1.1 Concomitants of Order Statistics

Order statistics play a very important role in statistical theory and practice and accordingly a remarkably large body of literature has been devoted to its study. It helps to develop special methods of statistical inference, which are valid with respect to a broad class of distributions. Specific properties of order statistics are used to identify probability distributions in the form of characterizations, see Arnold, et al. (1992), Balakrishnan and Rao (1998). In spite of the established role of order statistics in statistical theory, discussions are largely confined to the univariate case and comparatively lesser volume of work is available in a multivariate setup. This is due to the fact that there is no straightforward way of extending the concept of order statistics from the univariate case to the multivariate case. A survey of different attempts of introducing multivariate order statistics can be found in Barnett (1976). Most of the theoretical development in this area of research is on concomitants of order statistics. The concept of concomitants, when bivariate data are ordered by one of its components, was first introduced by David (1973) and almost simultaneously under the name of induced order statistics by Bhattacharya (1974).

Let \((X_i, Y_i)\), \(i = 1, 2, \ldots, n\) be a random sample from a bivariate distribution with cumulative distribution function (cdf) \(F(x,y)\). If we order the values of \(X_i\)'s in the increasing order of magnitude, then the corresponding \(Y_i\)'s need not have a
similar order among themselves. Unless X and Y are independent, the ordering of X's will affect the distribution of the associated Y's. The Y value associated with or paired with X_{r:n}, the r^{th} X order statistic is called the concomitant of X_{r:n} and will be denoted by Y_{[r:n]}. The ordering of concomitants is similar to that of the marginal variable X if \( \rho = 1 \) and completely reversed if \( \rho = -1 \).

The most important use of concomitants is identified in selection problems when \( k (<n) \) individuals are chosen on their X values. Then the corresponding Y values represent performance of an associated characteristic. For example, if the k out of n rams as judged by their genetic make up is selected for breeding, then \( Y_{[n-k+1:n]} \ldots Y_{[n:n]} \) might represent the quality of the wool of one of their female offspring. Or, X might be the score of a candidate on a screening test and Y the score on a later test. Concomitants have found a wide variety of applications in such applied fields as selection procedure (Yeo and David (1984)) ocean engineering (Castillo (1988)) inference problems (Do and Hall (1992), Yang (1981a,b)), prediction analysis (Gross (1973)) and double sampling plans (David (1996), O'Connell and David (1976)). An excellent review of work on concomitants of order statistics is available in David and Nagaraja (1998).

1.2 Basic Distribution Theory and a Brief Review

In a basic paper of concomitants, David (1973), considered the bivariate normal model in which the variable Y is linked with X through the regression model

\[
Y = \mu_y + \rho \sigma_y \left( \frac{X - \mu_x}{\sigma_x} \right) + Z
\]

where \( Z \sim N (0, \sigma_y^2 (1 - \rho^2)) \) and Z is independent of X. Under this model he has derived the finite and asymptotic distribution of the concomitants. Thus for \( r = 1, 2 \ldots n \)
Introduction

\[ Y_{[r:n]} = \mu_y + \rho \sigma_y \left( \frac{X_{r:n} - \mu_x}{\sigma_x} \right) + Z_{[r]}, \]  

(1.2.2)

where \( Z_{[r]} \) denotes the particular \( Z_r \) associated with \( X_{r:n} \). In view of the independence of \( X_r \) and the \( Z_r \) we see that the set of \( X_{r:n} \) is independent of \( Z_{[r]} \). Moreover \( Z_{[r]} \) are mutually independent and \( Z_r \sim Z \). A more general model of (1.2.1) is discussed in Kim and David (1990). Let \( Y_i = g( X_i, \varepsilon_i ) \) represent a general model for the regression of \( Y \) on \( X \), where neither the \( X_i \) nor the \( \varepsilon_i \) need be identically distributed (but still be independent). Then

\[ Y_{[r:n]} = g( X_{r:n}, \varepsilon_{[r]} ) \quad r = 1, 2 \ldots n \]  

(1.2.3)

from the mutual independence of the \( X_i \) and the \( \varepsilon_i \) it follows that \( \varepsilon_{[r]} \) has the same distribution as the \( \varepsilon_i \) accompanying \( X_{r:n} \) and that the \( \varepsilon_{[r]} \) are mutually independent.

They have shown that concomitants are associated random variables. Moreover concomitants satisfy a stronger form of dependence, multivariate total positivity of order 2 if each \( \varepsilon_{[r]} \) in the general linear model has a Polya frequency of order two (Karlin and Rinott (1980)).

The general distribution of concomitants may be derived from the following Theorem due to Bhattacharya (1974).

**Theorem 1.1**

For \( 1 \leq r_1 < r_2 \ldots < r_k \leq n \), the \( Y_{[r_1:n]} \) (i = 1, 2 \ldots n) are conditionally independent given \( X_{r_1:n} = x_i \) \( (i = 1, 2 \ldots k) \) with joint conditional density function \( \prod_{i=1}^{k} f(y_i | x_i) \).

