_{r:m:n}^{\mathbf{R}}$ and $X_{s:m:n}^{\mathbf{R}}$ $(1 \le r < s \le m \le n)$ as

$$f_{r,s:m:n}^{\mathbf{R}}(x_r, x_s) = c' \sum_{i=0}^{r-1} \sum_{k=0}^{s-r-1} c_{i,r-1}(R_1 + 1, \ldots, R_{r-1} + 1) c_{k,s-r-1}(R_{r+1}, \ldots, R_{s-1})$$

$$\times f(x_r)(1 - F(x_r))^{R_{ik}' - 1} f(x_s)(1 - F(x_s))^{R_k'' - 1}, -\infty < x_r < x_s < \infty \quad (2.22)$$

where $c' = n\ (n - R_1 - 1)\ldots(n - R_1 - \ldots - R_{s\text{-}1} - s + 1)$ and $R_s^* = n - s - R_1 - \ldots - R_{s\text{-}1}$. $R_{ik}' = \sum_{j=r-i}^{s-1-k}(R_j + 1)$ and $R_k'' = (R_s^* + 1) + \sum_{j=s-i}^{s-1}(R_j + 1)$.

In the next subsection the distributional properties of the conditionally ordered random vectors will be investigated under a Progressive Type II censoring scheme.

2.2.3 Progressively Type II Censored Order Statistics for Multivariate Observations

For a sequence of i.i.d. random vectors $\mathbf{X}_i = (X_i^1, X_i^2, \ldots, X_i^p)$, $i = 1, 2, \ldots, n$, it is considered the conditional ordering of these random vectors with respect to the magnitudes of $\mathbf{N}(\mathbf{X}_i)$, $i = 1, 2, \ldots, n$. Also the Progressive Type II right censoring for multivariate observations using conditional ordering is considered in the paper.

Let a system have n independent components each consisting of p arbitrarily dependent elements. Consider $\mathbf{X}_i = (X_i^1, X_i^2, \ldots, X_i^p)$, $i = 1, 2, \ldots, n$, where X_i^j denotes the life length of the jth element of the ith unit. Let $\mathbf{R}$ = (R_1, R_2,…, R_m), where

$\sum_{i}^{m} R_i + m = n$. It is assumed that $\mathbf{X}_i$, $i = 1, 2, ..., n$ are continuous random vectors with the p – variate distribution function $F(\mathbf{x})$ and probability density function $f(\mathbf{x})$, $\mathbf{x} \in IR^p$. The n units are placed on the test at time zero under Progressive Type II censoring scheme. $\mathbf{X}^{\mathbf{R}}_{1:m:n}$ is denoted as the vector of the life length corresponding to first failure time. Immediately following the first failure, R_1 surviving units are removed from the test at random. The second failure in the system (or first failure among the remaining n – R_1 – 1 items) occurs at time $\mathbf{N}(\mathbf{X}^{\mathbf{R}}_{2:m:n})$ and immediately R_2 items are randomly removed from the test. This process continues until, at time of the mth observed failure, the remaining $R_m = n - R_1 - \ldots - R_{m-1} - m$ units are all removed from the experiment. The times of failures are $\mathbf{N}(\mathbf{X}^{\mathbf{R}}_{1:m:n}), \mathbf{N}(\mathbf{X}^{\mathbf{R}}_{2:m:n}), ..., \mathbf{N}(\mathbf{X}^{\mathbf{R}}_{m:m:n})$ and the random variables $\mathbf{X}^{\mathbf{R}}_{1:m:n}, \mathbf{X}^{\mathbf{R}}_{2:m:n}, ..., \mathbf{X}^{\mathbf{R}}_{m:m:n}$ are called multivariate Progressively Type II censored conditionally N-ordered statistics.

If $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ is i.i.d. random vectors with distribution function $F(\mathbf{x})$ and probability density function $f(\mathbf{x})$, $\mathbf{x} \in IR^p$, then the joint probability density function of the Progressively Type II conditionally N-ordered statistics $\mathbf{X}^{\mathbf{R}}_{1:m:n}$, $\mathbf{X}^{\mathbf{R}}_{2:m:n}$, ..., $\mathbf{X}^{\mathbf{R}}_{m:m:n}$ is

$$f^{\mathbf{R}}_{1,...,m:m:n}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m) = c \prod_{i=1}^{m} f(\mathbf{x}_i)[1 - h(\mathbf{x}_i)]^{R_i}$$

$$\text{if } \mathbf{N}(\mathbf{x}_1) \le \mathbf{N}(\mathbf{x}_2) \le \cdots \le \mathbf{N}(\mathbf{x}_m) \qquad (2.23)$$

and $f^{\mathbf{R}}_{1,...,m:m:n}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m) = 0$; otherwise, where $c = n\,(n - R_1 - 1)\ldots(n - R_1 - \ldots - R_{m-1} - m + 1)$ and $h(\mathbf{x}) = P\{\mathbf{N}(\mathbf{X}) \le \mathbf{N}(\mathbf{x})\}$ (Bairamov, 2006).

The joint probability density function of $\mathbf{X}^{\mathbf{R}}_{1:m:n}, \mathbf{X}^{\mathbf{R}}_{2:m:n}, \ldots, \mathbf{X}^{\mathbf{R}}_{r:m:n}$ $(1 \le r \le m \le n)$ is

$$f^{\mathbf{R}}_{1,\ldots,r:m:n}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r) = c_r \prod_{i=1}^{r} f(\mathbf{x}_i)$$

$$\times \prod_{i=1}^{r-1} [1 - h(\mathbf{x}_i)]^{R_i} [1 - h(\mathbf{x}_r)]^{n - R_1 - R_2 - \cdots - R_{r-1} - r + 1} \tag{2.24}$$

where $c_r = n\,(n - R_1 - 1)\ldots(n - R_1 - \ldots - R_{r-1} - r + 1)$ (Bairamov, 2006).

2.3 Generalized Order Statistics

Generalized order statistics have been introduced by Kamps (1995a, 1995b) to unify several models of ordered random variables, e.g. ordinary order statistics, Progressively Type II censored order statistics, records, sequential order statistics. “In the distribution theoretical sense, all models, or appropriately restricted versions of these models, are contained in the model of generalized order statistics” (Kamps, 1995a, p. 49). The common approach makes it possible to define several distributional properties at once. These models can be effectively applied, e.g. in reliability theory.

“First, uniform generalized order statistics are defined via some joint probability density function on a cone of the IR^n. Using the quantile transformation we are able to define generalized order statistics based on an arbitrary distribution function $F(x)$” (Kamps, 1995a). Assuming an absolutely continuous distribution function $F(x)$ and choosing the parameters in the joint probability density of generalized order statistics appropriately, the densities corresponding to the models of ordered random variables mentioned above can be obtained.

2.3.1 Distribution Functions of Generalized Order Statistics

Let $n \in N$, k > 0, $m_1, \ldots, m_{n-1} \in IR$, $M_r = \sum_{j=r}^{n-1} m_j$, $1 \le r \le n-1$, be parameters such that $\gamma_r = k + n - r + M_r > 0$ for all $r \in \{1, \ldots, n-1\}$, and let $\tilde{m} = (m_1, m_2, \ldots, m_{n-1})$, if $n \ge 2$, $\tilde{m} \in IR$, if $n = 1$, arbitrary. If the random variables $U(r, n, \tilde{m}, k)$, $r = 1, \ldots, n$, possess a joint probability density function of the form

$$f^{U(1,n,\tilde{m},k),\ldots,U(n,n,\tilde{m},k)}(u_1,\ldots,u_n) = k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left(\prod_{i=1}^{n-1}(1-u_i)^{m_i}\right)(1-u_n)^{k-1} \tag{2.25}$$

on a cone of IR^n with the property $0 \le u_1 < \ldots < u_n \le 1$, then they are called uniform generalized order statistics (Kamps, 1995a).

If $F(x)$ is an arbitrary distribution function, then the random variables $X(r, n, \tilde{m}, k) = F^{-1}(U(r, n, \tilde{m}, k))$, $r = 1, 2, \ldots, n-1$, are called generalized order statistics based on $F(x)$. The joint probability density function of generalized order statistics $X(1, n, \tilde{m}, k), X(2, n, \tilde{m}, k), \ldots, X(n, n, \tilde{m}, k)$ is given by

$$f^{X(1,n,\tilde{m},k),\ldots,X(n,n,\tilde{m},k)}(x_1,\ldots,x_n) = k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i} f(x_i)\right)$$

$$\times (1-F(x_n))^{k-1} f(x_n) \tag{2.26}$$

on the cone $F^{-1}(0) < x_1 \le \ldots \le x_n < F^{-1}(1)$ (Kamps, 1995a).

If $c_{r-1}=\prod_{j=1}^{r}\gamma_j$, $1 \le i \le r \le n$ and $a_i(r)=\prod_{j=1, j\neq i}^{r}(1/(\gamma_j-\gamma_i))$ then the marginal probability density function and distribution function of $X(r,n,\widetilde{\mathrm{m}},k)$ is given by

$$f^{X(r,n,\widetilde{m},k)}(x)=c_{r-1}f(x)\sum_{i=1}^{r}a_i(r)(1-F(x))^{\gamma_i-1} \tag{2.27}$$

and

$$F^{X(r,n,\widetilde{m},k)}(x)=1-c_{r-1}\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}(1-F(x))^{\gamma_i} \tag{2.28}$$

where an empty product in $a_i(r)$ denotes 1, $\forall i \neq j$ and $\gamma_i \neq \gamma_j$.

Kamps (1995a) also gives the joint probability density function of $X(r,n,\widetilde{\mathrm{m}},k)$ and $X(s,n,\widetilde{\mathrm{m}},k)\,(1 \le r < s \le n)$ as

$$f^{X(r,n,\widetilde{m},k),X(s,n,\widetilde{m},k)}(x_r,x_s)=c_{s-1}\left(\sum_{j=r+1}^{s}a_j^{(r)}(s)\left(\frac{1-F(x_s)}{1-F(x_r)}\right)^{\gamma_j}\right)$$

$$\times\left(\sum_{i=1}^{r}a_i(r)\big(1-F(x_r)\big)^{\gamma_i}\right)\frac{f(x_r)}{(1-F(x_r))}\frac{f(x_s)}{(1-F(x_s))} \tag{2.29}$$

where $r+1 \le j \le s$ and $a_j^{(r)}(s)=\prod_{i=r+1, i\neq j}^{s}1/(\gamma_i-\gamma_j)$.

Kamps (1995a) introduces some notations for the sake of brevity to give different representation of the distribution function of generalized order statistics.

i) Let $c_{r-1} = \prod_{j=1}^{r} \gamma_j$, $r = 1,\ldots, n$ and $\gamma_n = k$, $M_n = 0$.

ii) The function g_m is defined by

$$g_m(x) = \begin{cases} \dfrac{1}{m+1}(1-(1-x)^{m+1}), \; m \neq -1 \\ \\ \log\left(\dfrac{1}{1-x}\right), \qquad\qquad m = -1 \end{cases} \qquad x \in [0,1) .$$

The following is a representation of the marginal distribution function $\Phi_{r,n}$ of the uniform generalized order statistics $U(r,n,\widetilde{\mathrm{m}},k)$ and by this the distribution function of the *r*th generalized order statistics $X(r,n,\widetilde{\mathrm{m}},k)$ based on an arbitrary distribution function $F(x)$ has been proved in Kamps (1995a).

Let $m_1 = \ldots = m_{r-1} = m$ and $r \in \{1,\ldots, n\}$. Then the marginal distribution function is given by

$$\Phi_{r,n}(x) = 1 - c_{r-1}(1-x)^{k+n-r+M_r} \sum_{i=0}^{r-1} \frac{1}{i!c_{r-i-1}} g_m^i(x), \; x \in (0,1) \tag{2.30}$$

and $F^{X(r,n,\widetilde{m},k)}(x) = \Phi_{r,n}(F(x))$.

A different representation of the distribution function of the *r*th generalized order statistics is given by Nasri-Roudsari (1996). Let $r \in \{1,\ldots, n\}$, $m_1 = \ldots = m_{r-1} = m > -1$ and let $F(x)$ be an arbitrary distribution function. Then the marginal distribution function of the *r*th generalized order statistics $X(r,n,\widetilde{\mathrm{m}},k)$ has the representation

$$F^{X(r,n,\widetilde{m},k)}(x) = I_{1-(1-F(x))^{m+1}}\left(r, \frac{k+n-r+M_r}{m+1}\right), \; x \in IR. \tag{2.31}$$

Some of the particular cases are as follows:

i) In the case $m_1 = \ldots = m_{r-1} = m > -1$, $M_r = \sum_{j=r}^{n-l} m_j = (n-r)m$ for all $r \in \{1, \ldots, n-1\}$. Thus equation (2.31) reads

$$F^{X(r,n,\widetilde{m},k)}(x) = I_{1-(1-F(x))^{m+1}}\left(r, n-r+\frac{k}{m+1}\right), \quad r \in \{1, 2\ldots, n\}. \tag{2.32}$$

For $r = n$, equation (2.32) becomes

$$F^{X(n,n,\widetilde{m},k)}(x) = I_{1-(1-F(x))^{m+1}}\left(n, \frac{k}{m+1}\right). \tag{2.33}$$

ii) $m_1 = \ldots = m_{n-1} = 0$ and $k = 1$, the generalized order statistics coincide with ordinary order statistics (Kamps, 1995a). Indeed, (2.31) reduces in this case to the well known result for the distribution function of the *r*th ordinary order statistic $F_{r:n}$ (David, 1981),

$$F_{r:n}(x) = F^{X(r,n,0,1)}(x) = I_{F(x)}(r, n-r+1), \; r \in \{1, 2\ldots, n\}. \tag{2.34}$$

Let $I_x(a,b) = (1/B(a,b))\int_0^x t^{a-1}(1-t)^{b-1}dt$ denote the incomplete beta function, and let $m_1 = m_2 = \ldots = m_{n-r} = m$. Furthermore, let $M_n = (1/(r-1))\sum_{j=n-r+1}^{n-1} m_j$ be the arithmetic mean of the parameters $m_{n-r+1}, \ldots, m_{n-1}$. If one of the conditions

C_1: $m > -1$ and $M_n \geq -1$ and

C_2 : $m > -1$ and $M_n < -1$ and $k > -(M_n + 1)\max(r-1, n-r)$

is satisfied, then for any $r \in \{2,\ldots, n\}$,

$$\Phi^{(m,k)}_{n-r+1,n}(x) = I_{G_m(x)}(N^* - R^*_n + 1, R^*_n), \tag{2.35}$$

where

$$R^*_n = \tfrac{k}{m+1} + \left(\tfrac{M_n+1}{m+1}\right)r - \left(\tfrac{M_n+1}{m+1}\right), \tag{2.36}$$

$$N^* = n + \tfrac{k}{m+1} + \tfrac{(M_n - m)r - (M_n + 1)}{m+1}, \tag{2.37}$$

and $G_m(x)$=1 – (1 – $(F(x))^{m+1}$ is a distribution function. Moreover, if $m > -1$ and $r = 1$, then (2.35) is satisfied with (2.36), (2.37) and $M_n = 0$.

In the particular case, when $m_1 = \ldots = m_{n-1} = m(> -1)$, then $M_n = m$. This leads to $R^*_n = k / (m + 1) + r - 1 = R$ and $N^* = n + k / (m + 1) - 1 = N$.

$$\Phi^{(m,k)}_{n-r+1,n}(x) = I_{G_m(x)}(n - r + 1, \frac{\gamma_{n-r+1}}{m+1}) = I_{G_m(x)}(N - R + 1, R). \tag{2.38}$$

The notation of generalized order statistics provides a unified approach to some distributional properties of ordered random variables. The model of generalized order statistics contains many models of ordered random variables as special cases.

