DISTRIBUTIONS OF LINEAR COMBINATIONS OF UNIFORM COVERAGES

DISTRIBUTIONS OF LINEAR COMBINATIONS OF UNIFORM COVERAGES

By

GEORGE ANTHONY WELLS, B.SC.

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

November 1973

MASTER OF SCIENCE (1973)
(Mathematics)

McMASTER UNIVERSITY
Hamilton, Ontario.

TITLE: Distribution of linear combinations of uniform coverages

AUTHOR: George Anthony Wells, B.Sc. (McMaster University)

SUPERVISOR: DR.E.R. Mead

NUMBER OF PAGES: 108, v

SCOPE AND CONTENTS:

This paper is concerned with the distributions of one and two linear combinations of uniform coverages with applications to the distribution of the sums of ordered intervals.

Acknowledgement

I would like to thank my instructor Dr. E. R. Mead, who through his guidance and patience has helped make this undertaking a reality.

I could not stop here without thanking my wife, parents, relatives and friends.

TABLE OF CONTENTS

Introduction

The aim of this paper is to derive generalized formulae for the density functions of one and two arbitrary linear combinations of the uniform coverages corresponding to the order statistics obtained from a random sample from the uniform distribution on the interval (0,1).

In Chapter I, the required background and a brief survey of some of the past papers on the subject are given. In addition, a few necessary results from the calculus of finite differences are listed. Then, in Chapter II, a formula is obtained for the density function of a single linear combination using a different procedure than that employed in other papers. Finally, in Chapter III, a formula for the joint density function of two linear combinations of the same set of coverages is derived with no restrictions whatsoever on the real constant coefficients.

Chapter I : Preliminaries

1.1 : Background

Consider a random sample $X_1, X_2, \ldots, X_n$ from the uniform distribution on the interval (0,1). Their joint density function is given by

$$f(x_1, x_2, \ldots, x_n) = 1 \qquad (0 \leq x_i \leq 1,\ i=1,2,\ldots,n) .$$

The joint density function of the order statistics $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, obtained by ordering $X_1, X_2, \ldots, X_n$ in ascending order of magnitude, is

$$f(x_{(1)}, x_{(2)}, \ldots, x_{(n)}) = n! \qquad (0 \leq x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)} \leq 1) .$$

Let $U_1, U_2, \ldots, U_n$ be the coverages, corresponding to the order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, defined by

$$U_i = X_{(i)} - X_{(i-1)} \qquad (i=1,2,\ldots,n)$$

where $X_{(0)} = 0$. Then

$$X_{(i)} = \sum_{j=1}^{i} U_j \qquad (i=1,2,\ldots,n)$$

and it can easily be shown that

$$U_i \geq 0 \quad (i=1,2,\ldots,n) ; \qquad \sum_{i=1}^{n} U_i \leq 1 .$$

Since the Jacobian $\frac{\partial(U_1,U_2,\ldots,U_n)}{\partial(X_{(1)},X_{(2)},\ldots,X_{(n)})} = 1$ it follows

that the joint density function of $U_1,U_2,\ldots,U_n$ is

$$f(u_1,u_2,\ldots,u_n) = n!$$

$$(u_i \geq 0,\ i=1,2,\ldots,n;\ \sum_{i=1}^{n} u_i \leq 1) .$$

By this transformation the joint distribution of

$X_{(1)},X_{(2)},\ldots,X_{(i)},\ldots,X_{(n)}$ will be that of

$$U_1,U_1+U_2,\ldots,\sum_{j=1}^{i} U_j,\ldots,\sum_{j=1}^{n} U_j \quad .$$

Thus the density function of the linear combination

$$Y = \sum_{i=1}^{n} a_i U_i$$

is the same as that of

$$X = \sum_{i=1}^{n} c_i X_{(i)}$$

where a_i and c_i are real constants and

$$a_i = \sum_{j=i}^{n} c_j \qquad (i=1,2,\ldots,n) .$$

The joint density function of the ordered coverages

$U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$, obtained by ordering $U_1, U_2, \ldots, U_n$ in ascending order of magnitude, is given by

$$f(u_{(1)}, u_{(2)}, \ldots, u_{(n)}) = n!\ n!$$

$$\left(0 \leq u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(n)} \leq 1;\ \sum_{i=1}^{n} u_{(i)} \leq 1\right).$$

To determine a transformation from the coverages to the ordered coverages consider the random variables $W_1, W_2, \ldots, W_n$ defined by

$$W_1 = nU_{(1)}$$

$$W_j = (n+1-j)(U_{(j)} - U_{(j-1)}) \qquad (j=2,3,\ldots,n) .$$

Then

$$U_{(i)} = \sum_{j=1}^{i} \frac{W_j}{n+1-j} \qquad (i=1,2,\ldots,n) ,$$

$$\frac{\partial(W_1, W_2, \ldots, W_n)}{\partial(U_{(1)}, U_{(2)}, \ldots, U_{(n)})} = n!$$

and the joint density function of $W_1, W_2, \ldots, W_n$ is given by

$$f(w_1, w_2, \ldots, w_n) = n!$$

$$\left(w_i \geq 0,\ i=1,2,\ldots,n;\ \sum_{i=1}^{n} w_i \leq 1\right)$$

which is identical with that of $U_1, U_2, \ldots, U_n$. By this transformation the joint distribution of $U_1, U_2, \ldots, U_i, \ldots, U_n$ will be that of

$$nU_{(1)},\ (n-1)(U_{(2)}-U_{(1)}), \ldots,\ (n+1-i)(U_{(i)}-U_{(i-1)}), \ldots,\ (U_{(n)}-U_{(n-1)}).$$

Thus the density function of the linear combination

$$U = \sum_{i=1}^{n} d_i U_{(i)}$$

is the same as that of

$$Y = \sum_{i=1}^{n} a_i U_i$$

where a_i and d_i are real constants and

$$a_i = \sum_{j=i}^{n} \frac{d_j}{n+1-j} \qquad (i=1,2,\ldots,n) .$$

It should also be noted, defining $U_{n+1} = X_{(n+1)} - X_{(n)}$ where $X_{(n+1)} = 1$, that the joint distribution of $U_1, U_2, \ldots, U_{n+1}$ is degenerate and their joint density function is given by

$$f(u_1, u_2, \ldots, u_{n+1}) = n!$$

$$(u_i \geq 0,\ i=1,2,\ldots,n+1;\ \sum_{i=1}^{n+1} u_i = 1)\ .$$

Then the joint density function of the ordered coverages $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n+1)}$, obtained by ordering $U_1,U_2,\ldots,U_{n+1}$ in ascending order of magnitude, is given by

$$f(u_{(1)},u_{(2)},\ldots,u_{(n+1)}) = (n+1)!\ n!$$

$$(0 \leq u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(n+1)} \leq 1;\ \sum_{i=1}^{n+1} u_{(i)}=1).$$

Consider the random variables $V_1,V_2,\ldots,V_{n+1}$ defined by

$$V_1 = (n+1)U_{(1)}$$

$$V_j = (n+2-j)(U_{(j)}-U_{(j-1)}) \qquad (j=2,3,\ldots,n+1).$$

Then $U_{(i)} = \sum_{j=1}^{i} \frac{V_j}{n+2-j}$ $(i=1,2,\ldots,n+1)$,

$$\frac{\partial(V_1,V_2,\ldots,V_{n+1})}{\partial(U_{(1)},U_{(2)},\ldots,U_{(n+1)})} = (n+1)!$$

and the joint density function of $V_1,V_2,\ldots,V_{n+1}$ is given by

$$f(v_1,v_2,\ldots,v_{n+1}) = n!$$

$$(v_i \geq 0,\ i=1,2,\ldots,n+1;\ \sum_{i=1}^{n+1} v_i = 1)$$

which is identical with that of $U_1, U_2, \ldots, U_{n+1}$. By this transformation the joint distribution of $U_1, U_2, \ldots, U_i, \ldots, U_{n+1}$ will be that of -

$$(n+1)U_{(1)},\ n(U_{(2)}-U_{(1)}), \ldots,\ (n+2-i)(U_{(i)}-U_{(i-1)}), \ldots,\ (U_{(n+1)}-U_{(n)}) .$$

Thus the density function of the linear combination

$$U = \sum_{i=1}^{n+1} d_i U_{(i)}$$

is the same as that of

$$Y = \sum_{i=1}^{n+1} a_i U_i$$

where

$$a_i = \sum_{j=1}^{n+1} \frac{d_j}{n+2-j} \qquad (i=1,2,\ldots,n+1).$$

Many papers in the past were concerned with the distributions of

$$X = \sum_{i=1}^{n} c_i X_{(i)} \qquad (1.1.1)$$

$$Y = \sum_{i=1}^{n+1} a_i U_i \qquad (1.1.2)$$

$$U = \sum_{i=1}^{n+1} d_i U_{(i)} \tag{1.1.3}$$

for varying degrees of generalities of the coefficients.

The papers by Mauldon (1951) and Barton and David (1955) were concerned with special cases of the distribution of the sum of ordered intervals. Mauldon considered the distribution of the sum of the k largest ordered intervals, that is, the distribution of the linear combination (1.1.3) with

$$d_i = 0 \qquad (i=1,2,\ldots,n+1-k)$$
$$d_i = 1 \qquad (i=n+2-k,n+3-k,\ldots,n+1).$$

Barton and David considered a slightly more general situation and derived the density function of the sum of (s-r+1) consecutively ordered intervals, that is, the density function of (1.1.3) with

$$d_i = 0 \qquad (i=1,\ldots,r-1,s+1,\ldots,n+1)$$
$$d_i = 1 \qquad (i=r,r+1,\ldots,s).$$

Then in 1957 Mauldon proved an inversion formula for a generalized transform and as an application of this result he obtained the distribution of (1.1.1) and (1.1.3) where the coefficients were arbitrary real numbers. This was the first paper which considered the distribution of such general linear combinations. All previous papers had considered only particular cases.

A paper by Dwass in 1961 considered the general distribution of the linear combination (1.1.2) and (1.1.3). As noted previously, with the proper transformation, the problem of finding the disribution of the linear combination (1.1.2) is equivalent to the one considered by Mauldon in 1957. However the proof by Dwass was different.

In 1968 Dempster and Kleyle considered distributions determined by cutting a simplex with hyperplanes. As an application of their results they obtained the distribution function of (1.1.1) for distinct a_i $(i=1,2,\ldots,n)$ defined by

$$a_i = \sum_{j=i}^{n} c_j \qquad (i=1,2,\ldots,n) .$$

Ali (1969) obtained the distribution function of (1.1.1) for any set of real constant coefficients. Even though he pointed out how the distribution function could be found in the case that some of the coefficients a_i $(i=1,2,\ldots,n)$ defined above coincide, no such distribution function was stated and the proofs were carried out based on the assumption of distinct a_i's.

Two years later in 1971, in a paper by Weisberg, the general distribution function of (1.1.1) for any set of coefficients was obtained.

In the realm of two or more linear combinations of the form of (1.1.1), (1.1.2), or (1.1.3), based on the same random sample $X_1, X_2, \ldots, X_n$, not many investigations have been undertaken. The 1955 paper by Barton and David considered the joint distribution of two linear combinations of the form (1.1.3) whose coefficients take on values 0 or 1. That is, they considered the joint density function of

$$G_{tv} = \sum_{i=t}^{v} U_{(i)} \qquad \text{and} \qquad G_{rs} = \sum_{i=r}^{s} U_{(i)}$$

for $v < r$. However they found the form too complicated to be expressed by a single formula. Niven (1963) derived the joint distribution of the sample mean

$$M_1 = \left(\sum_{i=1}^{n} X_{(i)}\right)/n$$

and the sample range

$$M_2 = X_{(n)} - X_{(1)}$$

for a sample of size $n=3,4$. Using a geometrical approach she obtained formulae for the density function in various regions but no general formula was derived. Dempster and Kleyle (1968) noted, using the geometrical approach of their paper, that the joint distribution of several linear combinations could be obtained in principle. Finally, in 1969, Ali and Mead investigated the joint distribution of several linear combinations of the form of (1.1.2) with $a_{n+1} = 0$. However certain restrictions were required on the coefficients. In 1972 S. W. Lim removed some of these restrictions in the case of two linear combinations.

1.2 : Some results from the calculus of finite differences

Since divided differences will be used throughout this paper certain results concerning such differences will be stated in this section. Such results can be found in Milne-Thomson (1933), for instance.

Let $g(z)$ be a function of the real variable z and let $z_1, z_2, \ldots, z_n$ be n distinct real numbers. Then the $(n-1)$st divided difference of $g(z)$ for the arguments $z_1, z_2, \ldots, z_n$ is denoted by

$$D^{n-1}(g(z):z_1, z_2, \ldots, z_n)$$

and defined by

$$D^{n-1}(g(z):z_1,z_2,\ldots,z_n)$$

$$= \frac{D^{n-2}(g(z):z_1,z_2,\ldots,z_{n-1}) - D^{n-2}(g(z):z_2,z_3,\ldots,z_n)}{z_1-z_n}$$

where

$$D^1(g(z):z_1,z_2) = \frac{g(z_1)-g(z_2)}{z_1-z_2}.$$

It can easily be shown that the interchange of any two of the arguments does not alter the value of the divided difference. Thus divided differences are symmetric functions of their arguments. Two more well known properties of divided differences are:

(1) $$D^{n-1}(g(z):z_1,z_2,\ldots,z_n) = \sum_{i=1}^{n} \frac{g(z_i)}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (z_i-z_j)};$$

(2) if $g(z)$ is a polynomial of degree n then

$$D^{n+r}(g(z):z_1,z_2,\ldots,z_{n+r+1}) = \begin{cases} \text{constant} & \text{if } r = 0 \\ 0 & \text{if } r \text{ is a positive integer} \end{cases}$$

In the case of divided differences with repeated arguments

$$D^{n-1}(g(z):z_1,z_1,\ldots,z_p,z_p)$$

$$= \frac{1}{(n_1-1)!\,(n_2-1)!\,\ldots\,(n_p-1)!}\,\frac{\partial^{n_1+n_2+\ldots+n_p-p}}{\partial_{z_1}^{n_1-1}\,\partial_{z_2}^{n_2-1}\cdots\partial_{z_p}^{n_p-1}}$$

$$\cdot\ D^{p-1}(g(z):z_1,z_2,\ldots,z_p)$$

where there are p distinct arguments and n_i arguments are equal to z_i $(i=1,2,\ldots,p)$.