It follows from the above Theorem that the joint density function of the concomitants \( Y_{[r_1:n]}, Y_{[r_2:n]}, \ldots, Y_{[r_k:n]} \),

\[ f_{Y_{[r_1:n]}, Y_{[r_2:n]}, \ldots, Y_{[r_k:n]}}(y_1, \ldots, y_k) = \int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{Y_{[r_1:n]}, Y_{[r_2:n]}, \ldots, Y_{[r_k:n]}}(x_1, \ldots, x_k) \prod_{h=1}^{k} f(y_h | x_h) dx_h \]  

(1.2.4)
Yang (1977) has shown that

\[ E[Y_{r:n}] = E[m(X_{r:n})] \]

\[ \text{Var} (Y_{r:n}) = \text{Var}(m(X_{r:n}))+E (\sigma^2 (X_{r:n})) \]

\[ \text{Cov}(X_{r:n}, Y_{s:n}) = \text{Cov}(X_{r:n}, m(X_{s:n})) \]

\[ \text{Cov} (Y_{r:n}, Y_{s:n}) = \text{Cov} (m(X_{r:n}), m(X_{s:n})) \quad r \neq s \]

where

\[ m(x) = E (Y|X = x) \]

and

\[ \sigma^2(x) = Var(Y | X = x). \quad (1.2.5) \]

Jha and Hossein (1986) noted that (1.2.5) continues to hold when \( X \) is absolutely continuous but \( Y \) is discrete. They have derived the following important recurrence relations connecting the moments of concomitants for an arbitrary specified function \( h(.) \) such that \( E[h(Y)] \) exists.

\[ (n-r) E(h(Y_{r:n})) + r E(h(Y_{r+1:n})) = n E(h(Y_{r:n-1})), \quad r = 1, 2 \ldots n-1 \quad (1.2.6) \]

\[ E(h(Y_{k:m})) = \binom{m}{k} \sum_{i=0}^{j} (-1)^s \begin{pmatrix} k \ Cr \begin{pmatrix} i \ Cs \begin{pmatrix} j \ Sm - j + s \begin{pmatrix} k \ S \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} E(h(Y_{k-i,m-i+s})), \quad i \leq k, \quad 1 \leq k \leq m \leq n, \quad (1.2.7) \]

\[ E(h(Y_{k:m})) = \binom{m}{k} \sum_{s=0}^{j} (-1)^s \begin{pmatrix} k \ Cs \begin{pmatrix} j \ Sm - j + s \begin{pmatrix} k \ S \end{pmatrix} \end{pmatrix} \end{pmatrix} E(h(Y_{k+s,m-j+s})), \quad 0 \leq j \leq m-k \quad (1.2.8) \]

\[ E(h(Y_{r:n})) = \sum_{i=r}^{n} \binom{i-1}{r-1} \binom{n}{i} (-1)^{i-r} E(h(Y_{i:i})), \quad i = 1, 2 \ldots n \quad (1.2.9) \]
\[(r-1) \mathbb{E}(Y_{[r:n]} Y_{[s:n]}) + (s-r) \mathbb{E}(Y_{[r-1:n]} Y_{[s:n]}) + (n-s+1) \mathbb{E}(Y_{[r-1:n]} Y_{[s-1:n]}) = n \mathbb{E}(Y_{[r-1:n-1]} Y_{[s-1:n-1]}), \quad 1 \leq r < s \leq n \quad (1.2.10)\]

A new recurrence relation emerges from the above results.

**Theorem 1.2**

For \(2 \leq i < j \leq n\),

\[
(i-1) \beta_{[i,j]} + (j-1) \beta_{[i-1,j]} + (n-j+1) \beta_{[i-1,j-1]} = \\
n\{ \beta_{[i-1,j-1,n-1]} + [\mu_{[i-1,j-1]} - \mu_{[i-1,j]}][\mu_{[j-1,n-1]} - \mu_{[j,n]}] \} \\
\]

where

\[
\beta_{[i,j]} = \text{Cov}(Y_{[i,n]}, Y_{[j,n]})
\]

and

\[
\mu_{[i,n]} = \mathbb{E}(Y_{[i,n]})
\]

**Proof:** We have from (1.2.10)

\[
(i-1) \beta_{[i,j]} + (j-1) \beta_{[i-1,j]} + (n-j+1) \beta_{[i-1,j-1]} = \\
n\beta_{[i-1,j-1,n-1]} - (i-1)\mu_{[i,n]}\mu_{[j,n]} - (j-1)\mu_{[i-1,n]}\mu_{[j,n]} - (n-j+1)\mu_{[i-1,n]}\mu_{[j-1,n]} + \\
n\mu_{[i-1,n-1]}\mu_{[j-1,n-1]} \} \\
\]

\[\text{(1.2.12)}\]

Now consider,

\[
(i-1) \mu_{[i,n]}\mu_{[j,n]} + (j-1)\mu_{[i-1,n]}\mu_{[j,n]} + (n-j+1)\mu_{[i-1,n]}\mu_{[j-1,n]} = \mu_{[j,n]}(i-1)\mu_{[i,n]} + \\
(n-i+1)\mu_{[i-1,n]} + (n-j+1)\mu_{[i-1,n]}(\mu_{[j-1,n]} - \mu_{[j,n]}) \\
= n\mu_{[j,n]}\mu_{[i-1,n-1]} + n\mu_{[i-1,n]}(\mu_{[j-1,n-1]} - \mu_{[j,n]}) \quad (1.2.13)
\]

since

\[
i\mu_{[i+1,n]} + (n-i)\mu_{[i,n]} = n\mu_{[i,n-1]}.
\]

Now using (1.2.13) on the right hand side of (1.2.12) and simplifying we get (1.2.11).