2.3.2 Special Cases of Generalized Order Statistics

1. In the case $m_1 = \ldots = m_{n-1} = 0$ and $k = 1$ (i.e. $\gamma_r = n - r + 1$, $1 \leq r \leq n - 1$), the model reduces to the joint probability density function of ordinary order statistics from n i.i.d. random variables $X_1, X_2, \ldots, X_n$ with distribution function $F(x)$ (David, 1981)

$$f^{X(1,n,0,1),\ldots,X(n,n,0,1)}(x_1,\ldots,x_n) = n!\left(\prod_{i=1}^{n} f(x_i)\right), -\infty < x_1 \leq \cdots \leq x_n < \infty . \qquad (2.39)$$

2. Choosing $m_i = R_i$ and $k = R_n + 1$ (i.e. $\gamma_r = n - r + 1$, $1 \leq r \leq n - 1$) the joint probability density function of Progressively Type II censored order statistics is obtained

$$f^{X(1,n,R,R_n+1),\ldots,X(n,n,R,R_n+1)}(x_1,\ldots,x_n) = c'\left(\prod_{i=1}^{n}(1-F(x_i))^{R_i} f(x_i)\right)$$

$$-\infty < x_1 \leq \cdots \leq x_n < \infty \qquad (2.40)$$

where $c' = n\,(n - R_1 - 1)\ldots(n - R_1 - \ldots - R_{n-1} - n + 1)$.

3. Choosing $m_1 = \ldots = m_{n-1} = 0$ and $k = \alpha - n + 1$ with $n - 1 < \alpha \in IR$ (i.e. $\gamma_r = \alpha - r + 1$, $1 \leq r \leq n - 1$) order statistics with non-integral sample size is

$$f^{X(1,n,0,\alpha-n+1),\ldots,X(n,n,0,\alpha-n+1)}(x_1,\ldots,x_n) = \left(\prod_{j=1}^{n}\alpha - j + 1\right)(1-F(x_n))^{\alpha-n}\prod_{i=1}^{n} f(x_i)$$

$$-\infty < x_1 \leq \cdots \leq x_n < \infty . \qquad (2.41)$$

4. When $m_i = (n-i+1)\,\alpha_i - (n-i)\,\alpha_{i+1} - 1$ where $\alpha_1,\ldots,\alpha_n$ are positive real numbers, $i = r = 1,\ldots, n-1$ and $k = \alpha_n$ (i.e. $\gamma_r = (n-r+1)\,\alpha_r$, $1 \le r \le n-1$), the joint probability density function of sequential order statistics

$$f^{X(1,n,\widetilde{m},\alpha_n),\ldots,X(n,n,\widetilde{m},\alpha_n)}(x_1,\ldots,x_n) = n!\left(\prod_{j=1}^{n}\alpha_j\right)\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i}f(x_i)\right)$$

$$\times(1-F(x_n))^{\alpha_n-1}f(x_n), -\infty < x_1 \le \cdots \le x_n < \infty \quad (2.42)$$

based on $F_r(t) = 1-(1-F(t))^{\alpha_r}$, $1 \le r \le n$, where $F(x)$ is an arbitrary, absolutely continuous distribution function is obtained.

5. In the case $m_1 = \ldots = m_{n-1} = -1$ and $k = 1$ (i.e. $\gamma_r = 1$, $1 \le r \le n-1$), the joint probability density function of the first r record values $X_{L(1)},\ldots, X_{L(r)}$

$$f^{X(1,n,-1,1),\ldots,X(r,n,-1,1)}(x_1,\ldots,x_r) = \left(\prod_{i=1}^{r-1}\frac{f(x_i)}{(1-F(x_i))}\right)\times f(x_r),$$

$$-\infty < x_1 \le \ldots \le x_r < \infty \quad (2.43)$$

based on the sequence $(X_i)_{i\in\mathrm{N}}$ of i.i.d. random variables with distribution function $F(x)$ is obtained.

6. In the case $m_1 = \ldots = m_{n-1} = -1$ and $k \in \mathrm{N}$ (i.e. $\gamma_r = k$, $1 \le r \le n-1$), the joint probability density function of the first n kth record values

$$f^{X(1,n,-1,k),\ldots,X(n,n,-1,k)}(x_1,\ldots,x_n) = k^n\left(\prod_{i=1}^{n-1}\frac{f(x_i)}{(1-F(x_i))}\right)(1-F(x_n))^{k-1}$$

$$\times f(x_n), -\infty < x_1 \le \cdots \le x_n < \infty \tag{2.44}$$

based on the sequence $(X_i)_{i\in \mathrm{N}}$ of i.i.d. random variables with distribution function $F(x)$ is obtained.

Table 2.1 The special cases of the generalized order statistics

Model	Vector of $\widetilde{\mathrm{m}}$	Parameter of k
1. Order statistics	$m_i = 0$ for $i = 1,2,...,n-1$	$k = 1$
2. Progressively Type II censored order statistics	$m_i = R_i$ for $i = 1,2,...,n-1$	$k = R_n + 1$
3. Order statistics with non-integral sample size	$m_i = 0$ for $i = 1,2,...,n-1$	$k = \alpha - n + 1$
4. Sequential order statistics	$m_i = (n-i+1)\alpha_i - (n-i)\alpha_{i+1} - 1$ for $i = 1,2,...,n-1$ and $\alpha_1,...,\alpha_n > 0$	$k = \alpha_n$
5. Records	$m_i = -1$ for $i = 1,2,...,n-1$	$k = 1$
6. *k*th record values	$m_i = -1$ for $i = 1,2,...,n-1$	$k \in \{1,2,...\}$

2.4 Concomitants of Order Statistics and Progressively Type II Censored Order Statistics

After giving the definition of concomitants, concomitants of order statistics are represented in Section 2.4.1. In the following section, concomitants of Progressively Type II censored order statistics are given.

Let $(X_1,Y_1),...,(X_n,Y_n)$ be a random sample of size *n* from a continuous bivariate distribution. If the X's are arranged in ascending order as $X_1 \le X_2 \le ... \le X_n$ then the Y's associated with these order statistics are denoted by $Y_{[1]}, Y_{[2]},...,Y_{[n]}$ and are called concomitants of order statistics. They are also known as induced order statistics in the literature. The concomitants are of interest in selection and prediction problems based on the ranks of the X's. For example, when $k(< n)$ individuals having the highest X-scores are selected, it may be wished to know the behavior of the corresponding Y-scores.

Properties of concomitants have been studied by many authors (e.g., Bhattacharya 1974; Sen, 1976, 1981; David, O'Connell, & Yang, 1977; Yang, 1977; Goel & Hall, 1994; Nagaraja & David, 1994). See David & Nagaraja (2003) for an overview. Applications of concomitants include their use in estimating correlation (Barnett, Green, & Robinson, 1976), in ranking and selection (Yeo & David, 1984; David, 1993), and ranked set sampling (Stokes, 1977). They are also of interest in a variety of estimation problems (Bhattacharya, 1984).

2.4.1 Concomitants of Order Statistics

Under the assumption that X and Y are linearly related, apart from an independent error term, the small sample theory of the concomitants of order statistics has been discussed in David (1973). The asymptotic distribution theory for the bivariate normal distribution has been investigated by David & Galambos (1974). Bhattacharya (1974), Sen (1976), and Yang (1977) obtained results for the general asymptotic distribution theory of concomitants of order statistics. Balasubramanian & Beg (1997) studied the concomitants of order statistics for Morgenstern type bivariate exponential distribution. Bairamov & Bekci (1999) have studied distribution and recurrence relation between moments of concomitants of order statistics in the bivariate Farlie–Gumbel–Morgenstern (FGM) with uniform marginals. Eryılmaz & Bairamov (2003) derived the distribution of the new sample rank of an order statistic and its concomitant. For a comprehensive review of this topic see Bhattacharya (1984) and David & Nagaraja (1998).

Let $f(y|x)$ denote the conditional probability density function of Y given $X = x$ and let $f_{r_1,\ldots,r_k}(x_1,\ldots,x_k)$ denote the joint probability density function of $X_{(r_1:n)},\ldots,X_{(r_k:n)}$ with $1 \le r_1 \le \ldots \le r_k \le n$. Then, as discussed in Yang (1977), the joint probability density function of the k-concomitants $(Y_{[r_1:n]},\ldots,Y_{[r_k:n]})$, $(1 \le k \le n)$ is

$$f_{Y_{[r_1]},\ldots,Y_{[r_k]}}(y_1,\ldots,y_k) = \int_{-\infty}^{+\infty}\int_{-\infty}^{x_k}\cdots\int_{-\infty}^{x_2}\prod_{i=1}^{k} f(y_i|x_i) f_{r_1,\ldots,r_k}(x_1,\ldots,x_k)\prod_{i=1}^{k} dx_i \ . \qquad (2.45)$$

From this the marginal probability density function of the rth concomitant $Y_{[r:n]}$ is obtained as

$$f_{Y_{[r:n]}}(y) = \int_{-\infty}^{+\infty} f(y|x) f_{r:n}(x)\, dx \ , \qquad (2.46)$$

where $f_{r:n}(x)$ is the probability density function $X_{(r)}$.

The joint probability density function of the rth order statistic $X_{r:n}$ and its concomitant $Y_{[r:n]}$ is represented as follows

$$f_{X_{r:n},Y_{[r:n]}}(x,y) = f(y|x) f_{r:n}(x) \ . \qquad (2.47)$$

Concomitants can also be defined in the case of generalized order statistics, or Progressively Type II censored order statistics. In Kamps (1995a) the distribution of concomitant is given for Pfeifer's record model.

2.4.2 Concomitants of Progressively Type II Censored Order Statistics

Recently, Bairamov & Eryılmaz (2006) gave recurrence relations between moments of concomitants for the model of Progressive Type II censoring. They unify and extend some results on spacings, exceedances and concomitants for the model of progressive censoring. The probability density function of *r*th Progressively Type II censored concomitant is given by

$$f^{Y_{r:n}^{\mathbf{R}}}(y) = \int_{-\infty}^{\infty} f^{X_{r:n}^{\mathbf{R}}}(x) f(y|x)dx \tag{2.48}$$

where $f(y\,|\,x)$ is the conditional probability density function of Y given $X = x$ and $f^{X_{r:n}^{\mathbf{R}}}(x)$ is the probability density function of the *r*th Progressively Type II censored order statistic. The probability density function of $Y_{[r:n]}^{\mathbf{R}}$ can be represented as follows

$$f^{Y_{[r:n]}^{\mathbf{R}}}(y) = c_{r-1}\sum_{i=1}^{r}\frac{a_{i,r}}{\gamma_i} f_{Y_{[1:\gamma_i]}}(y), \tag{2.49}$$

where $c_{r-1} = \prod_{j=1}^{r}\gamma_j$ with $\gamma_j = N - \sum_{i=1}^{j-1} R_i - j + 1$, $a_{i,r} = \prod_{j=1}^{r}\frac{1}{\gamma_j - \gamma_i}$, $N = n + \sum_{j=1}^{n} R_j$ and $f_{Y_{[1:\gamma_i]}}(y)$ is the probability density function of the first concomitant of ordinary statistic from a sample size γ_i. Then the *k*th moment of $Y_{[r:n]}^{\mathbf{R}}$ is given by Bairamov & Eryılmaz (2006)

$$\mu_{r,\mathbf{R}}^{(k)} = c_{r-1}\sum_{i=1}^{r}\frac{a_{i,r}}{\gamma_i}\mu_{\gamma_i}^{(k)} \tag{2.50}$$

where $\mu_{\gamma_i}^{(k)} = EY_{[1:\gamma_i]}^{k}$.

CHAPTER THREE
DISTRIBUTIONS OF NONADJACENT PROGRESSIVELY TYPE II RIGHT CENSORED ORDER STATISTICS

In this chapter, the joint probability density function of nonadjacent Progressively Type II censored order statistics is presented by using Lemma 3.1. After giving a lemma which is used to obtain the main theorem in Section 3.1, the main result which is the probability density function of nonadjacent Progressively Type II censored order statistics is presented in Theorem 3.1 in Section 3.2. In this section, some examples are given.

3.1 Distributional Properties of Progressively Type II Censored Order Statistics

The joint probability density function of nonadjacent Progressively Type II right censored order statistics, $X^{\mathbf{R}}_{r_1:m:n}, X^{\mathbf{R}}_{r_2:m:n}, \ldots, X^{\mathbf{R}}_{r_k:m:n}$, $1 \le k \le m$ from a continuous distribution is obtained. The result is a generalization of the work of Balakrishnan, Childs et al. (2002).

The following useful lemma which is established in Balakrishnan, Childs et al. (2002) paves a way to evaluate densities of distinct collections among Progressively Type II censored random variables that is essential for the result which will be given in equation (3.1).

Lemma 3.1 Let $f(x)$ and $F(x)$ denote the probability density function and the distribution function of an absolutely continuous random variable X, respectively and let $a_j > 0$ for $j = 1, 2, \ldots, r$. Then for $r \ge 1$, the following equality can be written

$$\int_T^{x_{r+1}} \cdots \int_T^{x_3} \int_T^{x_2} \prod_{j=1}^{r} f(x_j)\{1 - F(x_j)\}^{a_j - 1} dx_1 dx_2 \ldots dx_r$$

$$= \sum_{i=0}^{r} c_{i,r}(\mathbf{a}_r)\{1-F(x_{r+1})\}^{b_{i,r}(\mathbf{a}_r)}\{1-F(T)\}^{\sum_{j=1}^{r-i} a_j},$$

where $\mathbf{a}_r = (a_1, a_2, ..., a_r)$,

$$c_{i,r}(\mathbf{a}_r) = \frac{(-1)^i}{\left\{\prod_{j=1}^{i}\sum_{k=r-i+1}^{r-i+j} a_k\right\}\left\{\prod_{j=1}^{r-i}\sum_{k=j}^{r-i} a_k\right\}}, \quad b_{i,r}(\mathbf{a}_r) = \sum_{j=r-i+1}^{r} a_j$$

in which the usual conventions that $\prod_{j=1}^{0} d_j \equiv 1$ and $\sum_{j=i}^{i-1} d_j \equiv 0$ are adopted.

3.2 The Joint Distribution of Nonadjacent Progressively Type II Censored Order Statistics

The joint probability density function of $X_{r:m:n}^{\mathbf{R}}$ and $X_{s:m:n}^{\mathbf{R}}$ is given in Balakrishnan, Childs et al. (2002). This result is improved, and the joint probability density function for *k* Progressively Type II right censored order statistics is obtained.

Theorem 3.1 Let *f*(*x*) and *F*(*x*) denote the probability density function and the distribution function of an absolutely continuous random variable *X*. Let $f_{r_1,r_2,...,r_k:m:n}^{\mathbf{R}}(x_1, x_2, ..., x_k)$ denote the joint probability density function of the $X_{r_1:m:n}^{\mathbf{R}}, X_{r_2:m:n}^{\mathbf{R}}, ..., X_{r_k:m:n}^{\mathbf{R}}$. The probability density function of $X_{r_1:m:n}^{\mathbf{R}}, X_{r_2:m:n}^{\mathbf{R}}, ..., X_{r_k:m:n}^{\mathbf{R}}$ ($1 \le r_1 < r_2 < \ldots < r_k \le m \le n$) is then given by

$$f^{\mathbf{R}}_{r_1,r_2,\ldots,r_k:m:n}(x_1,x_2,\ldots,x_k)=c'f(x_1)f(x_k)\{1-F(x_1)\}^{R_{r_1}}$$

$$\times\{1-F(x_k)\}^{R^*_{r_k}}\left[\sum_{z=2}^{k-1}f(x_z)\{1-F(x_z)\}^{R_z}\right]$$

$$\times\prod_{z=1}^{k-1}\sum_{i=0}^{r_1-1}\sum_{t=0}^{r_{z+1}-r_z-1}c_{i,r_1-1}(R_1+1,\ldots,R_{r_1-1}+1)$$

$$\times c_{t,r_{z+1}-r_z-1}(R_{r_1+1},\ldots,R_{r_{z+1}-1})\{1-F(x_1)\}^{\sum_{j=r_1-i}^{r_1-1}(R_j+1)}$$

$$\times\{1-F(x_z)\}^{\sum_{j=r_z+1}^{r_{z+1}-t-1}(R_j+1)}\{1-F(x_{z+1})\}^{\sum_{j=r_{z+1}-t}^{r_{z+1}-1}(R_j+1)}$$

$$-\infty<x_1<\ldots<x_k<\infty\,, \tag{3.1}$$

where $c'=n(n-R_1-1)(n-R_1-R_2-2)\ldots(n-R_1-R_2-\ldots-R_{k-1}-k+1)$ and $R^*_k = n-k-R_1-R_2-\ldots-R_{k-1}$. The conventions $\prod_{j=1}^{0}d_j\equiv 1$ and $\sum_{j=i}^{i-1}d_j\equiv 1$ are adopted .