Chapter II : The distribution of one linear combination of uniform coverages

2.1 : Introduction

This chapter is primarily concerned with the distribution of a linear combination of the form

$$Y = a_1U_1 + a_2U_2 + \dots + a_nU_n$$

where the coefficients a_i $(i = 1,2,\dots,n)$ are real numbers.

In section 2.2 a formula which expresses divided differences in terms of definite integrals is applied to obtain an alternative derivation of the distribution function, density function, and characteristic function of a single linear combination of the uniform spacings with distinct coefficients. A generalization of this divided difference - definite integral formula is considered in section 2.3 and is subsequently used to acquire the desired functions in the general case which allows for the possibility of equal coefficients. Finally, an application of these results to the distribution of the sums of ordered intervals is given in section 2.4.

2.2 : The distribution of one linear combination with distinct coefficients

Milne-Thompson (1933) states a result due to Hermite which expresses divided differences by definite integrals.

That is, assuming $z_1, z_2, \ldots, z_n$ are distinct arguments and the (n-1)st derivative of a function g(z) is continuous

$$D^{n-1}(g(z):z_1,z_2,\ldots,z_n) = \int_0^1 dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{n-2}} g^{(n-1)}(v_n)dt_{n-1}$$

where $v_n = (1-t_1)z_1 + (t_1-t_2)z_2 + \ldots + (t_{n-2}-t_{n-1})z_{n-1} + t_{n-1}z_n$ and $t_1, t_2, \ldots, t_{n-1}$ are (n-1) independent variables.

The application of this result is required to obtain the distribution function of

$$Y = a_1U_1 + a_2U_2 + \ldots + a_nU_n$$

in the case of distinct coefficients. However, in its present state it cannot be applied directly since the function corresponding to $g^{(n-1)}(z)$ in the application will not be continuous but will in fact be a step function with a single discontinuity at zero. The following lemma establishes the above result in such a situation.

<u>Lemma 2.1</u> If $z_1, z_2, \ldots, z_n$ are distinct arguments then

$$D^{n-1}(g(z):z_1,z_2,\ldots,z_n) = \int_0^1 dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{n-2}} g^{(n-1)}(v_n)dt_{n-1}$$

where

$$v_n = z_1 + t_1(z_2-z_1) + \ldots + t_{n-1}(z_n-z_{n-1})$$

and

$$g^{(n-1)}(z) = \begin{cases} k & \text{if } z>0 \quad (k \text{ constant}) \\ 0 & \text{if } z\le 0 \end{cases}$$

Proof:

By the definition of divided differences and using the fact that $g^{(n-2)}(z) = \begin{cases} kz & \text{if } z>0 \\ 0 & \text{if } z\le 0 \end{cases}$ is continuous

$$D^{n-1}(g(z):z_1,z_2,\ldots,z_n)$$

$$= \frac{D^{n-2}(g(z):z_1,\ldots,z_{n-2},z_{n-1}) - D^{n-2}(g(z):z_1,\ldots,z_{n-2},z_n)}{z_{n-1} - z_n}$$

$$= \frac{1}{z_{n-1} - z_n}$$

$$\left[\int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-3}} g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))dt_{n-2}\right.$$

$$\left. -\int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-3}} g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))dt_{n-2}\right]$$

$$= \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-3}} \left\{\frac{1}{z_{n-1}-z_n}\left[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))\right.\right.$$

$$\left.\left. - g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))\right]\right\} dt_{n-2} \tag{2.2.1}$$

The lemma follows if it can be shown that

$$\frac{1}{z_{n-1}-z_n}\Big[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))$$

$$-g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))\Big] \qquad (2.2.2)$$

in (2.2.1) is equivalent to the integral

$$\int_0^{t_{n-2}} g^{(n-1)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-1}(z_n-z_{n-1}))\, dt_{n-1}. \qquad (2.2.3)$$

By definition

$$g^{(n-1)}(v_n) = \begin{cases} k & \text{if } v_n > 0 \\ 0 & \text{if } v_n \le 0 \end{cases} .$$

Assume $z_n > z_{n-1}$. Then $v_n > 0$ is equivalent to

$$z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}) > 0$$

that is

$$t_{n-1} > \frac{z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})}{z_{n-1}-z_n}$$

$$= \frac{v_{n-1}}{z_{n-1}-z_n} .$$

The limits of integration with respect to t_{n-1} are 0 and t_{n-2}. On the other hand the integrand is zero if

$$t_{n-1} \le \frac{v_{n-1}}{z_{n-1}-z_n} .$$

Thus the relationship between $\frac{v_{n-1}}{z_{n-1}-z_n}$, 0, and t_{n-2} will determine the specific ranges of integration for the integral (2.2.3) over which the integrand is nonzero.

This relationship determines three possible integrals:

(i) if $\frac{v_{n-1}}{z_{n-1}-z_n} \leq 0$ then $t_{n-1} > \frac{v_{n-1}}{z_{n-1}-z_n}$ in the range 0 to t_{n-2} and this implies that the integral is

$$\int_0^{t_{n-2}} g^{(n-1)}(v_n)dt_{n-1} \quad ;$$

(ii) if $0 < \frac{v_{n-1}}{z_{n-1}-z_n} < t_{n-2}$ then $t_{n-1} > \frac{v_{n-1}}{z_{n-1}-z_n}$ in the range $\frac{v_{n-1}}{z_{n-1}-z_n}$ to t_{n-2} and this implies that the integral is

$$\int_{\frac{v_{n-1}}{z_{n-1}-z_n}}^{t_{n-2}} g^{(n-1)}(v_n)dt_{n-1} \quad ;$$

(iii) if $\frac{v_{n-1}}{z_{n-1}-z_n} \geq t_{n-2}$ then $t_{n-1} \leq \frac{v_{n-1}}{z_{n-1}-z_n}$ for $0 < t_{n-1} < t_{n-2}$ which implies that the integral is zero.

The following shows that in all the above cases the integral (2.2.3) is equivalent to (2.2.2).

(i) Assume (2.2.3) corresponds to situation (i). Then

$$\int_0^{t_{n-2}} g^{(n-1)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}))dt_{n-1}$$

$$= \left. \frac{g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}))}{z_n-z_{n-1}} \right|_0^{t_{n-2}}$$

since $g^{(n-1)}(v_n)$ is continuous over the range of integration,

$$= \frac{1}{z_n-z_{n-1}}\Big[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-2}(z_n-z_{n-1}))$$

$$- g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+0(z_n-z_{n-1}))\Big]$$

$$= \frac{1}{z_{n-1}-z_n}\Big[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))$$

$$- g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))\Big]$$

which is identical with (2.2.2).

(2) Assume (2.2.3) corresponds to situation (ii). Then

$$\int_0^{t_{n-2}} g^{(n-1)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}))dt_{n-1}$$

$$= \int_{\frac{v_{n-1}}{z_{n-1}-z_n}}^{t_{n-2}} g^{(n-1)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}))dt_{n-1}$$

$$= \left. \frac{g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}))}{z_n-z_{n-1}} \right|_{\frac{v_{n-1}}{z_{n-1}-z_n}}^{t_{n-2}}$$

since $g^{(n-1)}(v_n)$ is continuous over the range of integration,

$$= \frac{1}{z_n - z_{n-1}}$$

$$\cdot\Big[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-2}(z_n-z_{n-1}))$$

$$- g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})-(\frac{z_n-z_{n-1}}{z_n-z_{n-1}})v_{n-1})\Big]$$

$$= \frac{1}{z_{n-1}-z_n}\Big[g^{(n-2)}(0) - g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))\Big]$$

$$= \frac{1}{z_{n-1}-z_n}\Big[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))$$

$$- g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))\Big]$$

which is identical with (2.2.2). The last statement follows from the fact that under situation (ii) $\frac{v_{n-1}}{z_{n-1}-z_n} > 0$. That is $v_{n-1} < 0$ since $z_{n-1} < z_n$ and thus

$$g^{(n-2)}(v_{n-1}) = g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})) = 0 = g^{(n-2)}(0).$$

(3) Assume (2.2.3) corresponds to situation (iii). Then

$$\int_0^{t_{n-2}} g^{(n-1)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})+t_{n-1}(z_n-z_{n-1}))dt_{n-1}$$

$$= 0$$

$$= \frac{1}{z_{n-1}-z_n}\Big[g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))$$

$$- g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}))\Big]$$

which is identical with (2.2.2). This last statement follows since under situation (iii) $\frac{v_{n-1}}{z_{n-1}-z_n} > t_{n-2} > 0$.

That is,

(a) $v_{n-1} < 0$ and therefore

$$g^{(n-2)}(v_{n-1}) = g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2}))=0$$

and

(b) $\frac{z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_{n-1}-z_{n-2})}{z_{n-1}-z_n} > t_{n-2}$

that is

$$z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2}) < 0$$

and therefore

$$g^{(n-2)}(z_1+t_1(z_2-z_1)+\ldots+t_{n-2}(z_n-z_{n-2})) = 0 \quad .$$

The assumption was made that $z_n > z_{n-1}$. If this is not the case then, using the symmetric property of divided differences, $z_1,z_2,\ldots,z_n$ can be relabelled by $z_1',z_2',\ldots,z_n'$ such that

$$z_{n-1} = z_n'$$

$$z_n = z_{n-1}'$$

$$z_i = z_i' \qquad (i=1,2,\ldots,n-2) \quad .$$

Then $z_n' > z_{n-1}'$ and the above procedure is repeated using $z_1',z_2',\ldots,z_n'$. ∎

The distribution function of Y is derived using the result of this lemma. Note that $a_0 = 0$.

For convenience define

$$(y-a_i)_+^n = \begin{cases} (y-a_i)^n & \text{if } y > a_i \\ 0 & \text{if } y \leq a_i \end{cases} .$$

<u>Theorem 2.2</u> The distribution function of

$$Y = a_1U_1 + a_2U_2 + \cdots + a_nU_n$$

is given by

$$F(y) = \sum_{i=0}^{n} \frac{(y-a_i)_+^n}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j-a_i)}$$

provided $a_i \neq a_j$ for all distinct i and j.

<u>Proof:</u>

By definition the distribution function of Y is

$$F(y) = Pr(Y < y) .$$

As noted in Chapter I, the distribution of

$$Y = \sum_{i=1}^{n} a_iU_i \qquad \text{and} \qquad X = \sum_{i=1}^{n} c_iX_{(i)}$$

are the same and thus

$$F(y) = \int\int \cdots \int\int_{\left\{\begin{array}{l} 0 \le x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} \le 1 \\ \sum_{i=1}^{n} c_i x_{(i)} < y \end{array}\right\}} n!\, dx_{(1)}\, dx_{(2)} \cdots dx_{(n-1)}\, dx_{(n)}$$

$$= n! \int_0^1 \int_0^{x_{(n)}} \cdots \int_0^{x_{(3)}} \int_0^{x_{(2)}} g^{(n)}\Big(y - \sum_{i=1}^{n} c_i x_{(i)}\Big) \cdot dx_{(1)}\, dx_{(2)} \cdots dx_{(n-1)}\, dx_{(n)}$$

where

$$g^{(n)}(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z \le 0 \end{cases}$$