The asymptotic distribution of concomitants in the simple linear model (1.2.1) is discussed in David and Galambos (1974). They have shown that for all $r$, under the conditions $n \to \infty$ and $\lim \beta_{r,n}^2 = 0$, the asymptotic distribution of $Z_{[r]} = Y_{[r,n]} - E[Y_{[r,n]}]$ is normal, $N(0, \sigma_j^2(1 - \rho^2))$, if $|\rho| < 1$. They have also proved two theorems concerning the asymptotic independence and asymptotic distribution of concomitants.

**Theorem 1.3**

For any fixed $k \geq 1$ and for any choice $1 \leq r_1 < r_2 < \ldots < r_k \leq n$ of subscripts,

$$
\lim_{n \to \infty} P\left( Y^{*}_{[r_1,n]} < x_1, \ldots, Y^{*}_{[r_k,n]} < x_k \right) = \prod_{i=1}^{k} \Phi(x_i | \sigma)
$$

(1.2.14)

where

$$
\sigma^2 = \sigma_j^2(1 - \rho^2)
$$

and

$$
Y^{*}_{[r,n]} = Y_{[r,n]} - E[Y_{[r,n]}]
$$

$\Phi$ is the cdf of a standard normal distribution.

**Theorem 1.4**

Let $X$ and $Z$ be independent random variables and let $Z$ has a continuous distribution function $F(x)$. Define $Y = X + Z$ and let $Y^{*}_{[r,n]} = X_{[r,n]} + Z_{[r]}$. 

Then if the distribution of $X$ is such that the variance $\beta^2_{r,n}$ of $X_{r,n}$ tends to zero as $n->\infty$, any fixed number of random variables $Y_{[r,n]}^* = Y_{[r,n]} - E[Y_{[r,n]}]$ are asymptotically independent, each with distribution function $F(x)$. A corollary of Theorem 1.3 is also stated here.

**Corollary 1.1**

Let $y_{i,n}, i = 1, 2, \ldots k$ real numbers such that, as $n->\infty \lim y_{i,n} = x_i$ exist. Then, with the notations of Theorem 1.3

$$\lim_{n->\infty} P[Y_{[r,n]}^* < y_{1,n}, \ldots, y_{[r,n]}^* < y_{k,n}] = \prod_{i=1}^{k} \Phi(x_i | \sigma)$$

A more general case is discussed in Yang (1977). He has proved a powerful theorem on the asymptotic distribution of concomitants.

**Theorem 1.5**

Let $1 \leq r_1 < r_2 < \ldots < r_k \leq n$ be sequences of integers such that, as $n->\infty$, $r_i/n \rightarrow \lambda_i$ with $0 < \lambda_i < 1$ ($i = 1, 2, \ldots k$).

Then $\lim_{n->\infty} P[Y_{[r_i,n]} < y_{r_i}, \ldots, y_{[r_i,n]} < y_{k,n}] = \prod_{i=1}^{k} P[Y_i \leq y_i | X_i = \lambda_i]$.

The extended version of the theorem in Galmbos (1978) by David (1994) gives a representation for the limit distribution of $Y_{[r,n]}$ for an arbitrary absolutely bivariate cdf $F(x,y)$.

**Theorem 1.6**

Let $F_X(x)$ satisfy one of the Von Mises condition and assume that the sequences of constants $a_n, b_n > 0$, are such that as $n->\infty$

$$\{ F_X(a_n + b_n x) \}^{n->\infty} G(x) \text{ for all } x$$

(1.2.16)
Further, suppose there exist constants $A_n$ and $B_n > 0$ such that
\[ F(y, A_n+B_n y | a_n+b_n x) \to H(y | x) \text{ for all } x \text{ and } y. \]
Then
\[ P[Y_{n-k+1:n} \leq A_n+B_n y] \to \int_{-\infty}^{\infty} H(y | x) dG(k)(x) \]
(1.2.17)
where $G(k)$, the kth lower record value from the extreme value cdf $G$. If (1.2.16) holds we say that $F_X$ is in the domain of attraction of $G$ it is well known that $G$ must be one of the three extreme value cdf’s which are of the following types, see Galambos (1987).

\[ G_1(x, \alpha) = 0 \quad x \leq 0 \]
\[ = \exp (-x^\alpha) \quad x > 0: \alpha > 0 \]
\[ G_2(x, \alpha) = \exp (-(-x^\alpha)) \quad x < 0: \alpha > 0 \]
\[ = 1 \quad x \geq 0 \]
\[ G_3(x) = \exp \left\{ -\exp (-x) \right\} ; -\infty < x < \infty. \]

Suresh (1993) has shown that the central concomitants and extreme concomitants are asymptotically independent. In selection problems, we use a very important statistic called the rank of the concomitants denoted by $R_{[r:n]}$. Here $R_{[r:n]}$ is the rank of $Y_{[r:n]}$ among the $n Y_i$’s. David et al. (1977) have derived the pdf and expected value $R_{[r:n]}$.