Proof. First, consider the case when $k = 3$, i.e. find the joint probability density function of $X^{\mathbf{R}}_{r:m:n}, X^{\mathbf{R}}_{k:m:n}, X^{\mathbf{R}}_{s:m:n}$, $r < k < s$. If the joint probability density function of $X^{\mathbf{R}}_{1:m:n}, X^{\mathbf{R}}_{2:m:n},\ldots, X^{\mathbf{R}}_{s:m:n}$ which is given in (2.20) is integrated out with respect to x_r, x_k, x_s then the joint probability density function of $X^{\mathbf{R}}_{r:m:n}, X^{\mathbf{R}}_{k:m:n}, X^{\mathbf{R}}_{s:m:n}$ can be given by

$$\int_0^{x_r}\cdots\int_0^{x_3}\int_0^{x_2}\int_{x_r}^{x_k}\cdots\int_{x_r}^{x_{r+3}}\int_{x_r}^{x_{r+2}}\int_{x_k}^{x_s}\cdots\int_{x_k}^{x_{k+3}}\int_{x_k}^{x_{k+2}}c'f(x_1)\ldots f(x_{r-1})f(x_r)$$

$$\times f(x_{r+1})...f(x_{k-1})\,f(x_k)\,f(x_{k+1})...f(x_{s-1})\,f(x_s)$$

$$\times\{1-F(x_1)\}^{R_1}...\{1-F(x_{r-1})\}^{R_{r-1}}\{1-F(x_r)\}^{R_r}$$

$$\times\{1-F(x_{r+1})\}^{R_{r+1}}...\{1-F(x_{k-1})\}^{R_{k-1}}\{1-F(x_k)\}^{R_k}$$

$$\times\{1-F(x_{k+1})\}^{R_{k+1}}...\{1-F(x_{s-1})\}^{R_{s-1}}\{1-F(x_s)\}^{R_s^*}\,d_{x_{k+1}}...d_{x_{s-1}}\,d_{x_{r+1}}...d_{x_{k-1}}\,d_{x_1}...d_{x_{r-1}}$$

$$x_1<\cdots<x_{r-1}<x_r<x_{r+1}<\cdots<x_{k-1}<x_k<x_{k+1}<\cdots<x_{s-1}<x_s\,.$$

If above expression is arranged then the probability density function is obtained by

$$f^{\mathbf{R}}_{r,k,s:m:n}(x_r,x_k,x_s)=c'f(x_r)\,f(x_k)\,f(x_s)\{1-F(x_r)\}^{R_r}\{1-F(x_k)\}^{R_k}\{1-F(x_s)\}^{R_s^*}$$

$$\int_0^{x_r}\cdots\int_0^{x_3}\int_0^{x_2}\int_{x_r}^{x_k}\cdots\int_{x_r}^{x_{r+3}}\int_{x_r}^{x_{r+2}}\int_{x_k}^{x_s}\cdots\int_{x_k}^{x_{k+3}}\int_{x_k}^{x_{k+2}}c'f(x_1)...f(x_{r-1})\,f(x_{r+1})...f(x_{k-1})$$

$$\times f(x_{k+1})...f(x_{s-1})\{1-F(x_1)\}^{R_1}...\{1-F(x_{r-1})\}^{R_{r-1}}\{1-F(x_{r+1})\}^{R_{r+1}}...\{1-F(x_{k-1})\}^{R_{k-1}}$$

$$\times\{1-F(x_{k+1})\}^{R_{k+1}}...\{1-F(x_{s-1})\}^{R_{s-1}}\,d_{x_{k+1}}...d_{x_{s-1}}\,d_{x_{r+1}}...d_{x_{k-1}}\,d_{x_1}...d_{x_{r-1}}\,. \qquad (3.2)$$

Using Lemma 2.1 and taking T as the lower endpoints to integrate out the variables $x_1, x_2, \ldots, x_{r-1}$ in (3.2), we have

$$\int_T^{x_r} \cdots \int_T^{x_3} \int_T^{x_2} \{1-F(x_1)\}^{R_1} \ldots \{1-F(x_{r-1})\}^{R_{r-1}} dF(x_1) \ldots dF(x_{r-1})$$

$$= \sum_{i=0}^{r-1} c_{i,r-1}(R_1+1,\ldots,R_{r-1}+1)\{1-F(x_r)\}^{\sum_{j=r-i}^{r-1}(R_j+1)} .$$

To integrate out the first part of the remaining variables x_{r+1}, $x_{r+2}, \ldots$, x_{k-1}, again Lemma 3.1 is used, this time with $T = x_r$ and a relabeling of the variables, i.e., it is used the fact that

$$\int_{x_r}^{x_k} \cdots \int_{x_r}^{x_{r+3}} \int_{x_r}^{x_{r+2}} \{1-F(x_{r+1})\}^{R_{r+1}} \ldots \{1-F(x_{k-1})\}^{R_{k-1}} dF(x_{r+1}) \ldots dF(x_{k-1})$$

$$= \sum_{t=0}^{k-r-1} c_{t,k-r-1}(R_{r+1},\ldots,R_{k-1})\{1-F(x_k)\}^{\sum_{j=k-t}^{k-1}(R_j+1)} \{1-F(x_r)\}^{\sum_{j=r+1}^{k-t-1}(R_j+1)} .$$

To integrate out the remaining variables x_{k+1}, $x_{k+2}, \ldots$, x_{s-1} by using Lemma 3.1, this time taking $T = x_k$, it is used the fact that

$$\int_{x_k}^{x_s} \cdots \int_{x_k}^{x_{k+3}} \int_{x_k}^{x_{k+2}} \{1-F(x_{k+1})\}^{R_{k+1}} \ldots \{1-F(x_{s-1})\}^{R_{s-1}} dF(x_{k+1}) \ldots dF(x_{s-1})$$

$$= \sum_{t=0}^{s-k-1} c_{t,s-k-1}(R_{k+1},\ldots,R_{s-1})\{1-F(x_s)\}^{\sum_{j=s-t}^{s-1}(R_j+1)} \{1-F(x_k)\}^{\sum_{j=k+1}^{s-t-1}(R_j+1)}$$

which completes the proof. The result is

$$f^{\mathbf{R}}_{r,k,s:m:n}(x_r, x_k, x_s) = c' f(x_r)\, f(x_k)\, f(x_s)\{1-F(x_r)\}^{R_r}\{1-F(x_k)\}^{R_k}\{1-F(x_s)\}^{R_s^*}$$

$$\times \sum_{i=0}^{r-1} c_{i,r-1}(R_1+1,\dots,R_{r-1}+1)\{1-F(x_r)\}^{\sum_{j=r-i}^{r-1}(R_j+1)}$$

$$\times \sum_{t=0}^{k-r-1} c_{t,k-r-1}(R_{r+1},\dots,R_{k-1})\{1-F(x_k)\}^{\sum_{j=k-t}^{k-1}(R_j+1)}\{1-F(x_r)\}^{\sum_{j=r+1}^{k-t-1}(R_j+1)}$$

$$\times \sum_{t=0}^{s-k-1} c_{t,s-k-1}(R_{k+1},\dots,R_{s-1})\{1-F(x_s)\}^{\sum_{j=s-t}^{s-1}(R_j+1)}\{1-F(x_k)\}^{\sum_{j=k+1}^{s-t-1}(R_j+1)}$$

generalized for the variables $X^{\mathbf{R}}_{r:m:n}, X^{\mathbf{R}}_{k:m:n}$ and $X^{\mathbf{R}}_{s:m:n}$, $r<k<s$.

It is clear that the general form for the joint probability density function of $X^{\mathbf{R}}_{r_1:m:n}, X^{\mathbf{R}}_{r_2:m:n}, \dots, X^{\mathbf{R}}_{r_k:m:n}$ is given in equation (3.1).

It is seen that equation (3.1) agrees with the formula for special case, when $k=2$, $r_1 = r$, $r_2 =$ s with the well known formulas on joint probability density function of order statistics

$$f^{\mathbf{R}}_{r_1,r_2:m:n}(x_r, x_s) = c' f(x_r) f(x_s)\{1-F(x_r)\}^{R_r}\{1-F(x_s)\}^{R_s^*}$$

$$\times \prod_{z=1}^{1}\sum_{i=0}^{r-1}\sum_{t=0}^{s-r-1} c_{i,r-1}(R_1+1,\dots,R_{r-1}+1)$$

$$\times c_{t,s-r-1}(R_{r+1},\dots,R_{s-1})\{1-F(x_r)\}^{\sum_{j=r-i}^{r-1}(R_j+1)}\{1-F(x_r)\}^{\sum_{j=r+1}^{s-t-1}(R_j+1)}$$

$$\times\{1-F(x_s)\}^{\sum_{j=s-t}^{s-1}(R_j+1)}, -\infty < x_r < x_s < \infty. \quad (3.3)$$

When the formula is simplified, the joint probability density function of $X^{\mathbf{R}}_{r:m:n}$ and $X^{\mathbf{R}}_{s:m:n}$ in (2.22) which is given by Balakrishnan, Childs et al. (2002) is obtained. In order to show that the theorem agrees with the ordinary order statistics in the special case, when $R_1 = R_2 =\ldots= R_m = 0$, some examples are given.

Example 3.1 Consider the probability density function of order statistics $X_{3:n}$.

$$f_{3:n}(x_3) = \frac{n(n-1)(n-2)}{2!} f(x_3) F(x_3)^2 \{1-F(x_3)\}^{n-3}, \ -\infty < x_3 < \infty. \quad (3.4)$$

From Theorem 3.1, probability density function of $X^{\mathbf{R}}_{3:m:n}$ is given

$$f^{\mathbf{R}}_{3:m:n}(x_3) = n(n-R_1-1)(n-R_1-R_2-2) f(x_3)\{1-F(x_3)\}^{n-3}$$

$$\times\sum_{i=0}^{2} c_{i,2}(R_1+1)\{1-F(x_3)\}^{\sum_{j=3-i}^{2}(R_j+1)}, -\infty < x_3 < \infty. \quad (3.5)$$

For the special case, when $R_1 = R_2 = 0$, the Progressive Type II right censoring scheme reduces to the case without censoring and the probability density function of usual order statistics is observed which is given in (3.4).

Example 3.2 The joint probability density function of $X_{2:n}$ and $X_{3:n}$ is given by

$$f_{2,3:n}(x_2, x_3) = n(n-1)(n-2) f(x_2) f(x_3) F(x_2)\{1-F(x_3)\}^{n-3},$$

$$-\infty < x_2 < x_3 < \infty. \quad (3.6)$$

From Theorem 3.1, the joint probability density function for $X^{\mathbf{R}}_{2:m:n}$ and $X^{\mathbf{R}}_{3:m:n}$ is given by

$$f^{\mathbf{R}}_{2,3:m:n}(x_2, x_3) = c'n(n-R_1-1)(n-R_1-R_2-2)\,f(x_2)\,f(x_3)$$

$$\times\{1-F(x_3)\}^{R_3^*}\sum_{i=0}^{1}c_{i,1}(R_1+1)\{1-F(x_2)\}^{\sum_{j=2-i}^{1}(R_j+1)}, -\infty < x_2 < x_3 < \infty \quad (3.7)$$

where $c' = n(n - R_1 - 1)(n - R_1 - R_2 - 2)$ and $R_3^* = n - R_1 - R_2 - 3$. For the special case, when $R_1 = R_2 = R_3 = 0$, the Progressive Type II right censoring scheme reduces to the case without censoring and joint probability density function in equation (3.6) for usual order statistics is observed.

Example 3.3 The joint probability density function of $X_{2:n}$ and $X_{4:n}$ is given by

$$f_{2,4:n}(x_2, x_4) = n(n-1)(n-2)(n-3)\,f(x_2)\,f(x_4)\,F(x_2)\{F(x_4)-F(x_2)\}$$

$$\times\{1-F(x_4)\}^{n-4}, -\infty < x_2 < x_4 < \infty\,. \quad (3.8)$$

From Theorem 3.1, the joint probability density function of $X^{\mathbf{R}}_{2:m:n}$ and $X^{\mathbf{R}}_{4:m:n}$ is

$$f^{\mathbf{R}}_{2,4:m:n}(x_2, x_4) = c'f(x_2)\,f(x_4)\{1-F(x_2)\}^{R_2}\{1-F(x_4)\}^{R_4^*}$$

$$\times\sum_{i=0}^{1}\sum_{t=0}^{1}c_{i,1}(R_1+1)c_{t,1}(R_3)\{1-F(x_2)\}^{\sum_{j=2-i}^{1}(R_j+1)}$$

$$\times\{1-F(x_2)\}^{\sum_{j=3}^{3-t}(R_j+1)}\{1-F(x_4)\}^{\sum_{j=4-t}^{3}(R_j+1)}, -\infty < x_2 < x_4 < \infty \quad (3.9)$$

where $c' = n(n - R_1 - 1)(n - R_1 - R_2 - 2)\ )(n - R_1 - R_2 - R_3 - 3)$ and $R_4^* = n - R_1 - R_2 - R_3 - 4$.

In the special case $R_1 = R_2 = R_3 = R_4 = 0$, the Progressive Type II right censoring scheme reduces to the case without censoring and in equation (3.8), the joint probability density function for usual order statistics is obtained.

Example 3.4 The joint probability density function of $X_{2:n}$, $X_{3:n}$ and $X_{4:n}$ is given by

$$f_{2,3,4:n}(x_2, x_3, x_4) = n(n-1)(n-2)(n-3)\, f(x_2)\, f(x_3)\, f(x_4)\, F(x_2)$$

$$\times \{1 - F(x_4)\}^{n-4},\ -\infty < x_2 < x_3 < x_4 < \infty\,. \tag{3.10}$$

From Theorem 3.1, the joint probability density function for $X_{2:m:n}^{\mathbf{R}}$, $X_{3:m:n}^{\mathbf{R}}$ and $X_{4:m:n}^{\mathbf{R}}$ is given

$$f_{2,3,4:m:n}^{\mathbf{R}}(x_2, x_3, x_4) = c' f(x_2) f(x_3) f(x_4) \{1 - F(x_2)\}^{R_2} \{1 - F(x_4)\}^{R_4^*}$$

$$\times \sum_{i=0}^{1} c_{i,1}(R_1 + 1)\{1 - F(x_2)\}^{\sum_{j=2-i}^{1}(R_j+1)}, -\infty < x_2 < x_3 < x_4 < \infty \tag{3.11}$$

where $c' = n(n - R_1 - 1)(n - R_1 - R_2 - 2)\ldots(n - R_1 - R_2 - R_3 - 3)$ and $R_4^* = n - R_1 - R_2 - R_3 - 4$. For the $R_1 = R_2 = R_3 = R_4 = 0$, the joint probability density function in (3.10) is obtained.

Example 3.5 The joint probability density function of $X_{2:n}$, $X_{4:n}$ and $X_{7:n}$ is given by

$$f_{2,4,7:n}(x_2,x_4,x_7)=c'f(x_2)f(x_4)f(x_7)F(x_2)\{F(x_4)-F(x_2)\}$$

$$\times\{F(x_7)-F(x_4)\}^2\{1-F(x_7)\}^{n-7}, -\infty<x_2<x_4<x_7<\infty \quad (3.12)$$

where $c'=\dfrac{n(n-1)(n-2)...(n-6)}{2}$.