Furthermore

$$F(y) = n! \int_0^1 \int_0^{x_{(n)}} \cdots \int_0^{x_{(2)}} g^{(n)}(s_{n+1})\, dx_{(1)} \cdots dx_{(n-1)}\, dx_{(n)}$$

where

$$s_{n+1} = (y-a_0) + \sum_{i=1}^{n-1} \big[(y-a_i) - (y-a_{i+1})\big] x_{(i)} + \big[(y-a_n)-(y-a_0)\big] x_{(n)}$$

since

$$y - \sum_{i=1}^{n} c_i x_{(i)} = y - \sum_{i=1}^{n-1} c_i x_{(i)} - c_n x_{(n)}$$

$$= y + \sum_{i=1}^{n-1} (a_{i+1}-a_i)x_{(i)} - a_n x_{(n)}$$

because $c_i = \sum_{j=i}^{n} c_j - \sum_{j=i+1}^{n} c_j = a_i - a_{i+1}$ $(i=1,2,\ldots,n-1)$

and

$$c_n = \sum_{j=n}^{n} c_j = a_n$$

Defining $x_{(i)} = t_{n-i+1}$ $(i=1,2,\ldots,n)$

$z_i = y-a_i$ $(i=0,1,\ldots,n)$

$$F(y) = n! \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} g^{(n)}(z_0+(z_n-z_0)t_1+(z_{n-1}-z_n)t_2+\ldots +(z_1-z_2)t_n)\, dt_n \ldots dt_2\, dt_1$$

since

$$s_{n+1} = (y-a_0) + \sum_{i=1}^{n-1} \left[(y-a_i)-(y-a_{i+1})\right] x_{(i)} + \left[(y-a_n)-(y-a_0)\right] x_{(n)}$$

$$= z_0 + \sum_{i=1}^{n-1} (z_i-z_{i+1})x_{(i)} + (z_n-z_0)x_{(n)}$$

$$= z_0 + \sum_{i=1}^{n-1} (z_i-z_{i+1})t_{n-i+1} + (z_n-z_0)t_1$$

$$= z_0 + (z_n - z_0)t_1 + (z_{n-1} - z_n)t_2 + \dots + (z_1 - z_2)t_n \ .$$

Then by Lemma 2.1

$$F(y) = n!\ D^n(g(z): z_0, z_n, z_{n-1}, \dots, z_1)$$

$$= n!\ D^n(g(z): y-a_0, y-a_n, y-a_{n-1}, \dots, y-a_1) \ .$$

Using the symmetry property of divided differences it follows that

$$F(y) = n!\ D^n(g(z): y-a_0, y-a_1, \dots, y-a_{n-1}, y-a_n)$$

$$= n! \sum_{i=0}^{n} \frac{g(y-a_i)}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (y-a_i-y+a_j)}$$

$$= n! \sum_{i=0}^{n} \frac{g(y-a_i)}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j-a_i)} \ .$$

Since

$$g^{(n)}(y-a_i) = \begin{cases} 1 & \text{if } y > a_i \\ 0 & \text{if } y \leq a_i \end{cases}$$

it follows that

$$g(y-a_i) = \begin{cases} \dfrac{(y-a_i)^n}{n!} & \text{if } y > a_i \\ 0 & \text{if } y \leq a_i \end{cases}$$

$$= \frac{(y-a_i)_+^n}{n!} \quad .$$

Thus

$$F(y) = \sum_{i=0}^{n} \frac{(y-a_i)_+^n}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j-a_i)} \quad . \blacksquare$$

Corollary 2.3 The density function of

$$Y = a_1U_1 + a_2U_2 + \dots + a_nU_n$$

is given by

$$f(y) = n \sum_{i=0}^{n} \frac{(y-a_i)_+^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j-a_i)}$$

provided $a_i \neq a_j$ for all distinct i and j.

Proof:

Since the conditions of Theorem 2.2 are satisfied, using the distribution function of Y obtained in that theorem

$$f(y) = \frac{d}{dy} F(y)$$

$$= \sum_{i=0}^{n} \frac{\frac{d}{dy}(y-a_i)_+^n}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_j-a_i)}$$

$$= n \sum_{i=0}^{n} \frac{(y-a_i)_+^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_j-a_i)} \quad . \; \|$$

Another consequence of Theorem 2.2 is the following result. Note that, here and throughout this paper, $i = \sqrt{-1}$.

Corollary 2.4 The characteristic function of

$$Y = a_1U_1 + a_2U_2 + \ldots + a_nU_n$$

is given by

$$\phi(t) = n!\,(it)^{-n} \sum_{i=0}^{n} \frac{e^{ita_i}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_i-a_j)}$$

provided $a_i \neq a_j$ for all distinct i and j.

Proof:

Since the conditions of Corollary 2.3 are satisfied the characteristic function of Y is

$$\emptyset(t) = \int_{-\infty}^{\infty} e^{ity} f(y)\, dy$$

$$= \int_{-\infty}^{\infty} e^{ity}\, n \sum_{i=0}^{n} \frac{(y-a_i)_+^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_j-a_i)}\, dy \quad .$$

For $y > \max_{j \in (o,1,\ldots,n)} a_j = a^*$

$$\sum_{i=0}^{n} \frac{(y-a_i)_+^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_i-a_j)} = \sum_{i=0}^{n} \frac{(y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_i-a_j)} = 0$$

since it is a divided difference of order n of a polynomial of degree (n-1). Thus

$$\emptyset(t) = \int_{-\infty}^{a^*} e^{ity}\, n \sum_{i=0}^{n} \frac{(y-a_i)_+^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_j-a_i)}\, dy$$

$$= n \sum_{i=0}^{n} \frac{1}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)} \int_{-\infty}^{a^*} e^{ity} (y-a_i)_+^{n-1} \, dy$$

$$= n \sum_{i=0}^{n} \frac{1}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)} \int_{a_i}^{a^*} e^{ity} (y-a_i)^{n-1} \, dy \, .$$

Let $w = y - a_i$. Then

$$\phi(t) = n \sum_{i=0}^{n} \frac{1}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)} \int_{0}^{a^*-a_i} e^{it(w+a_i)} w^{n-1} \, dw$$

$$= n \sum_{i=0}^{n} \frac{e^{ita_i}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)} \int_{0}^{a^*-a_i} e^{itw} w^{n-1} \, dw \, .$$

But using integration by parts

$$\int e^{itw} w^{n-1} \, dw = \sum_{k=0}^{n-1} \frac{\binom{n-1}{n-k-1} k! (-1)^k e^{itw} w^{n-k-1}}{(it)^{k+1}}$$

where $(0)^0$ is used to represent 1. Thus

$$\phi(t) = n \sum_{i=0}^{n} \frac{e^{ita_i}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)} \sum_{k=0}^{n-1} \frac{\binom{n-1}{n-k-1} k! \, (-1)^k \, e^{itw} \, w^{n-k-1}}{(it)^{k+1}} \Bigg|_{0}^{a^*-a_i}$$

$$= n \sum_{i=0}^{n} \frac{e^{ita_i}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)}$$

$$\cdot \left[\sum_{k=0}^{n-1} \frac{\binom{n-1}{n-k-1} k! \, (-1)^k \, e^{it(a^*-a_i)} \, (a^*-a_i)^{n-k-1}}{(it)^{k+1}} \right.$$

$$\left. - \frac{(n-1)! \, (-1)^{n-1}}{(it)^n} \right]$$

$$= n \sum_{k=0}^{n-1} \frac{\binom{n-1}{n-k-1} k! \, (-1)^{n+k} \, e^{ita^*}}{(it)^{k+1}} \left[\sum_{i=0}^{n} \frac{(a^*-a_i)^{n-k-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_i - a_j)} \right]$$

$$+ (-1)^n \, n! \, (it)^{-n} \sum_{i=0}^{n} \frac{e^{ita_i}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j - a_i)}$$

$$= n!\,(it)^{-n} \sum_{i=0}^{n} \frac{e^{ita_i}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_i-a_j)}$$

because all other terms are zero for each
$k \in (0,1,\ldots,n-1)$ since

$$\sum_{i=0}^{n} \frac{(a^*-a_i)^{n-k-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_i-a_j)}$$

is a divided difference of order n of a polynomial of degree less than n and is thus zero. ∎

The density function obtained in Corollary 2.3 can be written in a slightly different form by using the fact that

$$\sum_{i=0}^{n} \frac{(y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_i-a_j)} = 0$$

since this is a divided difference of order n of a polynomial of degree (n-1). Defining

$$e(x) = \begin{cases} 1 & \text{if } y > x \\ 0 & \text{if } y \leq x \end{cases}$$

one has

<u>Corollary 2.5</u> The density function of

$$Y = a_1U_1 + a_2U_2 + \dots + a_nU_n$$

is given by

$$f(y) = n \sum_{i=1}^{n} \frac{I(a_i)\,(y-a_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_j-a_i)}$$

where

$$I(a_i) = \begin{cases} -1 & \text{if } y \in (0,a_i],\ a_i > 0 \\ 0 & \text{if } y \notin (0,a_i],\ a_i > 0 \\ 0 & \text{if } y \notin (a_i,0],\ a_i < 0 \\ 1 & \text{if } y \in (a_i,0],\ a_i < 0 \end{cases}$$

provided $a_i \neq a_j$ for all distinct i and j.

<u>Proof:</u>

From

$$\sum_{i=0}^{n} \frac{(y-a_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{n} (a_i-a_j)} = 0$$

it follows that

$$\frac{(y-a_0)^{n-1}}{\prod_{j=1}^{n}(a_0-a_j)} = -\sum_{i=1}^{n}\frac{(y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n}(a_i-a_j)}$$

and since the conditions of Corollary 2.3 are satisfied

$$f(y) = n\sum_{i=0}^{n}\frac{(y-a_i)_+^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n}(a_j-a_i)}$$

$$= n\sum_{i=0}^{n}\frac{e(a_i)\ (y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n}(a_j-a_i)}$$

$$= n\sum_{i=1}^{n}\frac{e(a_i)\ (y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n}(a_j-a_i)} + \frac{(-1)^n n\ e(a_o)\ (y-a_0)^{n-1}}{\prod_{j=1}^{n}(a_0-a_j)}$$

$$= n\sum_{i=1}^{n}\frac{e(a_i)\ (y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n}(a_j-a_i)} - (-1)^n\ n\ e(a_0)\sum_{i=1}^{n}\frac{(y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n}(a_i-a_j)}$$

$$= n \sum_{i=1}^{n} \frac{(e(a_i)-e(a_0))\ (y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_j-a_i)}$$

$$= n \sum_{i=1}^{n} \frac{I(a_i)\ (y-a_i)^{n-1}}{\prod_{\substack{j=0\\ j\neq i}}^{n} (a_j-a_i)}$$

since a simple calculation shows that $I(a_i) = e(a_i)-e(a_0)$. ▌

2.3 : The distribution of one linear combination with repeated coefficients

The general situation in which the possibility of equal coefficients exists will now be considered. Suppose that only p+1 of the n+1 coefficients $a_0,a_1,\ldots,a_n$ are distinct. Denote these distinct coefficients by $b_0,b_1,\ldots,b_p$ and assume that n_i of the actual coefficients are equal to b_i for $i=0,1,\ldots,p$, Notice that the coefficient b_0 represents $a_0 = 0$ and n_0-1 other zero coefficients. For the derivation of the distribution function, density function, and characteristic function of the linear combination

$$Y = a_1U_1 + a_2U_2 + \ldots + a_nU_n$$

the following result given in Milne-Thompson (1933) is required.

<u>Lemma 2.6</u> If $g^{(n-1)}(z)$ is continuous

$$D^{n-1}(g(z):z_1,z_1,\ldots,z_p,z_p) = \int_0^1 dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{p-2}} \gamma(\underline{t})\, g^{(n-1)}(v_p)\, dt_{p-1}$$

where

$$v_p = z_1 + t_1(z_2-z_1) + \ldots + t_{p-1}(z_p-z_{p-1})$$

$$\gamma(\underline{t}) = \frac{(1-t_1)^{n_1-1}\,(t_1-t_2)^{n_2-1}\,\ldots\,(t_{p-2}-t_{p-1})^{n_{p-1}-1}\;t_{p-1}^{n_p-1}}{(n_1-1)!\,(n_2-1)!\,\ldots\,(n_p-1)!}$$

and z_i is repeated n_i times $(i=1,2,\ldots,p)$.

<u>Proof:</u>

Using the divided difference property given in section 1.2 for repeated arguments and the property

$$D^{p-1}(g(z):z_1,z_2,\ldots,z_p) = \int_0^1 dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{p-2}} g^{(p-1)}(v_p)\, dt_{p-1} \qquad (2.3.1)$$

where v_p is defined in the statement of this lemma it follows that

$$D^{n-1}(g(z):z_1,z_1,\ldots,z_p,z_p)$$

$$= \frac{1}{(n_1-1)!\,(n_2-1)!\,\ldots\,(n_p-1)!}\;\frac{\partial^{n_1+n_2+\ldots+n_p-p}}{\partial_{z_1}^{n_1-1}\ldots\partial_{z_p}^{n_p-1}}$$

$$\cdot \int_0^1 dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{p-2}} g^{(p-1)}(v_p)\, dt_{p-1} \quad .$$

Then using the chain rule for differentiation

$$\frac{\partial^{n_1+n_2+\ldots+n_p-p}}{\partial_{z_1}^{n_1-1}\ldots\,\partial_{z_p}^{n_p-1}}\, g^{(p-1)}(v_p)$$

$$= \frac{\partial^{n_1+n_2+\ldots+n_{p-1}-(p-1)}}{\partial_{z_1}^{n_1-1}\;\ldots\;\partial_{z_{p-1}}^{n_{p-1}-1}}\, g^{(n_p-2+p)}(v_p)\; t_{p-1}^{n_p-1}$$

and continuing in this manner

$$\frac{1}{(n_1-1)!\,(n_2-1)!\,\ldots\,(n_p-1)!}\;\frac{\partial^{n_1+n_2+\ldots+n_p-p}}{\partial_{z_1}^{n_1-1}\ldots\partial_{z_p}^{n_p-1}}\, g^{(p-1)}(v_p)$$

$$= \gamma(\underline{t})\; g^{(n_1+n_2+\ldots+n_p-p-1+p)}(v_p)$$

$$= \gamma(\underline{t})\; g^{(n-1)}(v_p)$$

and the lemma follows. ▮

Since, by Lemma 2.1, (2.3.1) is true for a step function the above result also holds for such a function.