We have
\[ R_{[r:n]} = \sum_{i=1}^{n} I(Y_{[r:n]} - Y_i), \]
where
\[ I(x) = 1 \text{ if } x \geq 0 \]
\[ = 0 \text{ if } x < 0. \]

Denote by
\[ \Pi_{r,s} = P[R_{[r:n]} = s] \]
and

\[
\pi_{r,s} = n \int \int \sum_{k = u} \theta_1^k \theta_2^k \theta_3^{r-k} \theta_4^{s-k} f(x, y) \, dx \, dy
\]

where

\[
u = \max(0, r + s - n - 1) \quad t = \min(r - 1, s - 1)
\]

\[
\theta_1(x, y) = P[X \leq x, Y \leq y] = F_{x,y}(x, y)
\]

\[
\theta_2(x, y) = P[X \leq x, Y > y]
\]

\[
\theta_3(x, y) = P[X > x, Y \leq y]
\]

\[
\theta_4(x, y) = P[X > x, Y > y]
\]

and

\[
C_k = \frac{(n-1)!}{k!(r-1-k)!(s-1-k)!(n-r-s+1+k)!}.
\]

They have discussed in detail the bivariate normal case. The expected value of R_{r:n} may be obtained directly by the characteristic order statistics argument.

They have shown that

\[
E[R_{r:n}] = 1 + n + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \theta_1(x, y) f(y | x) f_{r-1-n-1}(x) \, dx \, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \theta_3(x, y) f(y | x) f_{r-1-n-1}(x) \, dx \, dy.
\]

(1.2.19)

In the bivariate normal case David et al. (1977) have shown that

\[
\lim_{n \to \infty} \frac{E[R_{r:n}]}{n+1} = \Phi(\rho, \Phi^{-1}(\lambda)/(2 - \rho^2)^{1/2}),
\]

\[
\lim P[R_{r:n} \leq n\lambda] = \Phi([\Phi^{-1}(\lambda) - \rho, \Phi^{-1}(\lambda)]/(1 - \rho^2)^{1/2})
\]

(1.1.20)

where

\[
r / n+1 \to \lambda \quad (0 < \lambda < 1).
\]
For the general bivariate distribution Yang (1977) has shown that
\[
\lim P[R_{r:n} \leq nu] = P[Y \leq F^{-1}_Y(u) | X = F^{-1}_X(\lambda)].
\]
He has presented an interesting application of concomitants in a prediction problem. Spruill and Gatsworth (1996) have applied the above results in connection with employment problems of a professional couple.

Yeo and David (1984) consider the problem of choosing best k objects out of n when, instead of measurements Y_i of primary interest, only associated measurements X_i (i = 1, 2 ... n) are available. For example Y_i could be very expensive measurements but X_i's are inexpensive measurements. It is assumed that the n pairs (X_i, Y_i) be a random sample from a continuous population. The actual values of X_i's are not required, only their ranks. A general expression is developed for the probability \( \Pi \) that the s objects with the largest X values include the k objects (k \leq s) with the largest Y values. They have applied the formula for \( \Pi \) in the bivariate normal case. They have also developed the formula for selecting the best object based on the actual values of X, instead of ranks by using a computer program.

Nagaraja and David (1994) have developed an important statistic for selection problem. In their approach the statistic \( V_{k,n} = \max(Y_{[n-k+1:n]}, \ldots, Y_{[n:n]}) \), \( k = 1, 2, \ldots, n \) representing the best individual in a screening procedure with respect to the characteristic under study. They consider \( E[V_{k,n}] / E[Y_{[n:n]}] \) as a measure of effectiveness of the screening procedure. Both the finite and asymptotic theory of \( V_{k,n} \) are discussed by them. They have shown that the cdf of \( V_{k,n} \) is
\[
F_{k,n}(y) = P[V_{k,n} \leq y] = \int_{-\infty}^{\infty} \left[ F_{Y|X}(y|x) \right]^k f_{X_{[n:n]}}(x) \, dx
\]
where

$$F_{Y|X}^*(y|x) = P[Y \leq y|X > x].$$ \hspace{1cm} (1.2.21)

They have also derived the limit distribution of $V_{k,n}$ in the extreme and quantile case. When $k$ is held fixed, under some regularity conditions as $n$ increases

$$F_{k,n}(A_{n+B_n\cdot y}) = \int_{-\infty}^{\infty} \left[H_{Y|X}(y|x)\right]^k dG_{k+1}(x)$$ \hspace{1cm} (1.2.22)

where

$$dG_{k+1}(x) = [-\log G(x)]^k \frac{1}{k!} g(x).$$

In the quantile case, where $k = [np]$, $0<p<1$, under mild conditions, the limit distribution of $V_{k,n}$ coincides with the limit distribution of sample maximum from the cdf $F_{Y|X}^*(y|F_{Y|X}^{-1}(1-p))$. They have applied their results to some interesting situations, including the bivariate normal population and the simple linear regression model. Joshi and Nagaraja (1995) have derived the joint distribution of $V_{k,n}$ and $V_{k,n}^* = \max(Y_{[1:n]}, \ldots, Y_{[n-k:n]})$. They used their result to study the joint distribution of $V_{k,n}$ and $Y_{n:n}$, since $Y_{n:n} = \max(V_{k,n}, V_{k,n}^*)$. It can be used to choose $k$ such that $V_{k,n}/Y_{n:n}$ is close to $1$. LiXiande (1999) has established a sufficient condition for the convergence of concomitants of selected order statistics.