From Theorem 3.1, the joint probability density function for $X_{2:m:n}^{\mathbf{R}}$, $X_{4:m:n}^{\mathbf{R}}$ and $X_{7:m:n}^{\mathbf{R}}$ is given by

$$f_{2,4,7:m:n}^{\mathbf{R}}(x_2,x_4,x_7)=c'f(x_2)f(x_4)f(x_7)\{1-F(x_2)\}^{R_2}\{1-F(x_4)\}^{R_4}\{1-F(x_7)\}^{R_7^*}$$

$$\times\sum_{i=0}^{1}c_{i,1}(R_1+1)\{1-F(x_2)\}^{\sum_{j=2-i}^{1}(R_j+1)}\prod_{z=1}^{2}\sum_{t=0}^{r_{z+1}-r_z-1}c_{t,r_{z+1}-r_z-1}(R_3,...,R_6)$$

$$\times\{1-F(x_{r_z})\}^{\sum_{j=r_z+1}^{r_{z+1}-t-1}(R_j+1)}\{1-F(x_{r_{z+1}})\}^{\sum_{j=r_{z+1}-t}^{r_{z+1}-1}(R_j+1)}, -\infty<x_2<x_4<x_7<\infty \quad (3.13)$$

where $c'=n(n-R_1-1)(n-R_1-R_2-2)\dots(n-R_1-R_2-\dots-R_6-6)$ and $R_7^*=n-R_1-R_2-\dots-R_6-7$. If i, z and t take the related values in the equation (3.13), the results which are given in below expressions are obtained.

1. $\displaystyle\sum_{i=0}^{1}c_{i,1}(R_1+1)\{1-F(x_2)\}^{\sum_{j=2-i}^{1}(R_j+1)}$

for $i=0$,

$$c_{0,1}(R_1+1) = \frac{(-1)^0}{\left\{\prod_{j=1}^{0}\sum_{k=2}^{1+j}(R_k+1)\right\}\left\{\prod_{j=1}^{1}\sum_{k=j}^{1}(R_k+1)\right\}} = \frac{1}{R_1+1}$$

then for R_k, $R_j = 0$,

$$\frac{1}{R_1+1}\{1-F(x_2)\}^{\sum_{j=2}^{1}(R_j+1)} = 1\,.$$

For $i = 1$,

$$c_{1,1}(R_1+1) = \frac{(-1)^1}{\left\{\prod_{j=1}^{1}\sum_{k=1}^{j}(R_k+1)\right\}\left\{\prod_{j=1}^{0}\sum_{k=j}^{0}(R_k+1)\right\}} = -\frac{1}{R_1+1}$$

then for R_k, $R_j = 0$,

$$-\frac{1}{R_1+1}\{1-F(x_2)\}^{\sum_{j=1}^{1}(R_j+1)} = -\{1-F(x_2)\}\,.$$

By adding the results for $i = 0$ and $i = 1$, $\{1-\{1-F(x_2)\}\} = F(x_2)$ is obtained.

2. $\sum_{t=0}^{1} c_{t,1}(R_3)\{1-F(x_2)\}^{\sum_{j=3}^{3-t}(R_j+1)}\{1-F(x_4)\}^{\sum_{j=4-t}^{3}(R_j+1)}$.

For $i = 0$ and R_k, $R_j = 0$,

$$c_{0,1}(R_3)\{1-F(x_2)\}^{\sum_{j=3}^{3}(R_j+1)}\{1-F(x_4)\}^{\sum_{j=4}^{3}(R_j+1)} = \frac{1}{R_1+1}\{1-F(x_2)\}\,.$$

For $i = 1$ and R_k, $R_j = 0$,

$$c_{1,1}(R_3)\{1-F(x_2)\}^{\sum_{j=3}^{2}(R_j+1)}\{1-F(x_4)\}^{\sum_{j=3}^{3}(R_j+1)} = -\frac{1}{R_1+1}\{1-F(x_4)\}.$$

By combining these two results

$\{\{1-F(x_2)\}-\{1-F(x_4)\}\} = \{F(x_4)-F(x_2)\}$ is obtained.

3. $$\sum_{t=0}^{2} c_{t,2}(R_5,R_6)\{1-F(x_4)\}^{\sum_{j=5}^{6-t}(R_j+1)}\{1-F(x_7)\}^{\sum_{j=7-t}^{6}(R_j+1)}$$

for $i = 0$,

$$c_{0,2}(R_5,R_6) = \frac{(-1)^0}{\left\{\prod_{j=1}^{0}\sum_{k=3}^{2+j}(R_k+1)\right\}\left\{\prod_{j=1}^{2}\sum_{k=j}^{2}(R_k+1)\right\}} = \frac{1}{(R_1+R_2+2)(R_2+1)}$$

then for R_k, $R_j = 0$,

$$\tfrac{1}{2}\{1-F(x_4)\}^{R_5+R_6+2} = \tfrac{1}{2}\{1-F(x_4)\}^2.$$

For $i = 1$,

$$c_{1,2}(R_5,R_6) = \frac{(-1)^1}{\left\{\prod_{j=1}^{1}\sum_{k=2}^{1+j}(R_k+1)\right\}\left\{\prod_{j=1}^{1}\sum_{k=j}^{1}(R_k+1)\right\}} = -\frac{1}{(R_2+1)(R_1+1)}$$

then for R_k, $R_j = 0$,

$$-\{1-F(x_4)\}^{R_5+1}\{1-F(x_7)\}^{R_6+1} = -\{1-F(x_4)\}\{1-F(x_7)\}.$$

For $i = 2$,

$$c_{2,2}(R_5, R_6) = \frac{(-1)^2}{\left\{\prod_{j=1}^{2}\sum_{k=1}^{j}(R_k+1)\right\}\left\{\prod_{j=1}^{0}\sum_{k=j}^{0}(R_k+1)\right\}} = -\frac{1}{(R_2+1)(R_1+R_2+2)}$$

then for R_k, $R_j = 0$,

$$\tfrac{1}{2}\{1-F(x_7)\}^{(R_5+R_6+2)} = \tfrac{1}{2}\{1-F(x_7)\}^2 .$$

The three results are added and then the result is given by

$$\tfrac{1}{2}\{1-F(x_4)\}^2 - \{1-F(x_4)\}\{1-F(x_7)\} - \tfrac{1}{2}\{1-F(x_7)\}^2$$

$$= \tfrac{1}{2}[\{1-F(x_4)\}^2 - 2\{1-F(x_4)\}\{1-F(x_7)\} + \{1-F(x_7)\}^2]$$

$$= \tfrac{1}{2}[\{1-F(x_4)\} - \{1-F(x_7)\}]^2 = \tfrac{1}{2}[F(x_7) - F(x_4)]^2 .$$

When the expressions are combined, for $R_1 = R_2 = R_3 =\ldots= R_7 = 0$, the joint probability density function of ordinary order statistics in equation (3.12) is obtained.

As seen from the examples given above and from Theorem 3.1 in the special case, the joint probability density function of ordinary order statistics is obtained.

CHAPTER FOUR

CHARACTERIZATION OF DISTRIBUTIONS THROUGH THE CONDITIONAL EXPECTATIONS OF ORDER STATISTICS AND PROGRESSIVELY TYPE II CENSORED ORDER STATISTICS

After giving the important theorems which are used to obtain main results, characterizations through the conditional expectations of order statistics are represented in Section 4.1. In Section 4.2, before giving the characterizations through the conditional expectations of Progressively Type II censored order statistics, Theorem 4.3 which shows the reduction of Progressively Type II censored order statistics to order statistics is stated. It is introduced that the joint distribution of Progressively Type II censored order statistics can be reduced to the joint distribution of ordinary order statistics of a sample size *m*.

Let X_1, X_2,…, X_n be a random sample of a random variable *X* with a continuous distribution function *F*(*x*) which is strictly increasing over (*a*, *b*), $-\infty \leq a < b \leq \infty$, the support of *F*(*x*) and $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$, the corresponding order statistics. Using the condition $E(X_{i:n} \mid X_{i+1:n} = x) = \alpha x - \beta$ distributions characterized by Ferguson(1967). Beg and Kirmani(1974) characterized the distributions using the condition $E(X_i \mid X_{n:n} = x) = \alpha x - \beta$. In here α and β are some constants. Similar results can be found in Galambos & Kotz (1978), Azlarov & Volodin (1986), Beg & Balasubramanian (1990) and Balasubramanian & Beg (1992). And also Balasubramanian & Dey (1997) present certain results based on conditional expectations that lead to characterizations of several distributions and they use the harmonic and geometric means of the functions at the end points for obtaining their results.

For absolutely continuous probability density function using $E(h(X) \mid X \geq x)$ under certain conditions for function *h*(*x*), a representation is given by Hamdan (1972), Talwalker (1977), Shanbhag & Bhaskara Rao (1975), Gupta (1975) and Ouyang (1983, 1987). The above problem for random variables in general was studied by Kotz & Shanbhag (1980). Kotlarski (1972) shows that there exists a one to one map between functions $\phi(x) = E(h(X) | X \geq x)$ and continuous probability density function, while the set of functions $\phi(x)$ which are conditional expectations of continuous probability density function is characterized by Zoroa, Ruiz, & Mrin (1990).

In this thesis, for a general class of distributions some characterizations through the properties of conditional expectations of order statistics and Progressively Type II censored order statistics are given (Bairamov & Özkal, 2007). Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the order statistics of the sample $X_1, X_2, ..., X_n$ from a continuous distribution and $X_{1:m:n}^{\mathbf{R}}, X_{2:m:n}^{\mathbf{R}}, ..., X_{m:m:n}^{\mathbf{R}}$ be the Progressively Type II censored order statistics with the progressive Type II censoring scheme $\mathbf{R} = (R, R, ..., R)$. It is represented that in this case, the joint distribution of $X_{1:m:n}^{\mathbf{R}}, X_{2:m:n}^{\mathbf{R}}, ..., X_{m:m:n}^{\mathbf{R}}$ can be reduced to the joint distribution of ordinary order statistics of a sample size m from a continuous random variable.

Let $X_1, X_2, ..., X_n$ be independent random variables with a common absolutely continuous distribution function $F(x)$ and a probability density function $f(x)$. Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the corresponding order statistics. Characterizations of $F(x)$ based on the properties of conditional expectations with the constant sample size n have been discussed by many authors.

Ruiz & Navarro (1996) and also Navarro et al. (1998) gave a representation of any continuous distribution function $F(x)$ based on the doubly truncated mean function $m(x, y)$ defined by

$$m(x, y) = E(X \mid x \leq X \leq y) = \frac{1}{F(y) - F(x)} \int_x^y t dF(t), \tag{4.1}$$

whose domain of definition is the set $D = \{(x, y) \in IR^2$ such that $F(x) < F(y)\}$.

Ruiz et al. (1996) obtained the explicit expression of $F(x)$ starting from function $m(x, y)$ and define the relationship between $m(x, y)$ and order statistics as

$$E\left(\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} h(X_{i:n}) \mid X_{r:n} = x,\ X_{s:n} = y\right) = E(h(X) \mid x \le X \le y), \tag{4.2}$$

for all $(x, y) \in D$, if $1 \le r < s \le n$, where $X_{1:n} \le X_{2:n} \le \ldots \le X_{n:n}$ are the order statistics from the sample of the random variable X. Balasubramanian et al. (1992) presented similar results for a particular function $h(x)$. Gupta, Goswami, & Rao (1993) consider the case when $r = 1$ and $s = n$.

Some continuous distributions through the properties of conditional expectations of order statistics and Progressively Type II right censored order statistics are characterized in this study. Characterization theorems for a general class of distributions are presented in terms of the function

$$E\{g(X_{j:n}) \mid X_{j-p:n} = x,\ X_{j+q:n} = y\} = A(x, y)\,, \tag{4.3}$$

where p and q are positive integers such that $p + 1 \le j \le n - q$ and $g(.)$, $A(.,.)$ are real valued functions satisfying certain regularity conditions. The results obtained in this chapter are closely related to works by Kotlarski (1972) and Shandbag et al. (1975)**.**

4.1 Characterizations Through the Conditional Expectations of Order Statistics

Throughout the thesis, it is denoted by $a_F = \inf\{x : F(x) > 0\}$ and $b_F = \sup\{x : F(x) < 1\}$ the left and right extremities of $F(x)$, respectively where $F(x)$ is a distribution function of a random variable X. It is assumed that $F(x)$ is absolutely continuous and strictly increasing on (a_F, b_F).

The characterization results are based on the following theorem.

Theorem 4.1 Let $h(x)$ be a differentiable real valued function on [0,1] such that the condition

$$h'(y) \neq \frac{h(y)-h(x)}{y-x}, \tag{4.4}$$

holds for all $0 < x < y < 1$. Furthermore, let $G(x)$ be an absolutely continuous and strictly increasing distribution function with left and right extremities, $a_G = a_F$ and $b_G = b_F$, respectively. Then $F(x) = G(x)$ if and only if the representation

$$E\{h'(G(X)) \mid x \leq G(X) \leq y\} = \frac{h(y)-h(x)}{y-x}, \tag{4.5}$$

is valid for all $0 < x < y < 1$.

Proof. It is clear that (4.5) is equivalent to equation (4.6)

$$E\{h'(G(X)) \mid x \leq X \leq y\} = \frac{h(G(y))-h(G(x))}{G(y)-G(x)}, \tag{4.6}$$

for all $a_F < x < y < b_F$.

Necessity. If X has distribution function $G(x)$, $G'(x) = g(x)$, then

$$E\{h'(G(X)) \mid x \leq X \leq y\} = \frac{1}{G(y)-G(x)} \int_x^y h'(G(z))g(z)dz$$

$$= \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}. \tag{4.7}$$

Sufficiency. Let (4.5) hold. Denote by $F(x)$ and $f(x)$ the distribution function and probability density function of X, respectively. Then from (4.5)

$$\int_x^y h'(G(z)) f(z) dz = \frac{h(G(y)) - h(G(x))}{G(y) - G(x)} \{F(y) - F(x)\}. \tag{4.8}$$

Differentiating (4.8) with respect to y after some calculation, the following equation is obtained

$$[f(y)\{G(y) - G(x)\} - g(y)\{F(y) - F(x)\}] \times \left[h'(G(y)) - \frac{h(G(y)) - h(G(x))}{G(y) - G(x)} \right] = 0.$$

Therefore

$$\frac{f(y)}{F(y) - F(x)} = \frac{g(y)}{G(y) - G(x)}, \ \forall\, a_F < x < y < b_F.$$

Taking limit as $x \to a_F$ the following equality is obtained

$$\frac{f(y)}{F(y)} = \frac{g(y)}{G(y)}, \ \forall\, a_F < y < b_F.$$

Equality of hazard rate functions implies equality of $F(x)$ and $G(x)$.

Theorem 4.1 states that the representation (4.5) characterizes the distribution function $F(x)$ in the class F; all absolutely continuous and strictly increasing distribution functions with left and right extremities $l_F = a_F$ and $r_F = b_F$, respectively.

Remark 4.1 It is clear that, $\dfrac{h(y)-h(x)}{y-x}$ is the slope of the straight line joining points $(x, h(x))$, $(y, h(y))$ and $h'(y)$ is the slope of the tangent line to the graph of $h(.)$ at the point y. The condition (4.4) assumes these two slopes to be different. One can observe that for any convex or concave function $h(.)$, the property (4.4) is satisfied.

Let $X_1, X_2, \ldots, X_n$ be the independent copies of X, i.e. be i.i.d. random variables with distribution function $F(x)$ and probability density function $f(x)$. Denote by $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ the corresponding order statistics. The following fact shows that the conditional distribution of $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$ does not depend on the sample size n. Note that it depends on x, y and $F(x)$.