Among the coefficients $a_1, a_2, \ldots, a_n$ denote those equal to b_i by $a_{i1}, a_{i2}, \ldots, a_{in_i}$ $(i=1,2,\ldots,p)$ and those equal to b_0 by $a_{01}, a_{02}, \ldots, a_{0\,n_0-1}$. Denote the coverages having these coefficients by similar subscripts. That is,

$$U_{ij} \qquad (i=1,2,\ldots,p\ ;\ j=1,2,\ldots,n_i)$$

$$U_{0j} \qquad (j=1,2,\ldots,n_0-1) \quad .$$

Then

$$\begin{aligned} Y &= a_1U_1 + a_2U_2 + \ldots + a_nU_n \\ &= b_0(U_{01}+U_{02}+\ldots+U_{0\,n_0-1}) + b_1(U_{11}+U_{12}+\ldots+U_{1n_1}) \\ &\qquad + \ldots + b_p(U_{p1}+U_{p2}+\ldots+U_{pn_p}) \quad . \end{aligned}$$

Finally define $R_{(0)} = 0$ and

$$R_{(i)} = \sum_{k=1}^{i} \sum_{h=1}^{n_k} U_{kh} \qquad (i=1,2,\ldots,p) \quad .$$

In order to obtain the distribution function of Y in the case of repeated coefficients it will be required to know the probability element of the joint distribution of $R_{(1)}, R_{(2)}, \ldots, R_{(p)}$. This needed information is supplied by the following lemma.

Lemma 2.7 The probability element of the joint distribution of $R_{(1)}, R_{(2)}, \ldots, R_{(p)}$ is

$$n!\; B_p(\underline{r})\, dr_{(1)}\, dr_{(2)} \cdots dr_{(p)}$$

where

$$B_p(\underline{r}) = \frac{(1-r_{(p)})^{n_0-1}\,(r_{(p)}-r_{(p-1)})^{n_p-1} \cdots (r_{(2)}-r_{(1)})^{n_2-1}\, r_{(1)}^{n_1-1}}{(n_0-1)!\,(n_1-1)! \cdots (n_p-1)!}$$

for the domain $0 \le r_{(1)} \le r_{(2)} \le \cdots \le r_{(p)} \le 1$.

Proof:

Relabel the original coverages as follows.

$$V_{e_i+k} = U_{i+1\,k} \qquad (i=0,1,\ldots,p-1\ ;\ k=1,2,\ldots,n_{i+1})$$

$$V_{e_p+k} = U_{0k} \qquad (k=1,2,\ldots,n_0-1)$$

where $e_0, e_1, \ldots, e_p$ are defined by

$$e_0 = 0 \quad ; \quad e_i = \sum_{j=1}^{i} n_j \qquad (i=1,2,\ldots,p) .$$

Define the statistics, which are ordered in increasing magnitude,

$$G_{(i)} = \sum_{j=1}^{i} V_j \qquad (i=1,2,\ldots,n) .$$

It should be noted that $R_{(i)} = G_{(e_i)}$ $(i=1,2,\ldots,p)$.

It is known that (Sarhan and Greenberg (1962)) if $Z_{(1)}, Z_{(2)}, \ldots, Z_{(m)}$ are the order statistics corresponding to a random sample of size m from a population with the distribution function $H(z)$ and the density function $h(z)$ then the probability element of the joint distribution of the k order statistics $Z_{(m_1)}, Z_{(m_2)}, \ldots, Z_{(m_k)}$ where

$1 \le m_1 < m_2 < \ldots < m_k \le m$, is given by

$$\frac{m!}{\prod_{i=1}^{k+1} (m_i - m_{i-1} - 1)!} \prod_{i=1}^{k+1} \left[H(z_{(m_i)}) - H(z_{(m_{i-1})})\right]^{m_i - m_{i-1} - 1}$$

$$\cdot \prod_{i=1}^{k} h(z_{(m_i)})\, dz_{(m_i)}$$

for the domain $z_{(m_1)} \le z_{(m_2)} \le \ldots \le z_{(m_k)}$ where $m_0 = 0$, $m_{k+1} = m+1$, $z_{(m_0)} = -\infty$, and $z_{(m_{k+1})} = +\infty$.

Applying this to $R_{(i)} = G_{(e_i)}$ $(i=1,2,\ldots,p)$ it follows that the probability element of these p order statistics is

$$\frac{n!}{\prod_{i=1}^{p+1} (e_i - e_{i-1} - 1)!} \prod_{i=1}^{p+1} \left[r_{(i)} - r_{(i-1)}\right]^{e_i - e_{i-1} - 1} \prod_{i=1}^{p} dr_{(i)}$$

$$= \frac{n!}{(n_1-0-1)!\,(\sum_{j=1}^{2} n_j-n_1-1)!\,\cdots\,(\sum_{j=1}^{p} n_j-\sum_{j=1}^{p-1} n_j-1)!\,(n+1-\sum_{j=1}^{p} n_j-1)!}$$

$$\cdot(r_{(1)}-0)^{n_1-1}\,(r_{(2)}-r_{(1)})^{n_2-1}\cdots(r_{(p)}-r_{(p-1)})^{n_p-1}$$

$$\cdot(1-r_{(p)})^{n_0-1}\,dr_{(1)}\,dr_{(2)}\cdots dr_{(p)}$$

$$= n!\,B_p(\underline{r})\,dr_{(1)}\,dr_{(2)}\cdots dr_{(p)}$$

for the domain $0 \leq r_{(1)} \leq r_{(2)} \leq \cdots \leq r_{(p)} \leq 1$. ▌

The following theorem establishes the generalization of the result obtained in Theorem 2.2 and uses the notation defined earlier in this section.

<u>Theorem 2.8</u> The distribution function of

$$Y = a_1U_1 + a_2U_2 + \cdots + a_nU_n$$

is given by

$$F(y) = (-1)^n \prod_{s=0}^{p} \frac{1}{(n_s-1)!}\,\frac{\partial^{n-p}}{\partial_{b_0}^{n_0-1}\,\partial_{b_1}^{n_1-1}\cdots\partial_{b_p}^{n_p-1}}$$

$$\cdot \sum_{i=0}^{p} \frac{(y-b_i)_+^n}{\prod_{\substack{j=0\\ j\neq i}}^{p}(b_i-b_j)} \, .$$

Proof:

Defining $h_1, h_2, \ldots, h_p$ by

$$b_i = \sum_{j=i}^{p} h_j \qquad (i=1,2,\ldots,p),$$

Y can be expressed as the linear combination

$$Y = h_1 R_{(1)} + h_2 R_{(2)} + \cdots + h_p R_{(p)}$$

and using Lemma 2.7 it follows that

$$\begin{aligned} F(y) &= \Pr(Y < y) \\ &= \Pr\left(\sum_{i=1}^{p} h_i R_{(i)} < y\right) \\ &= \int\int_{\left\{\begin{array}{l} 0 \le r_{(1)} \le r_{(2)} \le \cdots \le r_{(p)} \le 1 \\ \sum_{i=1}^{p} h_i r_{(i)} < y \end{array}\right\}} \cdots \int\int n!\, B_p(\underline{r})\, dr_{(1)} \cdots dr_{(p-1)} dr_{(p)} \\ &= n! \int_0^1 \int_0^{r_{(p)}} \cdots \int_0^{r_{(3)}} \int_0^{r_{(2)}} g^{(n)}\left(y - \sum_{i=1}^{p} h_i r_{(i)}\right) \\ &\qquad \cdot B_p(\underline{r})\, dr_{(1)}\, dr_{(2)} \cdots dr_{(p-1)}\, dr_{(p)} \end{aligned}$$

where

$$g^{(n)}(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z \leq 0 \end{cases} .$$

Furthermore, similar to the situation in the proof of Theorem 2.2,

$$F(y) = n! \int_0^1 \int_0^{r_{(p)}} \cdots \int_0^{r_{(3)}} \int_0^{r_{(2)}} g^{(n)}(s_{p+1})\, B_p(\underline{r})$$

$$\cdot\, dr_{(1)}\, dr_{(2)} \cdots dr_{(p-1)}\, dr_{(p)}$$

where

$$s_{p+1} = (y-b_0) + \sum_{i=1}^{p-1} \left[(y-b_i)-(y-b_{i+1})\right] r_{(i)} + \left[(y-b_p)-(y-b_0)\right] r_{(p)} .$$

Define $r_{(i)} = t_{p-i+1}$ $(i=1,2,\ldots,p)$

$z_i = y-b_i$ $(i=0,1,\ldots,p)$.

Then, again similar to a situation in the proof of Theorem 2.2,

$$s_{p+1} = z_0 + (z_p-z_0)t_1 + (z_{p-1}-z_p)t_2 + \cdots + (z_1-z_2)t_p .$$

Also

$$B_p(\underline{r}) = \frac{(1-t_1)^{n_0-1}\,(t_1-t_2)^{n_p-1} \cdots (t_{p-1}-t_p)^{n_2-1}\, t_p^{n_1-1}}{(n_0-1)!\,(n_1-1)! \cdots (n_p-1)!} .$$

Denote the right side of the above line by $C_p(\underline{t})$. Then the distribution function of Y can be written as

$$F(y) = n! \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{p-1}} c_p(\underline{t})\, g^{(n)}(z_0+(z_p-z_0)t_1+(z_{p-1}-z_p)t_2$$

$$+\ldots+(z_1-z_2)t_p)\, dt_p \ldots dt_2\, dt_1 \;.$$

Lemma 2.6 now applies and

$$F(y) = n!\, D^n(g(z):z_0,z_0,\ldots,z_0,z_p,z_p,\ldots,z_1,z_1)$$

$$= \frac{n!}{(n_0-1)!\,(n_1-1)!\,\ldots\,(n_p-1)!} \frac{\partial^{n_0+n_1+\ldots+n_p-(p+1)}}{\partial_{z_0}^{n_0-1}\,\partial_{z_1}^{n_1-1}\cdots\partial_{z_p}^{n_p-1}}$$

$$.\, D^p(g(z):z_0,z_1,\ldots,z_p)$$

using the symmetric property of divided differences and the formula for repeated arguments given in Chapter I.

Then using the fact that $n_0+n_1+\ldots+n_p = n+1$ and the definition of z_i $(i=0,1,\ldots,p)$

$$F(y) = n! \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{y-b_0}^{n_0-1}\,\partial_{y-b_1}^{n_1-1}\cdots\partial_{y-b_p}^{n_p-1}}$$

$$.\, D^p(g(z):y-b_0,y-b_1,\ldots,y-b_p) \;.$$

Also

$$F(y) = n!\,(-1)^{n-p} \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_0}^{n_0-1}\partial_{b_1}^{n_1-1}\cdots\partial_{b_p}^{n_p-1}}$$

$$\cdot\, D^p(g(z):y-b_0,y-b_1,\ldots,y-b_p)$$

by the chain rule for differentiation,

$$= n!\,(-1)^{n-p} \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_0}^{n_0-1}\partial_{b_1}^{n_1-1}\cdots\partial_{b_p}^{n_p-1}}$$

$$\cdot \sum_{i=0}^{p} \frac{g(y-b_i)}{\prod_{\substack{j=0\\ j\neq i}}^{p} (y-b_i-y+b_j)}$$

$$= n!\,(-1)^{n} \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_0}^{n_0-1}\partial_{b_1}^{n_1-1}\cdots\partial_{b_p}^{n_p-1}}$$

$$\cdot \sum_{i=0}^{p} \frac{g(y-b_i)}{\prod_{\substack{j=0\\ j\neq i}}^{p} (b_i-b_j)}$$

$$= (-1)^n \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_0}^{n_0-1} \partial_{b_1}^{n_1-1} \cdots \partial_{b_p}^{n_p-1}}$$

$$\cdot \sum_{i=0}^{p} \frac{(y-b_i)_+^n}{\prod_{\substack{j=0 \\ j \neq i}}^{p} (b_i-b_j)} ,$$

since, as shown in the proof of Theorem 2.2,

$$g(y-b_i) = \begin{cases} \dfrac{(y-b_i)^n}{n!} & \text{if } y > b_i \\ 0 & \text{if } y \leq b_i \end{cases}$$

$$= \frac{(y-b_i)_+^n}{n!} .$$

The proof of the theorem is now complete. ∎

For convenience define

$$L_{np} = \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_0}^{n_0-1} \partial_{b_1}^{n_1-1} \cdots \partial_{b_p}^{n_p-1}} .$$

The following result is a consequence of the above theorem.

<u>Corollary 2.9</u> The density function of

$$Y = a_1U_1 + a_2U_2 + \cdots + a_nU_n$$

is given by

$$f(y) = (-1)^n \, n \, L_{np} \sum_{i=0}^{p} \frac{(y-b_i)_+^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)} . \qquad (2.3.2)$$

<u>Proof</u>:

Using Theorem 2.8

$$f(y) = \frac{d}{dy} F(y)$$

$$= (-1)^n L_{np} \sum_{i=0}^{p} \frac{\frac{d}{dy}(y-b_i)_+^{n}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

provided the differentiation operations can be interchanged. The result then follows. ∎

Similar to the case of distinct coefficients the following result is obtained.

<u>Corollary 2.10</u> The characteristic function of

$$Y = a_1U_1 + a_2U_2 + \cdots + a_nU_n$$

is given by

$$\phi(t) = n!\ (it)^{-n}\ L_{np} \sum_{i=0}^{p} \frac{e^{itb_i}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)} \quad .$$

Proof:

Using Corollary 2.9 the characteristic function of Y is given by

$$\phi(t) = \int_{-\infty}^{\infty} e^{ity}(-1)^n\ n\ L_{np} \sum_{i=0}^{p} \frac{(y-b_i)_+^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}\, dy \quad .$$

The remainder of the proof is similar to that of Corollary 2.4 provided the integration and differentiation operations can be interchanged. ▌

Using the fact that

$$0 = D^n((y-x)^{n-1}:a_0,a_1,\ldots,a_n) = L_{np} \sum_{i=0}^{p} \frac{(y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

formula (2.3.2) can be written in a different form.