Let $X, Y$ be the measurement of certain characteristic associated with the parent and offspring populations respectively. Suppose $k$ parents, ranked highest on $X$, are selected and the average $Y'_{[k:n]} = \frac{1}{k} \sum_{i=1}^{k} Y_{[n-i+1:n]}$ of the $y$ values associated with the offspring group due to the selection is the induced selection differential $D'_{[k,n]} = (Y'_{[k:n]} - \mu_y)/\sigma_y$, also known as response to selection. $D'_{[k,n]}$ measures the superiority of $Y$ of the $k$ individuals ranked highest on $X$. The asymptotic distribution this statistic, suitably standardized, is derived in extreme
and quantile cases by Nagaraja (1982). Asymptotic properties of \( D_{k,n} \) have also been investigated. Suresh and Kale (1994) discussed the induced selection percentiles and their properties. Yang (1981a, 1981b) and Sandstrom (1987) have studied the asymptotic properties of smooth linear functions of \( Y_{i:n} \). Yang (1981a) has considered general linear functions of the form

\[
L_{1n} = \frac{1}{n} \sum_{i=1}^{n} J\left( \frac{i}{n} \right) Y_{i:n}
\]

and

\[
L_{2n} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{i}{n} \right) \eta(X_{i:n}, Y_{i:n})
\]

where \( J \) is a smooth function which may depend on \( n \), and \( \eta \) is a real valued function. He has established the asymptotic normality of these statistics. These results are used to construct consistent estimators of quantiles associated with the conditional distribution of \( Y \) given \( X = x \).

Do and Hall (1992) used the Effron's (1990) technique to estimate the percentiles of the bootstrap distribution based on concomitants.

Let \( Y = X + \varepsilon \) where \( F_X \) is completely known and \( F_Y \) is to be estimated. The observation is

\[
(X_{i:n}, \varepsilon_{i:n}), 1 \leq i \leq n, \quad \varepsilon_{i:n} = Y_{i:n} - X_{i:n}
\]

They have suggested the estimator

\[
\hat{F}_{n,Y}(y) = \frac{1}{n} \sum_{i=1}^{n} I(F_X^{-1}(i/n) + \varepsilon_{i:n} \leq y)
\]

where \( I(.) \) represents the indicator function and established that if \( \varepsilon \)'s are sufficiently small, \( \hat{F}_{n,Y}(y) \) performs better than the classical estimator \( F_{n,Y}(y) \), the empirical cdf. Application of concomitants in double sampling is discussed in O'Connell and David (1976).
They suggested the simple linear estimator of $\mu_Y$ is

$$\bar{Y}_{[r:n]} = \mu_Y + \rho \frac{\sigma_y}{\sigma_x} (\bar{X}_{r:n} - \mu_X) + \tilde{e}_r$$  \hspace{1cm} (1.2.23)

where $\bar{X}_{r:n}$ and $\tilde{e}_r$ are the means of $X_{r:n}$ and $e_i$, $i = 1, 2 \ldots k$.

If $X$ has a symmetric distribution and the ranks are symmetrically chosen

\[ r_{k+1-i} = n+1-r_i, i = 1, 2 \ldots \left[\frac{k+1}{2}\right] \]

then $\bar{Y}_{[r:n]}$ is unbiased for $\mu_Y$. Also from (1.2.23)

$$\text{Var} (\bar{Y}_{[r:n]}/\sigma_Y) = \rho^2 \text{Var} (\bar{X}_{r:n}) + (1 - \rho^2) / k.$$  \hspace{1cm} (1.2.24)

Thus the ranks $r_i$ minimizing $\text{Var} (\bar{X}_{r:n})$ also minimize $\text{Var} (\bar{Y}_{[r:n]})$ for all values of $\rho$.

Waterson (1959) has considered the linear estimation of the parameters of a bivariate normal population under various forms of censoring. Harrell and Sen (1979) have used the method of likelihood in one of these situations, namely when $X_{1:n}, \ldots, X_{k:n}$ and $Y_{[1:n]}, \ldots, Y_{[k:n]}$ are available. They derive the test of independence of $X$ and $Y$.