From Theorem 2.1, if $p + 1 \le j \le n - q$, then

$$(X_{j:n} \mid X_{j-p:n} = x, X_{j+q:n} = y) \stackrel{d}{=} Z_{p:p+q-1}, \tag{4.9}$$

where $Z \stackrel{d}{=} (X \mid x < X < y)$.

Theorems 2.1 and 4.1 are used to characterize several distributions by properties of conditional expectations of order statistic $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$.

Remember, it is assumed that the distribution function $F(x)$ of the random variable X is absolutely continuous and strictly increasing on (a_F, b_F), where $a_F = \inf\{x : F(x) > 0\}$ and $b_F = \sup\{x : F(x) < 1\}$.

Theorem 4.2 Let $G(x)$ be an absolutely continuous and strictly increasing distribution function and the left and right extremities of $G(x)$ be $a_G = a_F$ and $b_G = b_F$, respectively. Then, $F(x) = G(x)$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[h'(G(X_{j:n})) \mid X_{j-p:n} = x, X_{j+k+1-p:n} = y] = \frac{h(G(y)) - h(G(x))}{G(y) - G(x)} \tag{4.10}$$

holds for all $a_F < x < y < b_F$, where $h(x)$ satisfies conditions of Theorem 4.1 and j, n, k are fixed integers such that

$$k + 1 \le j \le n - k. \tag{4.11}$$

Proof. If equation (4.9) is taken into account

$$\frac{1}{k}\sum_{p=1}^{k} E[h'(G(X_{j:n})) \mid X_{j-p:n} = x, X_{j+k+1-p:n} = y]$$

$$= E\left[\frac{1}{k}\sum_{p=1}^{k} h'(G(Z_{p:k}))\right] = E\left[\frac{1}{k}\sum_{i=1}^{k} h'(G(Z_i))\right]$$

$$= \frac{1}{k} kE[h'(G(Z))] = E[h'(G(Z))]$$

$$= E[h'(G(X)) \mid x \le X \le y].$$

then the proof is completed by using Theorem 4.1.

Now consider special cases when k takes the values 2, 3 and when p and q take the values 1, 2, 3. For these cases, Theorem 4.2 can be rewritten as follows:

Corollary 4.1 *i*) Let $k + 1 \leq j \leq n - k$, for $k = 2$, $3 \leq j \leq n - 2$ is obtained. A random variable X has absolutely continuous strictly increasing distribution function $G(x)$ if and only if the representation

$$\frac{1}{2}E\{X_{j:n} \mid X_{j-1:n} = x, X_{j+2:n} = y\} + \frac{1}{2}E\{X_{j:n} \mid X_{j-2:n} = x, X_{j+1:n} = y\}$$

$$= \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}$$

holds for all $(x, y) \in D$.

ii) For $k = 3$, $4 \leq j \leq n - 3$ is obtained. A random variable X has absolutely continuous strictly increasing distribution function $G(x)$ if and only if the representation

$$\frac{1}{3}E\{X_{j:n} \mid X_{j-1:n} = x, X_{j+3:n} = y\} + \frac{1}{3}E\{X_{j:n} \mid X_{j-2:n} = x, X_{j+2:n} = y\}$$

$$+ \frac{1}{3}E\{X_{j:n} \mid X_{j-3:n} = x, X_{j+1:n} = y\} = \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}$$

holds for all $(x, y) \in D$.

4.1.1 Characterizations for Some Special Distributions

The following characterizations for special distributions can be obtained from Theorem 4.2 by the appropriate choice of the function $h(x)$.

4.1.1.1 Uniform Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support (0,1) has Uniform distribution if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[X_{j:n}^{c} \mid X_{j-p:n} = x,\ X_{j+k+1-p:n} = y] = \frac{y^{c+1} - x^{c+1}}{(c+1)(y-x)}, \tag{4.12}$$

holds for all $0 < x < y < 1$. In (4.12) $j \in \{k + 1, k + 2, \ldots, n - k\}$ is fixed and j satisfies (4.11). This result follows from Theorem 4.2 by using $h(x) = \frac{x^{c+1}}{c+1}$, $c \geq 1$.

4.1.1.2 Weibull Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support $[0,\infty)$, has Weibull distribution $F(x) = 1 - \exp(-\alpha x^{\beta})$, $x \geq 0$, $\alpha > 0$, $\beta > 0$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[\alpha X_{j:n}^{\beta} \mid X_{j-p:n} = x,\ X_{j+k+1-p:n} = y]$$

$$= \frac{\exp(-\alpha x^{\beta})(\alpha x^{\beta} + 1) - \exp(-\alpha y^{\beta})(\alpha y^{\beta} + 1)}{\exp(-\alpha x^{\beta}) - \exp(-\alpha y^{\beta})} \tag{4.13}$$

holds for all $0 \leq x < y < \infty$. Here j satisfies (4.11). The above characterization with $\beta = 1$ turns into a characterization of Exponential distribution.

Special case. When $\beta = 1$, X has Exponential distribution $F(x) = 1 - \exp(-\alpha x)$, $x \geq 0$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[\alpha X_{j:n} \mid X_{j-p:n} = x,\ X_{j+k+1-p:n} = y]$$

$$= \frac{\exp(-\alpha x)(\alpha x + 1) - \exp(-\alpha y)(\alpha y + 1)}{\exp(-\alpha x) - \exp(-\alpha y)}, \tag{4.14}$$

holds for all $0 \leq x < y < \infty$. Here $h(x) = (1-x)(\ln(1-x) - 1)$. Then $h'(x) = -\ln(1-x)$.

4.1.1.3 Generalized Beta Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support $[a,b]$ has Generalized Beta distribution $F(x) = 1 - \frac{(b-x)^{\theta}}{(b-a)^{\theta}}$, $a \leq x \leq b$, $\theta > 0$, $-\infty < a < b < \infty$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[X_{j:n} \mid X_{j-p:n} = x,\ X_{j+k+1-p:n} = y] = b + \frac{(a-b)\theta}{\theta+1}\frac{a(x,y;\theta+1)}{a(x,y;\theta)}, \tag{4.15}$$

holds for all $a < x < y < b$ where $a(x,y;\theta) = \left(\frac{b-x}{b-a}\right)^{\theta} - \left(\frac{b-y}{b-a}\right)^{\theta}$. Here j satisfies (4.11). The result is obtained by using target function $h(x) = bx - \frac{(a-b)\theta}{\theta+1}(1-x)^{1+1/\theta}$, $h'(x) = b - (b-a)(1-x)^{1/\theta}$.

The above characterization with $a = 0$, $b = 1$ and $\theta = 1$ turns into a characterization of Uniform distribution.

4.1.1.4 Pareto Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support [γ, ∞) has Pareto distribution $F(x) = 1 - \frac{(\gamma+\delta)^{\theta}}{(x+\delta)^{\theta}}$, $x \geq \gamma$, $\theta > 0$, $\gamma + \delta > 0$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[X_{j:n} \mid X_{j-p:n} = x, X_{j+k+1-p:n} = y] = -\frac{\theta}{\theta-1}\frac{b(x,y;\theta-1)}{b(x,y;\theta)} - \delta, \tag{4.16}$$

holds for all $\gamma \leq x < y < \infty$ where $b(x,y;\theta) = \left(\frac{1}{x+\delta}\right)^{\theta} - \left(\frac{1}{y+\delta}\right)^{\theta}$, $\theta \neq 1$ and j satisfies (4.11). The result is obtained by using target function $h(x) = \frac{(\gamma+\delta)\theta}{\theta-1}(1-x)^{1-(1/\theta)} - \delta x$ and $h'(x) = -(\gamma+\delta)(1-x)^{-1/\theta} - \delta$.

4.2 Characterizations Through Conditional Expectations of Progressively Type II Censored Order Statistics

Let $X_1, X_2, \ldots, X_n$ be continuous i.i.d. random variables with distribution function $F(x)$ and probability density function $f(x)$. Denote by $X^{\mathbf{R}}_{1:m:n}, \ldots, X^{\mathbf{R}}_{m:m:n}$ the corresponding Progressively Type II censored order statistics with the censoring scheme $\mathbf{R} = (R_1, R_2, \ldots, R_m)$. It is considered the case $R_1 = R_2 = \ldots = R_m = R$ and is given some characterizations for a general class of distribution through the conditional expectations of Progressively Type II censored order statistics in this case. First, it is represented that when $\mathbf{R} = (R, R, \ldots, R)$, the joint distribution of $X^{\mathbf{R}}_{1:m:n}, \ldots, X^{\mathbf{R}}_{m:m:n}$ can be reduced to the joint distribution of usual order statistics of a sample size m from a continuous random variable.

Theorem 4.3 Let $X_1, X_2, \ldots, X_n$ be a sample of size n from a population with continuous distribution function $F(x)$ and probability density function $f(x)$ and let $X_{1:m:n}^{\mathbf{R}} \le X_{2:m:n}^{\mathbf{R}} \le \ldots \le X_{m:m:n}^{\mathbf{R}}$ be the Progressively Type II censored order statistics under the censoring scheme $\mathbf{R} = (R_1, R_2, \ldots, R_m)$. Then in the special case when $R_1 = R_2 = \ldots = R_m = R$, the following equality is given

$$(X_{1:m:n}^{\mathbf{R}}, \ldots, X_{m:m:n}^{\mathbf{R}}) \overset{d}{=} (Y_{1:m}, \ldots, Y_{m:m}), \tag{4.17}$$

where $Y_1, Y_2, \ldots, Y_m$ is the sample with distribution function $Q(x) = 1 - (1 - F(x))^{n/m}$ and $R + 1 = n / m$.

Proof. The joint probability density function of $(X_{1:m:n}^{\mathbf{R}}, X_{2:m:n}^{\mathbf{R}}, \ldots, X_{m:m:n}^{\mathbf{R}})$ with the censoring scheme $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ is given in equation (2.19) (Balakrishnan et al, 2000).

In equation (2.19), if $R_1 = R_2 = \ldots = R_m = R$, then $R + 1 = n / m$. Since

$$n(n - (R+1))(n - 2(R+1))\ldots(n - (m-1)(R+1))$$

$$= n(n - \frac{n}{m})(n - 2\frac{n}{m})\ldots(n - (m-1)\frac{n}{m}) = (R+1)^m m!,$$

the joint probability density function of m order statistics is given by

$$f_{1,2,\ldots,m}^{\mathbf{R}}(x_1, \ldots, x_m) = m!(R+1)^m (1 - F(x_1))^R \ldots (1 - F(x_m))^R f(x_1)\ldots f(x_m)$$

$$= m!\, g(x_1)\ldots g(x_m), \tag{4.18}$$

where $g(x) = (R + 1)f(x)(1 - F(x))^R$, $G(x) = 1 - (1 - F(x))^{R+1}$.

By using Theorem 4.2 and Theorem 4.3, characterizations for absolutely continuous distributions in terms of conditional expectations of Progressively Type II censored order statistics can be given.

From Theorem 2.1 and Theorem 4.3, equation (4.19) can be written

$$(X^{\mathbf{R}}_{j:m:n} \mid X^{\mathbf{R}}_{j-p:m:n} = x, X^{\mathbf{R}}_{j+q:m:n} = y) \equiv Y_{p:p+q-1} \tag{4.19}$$

and then the conditional probability density function of Progressively Type II censored order statistics $X^{\mathbf{R}}_{j:m:n}$ given $X^{\mathbf{R}}_{j-p:m:n} = x$ and $X^{\mathbf{R}}_{j+q:m:n} = y$ is given by

$$f^{\mathbf{R}}_{j,j-p,j+q:m:n}(x_{j:m:n} \mid x_{j-p:m:n} = x, x_{j+q:m:n} = y) = \frac{(p+q-1)!}{(p-1)!(q-1)!}\frac{g(t)}{G(y)-G(x)}$$

$$\times\left[\frac{G(t)-G(x)}{G(y)-G(x)}\right]^{p-1}\left[1-\frac{G(t)-G(x)}{G(y)-G(x)}\right]^{q-1} \equiv f_{p:p+q-1}(y) \tag{4.20}$$

where $\left(\dfrac{G(t)-G(x)}{G(y)-G(x)}\right)' = \dfrac{g(t)}{G(y)-G(x)}$ and $G(t) = 1-(1-F(t))^{R+1}$, $g(t) = G'(t) = (1-(1-F(t))^{R+1})' = (R+1)(1-F(t))^R f(t)$.

To prove this result Example (4.1) is represented.

Example 4.1 The conditional probability density function of $X^{\mathbf{R}}_{j:m:n}$ given $X^{\mathbf{R}}_{j-p:m:n}$ and $X^{\mathbf{R}}_{j+q:m:n}$ is

$$f^{\mathbf{R}}_{j,\,j-p,\,j+q:m:n}(x_{j:m:n} \mid x_{j-p:m:n}, x_{j+q:m:n})$$

$$= \{c' f(x_{j-p})\, f(x_j)\, f(x_{j+q}) \{1-F(x_{j-p})\}^{R_{j-p}} \{1-F(x_j)\}^{R_j}$$

$$\times \{1-F(x_{j+q})\}^{R^*_{j+q}} \sum_{i=0}^{j-p-1} c_{i,\,j-p-1}(R_1+1,\ldots,R_{j-p-1}+1)\{1-F(x_{j-p})\}^{\sum_{l=j-p-i}^{j-p-1}(R_l+1)}$$

$$\times \sum_{t=0}^{p-1} c_{t,\,p-1}(R_{j-p+1},\ldots,R_{j-1})\{1-F(x_j)\}^{\sum_{l=j-t}^{j-1}(R_l+1)} \{1-F(x_{j-p})\}^{\sum_{l=j-p+1}^{j-t-1}(R_l+1)}$$

$$\times \sum_{t=0}^{q-1} c_{t,\,q-1}(R_{j-1},\ldots,R_{j+q-1})\{1-F(x_{j+q})\}^{\sum_{l=j+q-t}^{j+q-1}(R_l+1)} \{1-F(x_j)\}^{\sum_{l=j+1}^{j+q-t-1}(R_l+1)}\}$$

$$/\{c'' f(x_{j-p})\, f(x_{j+q}) \{1-F(x_{j-p})\}^{R_{j-p}} \{1-F(x_{j+q})\}^{(R^*_{j+q})'}$$

$$\times \sum_{i=0}^{j-p-1} c_{i,\,j-p-1}(R_1+1,\ldots,R_{j-p-1}+1)\{1-F(x_{j-p})\}^{\sum_{l=j-p-i}^{j-p-1}(R_l+1)}$$

$$\times \sum_{t=0}^{p+q-1} c_{t,\,p+q-1}(R_{j-p+1},\ldots,R_{j+q-1})\{1-F(x_{j+q})\}^{\sum_{l=j+q-t}^{j+q-1}(R_l+1)} \{1-F(x_{j-p})\}^{\sum_{l=j-p+1}^{j+q-t-1}(R_l+1)}\}, \tag{4.21}$$

where $p > q$, $c' = c''$ and $R^*_{j+q} = (R^*_{j+q})'$. Also the coefficient of $c_{i,r}(R_r)$ is given by the following formula

$$c_{i,r}(R_r) = \frac{(-1)^i}{\left\{\prod_{j=1}^{i} \sum_{k=r-i+1}^{r-i+j} (R_k + 1)\right\}\left\{\prod_{j=1}^{r-i} \sum_{k=j}^{r-i} (R_k + 1)\right\}} . \tag{4.22}$$