Corollary 2.11 The density function of

$$Y = a_1U_1 + a_2U_2 + \ldots + a_nU_n$$

is given by

$$f(y) = (-1)^n \, n \, L_{np} \sum_{i=1}^{p} \frac{I(b_i)\,(y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)} \qquad (2.3.3)$$

where

$$I(b_i) = \begin{cases} -1 & \text{if } y \in (0,b_i] \, , \; b_i > 0 \\ 0 & \text{if } y \notin (0,b_i] \, , \; b_i > 0 \\ 0 & \text{if } y \notin (b_i,0] \, , \; b_i < 0 \\ 1 & \text{if } y \in (b_i,0] \, , \; b_i < 0 \end{cases}$$

Proof:

From

$$L_{np} \sum_{i=0}^{p} \frac{(y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)} = 0$$

it follows that

$$L_{np} \frac{(y-b_0)^{n-1}}{\prod_{j=1}^{p} (b_0-b_j)} = -L_{np} \sum_{i=1}^{p} \frac{(y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)} \, .$$

This fact is used in the following.

By Theorem 2.8 and recalling the definition of $e(x)$

$$f(y) = (-1)^n \, n \, L_{np} \sum_{i=0}^{p} \frac{e(b_i) \; (y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$= (-1)^n \, n \, L_{np} \sum_{i=1}^{p} \frac{e(b_i) \; (y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$+ (-1)^n \, n \, L_{np} \frac{e(b_0) \; (y-b_0)^{n-1}}{\prod_{j=1}^{p} (b_0-b_j)}$$

$$= (-1)^n \, n \, L_{np} \sum_{i=1}^{p} \frac{e(b_i) \; (y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$- (-1)^n \, n \, L_{np} \, e(b_0) \sum_{i=1}^{p} \frac{(y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$= (-1)^n \, n \, L_{np} \sum_{i=1}^{p} \frac{(e(b_i)-e(b_0)) \, (y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$= (-1)^n \, n \, L_{np} \sum_{i=1}^{p} \frac{I(b_i) \, (y-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

since it can easily be shown that $I(b_i) = e(b_i)-e(b_0)$. ▮

2.4 : Application to the distribution of the sums of ordered intervals

The application of the results obtained in the previous section is illustrated by a paper due to Barton and David (1955) . In this paper they were concerned with the analysis of the intervals which arise between events occurring randomly in time.

A line of unit length is randomly divided by n points $X_1, X_2, \ldots, X_n$ into (n+1) intervals. The (n+1) intervals between the successive points, including the endpoints, are ordered in magnitude:

$$0 \leq H_1 \leq H_2 \leq \ldots \leq H_{n+1} \leq 1 \quad , \quad \sum_{i=1}^{n+1} H_i = 1 \ .$$

At one point of their paper they derived the density function of the sum of (s-r+1) consecutively ordered intervals

$$G_{rs} = \sum_{i=r}^{s} H_i$$

to be

$$f(g_{rs}) = \frac{n\,(n+1)!}{(n+1-s)!\,(s-r)!\,(r-1)!\,(n+1-s)^{s-r-1}} \cdot \sum_{k=0}^{r-1} \sum_{h=0}^{s-r-1} (-1)^{k+h} \binom{r-1}{k} \binom{s-r}{h} \frac{(s-r-h)^{s-r}\,\Delta_{kh}}{(h+1)(n+1-s)-k(s-r-h)}$$

(2.4.1)

where

$$\Delta_{kh} = \left[H\left(1-g_{rs}\,\frac{n+2-r+k}{s-r+1}\right)\right]^{n-1} - \left[H\left(1-g_{rs}\,\frac{n+1-r-h}{s-r-h}\right)\right]^{n-1}$$

and

$H(z) = 1/2\ (|z| + z).$

They noted that if

$$(h+1)(n+1-s) \neq k(s-r-h)$$

then it is a factor of Δ_{kh} . Also if there are any terms for which

$$(h+1)(n+1-s) = k(s-r-h)\ ,$$

and thus an indeterminate form appears in the above formula, the factor can be divided out before the equality is substituted.

In order to use the approach of section 2.3 it is required to express G_{rs} not as the sum of the ordered intervals but rather as the sum of the coverages corresponding to the order statistics determined by $X_1, X_2, \ldots, X_n$. That is, let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be the order statistics corresponding to $X_1, X_2, \ldots, X_n$ and define the coverages $U_1, U_2, \ldots, U_{n+1}$ in the usual fashion. Ordering the coverages in ascending order of magnitude the ordered intervals $U_{(1)}, U_{(2)}, \ldots, U_{(n+1)}$, which were previously denoted by $H_1, H_2, \ldots, H_{n+1}$ respectively, are obtained. It was found in Chapter I that

$$U_{(i)} = \sum_{j=1}^{i} \frac{U_j}{n+2-j} \qquad (i=1,2,\ldots,n+1) .$$

Using this transformation G_{rs} can be expressed as a linear combination of the coverages as follows:

$$\begin{aligned} G_{rs} &= \sum_{i=r}^{s} U_{(i)} \\ &= \sum_{i=r}^{s} \sum_{j=1}^{i} \frac{U_j}{n+2-j} \end{aligned}$$

$$= (s+1-r) \sum_{i=1}^{r} \frac{U_i}{n+2-i} + \sum_{i=r+1}^{s} \frac{s+1-i}{n+2-i} U_i \ .$$

Before finding the density function of G_{rs} a few results pertaining to the coefficients of the coverages in the above linear combination are required.

Four types of coefficients can arise.

Type I	$a_i = 0$	$(i=0,s+1,\ldots,n)$
Type II	$a_i = \dfrac{s+1-r}{n+2-i}$	$(i=1,2,\ldots,r-1)$
Type III	$a_i = \dfrac{s+1-r}{n+2-r}$	$i=r$
Type IV	$a_i = \dfrac{s+1-i}{n+2-i}$	$(i=r+1,r+2,\ldots,s)$

For the application of formula (2.3.3) it is required to know what, if any, coefficients are identical. In this regard it can easily be shown that:

(1) Type I coefficients constitute a set of $n-s+1$ equal coefficients and type II, III, and IV coefficients are distinct from type I.

(2) Type II coefficients are distinct from one another and the magnitude of these coefficients varies directly with the subscript $i \in (1,2,\ldots,r-1)$.

(3) Type IV coefficients are distinct from one another and their magnitude varies indirectly with the subscript $i \in (r+1,r+2,\ldots,s)$.

(4) The type III coefficient is distinct from all other coefficients and is the largest coefficient.

(5) The possibility exists that type II and IV coefficients are identical. More about this fact will be considered in Lemma 2.13 .

To derive the general form of the density function of G_{rs} assume that only p+1 of the coefficients are distinct. Denote these distinct coefficients by $b_0, b_1, \ldots, b_p$ where

$$b_0 = a_i \qquad (i=0,s+1,\ldots,n)$$

$$b_i = a_i \qquad (i=1,2,\ldots,r)$$

$$b_i = a_{p_i} \qquad (i=r+1,r+2,\ldots,p)$$

and a_{p_i} $(i=r+1,r+2,\ldots,p)$ are the type IV coefficients which are distinct from the type II coefficients. Let n_i be the number of actual coefficients identical with b_i $(i=0,1,\ldots,p)$. It is known that

$$n_0 = n-s+1$$

$$n_i = 1 \text{ or } 2 \qquad (i=1,2,\ldots,r-1)$$

$$n_i = 1 \qquad (i=r,r+1,\ldots,p) .$$

Finally, define $m_i = \max(r,i)$ and let $D_i(g_{rs})$ be the smallest positive integer w such that

$$\frac{s+1-r}{n+2-w} \geq g_{rs}$$

and

$D_2(g_{rs})$ be the largest positive integer w such that

$$\frac{s+1-w}{n+2-w} \geq g_{rs} \; .$$

<u>Theorem 2.12</u> The density function of

$$G_{rs} = \sum_{i=r}^{s} U_{(i)}$$

is given by

$$f(g_{rs}) = n \sum_{i=D_1(g_{rs})}^{D_2(g_{rs})} \frac{\left(1-\frac{g_{rs}}{b_i}\right)^{n-1} J_i}{b_i^{2-s} \prod_{\substack{j=1 \\ j \neq i}}^{p} (b_i-b_j)^{n_j}} \tag{2.4.2}$$

where

$$J_i = (1-n)^{n_i-1} (g_{rs}-b_i)^{1-n_i} + (1-n_i) \sum_{\substack{w=0 \\ w \neq i}}^{p} n_w (b_i-b_w)^{-1}$$

$$b_0 = 0$$

$$b_i = \frac{s+1-m_i}{n+2-i} \qquad (i=1,2,\ldots,p) \; .$$

Proof:

By Corollary 2.11 the density function of G_{rs} is

$$f(g_{rs}) = (-1)^n \, n \prod_{k=0}^{p} \frac{1}{(n_k-1)!} \frac{\partial^{n-p}}{\partial b_0^{n_0-1} \, \partial b_1^{n_1-1} \cdots \partial b_p^{n_p-1}} \cdot \sum_{i=1}^{p} \frac{I(b_i)\,(g_{rs}-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j \neq i}}^{p} (b_i-b_j)}$$

$$= \frac{(-1)^n \, n}{(n-s)!} \frac{\partial^{n-p}}{\partial b_0^{n-s} \, \partial b_1^{n_1-1} \cdots \partial b_{r-1}^{n_{r-1}-1}} \cdot \sum_{i=1}^{p} \frac{I(b_i)\,(g_{rs}-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j \neq i}}^{p} (b_i-b_j)} \, . \qquad (2.4.3)$$

For any $i \in (1,2,\ldots,p)$ there is a possible term

$$\frac{(-1)^n n}{(n-s)!} \frac{\partial^{n-p}}{\partial_{b_0}^{n-s} \partial_{b_1}^{n_1-1} \cdots \partial_{b_{r-1}}^{n_{r-1}-1}} \frac{(g_{rs}-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$= \frac{(-1)^n n}{(n-s)!} \frac{\partial^{s-p}}{\partial_{b_1}^{n_1-1} \cdots \partial_{b_{r-1}}^{n_{r-1}-1}} \frac{(n-s)!\,(g_{rs}-b_i)^{n-1}}{(b_i-b_0)^{n-s+1} \prod_{\substack{j=1 \\ j\neq i}}^{p} (b_i-b_j)}$$

$$= (-1)^n n \frac{\partial^{n_i-1}}{\partial_{b_i}^{n_i-1}} \frac{(g_{rs}-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)^{n_j}} \quad . \qquad (2.4.4)$$

If $n_i = 2$ the above line is equal to

$$(-1)^n n \left[\frac{-(n-1)(g_{rs}-b_i)^{n-2}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)^{n_j}} - (g_{rs}-b_i)^{n-1} \sum_{\substack{w=0 \\ w\neq i}}^{p} \frac{n_w(b_i-b_w)^{-(n_w+1)}}{\prod_{\substack{j=0 \\ j\neq i,w}}^{p} (b_i-b_j)^{n_j}} \right]$$

$$= (-1)^n n \left[\frac{-(n-1)(g_{rs}-b_i)^{n-2} - (g_{rs}-b_i)^{n-1} \sum_{\substack{w=0 \\ w\neq i}}^{p} n_w (b_i-b_w)^{-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)^{n_j}} \right]$$

$$= (-1)^n n \left[(1-n)^{n_i-1} (g_{rs}-b_i)^{n-n_i} + (1-n_i)(g_{rs}-b_i)^{n-1} \sum_{\substack{w=0 \\ w\neq i}}^{p} n_w (b_i-b_w)^{-1} \right] \frac{1}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)^{n_j}}$$

$$= (-1)^n n \frac{(g_{rs}-b_i)^{n-1} \quad J_i}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)^{n_j}}$$

On the other hand if $n_i = 1$ then (2.4.4) becomes

$$(-1)^n n \frac{(g_{rs}-b_i)^{n-1}}{\prod_{\substack{j=0 \\ j\neq i}}^{p} (b_i-b_j)^{n_j}}$$

$$= (-1)^{n}\, n \;\frac{(g_{rs}-b_i)^{n-1}\; J_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p} (b_i-b_j)^{n_j}}$$

since $J_i = 1$ for $n_i = 1$.