An unbiased estimator of the regression coefficient is considered in Barton and Casley (1958). The estimator

$$B' = \frac{\bar{Y}'_{[k:n]} - \bar{Y}_{[k:n]}}{\bar{X}'_{k:n} - \bar{X}_{k:n}}$$

where

$$\bar{Y}'_{[k:n]} = \frac{1}{k} \sum_{i=1}^{k} Y_{[n-i+1:n]}, \quad \bar{Y}_{[k:n]} = \frac{1}{k} \sum_{i=1}^{k} Y_{[r:n]}$$
\[ \tilde{X}_{k,n} = \frac{1}{k} \sum_{i=n}^{k} X_{n-i+1,n}, \quad \tilde{X}_{k,n} = \frac{1}{k} \sum_{i=1}^{k} X_{i,n}, \]

does not use the i.i.d property and has efficiency of 75-80% when (X,Y) is bivariate normal, provided k is chosen about 0.27n. Tuskibayashi (1962) has suggested an estimator
\[
\hat{\rho} = \frac{\bar{Y}_{[n:n]} - \bar{Y}_{[1:n]}}{\bar{Y}_{n:n} - \bar{Y}_{1:n}}
\]
of \( \rho \), the correlation coefficient. He points out that \( \hat{\rho} \) can be calculated even if only the ranks of \( X_i \)'s are available. Interesting results related to the distribution of \( \hat{\rho} \) are developed in Tuskibayashi (1998). Barnett et.al.(1976) have discussed the estimation of \( \rho \) using concomitants.

Multivariate generalization of concomitants is first discussed in David (1973). Suppose that associated with each \( X \) there are \( t \) variates \( Y_j \) (\( j = 1, 2 \ldots t \)) the \( t+1 \) variates following a multivariate normal distribution with covariance matrix
\[
\Sigma = \begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{bmatrix}
\]

where
\[
\Sigma_{11} = \sigma^2, \quad \Sigma_{12} = (\text{Cov}(X,Y_j))_{j,t} = \Sigma_{21}^T
\]
and
\[
\Sigma_{22} = (\text{Cov}(Y_j,Y_j))_{j,t}.
\]

If \( Y_{[r:n]}^* \) denotes the \( t \times 1 \) column vector of the \( Y_{j,[r:n]}^* \),

where
\[
Y_{j,[r:n]}^* = Y_{j,[r:n]} - E \{ Y_{j,[r:n]} \},
\]
the vectors $Y^*_{[r,n]}, i = 1, 2, \ldots, k$ are asymptotically independent, identically distributed $N(0, \Sigma_{22.1})$ variates,

and

\[
\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}
\]

\[
= \Sigma_{22} - \frac{\Sigma_{21} \Sigma_{12}}{\sigma_x^2}
\]

(1.2.25)

The general multivariate case and some applications are discussed in Balakrishnan (1993). Suppose we have \(n\) sets of variates \((X_i, Y_{1i}, \ldots, Y_{ti})\). Setting \(m_j(x_i) = \text{E}(Y_{ji} | x_i)\) and writing \(Y_{ji[r:n]}\) for that \(Y_{ji}\) paired with \(X_{r:n}\), we have

\[
\text{E}(Y_{ji[r:n]}) = \text{E}(m_j(X_{r:n}))
\]

and

\[
\text{Cov}(Y_{ji[r:n]}, Y_{kj[r:n]}) = \text{Cov}(m_j(X_{r:n}), m_k(X_{r:n})) + \text{E}\sigma_{jk}(X_{r:n})\quad j \neq k
\]

(1.2.26)

where

\[
\sigma_{jk}(x_i) = \text{Cov}(Y_{ji}, Y_{ki} | x_i).
\]

In the multivariate normal case \(\sigma_{jk}(x_i)\) does not depend on \(x_i\) and may be obtained from standard theory

\[
\sigma_{jk}(x_i) = \sigma_{jk} - \sigma_j \sigma_k / \sigma_x^2.
\]

(1.2.27)

Then

\[
Y_{ji[r:n]} = \mu_j + \rho_j \sigma_j (X_{r:n} - \mu_x) / \sigma_x + \varepsilon_{ji[r]}
\]

(1.2.28)

where

\[
\mu_j = \text{E}(Y_{ji}), \sigma_j^2 = \text{Var}(Y_{ji})
\]

and

\[
\rho_j = \text{Cov}(X, Y_{ji}).
\]

Also noting that \(\varepsilon_{ji[r]}\) and \(\varepsilon_{ki[s]}\) are independent unless \(r = s\).
They have shown that
\[
\text{Cov}(Y_{j[r:n]}, X_{k[r:n]}) = \rho_j \sigma_j \rho_k \sigma_k \beta_{rr,n} + \sigma_{jk}(x)
\]
\[
= \sigma_{jk} - \rho_j \sigma_j \rho_k \sigma_k (1 - \beta_{rr,n})
\] (1.2.29)
\[
\text{Cov}(Y_{j[r:n]}, Y_{k[r:n]}) = \rho_j \sigma_j \rho_k \sigma_k \beta_{rr,n}
\] (1.2.30)
where,
\[\beta_{rr,n} = \text{Cov}(X_{r:n}, X_{s:n})\].

Balakrishnan (1993) has introduced the multivariate order statistics induced by ordering linear combinations of the components observed in \(n\) independent observations from a multivariate normal distribution. The concept of induced bivariate order statistics is explained below.