After simplifying equation (4.21), the probability density function is given

$$f^{\mathbf{R}}_{j,\,j-p,\,j+q:m:n}(x_{j:m:n} \mid x_{j-p:m:n}, x_{j+q:m:n}) = \{f(x_j)\{1-F(x_j)\}^{R_j}$$

$$\times \sum_{t=0}^{p-1} c_{t,\,p-1}(R_{j-p+1},\dots,R_{j-1})\,\{1-F(x_j)\}^{\sum_{l=j-t}^{j-1}(R_l+1)}\{1-F(x_{j-p})\}^{\sum_{l=j-p+1}^{j-t-1}(R_l+1)}$$

$$\times \sum_{t=0}^{q-1} c_{t,\,q-1}(R_{j-1},\dots,R_{j+q-1})\,\{1-F(x_{j+q})\}^{\sum_{l=j+q-t}^{j+q-1}(R_l+1)}\{1-F(x_j)\}^{\sum_{l=j+1}^{j+q-t-1}(R_l+1)}\}$$

$$/\{\sum_{t=0}^{p+q-1} c_{t,\,p+q-1}(R_{j-p+1},\dots,R_{j+q-1})\,\{1-F(x_{j+q})\}^{\sum_{l=j+q-t}^{j+q-1}(R_l+1)}\{1-F(x_{j-p})\}^{\sum_{l=j-p+1}^{j+q-t-1}(R_l+1)}\} . \tag{4.23}$$

In the case $p = 3$, $j = 5$ and $q = 2$, the numerator of equation (4.23) is given by

$$f(x_5)\{1-F(x_5)\}^{R_5}\sum_{t=0}^{2} c_{t,2}(R_3,R_4)\{1-F(x_5)\}^{\sum_{l=5-t}^{4}(R_l+1)}\{1-F(x_2)\}^{\sum_{l=3}^{4-t}(R_l+1)}$$

$$\times \sum_{t=0}^{1} c_{t,1}(R_4,R_5,R_6)\{1-F(x_7)\}^{\sum_{l=7-t}^{6}(R_l+1)}\{1-F(x_5)\}^{\sum_{l=6}^{6-t}(R_l+1)} . \tag{4.24}$$

Also the left member of the multiplication in (4.24) is introduced as

$$\sum_{t=0}^{2} c_{t,2}(R_3,R_4)\{1-F(x_5)\}^{\sum_{l=5-t}^{4}(R_l+1)}\{1-F(x_2)\}^{\sum_{l=3}^{4-t}(R_l+1)}. \tag{4.25}$$

i) If $t=0$ in (4.25),

$$c_{0,2}(R_3,R_4)=\tfrac{1}{(R_1+R_2+2)(R_2+1)}\{1-F(x_5)\}^0\{1-F(x_2)\}^{(R_3+R_4+2)} \tag{4.26}$$

ii) If $t=1$ in (4.25),

$$c_{1,2}(R_3,R_4)=-\tfrac{1}{(R_2+1)(R_1+1)}\{1-F(x_5)\}^{R_4+1}\{1-F(x_2)\}^{(R_3+1)} \tag{4.27}$$

iii) If $t=2$ in (4.25),

$$c_{2,2}(R_3,R_4)=\tfrac{1}{(R_2+1)(R_1+R_2+2)}\{1-F(x_5)\}^{(R_3+R_4+2)}\{1-F(x_2)\}^0, \tag{4.28}$$

then expressions in (4.26), (4.27) and (4.28) are obtained.

And the right member of multiplication in (4.24) is given by

$$\sum_{t=0}^{1} c_{t,1}(R_4,R_5,R_6)\{1-F(x_7)\}^{\sum_{l=7-t}^{6}(R_l+1)}\{1-F(x_5)\}^{\sum_{l=6}^{6-t}(R_l+1)}. \tag{4.29}$$

i) If $t=0$ in (4.29),

$$c_{0,1}(R_4,R_5,R_6)=\tfrac{1}{R_1+1}\{1-F(x_7)\}^0\{1-F(x_5)\}^{R_6+1} \tag{4.30}$$

ii) If $t=1$ in (4.29),

$$c_{1,1}(R_4,R_5,R_6)=-\tfrac{1}{R_1+1}\{1-F(x_7)\}^{R_6+1}\{1-F(x_5)\}^0, \tag{4.31}$$

then expressions in (4.30), (4.31) are obtained.

Then the numerator of equation (4.23) is obtained by

$$f(x_5)\{1-F(x_5)\}^{R_5}\times\left\{\frac{1}{(R_2+1)(R_1+R_2+2)}\{1-F(x_2)\}^{(R_3+R_4+2)}-\frac{1}{(R_2+1)(R_1+1)}\right.$$

$$\left.\times\{1-F(x_5)\}^{R_4+1}\{1-F(x_2)\}^{R_3+1}\times\left\{\frac{1}{(R_1+1)}\{1-F(x_5)\}^{R_6+1}-\frac{1}{(R_1+1)}\{1-F(x_7)\}^{R_6+1}\right\}\right\}. \tag{4.32}$$

The denominator of equation (4.23) is given by

$$\sum_{t=0}^{4}c_{t,4}(R_2,...,R_6)\{1-F(x_7)\}^{\sum_{6-t}^{5}(R_l+1)}\{1-F(x_2)\}^{\sum_{3}^{6-t}(R_l+1)}. \tag{4.33}$$

i) If $t = 0$ in (4.33),

$$c_{0,4}(R_2,...,R_6)=\frac{1}{(R_1+...+R_4+4)(R_2+R_3+R_4+3)(R_3+R_4+2)(R_4+1)}$$

$$\times\{1-F(x_7)\}^0\{1-F(x_2)\}^{(R_3+R_4+R_5+R_6+4)} \tag{4.34}$$

ii) If $t = 1$ in (4.33),

$$c_{1,4}(R_2,...,R_6) = \frac{1}{(R_1 + R_2 + R_3 + 3)(R_2 + R_3 + 2)(R_3 + 1)(R_4 + 1)}$$

$$\times \{1 - F(x_7)\}^{(R_5+1)} \{1 - F(x_2)\}^{(R_3+R_4+R_5+3)} \tag{4.35}$$

iii) If $t = 2$ in (4.33),

$$c_{2,4}(R_2,...,R_6) = \frac{1}{(R_3 + 1)(R_3 + R_4 + 2)(R_1 + R_2 + 2)(R_2 + 1)}$$

$$\times \{1 - F(x_7)\}^{(R_4+R_5+2)} \{1 - F(x_2)\}^{(R_3+R_4+2)} \tag{4.36}$$

iv) If $t = 3$ in (4.33),

$$c_{3,4}(R_2,...,R_6) = \frac{1}{(R_2 + 1)(R_2 + R_3 + 2)(R_2 + R_3 + R_4 + 3)(R_1 + 1)}$$

$$\times \{1 - F(x_7)\}^{(R_3+R_4+R_5+3)} \{1 - F(x_2)\}^{(R_3+1)} \tag{4.37}$$

v) If $t = 4$ in (4.33),

$$c_{4,4}(R_2,...,R_6) = \frac{1}{(R_1 + 1)(R_1 + R_2 + 2)(R_1 + R_2 + R_3 + 3)(R_1 + R_2 + R_3 + R_4 + 4)}$$

$$\times \{1 - F(x_7)\}^{(R_2+R_3+R_4+R_5+4)} \{1 - F(x_2)\}^{0}, \tag{4.38}$$

then expressions in (4.34), (4.35), (4.36), (4.37) and (4.38) are obtained.

By unifying these results, the denominator of equation (4.23) is obtained by

$$\frac{\{1-F(x_2)\}^{(R_3+R_4+R_5+R_6+4)}}{(R_4+1)(R_1+R_2+R_3+R_4+4)(R_3+R_4+2)(R_2+R_3+R_4+3)}$$

$$-\frac{\{1-F(x_7)\}^{(R_5+1)}\{1-F(x_2)\}^{(R_3+R_4+R_5+3)}}{(R_1+R_2+R_3+3)(R_2+R_3+2)(R_3+1)(R_4+1)}$$

$$+\frac{\{1-F(x_7)\}^{(R_4+R_5+2)}\{1-F(x_2)\}^{(R_3+R_4+2)}}{(R_3+1)(R_3+R_4+2)(R_1+R_2+2)(R_2+1)}$$

$$-\frac{\{1-F(x_7)\}^{(R_3+R_4+R_5+3)}\{1-F(x_2)\}^{(R_3+1)}}{(R_2+1)(R_2+R_3+2)(R_2+R_3+R_4+3)(R_1+1)}$$

$$+\frac{\{1-F(x_7)\}^{(R_2+R_3+R_4+R_5+4)}\{1-F(x_2)\}^{0}}{(R_1+1)(R_1+R_2+2)(R_1+R_2+R_3+3)(R_1+R_2+R_3+R_4+4)}. \tag{4.39}$$

In the case $p=3$, $j=5$ and $q=2$ when the censoring scheme $\mathbf{R}=(R,R,\ldots,R)$, equation (4.23) becomes

$$f^{\mathbf{R}}_{2,5,7:m:n}(x_{5:m:n}\mid x_{2:m:n},x_{7:m:n})=f(x_5)\{1-F(x_5)\}^{R}\times\left\{\left\{\frac{1}{(R+1)(2R+2)}\{1-F(x_2)\}^{2R+2}\right.\right.$$

$$\left.-\frac{1}{(R+1)^2}\{1-F(x_5)\}^{R+1}\{1-F(x_2)\}^{R+1}+\frac{1}{2(R+1)^2}\{1-F(x_5)\}^{2R+2}\right\}$$

$$\left.\times\left\{\frac{1}{(R+1)}\{1-F(x_5)\}^{R+1}-\frac{1}{(R+1)}\{1-F(x_7)\}^{R+1}\right\}\right\}$$

$$/\left\{\frac{\{1-F(x_2)\}^{4(R+1)}}{(R+1)4(R+1)2(R+1)3(R+1)}-\frac{\{1-F(x_7)\}^{R+1}\{1-F(x_2)\}^{3(R+1)}}{3(R+1)2(R+1)(R+1)^2}\right.$$

$$+\frac{\{1-F(x_7)\}^{2(R+1)}\{1-F(x_2)\}^{2(R+1)}}{(R+1)2(R+1)2(R+1)(R+1)}-\frac{\{1-F(x_7)\}^{3(R+1)}\{1-F(x_2)\}^{R+1}}{(R+1)2(R+1)3(R+1)(R+1)}$$

$$\left.+\frac{\{1-F(x_7)\}^{4(R+1)}}{(R+1)2(R+1)3(R+1)4(R+1)}\right\} \tag{4.40}$$

$$f^{\mathbf{R}}_{2,5,7:m:n}(x_{5:m:n} \mid x_{2:m:n}, x_{7:m:n}) = f(x_5)\{1-F(x_5)\}^R \times \left\{\left\{\frac{1}{2(R+1)^2}\{1-F(x_2)\}^{2(R+1)}\right.\right.$$

$$\left.-\frac{1}{(R+1)^2}\{1-F(x_5)\}^{R+1}\{1-F(x_2)\}^{R+1}+\frac{1}{2(R+1)^2}\{1-F(x_5)\}^{2(R+1)}\right\}$$

$$\left.\times\left\{\frac{1}{(R+1)}\{1-F(x_5)\}^{R+1}-\frac{1}{(R+1)}\{1-F(x_7)\}^{R+1}\right\}\right\}$$

$$/\left\{\frac{\{1-F(x_2)\}^{4(R+1)}}{4!(R+1)^4}(R+1)-\frac{\{1-F(x_7)\}^{R+1}\{1-F(x_2)\}^{3(R+1)}}{3!(R+1)^4}\right.$$

$$+\frac{\{1-F(x_7)\}^{2(R+1)}\{1-F(x_2)\}^{2(R+1)}}{4(R+1)^4}-\frac{\{1-F(x_7)\}^{3(R+1)}\{1-F(x_2)\}^{R+1}}{3!(R+1)^4}$$

$$\left.+\frac{\{1-F(x_7)\}^{4(R+1)}}{4!(R+1)^4}\right\}. \tag{4.41}$$

Due to simplification, some expression in equation (4.41) can be coded. If $\{1-F(x_2)\}^{R+1}=a$, $\{1-F(x_5)\}^{R+1}=b$ and $\{1-F(x_7)\}^{R+1}=c$, then the equation becomes

$$f^{\mathbf{R}}_{2,5,7:m:n}(x_{5:m:n}\mid x_{2:m:n},x_{7:m:n})=\left\{f(x_5)\{1-F(x_5)\}^R\left\{\frac{(a-b)^2}{2(R+1)^2}\frac{(b-c)}{R+1}\right\}\right\}$$

$$/\left\{\frac{a^4}{4!(R+1)^4}-\frac{ca^3}{3!(R+1)^4}+\frac{c^2a^2}{4(R+1)^4}-\frac{c^3a}{3!(R+1)^4}+\frac{c^4}{4!(R+1)^4}\right\}$$

$$=\left\{f(x_5)\{1-F(x_5)\}^R\left\{\frac{(a-b)^2}{2(R+1)^2}\frac{(b-c)}{R+1}\right\}\right\}/\left\{\frac{a^4-4ca^3+6c^2a^2-4c^3a+c^4}{4!(R+1)^4}\right\}$$

$$=\left\{f(x_5)\{1-F(x_5)\}^R\left\{\frac{(a-b)^2}{2(R+1)^2}\frac{(b-c)}{R+1}\right\}\right\}/\left\{\frac{(a-c)^4}{4!(R+1)^4}\right\}$$

$$=\left\{f(x_5)\{1-F(x_5)\}^R\left\{\frac{(a-b)^2(b-c)}{2}\right\}\right\}/\left\{\frac{(a-c)^4}{4!(R+1)}\right\}$$

$$=12(R+1)\frac{f(x_5)\{1-F(x_5)\}^R(a-b)^2(b-c)}{(a-c)^4}. \tag{4.42}$$

If 12(R+1) = k, then the conditional probability density function of $X^{\mathbf{R}}_{j:m:n}$ given $X^{\mathbf{R}}_{j-p:m:n}$ and $X^{\mathbf{R}}_{j+q:m:n}$ is

$$f^{\mathbf{R}}_{j,j-p,j+q:m:n}(x_{j:m:n}\mid x_{j-p:m:n},x_{j+q:m:n})=k\{f(x_j)\{1-F(x_j)\}^R$$

$$\times\{\{1-F(x_{j-p})\}^{R+1}-\{1-F(x_j)\}^{R+1}\}^2\{\{1-F(x_j)\}^{R+1}-\{1-F(x_{j+q})\}^{R+1}\}\}$$

$$/\{\{1-F(x_{j-p})\}^{R+1}-\{1-F(x_{j+q})\}^{R+1}\}^{4}. \tag{4.43}$$

Since $p = 3, j = 5$ and $q = 2$,

$$f^{\mathbf{R}}_{2,5,7:m:n}(x_{5:m:n} \mid x_{2:m:n}, x_{7:m:n})$$

$$= k\frac{f(x_5)\{1-F(x_5)\}^{R}\{\{1-F(x_2)\}^{R+1}-\{1-F(x_5)\}^{R+1}\}^{2}\{\{1-F(x_5)\}^{R+1}-\{1-F(x_7)\}^{R+1}\}}{\{\{1-F(x_2)\}^{R+1}-\{1-F(x_7)\}^{R+1}\}^{4}} \tag{4.44}$$

where $k = 12(R+1)$.

In order to obtain the general expression for $j = u$, $j + q = y$ and $j - p = x$, and also $p - 1 = 2$ and $q - 1 = 1$, equation (4.44) can be written like the expression in equation (4.20).

$$= k\frac{f(t)}{F(y)-F(x)}\left\{\frac{F(t)-F(x)}{F(y)-F(x)}\right\}^{2}\left\{1-\frac{F(t)-F(x)}{F(y)-F(x)}\right\}. \tag{4.45}$$

Due to Theorem 4.3, any property of order statistics can be easily transformed to the corresponding property of Progressively Type II censored order statistics with censoring scheme (R, R,…, R). In particular, any characterization of a probability distribution in terms of order statistics can be rewritten as a characterization in terms of Progressively Type II censored order statistics with censoring scheme (R, R,…, R). Theorem 4.4 can serve as an example.