Thus (2.4.3) becomes

$$f(g_{rs}) = (-1)^{n}\, n \sum_{i=1}^{p} \frac{I(b_i)\; (g_{rs}-b_i)^{n-1}\; J_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p} (b_i-b_j)^{n_j}}$$

$$= (-1)\, n \sum_{i=1}^{p} \frac{I(b_i)\left(1-\frac{g_{rs}}{b_i}\right)^{n-1} J_i}{b_i^{2-s} \prod\limits_{\substack{j=1\\ j\neq i}}^{p} (b_i-b_j)^{n_j}} \;. \qquad (2.4.5)$$

Since each of the coefficients

$$b_i = \frac{s+1-m_i}{n+2-i} > 0 \qquad (i=1,2,\ldots,p)$$

it follows that

$$I(b_i) = \begin{cases} -1 & \text{if } g_{rs} \in (0,b_i] \\ 0 & \text{if } g_{rs} \notin (0,b_i] \end{cases}$$

Furthermore, since $b_1 < b_2 < \ldots < b_r$, if for a given g_{rs}

$g_{rs} \notin (0,b_i]$ for some $i \in (1,2,\ldots,r)$ then $g_{rs} \notin (0,b_w]$ for $w=1,2,\ldots,i-1$. Similarly, since $b_r > b_{r+1} > \ldots > b_p$, if $g_{rs} \notin (0,b_i]$ for some $i \in (r,r+1,\ldots,p)$ then $g_{rs} \notin (0,b_w]$ for $w=i+1,i+2,\ldots,p$. Thus it is only required in (2.4.5) to sum over type II coefficients b_i, $i \geq i_1$ where b_{i_1} is the smallest type II coefficient for which $g_{rs} \in (0,b_{i_1}]$ and over the distinct type IV coefficients b_i, $i \leq i_2$ where b_{i_2} is the smallest distinct type IV coefficient for which $g_{rs} \in (0,b_{i_2}]$. Defining $D_1(g_{rs})$ and $D_2(g_{rs})$ as given in the statement of the theorem the result follows. ▮

In Barton and David special cases are considered in which formula (2.4.1) reduces to a relatively simpler form. In the following, two special cases are considered. In addition to stating the formula obtained by Barton and David for each case, the implications and results using the approach of this paper are given.

The first such case considered assumes that

$$(h+1)(n+1-s) \neq k(s-r-h)$$

for all $k \in (0,1,\ldots,r-1)$ and $h \in (0,1,\ldots,s-r-1)$. Formula (2.4.1), which involves a double sum, then reduces to a formula just involving two single sums. That is

$$f(g_{rs}) = K \sum_{h=0}^{s-r-1} (s-r-h)^{s-1} \binom{s-r}{h} (-1)^{r-h} A_h \left[H(1-g_{rs}\frac{n+1-r-h}{s-r-h}) \right]^{n-1}$$

$$+ K' \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} B_k \left[H(1-g_{rs}\frac{n-r+k+2}{s-r+1}) \right]^{n-1} \qquad (2.4.6)$$

where

$$K = \frac{s!\, n \binom{n+1}{s}}{(s-r)!\, (n+1-s)^{s-r-1}}$$

$$K' = \frac{s!\, n \binom{n+1}{s} (s-r+1)^{s-r-1}}{(r-1)!}$$

$$A_h^{-1} = \prod_{t=0}^{r-1} \left[(h+1)(n+1-s) - t(s-r-h) \right]$$

$$B_k^{-1} = \prod_{t=0}^{s-r-1} \left[(s-r+1)(n+1-s) - (t+1)(n+1-s+k) \right] .$$

The following lemma gives the implications of the above condition for this special case in our approach.

<u>Lemma 2.13</u> If $(h+1)(n+1-s) \neq k(s-r-h)$ for all $k \in (0,1,\ldots,r-1)$ and $h \in (0,1,\ldots,s-r-1)$ then no type II and IV coefficients are equal.

Proof:

Suppose a type II coefficient $\frac{s+1-r}{n+2-i}$, for some $i \in (1,2,\ldots,r-1)$, and a type IV coefficient $\frac{s+1-j}{n+2-j}$, for some $j \in (r+1,r+2,\ldots,s)$, are equal. Then

$$\frac{s+1-r}{n+2-i} = \frac{s+1-j}{n+2-j}$$

which is equivalent to

$$(s+1-r)(n+2-j) = (s+1-j)(n+2-i) .$$

That is

$$i = \frac{(s+1-j)(n+2) - (s+1-r)(n+2-j)}{s+1-j}$$

$$= \frac{r(n+2) - j(n+1-s+r)}{s+1-j} .$$

Defining $j=r+w$, $w \in (1,2,\ldots,s-r)$

$$i = \frac{r(n+2) - (r+w)(n+1-s+r)}{s+1-r-w}$$

$$= r - \frac{w(n+1-s)}{s+1-r-w} . \qquad (2.4.7)$$

Thus for the type II coefficient $\frac{s+1-r}{n+2-i}$ and the type IV coefficient $\frac{s+1-j}{n+2-j}$, where $j = r+w$, to be equal it is required that

$$\frac{w(n+1-s)}{s+1-r-w}$$

be an integer belonging to $(1,2,\dots,r-1)$.

Since $w = h+1 \in (1,2,\dots,s-r)$,

$$(h+1)(n+1-s) \neq k(s-r-h)$$

is equivalent to

$$k \neq \frac{w(n+1-s)}{s+1-r-w} .$$

Thus if $(h+1)(n+1-s) \neq k(s-r-h)$ for all $k \in (0,1,\dots,r-1)$ and $h \in (0,1,\dots,s-r-1)$ it follows, in particular, that $\frac{w(n+1-s)}{s+1-r-w}$ is not an integer in $(1,2,\dots,r-1)$ and the necessary condition (2.4.7) for a type II and IV coefficient to be equal is not satisfied. ▮

Using Theorem 2.12 the following corollary is obtained.

<u>Corollary 2.14</u> If $(h+1)(n+1-s) \neq k(s-r-h)$ for all $k \in (0,1,\dots,r-1)$ and $h \in (0,1,\dots,s-r-1)$ then

$$f(g_{rs}) = n\ s! \binom{n+1}{s} \sum_{i=D_1(g_{rs})}^{D_2(g_{rs})} \frac{(s+1-m_i)^{s-2}\left(1-g_{rs}\dfrac{n+2-i}{s+1-m_i}\right)^{n-1}}{\prod\limits_{\substack{j=1\\ j\neq i}}^{s}\left[(i-j)(s+1)+m_j(n+2-i)-m_i(n+2-j)\right]} \qquad (2.4.8)$$

<u>Proof:</u>

By Lemma 2.13 it follows that no type II and IV coefficients are equal. Thus $p = s$ and $n_i = 1$ $(i=1,2,\ldots,s)$. Substituting these values in (2.4.2) it follows that

$f(g_{rs})$

$$= n \sum_{i=D_1(g_{rs})}^{D_2(g_{rs})} \frac{\left(1-g_{rs}\,\frac{n+2-i}{s+1-m_i}\right)^{n-1}}{\left(\frac{s+1-m_i}{n+2-i}\right)^{2-s} \prod_{\substack{j=1\\ j\neq i}}^{s} \left(\frac{s+1-m_i}{n+2-i} - \frac{s+1-m_j}{n+2-j}\right)}$$

$$= n \sum_{i=D_1(g_{rs})}^{D_2(g_{rs})} \frac{\left(1-g_{rs}\,\frac{n+2-i}{s+1-m_i}\right)^{n-1} (n+2-i) \prod_{\substack{j=1\\ j\neq i}}^{s} (n+2-j)}{(s+1-m_i)^{2-s} \prod_{\substack{j=1\\ j\neq i}}^{s} \left[(i-j)(s+1)+m_j(n+2-i)-m_i(n+2-j)\right]}$$

$$= \frac{n(n+1)!}{(n+1-s)!} \sum_{i=D_1(g_{rs})}^{D_2(g_{rs})} \frac{(s+1-m_i)^{s-2} \left(1-g_{rs}\,\frac{n+2-i}{s+1-m_i}\right)^{n-1}}{\prod_{\substack{j=1\\ j\neq i}}^{s} \left[(i-j)(s+1)+m_j(n+2-i)-m_i(n+2-j)\right]} .$$

Formula (2.4.8) then follows from the above line by simply noting that

$$\frac{(n+1)!}{(n+1-s)!} = s! \binom{n+1}{s} . \blacksquare$$

It can be shown that Barton and David's formula (2.4.6) and formula (2.4.8) are equivalent.

The second special case is based on the assumption that r = 1. In such a case Barton and David state that formula (2.4.1) simplifies to

$$f(g_{1s}) = \frac{n\binom{n+1}{s}}{(n+1-s)^{s-1}} \sum_{k=1}^{s} \binom{s}{k} (-1)^{k-1} k^{s-1} \left[H\left(1-g_{1s} \frac{n+1-s+k}{k}\right) \right]^{n-1} . \tag{2.4.9}$$

In our approach if r = 1 then there are no type II coefficients and the following corollary is obtained.

<u>Corollary 2.15</u> The density function of G_{1s} is given by

$$f(g_{1s}) = \frac{n\binom{n+1}{s}}{(n+1-s)^{s-1}} \sum_{i=1}^{D_2(g_{1s})} \binom{s}{i} i (-1)^{i-1} (s+1-i)^{s-2} \cdot \left(1-g_{1s} \frac{n+2-i}{s+1-i}\right)^{n-1} \tag{2.4.10}$$

<u>Proof</u>:

Since r = 1 there are no type II coefficients and thus all coefficients, except for type I, are distinct. Therefore p = s and $n_i = 1$ (i=1,2,...,s). Furthermore $m_i = \max(r,i) = i$ (i=1,2,...,s).

Using Theorem 2.12, with these values, it follows that

$$f(g_{1s}) = n \sum_{i=1}^{D_2(g_{1s})} \frac{\left(1-g_{1s}\,\frac{n+2-i}{s+1-i}\right)^{n-1}}{\left(\frac{s+1-i}{n+2-i}\right)^{2-s} \prod\limits_{\substack{j=1\\ j\neq i}}^{s} \left(\frac{s+1-i}{n+2-i} - \frac{s+1-j}{n+2-j}\right)}$$

$$= n \sum_{i=1}^{D_2(g_{1s})} \frac{\left(1-g_{1s}\,\frac{n+2-i}{s+1-i}\right)^{n-1}}{\left(\frac{s+1-i}{n+2-i}\right)^{2-s} \prod\limits_{\substack{j=1\\ j\neq i}}^{s} \left[\frac{(j-i)\ (n+1-s)}{(n+2-i)(n+2-j)}\right]}$$

$$= n \sum_{i=1}^{D_2(g_{1s})} \frac{\left(1-g_{1s}\,\frac{n+2-i}{s+1-i}\right)^{n-1} (n+2-i) \prod\limits_{\substack{j=1\\ j\neq i}}^{s} (n+2-j)}{(s+1-i)^{2-s}\ (n+1-s)^{s-1} \prod\limits_{\substack{j=1\\ j\neq i}}^{s} (j-i)}$$

$$= \frac{n\ (n+1)!}{(n+1-s)^{s-1}(n+1-s)!} \sum_{i=1}^{D_2(g_{1s})} \frac{\left(1-g_{1s}\,\frac{n+2-i}{s+1-i}\right)^{n-1}}{(i-1)!\ (s-i)!}$$

$$\cdot\ (-1)^{i-1}\ (s+1-i)^{s-2}$$

Formula (2.4.10) follows from the above line by noting that

$$\frac{(n+1)!}{(n+1-s)!\ (i-1)!\ (s-i)!} = i \binom{n+1}{s} \binom{s}{i} \quad .$$

Although formulae (2.4.9) and (2.4.10) should be equivalent, it can be shown that they differ by a factor $(-1)^{s+1}$. However deriving the density function for $r = 1$ from Barton and David's general formula (2.4.1) one obtains the expression equivalent to (2.4.10). Thus formula (2.4.10) is assumed to be the correct one.

Chapter III : The joint distribution of two linear combinations of uniform coverages

3.1 : Introduction

The problem of finding the joint density function for two linear combinations

$$Y_1 = a_{11}U_1 + a_{12}U_2 + \dots + a_{1n}U_n$$

$$Y_2 = a_{21}U_1 + a_{22}U_2 + \dots + a_{2n}U_n$$

of the uniform coverages, based on the same random sample from the uniform distribution on the interval (0,1), is now investigated in this chapter.

Previously such a density function has been derived but varying degrees of restrictions on the coefficients

$$a_{ij} \qquad (j=1,2,\dots,n \; ; \; i=1,2)$$

were required.

Ali and Mead (1969) considered the joint distribution for several linear combinations of the uniform coverages. In particular for the case of these two linear combinations they found, defining $a_{10} = a_{20} = 0$, that the joint density function was given by

$$f(y_1,y_2) = \frac{n!}{(n-2)!} \sum_{i=0}^{n} \sum_{\substack{j=0\\ j\neq i}}^{n} \lambda(y_1-a_{1i})\lambda\left[\frac{\begin{vmatrix} 1 & a_{1i} & a_{2i} \\ 1 & a_{1j} & a_{2j} \\ 1 & y_1 & y_2 \end{vmatrix}}{\begin{vmatrix} 1 & a_{1i} \\ 1 & a_{1j} \end{vmatrix}}\right]$$

$$\cdot \frac{\begin{vmatrix} 1 & a_{1i} & a_{2i} \\ 1 & a_{1j} & a_{2j} \\ 1 & y_1 & y_2 \end{vmatrix}^{n-2}}{\prod\limits_{\substack{k=0\\ k\neq i,j}}^{n} \begin{vmatrix} 1 & a_{1i} & a_{2i} \\ 1 & a_{1j} & a_{2j} \\ 1 & a_{1k} & a_{2k} \end{vmatrix}} \qquad (3.1.1)$$

provided

$$\begin{vmatrix} 1 & a_{1i} & a_{2i} \\ 1 & a_{1j} & a_{2j} \\ 1 & a_{1k} & a_{2k} \end{vmatrix} \neq 0 \qquad (3.1.2)$$

and

$$\begin{vmatrix} 1 & a_{1i} \\ 1 & a_{1j} \end{vmatrix} \neq 0 \qquad (3.1.3)$$

for all distinct i, j, k. The function λ is defined by

$$\lambda(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases} .$$

They obtained this result by initially finding the characteristic function and then inverting it in the usual fashion to obtain the density function.