Let \((X_i, Y_i), i = 1, 2 \ldots n\) be independent observations from the bivariate normal distribution \(\text{BVN} (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho \sigma_x \sigma_y)\). Let \(a\) and \(b\) be non zero constants and
\[
S_j = aX_j + bY_j, \quad j = 1, 2 \ldots n.
\] (1.2.31)

Let \(S_{1:n} \leq S_{2:n} \ldots \leq S_{n:n}\) be the order statistics of \(S_1, S_2, \ldots S_n\) defined in (1.2.31). Then the bivariate order statistics induced by the order statistics \(S_{k:n}\) as follows \(X_{[k:n]} = X_j\) and \(Y_{[k:n]} = Y_j\) whenever \(S_{k:n} = S_j\). In other words \((X_{[k:n]}, Y_{[k:n]}\) is that \((X, Y)\) pair which corresponds to the smallest value among \(S_j\)'s in (1.2.31). He has derived explicit expression for the means, variances and co-variances of the induced bivariate order statistics. He also extended the bivariate induced order statistics to the multivariate case and derived explicit expressions for means variances and co-variances in the \(p\)-variate normal case.

Balasubramanian and Balakrishnan (1995) have provided a method of constructing a general class of distributions which is closed under marginal,
Introduction

conditional and concomitance of order statistics. They have constructed the bivariate member of the class defined by

\[ h_2 [x_1 , x_2 ; a_1 , a_2 , a_{12}] = f(x_1) f(x_2) \{ 1 + a_1 g(x_1) + a_2 g(x_2) + a_{12} g(x_1) g(x_2) \} \]

(1.2.32)

where \( f(x) \) is a density function and \( g(x) \) is an orthogonal function such that \( E[g(X)] = 0 \) and the parameters \( a_1, a_2 \) and \( a_{12} \) satisfy the conditions

\[
\begin{align*}
1 + a_1 + a_2 + a_{12} & \geq 0 \\
1 + a_1 - a_2 - a_{12} & \geq 0 \\
1 - a_1 + a_2 - a_{12} & \geq 0 \\
1 - a_1 - a_2 + a_{12} & \geq 0.
\end{align*}
\]

(1.2.33)

They have shown that the concomitants belong to the univariate member of the family

\[ h(x, a) = f(x) \{ 1 + a g(x) \}, a \in [-1, 1], \]

and have extended the method to the multivariate case and discussed some interesting properties of this class.

1.3 Morgenstern Distributions

In modelling problems, one general approach is to first choose a family of distributions and then select a member that is appropriate to describe the observation. Of the desiderata for choice of the family, the most important one is that the family should be flexible, in the sense that it should contain a wide variety of models capable of representing any data situation. Another consideration is the sort of prior information available in the choice of the model. In problems involving several random variables, the analyst may make reasonable assumptions about the marginal distributions. Then the question is to construct a joint distribution function with a set of given marginals. The Morgenstern family of distributions assumes importance in such contexts as a highly flexible system. Accordingly in the present study we deal with the distribution theory and applications of concomitants from the Morgenstern family of bivariate distributions.
The Morgenstern system of bivariate distributions includes all cumulative distribution functions of the form

\[ F_{X,Y}(x,y) = F_X(x) F_Y(y)[1+\alpha(1- F_X(x))(1- F_Y(y))], \quad -1 \leq \alpha \leq 1. \]  \hspace{1cm} (1.3.1)

The system provides a very general expression of a bivariate distribution from which members can be derived by substituting expressions of any desired set of marginal distributions. The joint density is given by

\[ f_{X,Y}(x,y) = f_X(x)f_Y(y)[1+\alpha(1-2 F_X(x))(1-2 F_Y(y))], \quad -1 \leq \alpha \leq 1. \]  \hspace{1cm} (1.3.2)

Since both the bivariate distribution function and density are given in terms of marginals, it is easy to generate a random sample from a Morgenstern distribution. Thus members of this family can be used in simulation studies, especially when weak dependence between variates is of interest. It follows that the conditional density of \( X \) given \( Y=y \) is

\[ f_{X|Y}(x,y) = f_Y(y)[1+\alpha(1-2 F_X(x))(1-2 F_Y(y))], \quad -1 \leq \alpha \leq 1. \]  \hspace{1cm} (1.3.3)

When \( y = \text{median (Y)} \), the conditional density of \( X \) given \( Y = y \) is the same as the marginal density of \( X \). The regression curve of \( X \) given \( Y = y \) is

\[ E [X|Y=y] = E[X] + \alpha (1-2 F_Y(y)) \int x (1-2 F_X(x)) f_X(x) \, dx \]  \hspace{1cm} (1.3.4)

which is linear in \( F_Y(y) \).

A number of properties results from the simple analytic form of Morgenstern distributions. If the marginal distributions of \( X \) and \( Y \) are symmetric, the joint distribution is also symmetric. Random variables having a bivariate Morgenstern distribution are exchangeable whenever the marginal distributions are identical. The Morgenstern system is closed with respect to monotonic increasing functions of random variables. Also the system is closed with respect to mixtures of bivariate Morgenstern distributions having the same marginal distributions.
The Morgenstern family is characterized by its "closeness" to the
distribution of independent random variables. The following characterization is
discussed in Nelson (1994).

**Theorem 1.5**

If $|\rho_0| \leq 1/3$, the one whose joint density is closest to the product density of independent random variables (in the sense of minimizing $\Psi^2$-divergence) is the Morgenstern distribution with parameter $\alpha = 3\rho_0$, where $\rho_0$ is the Spearman's rank correlation coefficient.