Theorem 4.4 Assume that the random variable X has absolutely continuous and strictly increasing distribution function $F(x)$ with left and right extremities a_F and b_F, respectively. Let $G(x)$ be also an absolutely continuous and strictly increasing distribution function with left and right extremities $a_G = a_F$ and $b_G = b_F$. Then $F(x) = 1-(1-G(x))^{n/m}$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[h'(G(X_{j:m:n}^{\mathbf{R}})) \mid X_{j-p:m:n}^{\mathbf{R}} = x,\ X_{j+k+1-p:m:n}^{\mathbf{R}} = y] = \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}$$

(4.46)

holds for all $a_X < x < y < b_X$. $\mathbf{R} = (R, R, ..., R)$ and the number j, n, k are fixed and satisfies (4.11). Proof of Theorem 4.4 follows easily from Theorems 4.2 and 4.3. Consider some special cases.

Corollary 4.2 *i*) Let j be a positive number satisfying $3 \le j \le n - 2$. A random variable X has absolutely continuous strictly increasing distribution function $G(x)$ if and only if the representation

$$\frac{1}{2}E\{X_{j:m:n} \mid X_{j-1:m:n} = x, X_{j+2:m:n} = y\} + \frac{1}{2}E\{X_{j:m:n} \mid X_{j-2:m:n} = x, X_{j+1:m:n} = y\}$$

$$= \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}$$

holds for all $(x, y) \in D$.

ii) Let $4 \le j \le n - 3$. A random variable X has absolutely continuous and strictly increasing distribution function $G(x)$ if and only if the representation

$$\frac{1}{3}E\{X_{j:m:n} \mid X_{j-1:m:n} = x, X_{j+3:m:n} = y\} + \frac{1}{3}E\{X_{j:m:n} \mid X_{j-2:m:n} = x, X_{j+2:m:n} = y\}$$

$$+\frac{1}{3}E\{X_{j:m:n} \mid X_{j-3:m:n} = x, X_{j+1:m:n} = y\} = \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}$$

holds for all $(x, y) \in D$.

The characterization of Generalized Beta distribution for order statistics and Theorem 4.4 immediately implies the characterization of this distribution in terms of Progressively Type II censored order statistics. Similarly, one can transform characterizations of Uniform, Weibull (Exponential) and Pareto distributions given in terms of order statistics into characterizations in terms of Progressively Type II censored order statistics.

4.2.1 Characterizations for Some Special Distributions

4.2.1.1 Uniform Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support (0,1) has Uniform distribution over this interval if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[X^{\mathbf{R}}_{j:m:n} \mid X^{\mathbf{R}}_{j-p:m:n} = x, X^{\mathbf{R}}_{j+k+1-p:m:n} = y]$$

$$= 1 - \frac{R+1}{R+2}\frac{c(x,y;R+2)}{c(x,y;R+1)},\ k+1 \le j \le n-k \tag{4.47}$$

holds for all $0 < x < y < 1$, where $c(x,y;R+1) = (1-x)^{R+1} - (1-y)^{R+1}$. The result follows from the Theorem 4.4 by a choice of $h(x) = x + \frac{R+1}{R+2}(1-x)^{(R+2)/(R+1)} - \frac{R+1}{R+2}$, $h'(x) = 1-(1-x)^{1/(R+1)}$, $G(x) = 1-(1-x)^{R+1}$.

4.2.1.2 Weibull Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support $[0,\infty)$ has Weibull distribution $F(x)=1-\exp(-\alpha x^{\beta})$, $x\geq 0$, $\alpha>0$, $\beta>0$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k}E[X_{j:m:n}^{\mathbf{R}}\mid X_{j-p:m:n}^{\mathbf{R}}=x, X_{j+k+1-p:m:n}^{\mathbf{R}}=y]$$

$$=\frac{1}{R+1}+\frac{\alpha[x^{\beta}(\exp(-\alpha x^{\beta}))^{R+1}-y^{\beta}(\exp(-\alpha y^{\beta}))^{R+1}]}{(\exp(-\alpha x^{\beta}))^{R+1}-(\exp(-\alpha y^{\beta}))^{R+1}} \tag{4.48}$$

holds for all $0\leq x<y<\infty$ and $k+1\leq j\leq n-k$, here $h(x)=-\ln(1-x)^{1/(R+1)}x+\frac{x}{R+1}+\frac{\ln(1-x)}{R+1}-\frac{1}{R+1}$, $h'(x)=-\ln(1-x)^{1/(R+1)}$ and $G(x)=1-(\exp(-\alpha x^{\beta}))^{R+1}$.

Special case. When $\beta=1$, X has Exponential distribution, then the following characterization for the Exponential distribution can be given. X has Exponential distribution $F(x)=1-\exp(-\alpha x)$, $x\geq 0$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k}E[X_{j:m:n}^{\mathbf{R}}\mid X_{j-p:m:n}^{\mathbf{R}}=x, X_{j+k+1-p:m:n}^{\mathbf{R}}=y]$$

$$=\frac{1}{R+1}+\frac{\alpha[x(\exp(-\alpha x))^{R+1}-y(\exp(-\alpha y))^{R+1}]}{(\exp(-\alpha x))^{R+1}-(\exp(-\alpha y))^{R+1}} \tag{4.49}$$

holds for all $0\leq x<y<\infty$.

4.2.1.3 Generalized Beta Distribution

The absolutely continuous random variable X with strictly increasing distribution function having support $[a,b]$ has Generalized Beta distribution $F(x) = 1 - \frac{(b-x)^{\theta}}{(b-a)^{\theta}}$, $a \le x \le b$, $\theta > 0$, $-\infty < a < b < \infty$, if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[X_{j:m:n}^{\mathbf{R}} \mid X_{j-p:m:n}^{\mathbf{R}} = x, X_{j+k+1-p:m:n}^{\mathbf{R}} = y]$$

$$= b + \frac{\theta(R+1)(a-b)}{\theta(R+1)+1}\frac{d(x,y;\theta+1)}{d(x,y;\theta)}, \tag{4.50}$$

holds for all $a \le x < y < b$ and $k+1 \le j \le n-k$, where $d(x,y;\theta) = \left(\frac{b-x}{b-a}\right)^{\theta(R+1)} - \left(\frac{b-y}{b-a}\right)^{\theta(R+1)}$. The result is obtained by using target function $h(x) = bx - \frac{(a-b)(1-x)^{1+1/(\theta(R+1))}}{1+1/(\theta(R+1))}$ and $h'(x) = b - (b-a)(1-x)^{1/(\theta(R+1))}$ and $G(x) = 1 - \frac{(b-x)^{\theta(R+1)}}{(b-a)^{\theta(R+1)}}$.

The above characterization with a = 0, b = 1, θ = 1 turns into a characterization of Uniform distribution.

4.2.1.4 Pareto Distribution

The absolutely continuous random variable X has strictly increasing distribution function having support [γ, ∞) has Pareto distribution $F(x) = 1 - \frac{(\gamma+\delta)^{\theta}}{(x+\delta)^{\theta}}$, $x \geq \gamma$, $\theta > 0$, $\gamma + \delta > 0$ if and only if the representation

$$\frac{1}{k}\sum_{p=1}^{k} E[X_{j:m:n}^{\mathbf{R}} \mid X_{j-p:m:n}^{\mathbf{R}} = x, X_{j+k+1-p:m:n}^{\mathbf{R}} = y]$$

$$= -\frac{\theta(R+1)}{\theta(R+1)-1}\frac{e(x,y;\theta-1)}{e(x,y;\theta)} - \delta \tag{4.51}$$

holds for all $\gamma \leq x < y < \infty$ and $k+1 \leq j \leq n-k$, where $e(x,y;\theta) = \left(\frac{1}{x+\delta}\right)^{\theta(R+1)} - \left(\frac{1}{y+\delta}\right)^{\theta(R+1)}$. Here the target function $h(x) = \frac{(\gamma+\delta)}{(-1/(\theta(R+1)))+1}(1-x)^{(-1/(\theta(R+1)))+1} - \delta x$, $h'(x) = -(\gamma+\delta)(1-x)^{-1/(\theta(R+1))} - \delta$ and $G(x) = 1 - \frac{(\gamma+\delta)^{\theta(R+1)}}{(x+\delta)^{\theta(R+1)}}$.

CHAPTER FIVE

APPLICATIONS TO GOODNESS-OF-FIT TESTS

In this chapter, the applications of characterization theorems involving order statistics and Progressively Type II censored order statistics to goodness-of-fit tests are discussed. The idea on how to use characterization theorems in goodness-of-fit tests is described in Balasubramanian & Dey (1997).

"The history of goodness-of-fit tests started with the seminal paper of Karl Pearson in 1900 on chi-squared tests" Balakrishnan & Lin (2003).

> Balakrishnan, Ng, & Kannan (2004) describe the goodness-of-fit test as *"In statistics, the problem of testing for the adequacy of model specifications is extremely important. The validity of inferential procedures depends in great part on the assumption of a particular underlying distribution for the data. There have been numerous procedures developed in the literature for determining whether a random sample comes from some specific distribution. All these procedures have been classified as 'goodness-of-fit' tests."*

The general procedure consists of defining a test statistic, which is some function of the data measuring the distance between the hypothesis and the data, and then calculating the probability of obtaining data which have a still larger value of this test statistic than the value observed, assuming the hypothesis is true. This probability is called the size of the test or confidence level. Small probabilities indicate a poor fit. Especially high probabilities correspond to a fit which is too good to happen very often and may indicate a mistake in the way the test was applied, such as treating data as independent when they are correlated. The most common tests for goodness-of-fit are the chi-square test, Kolmogorov test, Smirnov test and Cramer – von Mises test.

Suppose a random experiment with a random variable X of interest. Assume additionally that X is absolutely continuous with probability density function $f(x)$ on a finite set S. The experiment is repeated n times to generate a random sample of size n from the distribution of X : $X_1, X_2, ..., X_n$. These are independent random variables, each with the distribution of X.

In this chapter, it is assumed that the distribution of X is unknown. For a given distribution function $F_0(x)$, the hypotheses will be tested

$$H_0: F(x) = F_0(x) \text{ versus } H_1: F(x) \neq F_0(x).$$

The test that will be constructed is known as the goodness-of-fit test for the conjectured density $F_0(x)$. As usual, the challenge in developing the test is to find a good test statistic that gives information about the hypotheses and whose distribution, under the null hypothesis, is known, at least approximately.

Balakrishnan et al. (2002) proposed a goodness-of-fit test for the exponential distribution when the available sample is Progressively Type II censored. Balakrishnan et al. (2002, 2004) proposed goodness-of-fit tests for testing exponentiality and a general location scale distribution, based on Progressively Type II censored samples. Balasubramanian et al. (1997) presented results that lead to the characterization of several distributions using conditional expectations. They considered both absolutely continuous random variables and discrete random variables. In the case of discrete random variables, the distribution is almost uniquely determined under the stated conditions while in the case of absolutely continuous random variables, the results lead to the characterization of a family of distributions.

5.1 Applications to Goodness-of-fit Tests for Order Statistics

Assume that $X_1,...,X_n$ is a random sample of *n* i.i.d. observations with an unknown distribution *F*(*x*). By using Theorem 4.2, several distributions are characterized through order statistics. It is introduced that the characterizations for Uniform, Weibull, Generalized Beta and Pareto distribution can be used to test whether or not *F*(*x*) belongs to the these distributions.

Under hypothesis H_0, it is expected that $E(S)=0$. Hence large values of $|S|$ lead to the rejection of the hypothesis H_0. As this test is based on a characterization property, it is likely to have large power. Characterizations are made very useful by these inferential aspects.

5.1.1 Goodness-of-fit test for Uniform Distribution

Consider the problem of testing H_0 : *F*(*x*) is Uniform.

Under H_0 : *F*(*x*) is Uniform

$$S=\frac{1}{k}\sum_{p=1}^{k} X_{j:n}^{c}-\frac{(X_{j+k+1-p:n})^{c+1}-(X_{j-p:n})^{c+1}}{(c+1)(X_{j+k+1-p:n}-X_{j-p:n})} \tag{5.1}$$

holds for all $0<x<y<1$. In (5.1) $j\in\{k+1,\ k+2,...,\ n-k\}$ is fixed and $c\geq 1$.

The test : H_0 would be rejected

$|S|>0$.

5.1.2 Goodness-of-fit Test for Weibull Distribution

Consider the problem of testing $H_0 : F(x)$ is Weibull.

Under $H_0 : F(x)$ is Weibull

$$S = \frac{1}{k}\sum_{p=1}^{k} \alpha X_{j:n}^{\beta}$$

$$- \frac{\exp(-\alpha X_{j-p:n}^{\beta})(\alpha X_{j-p:n}^{\beta} + 1) - \exp(-\alpha X_{j+k+1-p:n}^{\beta})(\alpha X_{j+k+1-p:n}^{\beta} + 1)}{\exp(-\alpha X_{j-p:n}^{\beta}) - \exp(-\alpha X_{j+k+1-p:n}^{\beta})} \tag{5.2}$$

holds for all $0 \le x < y < \infty$. Here j satisfies (4.11).

The test : H_0 would be rejected

$$|S| > 0 .$$

5.1.3 Goodness-of-fit Test for Generalized Beta Distribution

Consider the problem of testing $H_0 : F(x)$ is Generalized Beta.

Under $H_0 : F(x)$ is Generalized Beta

$$S = \frac{1}{k}\sum_{p=1}^{k} X_{j:n} - b + \frac{(a-b)\theta}{\theta+1}\frac{a(X_{j-p:n}, X_{j+k+1-p:n}; \theta+1)}{a(X_{j-p:n}, X_{j+k+1-p:n}; \theta)} \tag{5.3}$$

holds for all $a < x < y < b$ where $a(x, y; \theta) = \left(\frac{b-x}{b-a}\right)^{\theta} - \left(\frac{b-y}{b-a}\right)^{\theta}$. Here j satisfies (4.11).

The test : H_0 would be rejected

$$|S| > 0.$$

5.1.4 Goodness-of-fit Test for Pareto Distribution

Consider the problem of testing $H_0 : F(x)$ is Pareto.

Under $H_0 : F(x)$ is Pareto

$$S = \frac{1}{k}\sum_{p=1}^{k} X_{j:n} - \left[-\frac{\theta}{\theta - 1} \frac{b(X_{j-p:n}, X_{j+k+1-p:n}; \theta - 1)}{b(X_{j-p:n}, X_{j+k+1-p:n}; \theta)} - \delta \right] \tag{5.4}$$

holds for all $\gamma \le x < y < \infty$ where $b(x, y; \theta) = \left(\frac{1}{x+\delta}\right)^{\theta} - \left(\frac{1}{y+\delta}\right)^{\theta}$, $\theta \ne 1$ and j satisfies (4.11).

The test : H_0 would be rejected

$$|S| > 0.$$

Theorem 4.4 can also be used similarly.

5.2 Applications to Goodness-of-fit Tests for Progressively Type II Censored Order Statistics

Assume that $X_1, X_2 ..., X_n$ is a random sample of n i.i.d. observations with an unknown distribution $F(x)$. By using Theorem 4.4, several distributions are characterized through Progressively Type II censored order statistics. It is introduced that the characterizations for Uniform, Weibull, Generalized Beta and Pareto distribution can be used to test whether or not $F(x)$ belongs to the these distributions.

5.2.1 Goodness-of-fit Test for Uniform Distribution

Consider the problem of testing H_0 : $F(x)$ is Uniform.

Under H_0 : $F(x)$ is Uniform

$$S = \frac{1}{k}\sum_{p=1}^{k} X_{j:m:n}^{\mathbf{R}} - 1 - \frac{R+1}{R+2}\frac{c(X_{j-p:m:n}^{\mathbf{R}}, X_{j+k+1-p:m:n}^{\mathbf{R}}; R+2)}{c(X_{j-p:m:n}^{\mathbf{R}}, X_{j+k+1-p:m:n}^{\mathbf{R}}; R+1)} \tag{5.5}$$

holds for all $k+1 \le j \le n-k$, $0 < x < y < 1$ where $c(x, y; R+1) = (1-x)^{R+1} - (1-y)^{R+1}$.