The restrictions (3.1.2) and (3.1.3) on the coefficients were required to avoid indeterminate forms in the density function given by formula (3.1.1). Considering the coefficients for corresponding coverages in the two linear combinations

$$Y_1 = a_{11}U_1 + a_{12}U_2 + \dots + a_{1n}U_n$$
$$Y_2 = a_{21}U_1 + a_{22}U_2 + \dots + a_{2n}U_n$$

as coordinates of a point in the two-dimensional Euclidean space, the abscissa and ordinate of which are from the first and second linear combination respectively, the restriction (3.1.2) is equivalent to the geometric restiction that no points are coincident and no three points are colinear and the restriction (3.1.3) is equivalent to the geometric restriction that no two points are vertical.

Mead (1969), using a geometrical approach to replace in (3.1.1) the function λ defined above, reduced the number of sums in the joint density function (3.1.1) and obtained the equivalent formula

$$f(y_1,y_2) = \frac{n!}{(n-2)!} \sum_{i=2}^{n} \sum_{j=1}^{i-1} \lambda_{ij} \operatorname{sgn} \begin{vmatrix} a_{1i} & a_{2i} \\ a_{1j} & a_{2j} \end{vmatrix} \frac{\begin{vmatrix} 1 & a_{1i} & a_{2i} \\ 1 & a_{1j} & a_{2j} \\ 1 & y_1 & y_2 \end{vmatrix}^{n-2}}{\prod\limits_{\substack{k=0 \\ k \neq i,j}}^{n} \begin{vmatrix} 1 & a_{1i} & a_{2i} \\ 1 & a_{1j} & a_{2j} \\ 1 & a_{1k} & a_{2k} \end{vmatrix}} \tag{3.1.4}$$

where

$$\lambda_{ij} = \begin{cases} 1 & \text{if } (y_1,y_2) \text{ lies in the triangle with vertices } (0,0),\ (a_{1i},a_{2i}), \text{ and } (a_{1j},a_{2j}) \\ 0 & \text{otherwise} \end{cases}$$

and $(a_{10},a_{20}) = (0,0)$.

The first attempt to remove the restrictions on the coefficients was undertaken by S. W. Lim (1972). She was primarily concerned with the removal of the restriction that there be no coincident points. She adjusted the coincident points in such a way that they were no longer coincident and thus for this modified problem formula (3.1.4) could be used. She then considered the limit as the adjusted points approached the original ones. The resulting density function was

$$f(y_1,y_2) = \frac{n!}{(n-2)!} \sum_{i=3}^{p} \sum_{j=2}^{i-1} \lambda_{k_i k_j} \operatorname{sgn} \begin{vmatrix} a_{1k_i} & a_{2k_i} \\ a_{1k_j} & a_{2k_j} \end{vmatrix}$$

$$\cdot \frac{1}{\prod_{i=1}^{p} (n_i-1)! \begin{vmatrix} 1 & a_{2k_i} \\ 1 & a_{2k_j} \end{vmatrix}^{(n+1)-p}} \frac{\partial^{(n+1)-p}}{\partial^{n_1-1}_{a_{1k_1}} \cdots \partial^{n_p-1}_{a_{1k_p}}}$$

$$\cdot \frac{\begin{vmatrix} 1 & a_{1k_i} & a_{2k_i} \\ 1 & a_{1k_j} & a_{2k_j} \\ 1 & y_1 & y_2 \end{vmatrix}^{n-2}}{\prod_{\substack{h=1 \\ h\neq i,j}}^{p} \begin{vmatrix} 1 & a_{1k_i} & a_{2k_i} \\ 1 & a_{1k_j} & a_{1k_j} \\ 1 & a_{1k_h} & a_{2k_h} \end{vmatrix}}$$

where $(n+1)$ is the total number of points (a_{1i},a_{2i}) $(i=0,1,\ldots,n)$, p is the total number of distinct points (a_{1k_i},a_{2k_i}) $(i=1,2,\ldots,p)$, and n_i is the number of actual points coinciding at the ith distinct point $(i=1,2,\ldots,p)$.

This formula was derived under the assumption that no three points were colinear. However she also noted that with modifications it could be used for certain cases involving colinear points.

In this chapter the joint density function of Y_1 and Y_2 is found with no restrictions whatsoever. The method employed is basically the same as that of Ali and Mead (1969) except now the more general characteristic function, derived in the next section, is used.

3.2 : The characteristic function of two linear combinations

As a consequence of Corollary 2.10 the general characteristic function for two linear combinations of the uniform coverages can be obtained.

Consider the linear combinations

$$Y_1 = a_{11}U_1 + a_{12}U_2 + \dots + a_{1n}U_n$$

$$Y_2 = a_{21}U_1 + a_{22}U_2 + \dots + a_{2n}U_n$$

and the n+1 sums

$$a_{1i}t_1 + a_{2i}t_2 \qquad (i=0,1,\dots,n)$$

where $a_{10} = a_{20} = 0$ and t_1 and t_2 are the arguments of the characteristic function of Y_1 and Y_2.

Suppose there are only p+1 distinct sums, denoted by, $b_{1i}t_1 + b_{2i}t_2$ $(i=0,1,\dots,p)$ and that n_i of the actual sums are equal to $b_{1i}t_1 + b_{2i}t_2$ $(i=0,1,\dots,p)$.

Lemma 3.1 The joint characteristic function of Y_1 and Y_2 is

$$\phi(t_1,t_2) = \frac{n!}{i^n} \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_{10}t_1+b_{20}t_2}^{n_0-1} \cdots \partial_{b_{1p}t_1+b_{2p}t_2}^{n_p-1}}$$

$$\cdot \sum_{i=0}^{p} \frac{e^{i(b_{1i}t_1+b_{2i}t_2)}}{\prod_{\substack{j=0\\ j\neq i}}^{p} \left[(b_{1i}-b_{1j})t_1 + (b_{2i}-b_{2j})t_2\right]} .$$

Proof:

Letting 'E' denote the expectation operator,

$$\phi(t_1,t_2) = E\left[e^{i(t_1Y_1+t_2Y_2)}\right]$$

$$= E\left[e^{i[t_1(a_{11}U_1+\cdots+a_{1n}U_n) + t_2(a_{21}U_1+\cdots+a_{2n}U_n)]}\right]$$

$$= E\left[e^{i[U_1(a_{11}t_1+a_{21}t_2) + \cdots + U_n(a_{1n}t_1+a_{2n}t_2)]}\right]$$

$$= \frac{n!}{i^n} \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial_{b_{10}t_1+b_{20}t_2}^{n_0-1} \cdots \partial_{b_{1p}t_1+b_{2p}t_2}^{n_p-1}}$$

$$\cdot \sum_{i=0}^{p} \frac{e^{i(b_{1i}t_1+b_{2i}t_2)}}{\prod_{\substack{j=0\\ j\neq i}}^{p} \left[(b_{1i}t_1+b_{2i}t_2) - (b_{1j}t_1+b_{2j}t_2)\right]} .$$

This line follows from Corollary 2.10 which gives the characteristic function for a single linear combination of the uniform coverages. The form of the characteristic function given in the lemma is obtained by a simple rearrangement of the terms in the denominator of the last line. ∎

The proof of the following lemma is identical to the one in the paper by Ali and Mead (1969).

<u>Lemma 3.2</u> The expression for $\phi(t_1,t_2)$ in Lemma 3.1 is analytic everywhere in the complex two-dimensional space.

As a result of this lemma, if the Inversion Formula for characteristic functions is used to find the density function, the integrand of such a Formula will also be analytic everywhere in the complex two-dimensional space. Thus the density function of Y_1 and Y_2 can be evaluated by successively integrating with respect to t_1 and t_2 using partial fraction decompositions and distorting finite portions of the contour of integration in such a way so as to avoid the singularities that occur if the terms of the integrand are considered individually. Since the integrand is analytic in t_2 after the first integration the contours can again be distorted.

3.3 The joint density function of two linear combinations

The joint density function of Y_1 and Y_2, denoted by $f(y_1,y_2)$, is obtained by basically the same method used by

Ali and Mead. That is, the density function is now found by inverting the general characteristic function obtained in the previous section. To evaluate the resulting integrals the following results, obtained from the theory of residues, are required. For constants 'a' and 'b', with 'a' real,

$$\int_{\Gamma_1} \frac{e^{-iaz}}{z^n}\, dz = \begin{cases} \dfrac{(-1)^n\, 2\pi i^n\, a^{n-1}}{(n-1)!} & \text{if } a \geq 0 \\ 0 & \text{if } a < 0 \end{cases}$$

where the contour Γ_1 consists of the real axis, except for a semi-circle above the real axis from $-c$ to c $(c > 0)$ and

$$\int_{\Gamma_2} \frac{e^{-iaz}}{z+b}\, dz = \begin{cases} -2\pi i\, e^{iab} & \text{if } a \geq 0 \\ 0 & \text{if } a < 0 \end{cases}$$

where Γ_2 is similar to Γ_1 , except now the semi-circle is such that $-b$ lies below it.

Let

(1) $n+1$ be the total number of points (a_{1i}, a_{2i})

$(i=0,1,\ldots,n)$.

(2) $p+1$ be the total number of distinct points (b_{1i}, b_{2i})

$(i=0,1,\ldots,p)$

(3) n_i be the number of actual points coinciding at the

i th distinct point $(i=0,1,\ldots,p)$.

Define for each distinct point (b_{1i}, b_{2i}) $(i=0,1,\ldots,p)$ the set

$$S_i = (i_1, i_2, \ldots, i_{p_i} ; p_{ii_1}, p_{ii_2}, \ldots, p_{ii_{p_i}} ; p_i) .$$

This is a set of p_i elements each of which is a subscript corresponding to a distinct point which is colinear with (b_{1i}, b_{2i}) and is taken as a representative of all other points colinear with (b_{1i}, b_{2i}) and the point chosen. That is, all possible sets colinear with (b_{1i}, b_{2i}) are determined and from each set one point is chosen to represent the other points in the colinear set. The numbers $p_{ii_1}, p_{ii_2}, \ldots, p_{ii_{p_i}}$ are the number of points in the set which is colinear with (b_{1i}, b_{2i}) and is represented by the point corresponding to the subscript $i_1, i_2, \ldots, i_{p_i}$ respectively.

Hence

(1) i_j is the subscript corresponding to the representative point for the colinear set j $(j=1,2,\ldots,p_i)$

(2) p_i is the number of subscripts in S_i, that is, the number of distinct sets colinear with (b_{1i}, b_{2i})

(3) p_{ii_j} is the number of points represented by the point corresponding to the subscript i_j $(j=1,2,\ldots,p_i)$.

Furthermore, for $i=0,1,\ldots,p$, define:

(1) V_i to be the set of points having the same abscissa

as the point (b_{1i}, b_{2i});

(2) S_i' to be the set S_i with the subscript corresponding to the representative point of the vertical colinear set omitted.

(3) p_i' to be the number of elements in the set S_i'.

Notice that if there are no vertical points then S_i' is identical with S_i and $p_i' = p_i$.

Finally, for convenience, define

$$(i\ j\ k) = \begin{vmatrix} 1 & b_{1i} & b_{2i} \\ 1 & b_{1j} & b_{2j} \\ 1 & b_{1k} & b_{2k} \end{vmatrix}, \quad (i\ j\ y) = \begin{vmatrix} 1 & b_{1i} & b_{2i} \\ 1 & b_{1j} & b_{2j} \\ 1 & y_1 & y_2 \end{vmatrix},$$

$$K_{np} = \prod_{s=0}^{p} \frac{1}{(n_s-1)!} \frac{\partial^{n-p}}{\partial^{n_0-1} b_{20}\, \partial^{n_1-1} b_{21} \cdots \partial^{n_p-1} b_{2p}},$$

$$D_i = \frac{\prod_{w \in S_i'} (b_{1i}-b_{1w})^{p_{1w}}}{\prod_{\substack{r=0 \\ r\neq i \\ r \notin V_i}}^{p} (b_{1i}-b_{1r}) \prod_{h \in V_i} (b_{2i}-b_{2h})} \cdot \prod_{k \in S_i'} \frac{1}{(p_{1k}-1)!} \frac{\partial^{p_{1k}-1}}{\partial^{p_{1k}-1} b_{2k}}.$$

where $\prod_{k\in S_i'} \frac{\partial^{p_{ik}-1}}{\partial^{p_{ik}-1} b_{2k}}$ represents the appropriate partial derivatives omitting the partial derivative with respect to the ordinate of the vertical representative point.

In the proof of Theorem 3.4, in which the joint density function of Y_1 and Y_2 is obtained, the joint characteristic function of Y_1 and Y_2 will be required. However before proceeding to this theorem a more convenient form of the characteristic function, given in Lemma 3.1, is now derived.