The Morgenstern distributions are specially suited to data situations describing weak dependence between the random variables $X$ and $Y$. Measures of dependence vary over a smaller range than for some other general classes of bivariate distributions. Schucany et al. showed that for this family the Pearson's correlation coefficient lies between $-1/3$ and $1/3$ (see Convey (1983)).

**1.4 The Present Work**

The present work is organized into five chapters. Chapter 1 contains a brief description of the basic distribution theory and a quick review of the existing literature. In this chapter we derive a new recurrence relation connecting the product moments of concomitants. We also introduce the concept of bivariate Morgenstern family of distributions and its basic properties.

Chapter 2 deals with the distribution theory of concomitants from the Morgenstern family. It also contains some interesting recurrence relations connecting the moments of concomitants. In this chapter we specialize the results to some well known members of the family, viz, bivariate exponential, bivariate uniform, bivariate logistic and bivariate gamma distributions. We also provide quick estimators for the parameters of the exponential, uniform and logistic models.
In Chapter 3 we deal with the distribution theory of the statistic $V_{k,n}$ discussed in the previous section from the Morgenstern family and obtain certain characteristics that could be useful in selection problems. We also derive the limiting distribution of $V_{k,n}$ and provide illustrative tables of the values of $e_{k,n}$ for the bivariate uniform, bivariate exponential and bivariate logistic models.

Let $X_1, X_2, \ldots$ be an infinite sequence of independent and identically distributed random variables having the same absolutely continuous distribution function $F(x)$. An observation $X_j$ will be called an upper record (or simply a record) if its value exceeds that of all previous observations. Thus $X_j$ is a record if $X_j > X_i$ for every $i < j$. An analogous definition deals with lower record values. A comprehensive study on record values is presented in Arnold Balakrishnan and Nagaraja (1998).

Let $(X_i, Y_i)$, $i=1, 2, \ldots$ be a sequence of i.i.d random variables from an absolutely continuous distribution with distribution function $F(x,y)$ and density function $f(x,y)$. Let $R_n$ denote the $n^{th}$ record value in the sequence of the $X$'s. The corresponding random variable $Y$, i.e. the $Y$-value paired with the $X$-value $R_n$ is called the $n^{th}$ record concomitant and will be denoted by $R_{[n]}$. The distribution theory of record concomitants from the Morgenstern family of bivariate distributions is discussed in Chapter 4. We also discuss the distribution theory of record concomitants from some important members of the family bivariate exponential, bivariate uniform and bivariate logistic distributions.

The procedure of ranked set sampling was suggested by McIntyre (1952) for improving the precision of $\overline{Y}$ as an estimator of the population mean. This method is applicable for situations where the primary variable of interest, $Y$, is
Introduction

difficult or expensive to measure, but where ranking in small sets is easy. The process involves selecting m samples, each of size m, and ordering each of the samples by eye or some relatively inexpensive means, without actual measurement of the individual, see David and Levine (1972) Stokes (1977). The smallest observation from the first sample is chosen for measurement, as is the second smallest observation from the second sample. The process continues in this way until the largest observation from the nth sample is measured, producing a total of n measured observations one from each order class.

Motivated from the ranked set sampling we use the following sampling method for selection of primary variable. Suppose there are two correlated variables Y and X, where Y is difficult to measure or to rank. Consider a bivariate sample of size n = mk, where k is an integer. Randomly subdivide the sample in to k sub samples (groups) each of size m . In each sub sample we measured only the Y-value corresponding to the rth order statistic X_{r:m} . Then the Y-value measured in the ith sample is the rth concomitant will be denoted by Y_{r:m,i} i=1,2,...k . The Y_{r:m,i} are independent random variables having the same marginal distribution as Y_{r:m}.

Let \( M_{k,r:m} = \max[Y_{r:m,1}, Y_{r:m,2}, \ldots, Y_{r:m,k}] \)
and
\( m_{k,r:m} = \min[Y_{r:m,1}, Y_{r:m,2}, \ldots, Y_{r:m,k}] \)
denote the largest and smallest among the selected concomitants. Thus \( M_{k,r:m} \) and \( m_{k,r:m} \) are the extremes of the selected expensive measurements in the samples. In particular \( M_{k,[m:m]} \) is the largest observation of the concomitants of maximum of order statistics in the sub samples. Then the ratio \( \frac{E[M_{k,[m:m]}]}{E[Y_{nn}]} \) which clearly
increases to 1 with k, is a measure of effectiveness of the selection procedure. One may wish to choose the value of the number of subdivisions (populations), k, to make this ratio sufficiently close to 1. In Chapter 5 we discuss the general distribution theory of $M_{k,[r,m]}$ and $m_{k,[r,m]}$ from the Morgenstern family of distributions and discuss some applications in inference, estimation of the parameter of the marginal variable $Y$ in the Morgenstern type uniform distributions. We also apply the results to the selection problem discussed earlier. The work concludes with the distribution theory of the rank of the rth concomitant $R_{[r,n]}$. We also provide illustrative tables for values of $\Pi_{r,s} = P[R_{[r,n]} = s]$ and $E[R_{[r,n]}]$. 