The test : H_0 would be rejected

$$|S| > 0\,.$$

5.2.2 Goodness-of-fit Test for Weibull Distribution

Consider the problem of testing $H_0 : F(x)$ is Weibull.

Under $H_0 : F(x)$ is Weibull

$$S = \frac{1}{k}\sum_{p=1}^{k} X^{\mathbf{R}}_{j:m:n} - \frac{1}{R+1}$$

$$+\frac{\alpha[(X^{\mathbf{R}}_{j-p:m:n})^{\beta}(\exp(-\alpha(X^{\mathbf{R}}_{j-p:m:n})^{\beta}))^{R+1}-(X^{\mathbf{R}}_{j+k+1-p:m:n})^{\beta}(\exp(-\alpha(X^{\mathbf{R}}_{j+k+1-p:m:n})^{\beta}))^{R+1}]}{(\exp(-\alpha(X^{\mathbf{R}}_{j-p:m:n})^{\beta}))^{R+1}-(\exp(-\alpha(X^{\mathbf{R}}_{j+k+1-p:m:n})^{\beta}))^{R+1}} \tag{5.6}$$

holds for all $0 < x < y < \infty$ and $k+1 \le j \le n-k$.

The test : H_0 would be rejected

$$|S| > 0.$$

5.2.3 Goodness-of-fit Test for Generalized Beta Distribution

Consider the problem of testing $H_0 : F(x)$ is Generalized Beta.

Under $H_0 : F(x)$ is Generalized Beta

$$S = \frac{1}{k}\sum_{p=1}^{k} X^{\mathbf{R}}_{j:m:n} - b + \frac{\theta(R+1)(a-b)}{\theta(R+1)+1}\frac{d(X^{\mathbf{R}}_{j-p:m:n}, X^{\mathbf{R}}_{j+k+1-p:m:n};\theta+1)}{d(X^{\mathbf{R}}_{j-p:m:n}, X^{\mathbf{R}}_{j+k+1-p:m:n};\theta)} \tag{5.7}$$

holds for all $a \le x < y < b$ and $k+1 \le j \le n-k$ where $d(x, y; \theta) = \left(\frac{b-x}{b-a}\right)^{\theta(R+1)} - \left(\frac{b-y}{b-a}\right)^{\theta(R+1)}$.

The test : H_0 would be rejected

$$|S| > 0.$$

5.2.4 Goodness-of-fit Test for Pareto Distribution

Consider the problem of testing $H_0 : F(x)$ is Pareto.

Under $H_0 : F(x)$ is Pareto

$$S = \frac{1}{k}\sum_{p=1}^{k} X_{j:m:n}^{\mathbf{R}} - \left[-\frac{\theta(R+1)}{\theta(R+1)-1} \frac{e(X_{j-p:m:n}^{\mathbf{R}}, X_{j+k+1-p:m:n}^{\mathbf{R}}; \theta-1)}{e(X_{j-p:m:n}^{\mathbf{R}}, X_{j+k+1-p:m:n}^{\mathbf{R}}; \theta)} - \delta \right] \tag{5.8}$$

holds for all $\gamma < x < y < \infty$ and $k+1 \le j \le n-k$ where $e(x, y; \theta) = \left(\frac{1}{x+\delta}\right)^{\theta(R+1)} - \left(\frac{1}{y+\delta}\right)^{\theta(R+1)}$.

The test : H_0 would be rejected

$$|S| > 0.$$

CHAPTER SIX
CONCLUSIONS

In this thesis, some general models of ordered random variables which are order statistics, Progressively Type II censored order statistics and generalized order statistics are investigated. Balakrishnan, Childs et al. (2002) have introduced the probability density function of two nonadjacent Progressively Type II censored order statistics in explicit form. As an extension of this result, the probability density function of k nonadjacent Progressively Type II censored order statistics is obtained by using explicit form.

It is shown that for a particular censoring scheme $\mathbf{R} = (R, R, ..., R)$, the joint distribution of Progressively Type II censored order statistics can be reduced to the joint distribution of ordinary order statistics with the sample size m. Using this result, some characterization theorems for important classes of distributions are given. These theorems involve the conditional expected values of ordinary order statistics and Progressively Type II censored order statistics.

Also, applications in goodness-of-fit tests are discussed. Balasubramanian & Dey (1997) introduced the goodness-of-fit tests for order statistics by using characterization theorems. In the same way, the goodness-of-fit test statistics are described for some general classes of distributions.

Throughout the thesis the distribution functions are assumed to be absolutely continuous and strictly increasing. The relation that gives the reduction of Progressively Type II censored order statistics to order statistics when the censoring scheme $\mathbf{R} = (R, R, ..., R)$ is also valid for discrete distributions. Balakrishnan & Dembinska (2007) investigated whether this relation holds true for discontinuous $F(x)$ and they have proposed the theorem giving an affirmative answer to this question. This reduction theorem enables to extend distributional properties of order statistics into properties of Progressively Type II censored order statistics with censoring scheme $\mathbf{R} = (R, R, ..., R)$.

REFERENCES

Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A first course in order statistics*. New York: John Wiley.

Azlarov, T. A., & Volodin, N. A. (1986). *Characterization problems associated with the exponential distribution*. New York: Springer-Verlag.

Bairamov, I. (2006) Progressive Type II censored order statistics for multivariate observations. *Journal of Multivariate Analysis*, 97, 797–809.

Bairamov, I. G., & Bekci, M. (1999). Concomitant of order statistics in FGM type bivariate uniform distributions. *Istatistik*, *Journal of Turkish Statistical Association*, 2, 135–144.

Bairamov, I. G., & Eryılmaz, S. (2006) Spacings, exceedances and concomitants in progressive Type II censoring scheme. *Journal of Statistical Planning and Inference*, 136, 527-536.

Bairamov, I. G., & Gebizlioglu, O. L. (1997). On the ordering of random vectors in a norm sense. *Journal of Applied Statistical Science*, 6(1), 77–86.

Bairamov, I., & Özkal, T. (2007) On characterization of distributions through the properties of conditional expectations of order statistics. *Communications in Statistics: Theory and Methods*, 36, 1319-1326.

Balakrishnan, N., & Aggarwala, R. (2000). *Progressive censoring: Theory, methods and applications*. Boston: Birkhauser.

Balakrishnan, N., Childs, A., & Chandrasekar, B. (2002). An efficient computational method for moments of order statistics under progressive censoring. *Statistics & Probability Letters*, 60, 359-365.

Balakrishnan, N., Cramer, E., Kamps, U., & Schenk, N. (2001). Progressive type II censored order statistics from exponential distributions. *Statistics*, 35, 537-556.

Balakrishnan, N., & Dembinska, A. (2007). Progressively Type-II right censored order statistics from discrete distributions. *Journal of Statistical Planning and Inference*, article in press.

Balakrishnan, N., Kannan, N., Lin, C. -T., & Ng, H. K. T. (2003). Point & interval estimation for Gaussian distribution, based on progressively Type-II censored samples. *IEEE Transactions on Reliability*, 52, 90-95.

Balakrishnan, N., & Lin, C.-T. (2002). Exact linear inference and prediction for exponential distributions based on general progressively Type-II censored samples. *Journal of Statistical Computation and Simulation*, 72, 677-686.

Balakrishnan, N., & Lin, C.-T. (2003). On the distribution of a test for exponentiality based on progressively Type-II right censored spacings. *Journal of Statistical Computation and Simulation*, 73, 277-283.

Balakrishnan, N., Ng, H. K. T., Kannan, N. (2002). A test of exponentiality based on spacings for progressively Type-II censored data. In C. Huber-Carol, N. Balakrishnan, M. S. Nikulin, & M. Mesbah, (Eds.). *Goodness-of-fit tests and model validity* (89-111). Boston: Birkhauser.

Balakrishnan, N., Ng, H. K. T., Kannan, N. (2004). Goodness-of-fit tests based on spacings for progressively Type-II censored data from a general location-scale distribution. *IEEE Transactions on Reliability*, 53, 349-356.

Balakrishnan, N., & Sandhu, R. A. (1995). A simple simulational algorithm for generating progressive Type II censored samples. *American Statistician*, 49, 229-230.

Balakrishnan, N., & Sandhu, R. A. (1996). Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive Type II censored samples. *Sankhy*ã, *Ser. B*, 1 (58), 1-9.

Balasooriya, U., & Balakrishnan, N. (2000). Reliability sampling plans for lognormal distribution, based on progressively censored samples. *IEEE Transactions on Reliability*, 49, 199-203.

Balasooriya, U., Saw, S. L. C., & Gadag, V. (2000). Progressively censored reliability sampling plans for the Weibull distribution. *Technometrics*, 42, 160–168.

Balasubramanian, K., & Beg, M. I. (1992). Distributions determined by conditioning on a pair of order statistics. *Metrika*, 39,107-112.

Balasubramanian, K., & Beg, M. I. (1997). Concomitant of order statistics in Morgenstern type bivariate exponential distribution. *Journal of Applied Statistical Science*, 5, 233-245.

Balasubramanian, K., & Dey, A. (1997). Distributions characterized through conditional expectations. *Metrika*, 45, 189-196.

Barnett, V., Green, P. J., & Robinson, A. (1976). Concomitants and correlation estimates. *Biometrika*, 63, 323-328.

Beg, M. I., & Balasubramanian, K. (1990). Distributions determined by conditioning on a single order statistic. *Metrika*, 37 (1), 37-43.

Beg, M. I., & Kirmani, S. N. U. A. (1974). On a characterization of exponential and related distributions. *Australian Journal of Statistics*, 16, 163-166.

Bhattacharya, P. K. (1974). Covergence of sample paths of normalized sums of induced order statistics. *Annals of Statistics*, 5, 1034-1039.

Bhttacharya, P. K. (1984). Induced order statistics. In P. R. Krishnaiah, & P. K. Sen (Eds.). *Handbook of Statistics* (Vol. 4) (383-403).

Bordes, L. (2004). Non-parametric estimation under progressive censoring. *Journal of Statistical Planning and Inference*, 119, 171-189.

Cohen, A. C. (1963). Progressively censored samples in life testing. *Technometrics*, 5, 327-329.

Cohen, A. C. (1991). *Truncated and censored samples*. New York: Marcel Dekker.

David, H. A. (1973). Concomitants of order statistics. *Bulletin of the International Statistical Institute*, 45, 295-300.

David, H. A. (1981). *Order statistics* (2nd ed.). New York: Wiley.

David, H. A. (1993). Concomitant of order statistics: Review and recent developments. In F. M. Hoppe, (Ed.). *Multiple comparisons, selection and application in biometry* (507-518). New York: Dekker.

David, H. A., & Galambos, J. (1974). The asymptotic theory of concomitants of order statistics. *Journal of Applied Probability*, 11, 762–770.

David, H. A., & Nagaraja, H. N. (1998). Concomitants of order statistics. In N. Balakrishnan, & C. R. Rao, (Eds.). *Handbook of statistics, order statistics: Theory and methods* (Vol. 16) (487-513). New York: Elsevier.

David, H. A., & Nagaraja, H. N. (2003). *Order statistics* (3rd ed.). New Jersey: John Wiley & Sons.

David, H. A., O'Connell, M. J., & Yang, S. S. (1977). Distribution and expected value of the rank of a concomitant of an order statistics. *The Annals of Statistics*, 5, 216–223.

Eryılmaz, S., & Bairamov, I. (2003). On a new sample rank of an order statistics and its concomitant. *Statistical Probability Letters*, 63, 123–131.

Ferguson, T. S. (1967). On characterizations of distributions by properties of order statistics. *Sankhy*ã, Ser. A, 29, 265-278.

Fernandez, A. J. (2004). On estimating exponential parameters with general Type II progressive censoring. *Journal of Statistical Planning and Inference*, 121, 135-147.

Galambos, J., & Kotz, S. (1978). *Characterization of probability distributions, lecture notes in mathematics* (675). New York: Springer-Verlag.

Goel, P. K., & Hall, P. (1994). On the average difference between concomitants and order statistics. *Annals of Probability*, 22 (1), 126-144.

Gupta, R. C. (1975). On characterization of distribution by conditional expectations. *Communications in Statistics*, 4 (1), 99-103.

Gupta, S. D., Goswami, A., & Rao, B. V. (1993). On a characterization of uniform distribution. *Journal of Multivariate Analysis*, 44 (1), 102-114.

Hamdan, M. A. (1972). On a characterization by conditional expectations. *Technometrics*, 14, 497-499.

Kamps, U. (1995a). *A concept of generalized order statistics*. Stuttgart: Teubner.

Kamps, U. (1995b). A concept of generalized order statistics. *Journal of Statistical Planning and Inference*, 48,1-23.

Kotlarski, I. I. (1972). On a characterization of some probability distributions by conditional expectations. *Sankhy*ã, Ser. A., 34, 461-466.

Kotz, S., & Shanbhag, D. N. (1980). Some new approaches to probability distributions. *Advances in Applied Probability*, 12, 903-921.

Montanari, G. C., & Cacciari, M. (1988). Progressively-censored aging tests on XLPE-insulated cable models. *IEEE Trans. Electrical Insul*., 23, 365–372.

Nagaraja, H., & David, H. A. (1994). Distribution of the maximum of concomitants of selected order statistics. *The Annals of Statistics*, 22 (1), 478-494.

Nasri-Roudsari, D. (1996). Extreme value theory of generalized order statistics. *Journal of Statistical Planning and Inference*, 55, 281-297.

Navarro, J., Ruiz, J. M., & Zoroa, N. (1998). A unified approach characterization problems using conditional expectations. *Journal of Statistical Planning and Inference*, 69, 193-207.

Ng, H. K. T., Chan, P. S., & Balakrishnan, N. (2002). Estimation of parameters from progressively censored data using EM algorithm. *Computational Statistics and Data Analysis*, 39, 371-386.

Ng, H. K. T., Chan, P. S., & Balakrishnan, N. (2004). Optimal progressive censoring plans for the Weibull distribution. *Technometrics*, 46, 470-481.

Ouyang, L. Y. (1983). On characterization of distributions by conditional expectations. *Tamkang Journal of Management Science*, 4(1), 13-21.

Ouyang, L. Y. (1987). On characterizations of probability distributions based on conditional expected values. *Tamkang Journal of Management Science*, 18(1), 113-122.

Ruiz, J. M., & Navarro, J. (1996). Characterizations based on conditional expectations of the doubly truncated distribution. *Annals of the Institute of Statistical Mathematics*, 48, 563-572.

Sen, P. K. (1976). A note on invariance principles for induced order statistics. *Annals of Probability*, 4, 474-479.

Sen, P. K. (1981). The Cox regression model, invariance principles for some induced quantile processes and some repeated significance tests. *Annals of Statistics*, 9(1), 109-121.

Sen, P. K. (1986). Progressive censoring schemes. In S. Kotz, & N. L. Johnson, (Eds.). *Encyclopedia of Statistical Sciences* (Vol. 7) (296-299). New York: Wiley.

Shanbhag, D. N., & Bhaskara Rao, M. (1975). A note on characterizations of probability distributions based on conditional expected values. *Sankhy*ã, Ser. A., 37, 297-300.

Stokes, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics*: *Theory and Methods*, 6, 1207-1211.

Talwalker, S. (1977). A note on characterization by the conditional expectation. *Metrika*, 24, 129-136.

Thomas, D. R., & Wilson, W. M. (1972). Linear order statistic estimation for the two-parameter Weibull and extreme-value distributions under Type II progressively censored samples. *Technometrics*, 14, 679–691.

Viveros, R., & Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data. *Technometrics*, 36, 84-91.

Yang, S. S. (1977). General distribution theory of the concomitants of order statistics. *Annals of Statistics*, 5, 996-1002.

Yeo, W. B., & David, H. A. (1984). Selection through an associated characteristic with applications to the random effects model. *Journal of the American Statistical Association*, 79, 399-405.

Zoroa, P., Ruiz, J. M., & Mrin, J. (1990). A characterization based on conditional expectations. *Communications in Statistics: Theory and Methods*, 19 (8), 3127-3135.