Lemma 3.3 The joint characteristic function of

$$Y_1 = a_{11}U_1 + a_{12}U_2 + \dots + a_{1n}U_n$$

$$Y_2 = a_{21}U_1 + a_{22}U_2 + \dots + a_{2n}U_n$$

is given by

$$\phi(t_1,t_2) = \frac{n!}{i^n} K_{np} \sum_{i=0}^{p} \frac{D_i}{t_2^{n-1}} \sum_{j\in S_i'} \frac{e^{i(b_{1i}t_1+b_{2i}t_2)} (b_{1i}-b_{1j})^{p_i'-2}}{\prod_{\substack{g\in S_i' \\ g\neq j}} (i\ j\ g)} \cdot \frac{1}{t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2} .$$

Proof:

Using Lemma 3.1 the characteristic function for Y_1 and Y_2 is

$$\frac{n!}{i^n}\prod_{s=0}^{p}\frac{1}{(n_s-1)!}\frac{\partial^{n-p}}{\partial^{n_0-1}_{b_{10}t_1+b_{20}t_2}\cdots\partial^{n_p-1}_{b_{1p}t_1+b_{2p}t_2}}$$

$$\cdot\sum_{i=0}^{p}\frac{e^{i(b_{1i}t_1+b_{2i}t_2)}}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p}\left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]}$$

$$=\frac{n!}{i^n}\prod_{s=0}^{p}\frac{1}{(n_s-1)!\,t_2^{n_s-1}}\frac{\partial^{n-p}}{\partial^{n_0-1}_{b_{20}}\,\partial^{n_1-1}_{b_{21}}\cdots\partial^{n_p-1}_{b_{2p}}}$$

$$\cdot\sum_{i=0}^{p}\frac{e^{i(b_{1i}t_1+b_{2i}t_2)}}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p}\left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]}$$

$$= \frac{n!}{t_1^{p}\, t_2^{n-p}} K_{np} \sum_{i=0}^{p} \frac{e^{i(b_{1i}t_1+b_{21}t_2)}}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p} \left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]} . \qquad (3.3.1)$$

For $i=0,1,\ldots,p$ the product

$$\frac{1}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p} \left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]}$$

can be written

$$\frac{1}{\prod\limits_{\substack{j=0\\ j\neq i\\ j\notin V_i}}^{p} \left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right] \prod\limits_{h\in V_i} \left[(b_{1i}-b_{1h})t_1+(b_{2i}-b_{2h})t_2\right]}$$

$$= \frac{1}{\prod\limits_{\substack{j=0\\ j\neq i\\ j\notin V_i}}^{p} \left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right] \prod\limits_{h\in V_i} \left[(b_{2i}-b_{2h})t_2\right]}$$

$$= \frac{1}{\prod\limits_{\substack{r=0\\ r\neq i\\ r\notin V_i}}^{p} (b_{1i}-b_{1r}) \prod\limits_{\substack{j=0\\ j\neq i\\ j\notin V_i}}^{p} \left[t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2\right] \prod\limits_{h\in V_i} \left[(b_{2i}-b_{2h})t_2\right]}$$

$$= \frac{1}{\prod_{\substack{r=0\\ r\neq i\\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod_{j\in S_i'} \left[t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2\right]^{p_{ij}} \prod_{h\in V_1}\left[(b_{2i}-b_{2h})t_2\right]}$$

$$= \frac{1}{\prod_{\substack{r=0\\ r\neq i\\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod_{j\in S_i'} (t_1-b_{ij}')^{p_{ij}} \prod_{h\in V_1}\left[(b_{2i}-b_{2h})t_2\right]}$$

where $b_{ij}' = -\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}t_2$,

$$= \frac{1}{\prod_{\substack{r=0\\ r\neq i\\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod_{h\in V_1}\left[(b_{2i}-b_{2h})t_2\right]} \prod_{j\in S_i'} \frac{1}{(p_{ij}-1)!} \frac{\partial^{p_{ij}-1}}{\partial_{b_{ij}'}^{p_{ij}-1}}\left(\frac{1}{t_1-b_{ij}'}\right)$$

$$= \frac{1}{\prod_{\substack{r=0\\ r\neq i\\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod_{h\in V_1}\left[(b_{2i}-b_{2h})t_2\right]} \prod_{k\in S_i'} \frac{1}{(p_{ik}-1)!} \frac{\partial^{p_{ik}-1}}{\partial_{b_{ik}'}^{p_{ik}-1}}$$

$$\cdot \frac{1}{\prod_{j\in S_i'} (t_1-b_{ij}')}$$

$$= \frac{1}{\prod\limits_{\substack{r=0\\ r\neq 1\\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod\limits_{h\in V_1} \left[(b_{2i}-b_{2h})t_2\right]}$$

$$\cdot \prod_{k\in S_1'} \frac{1}{(p_{1k}-1)!} \frac{(b_{1i}-b_{1k})^{p_{1k}-1}}{t_2^{p_{1k}-1}} \frac{\partial^{p_{1k}-1}}{\partial^{p_{1k}-1} b_{2k}} \frac{1}{\prod\limits_{j\in S_1'} (t_1-b_{1j}')}$$

$$= \frac{\prod\limits_{w\in S_1'} (b_{1i}-b_{1w})^{p_{1w}-1}}{\prod\limits_{\substack{r=0\\ r\neq 1\\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod\limits_{h\in V_1} \left[(b_{2i}-b_{2h})t_2\right] t_2^{p-p_{1V_1}-p_1'}}$$

$$\cdot \prod_{k\in S_1'} \frac{1}{(p_{1k}-1)!} \frac{\partial^{p_{1k}-1}}{\partial^{p_{1k}-1} b_{2k}} \frac{1}{\prod\limits_{j\in S_1'} \left[t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right) t_2\right]}$$

where p_{1V_1} is the number of points vertical to (b_{1i}, b_{2i}),

$$= \frac{\prod_{w\in S_1'} (b_{1i}-b_{1w})^{p_{1w}-1} \prod_{w'\in S_1'} (b_{1i}-b_{1w'})}{\prod_{\substack{r=0 \\ r\neq 1 \\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod_{h\in V_1} (b_{2i}-b_{2h})\, t_2^{p-p_1'}} \prod_{k\in S_1'} \frac{1}{(p_{1k}-1)!} \frac{\partial^{p_{1k}-1}}{\partial^{p_{1k}-1} b_{2k}}$$

$$\cdot \frac{1}{\prod_{j\in S_1'} \left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]}$$

$$= \frac{\prod_{w\in S_1'} (b_{1i}-b_{1w})^{p_{1w}}}{\prod_{\substack{r=0 \\ r\neq 1 \\ r\notin V_1}}^{p} (b_{1i}-b_{1r}) \prod_{h\in V_1} (b_{2i}-b_{2h})\, t_2^{p-p_1'}} \prod_{k\in S_1'} \frac{1}{(p_{1k}-1)!} \frac{\partial^{p_{1k}-1}}{\partial^{p_{1k}-1} b_{2k}}$$

$$\cdot \frac{1}{\prod_{j\in S_1'} \left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]} \qquad (3.3.2)$$

Furthermore, since $\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix} \neq 0$ for $i=0,1,\ldots,p$, $j\in S_1'$, using the partial fraction expansion,

$$\frac{1}{\prod_{i=0}^{n}(x-g_i)} = \sum_{i=0}^{n} \frac{1}{\prod_{\substack{j=0\\ j\neq i}}^{n}(g_i-g_j)} \cdot \frac{1}{x-g_i}$$

where $g_0, g_1, \ldots, g_n$ are distinct, with t_1 as the variable

$$\frac{1}{\prod_{j\in S_i'}\left[(b_{1i}-b_{1j})t_1+(b_{2i}-b_{2j})t_2\right]}$$

$$= \frac{1}{\prod_{q\in S_i'}(b_{1i}-b_{1q}) \prod_{j\in S_i'}\left[t_1-\left(-\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2\right]}$$

$$= \frac{1}{\prod_{q\in S_i'}(b_{1i}-b_{1q})} \sum_{j\in S_i'} \frac{1}{\prod_{\substack{g\in S_i'\\ g\neq j}}\left[\left(-\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}+\frac{b_{2i}-b_{2g}}{b_{1i}-b_{1g}}\right)t_2\right]}$$

$$\cdot \frac{1}{t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2}$$

$$= \frac{1}{\prod\limits_{q \in S_i'} (b_{1i}-b_{1q})} \sum_{j \in S_i'} \frac{1}{\prod\limits_{\substack{g \in S_i' \\ g \neq j}} \left[\frac{(i\ j\ g)\ t_2}{(b_{1i}-b_{1j})\ (b_{1i}-b_{1g})} \right]} \cdot \frac{1}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}} \right) t_2}$$

$$= \frac{1}{\prod\limits_{q \in S_i'} (b_{1i}-b_{1q})} \sum_{j \in S_i'} \frac{(b_{1i}-b_{1j})^{p_i'-1} \prod\limits_{\substack{v \in S_i' \\ v \neq j}} (b_{1i}-b_{1v})}{\prod\limits_{\substack{g \in S_i' \\ g \neq j}} [(i\ j\ g)\ t_2]} \cdot \frac{1}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}} \right) t_2}$$

$$= \sum_{j \in S_i'} \frac{(b_{1i}-b_{1j})^{p_i'-2}}{\prod\limits_{\substack{g \in S_i' \\ g \neq j}} [(i\ j\ g)\ t_2]} \cdot \frac{1}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}} \right) t_2} .$$

Substituting this result in (3.3.2) it follows that

$$\frac{1}{\prod\limits_{\substack{j=0\\ j\neq i}}^{p}\left[(b_{1i}-b_{1j})t_1-(b_{2i}-b_{2j})t_2\right]}$$

$$=\frac{\prod\limits_{w\in S_i'}(b_{1i}-b_{1w})^{p_{iw}}}{\prod\limits_{\substack{r=0\\ r\neq i\\ r\notin V_i}}^{p}(b_{1i}-b_{1r})\prod\limits_{h\in V_i}(b_{2i}-b_{2h})\; t_2^{\,p-p_i'}}\prod_{k\in S_i'}\frac{1}{(p_{ik}-1)!}\frac{\partial^{p_{ik}-1}}{\partial^{p_{ik}-1} b_{2k}}$$

$$\cdot\sum_{j\in S_i'}\frac{(b_{1i}-b_{1j})^{p_i'-2}}{\prod\limits_{\substack{g\in S_i'\\ g\neq j}}\left[(i\ j\ g)\ t_2\right]}\cdot\frac{1}{t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2}$$

$$=\frac{\prod\limits_{w\in S_i'}(b_{1i}-b_{1w})^{p_{iw}}}{\prod\limits_{\substack{r=0\\ r\neq i\\ r\notin V_i}}^{p}(b_{1i}-b_{1r})\prod\limits_{h\in V_i}(b_{2i}-b_{2h})\; t_2^{\,p-1}}\prod_{k\in S_i'}\frac{1}{(p_{ik}-1)!}\frac{\partial^{p_{ik}-1}}{\partial^{p_{ik}-1} b_{2k}}$$

$$\cdot\sum_{j\in S_i'}\frac{(b_{1i}-b_{1j})^{p_i'-2}}{\prod\limits_{\substack{g\in S_i'\\ g\neq j}}(i\ j\ g\ )}\cdot\frac{1}{t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2}$$

$$= \frac{D_i}{t_2^{p-1}} \sum_{j \in S_i'} \frac{(b_{1i}-b_{1j})^{p_i'-2}}{\prod\limits_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g)} \cdot \frac{1}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right) t_2} .$$

Using this in (3.3.1) the characteristic function can be written

$$\frac{n!}{\underline{i}^n\, t_2^{n-p}} K_{np} \sum_{i=0}^{p} \frac{D_i}{t_2^{p-1}} \sum_{j \in S_i'} \frac{e^{\underline{i}(b_{1i}t_1+b_{2i}t_2)}\,(b_{1i}-b_{1j})^{p_i'-2}}{\prod\limits_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g)}$$

$$\cdot \frac{1}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right) t_2}$$

$$= \frac{n!}{\underline{i}^n} K_{np} \sum_{i=0}^{p} \frac{D_i}{t_2^{n-1}} \sum_{j \in S_i'} \frac{e^{\underline{i}(b_{1i}t_1+b_{2i}t_2)}\,(b_{1i}-b_{1j})^{p_i'-2}}{\prod\limits_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g)}$$

$$\cdot \frac{1}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right) t_2}$$

and the proof is complete. ∎

Define

$$\lambda_i = \begin{cases} 1 & \text{if } y-b_{1i} \geq 0 \\ 0 & \text{if } y-b_{1i} < 0 \end{cases}$$

and

$$\lambda_{ij} = \begin{cases} 1 & \text{if } \dfrac{(i\ j\ y)}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}} \geq 0 \\[2ex] 0 & \text{if } \dfrac{(i\ j\ y)}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}} < 0 \end{cases}$$

The following theorem gives the joint density function of Y_1 and Y_2 .

Theorem 3.4 The joint density function of

$$Y_1 = a_{11}U_1 + a_{12}U_2 + \dots + a_{1n}U_n$$

$$Y_2 = a_{21}U_1 + a_{22}U_2 + \dots + a_{2n}U_n$$

is given by

$$f(y_1,y_2) = \frac{n!}{(n-2)!} K_{np} \sum_{i=0}^{p} \lambda_i D_i \sum_{j \in S_i'} \lambda_{ij} \frac{(i\ j\ y)^{n-2}}{(b_{1i}-b_{1j})^{n-p_i'}} \cdot \left[\prod_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g) \right]^{-1} . \qquad (3.3.3)$$

Proof:

By Lemma 3.3 and the Inversion Formula for characteristic functions

$f(y_1,y_2)$

$$= \frac{n!}{(2\pi)^2 i^n} K_{np} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\sum_{i=0}^{p} \frac{D_i}{t_2^{n-1}} \sum_{j\in S_i'} \frac{e^{i(b_{1i}t_1+b_{2i}t_2)} (b_{1i}-b_{1j})^{p_i'-2}}{\prod_{\substack{g\in S_i' \\ g\neq j}} (i\ j\ g)} \cdot \frac{e^{-it_1y_1}\, e^{-it_2y_2}}{t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2} \right] dt_1\, dt_2$$

$$= \frac{n!}{(2\pi)^2 i^n} K_{np} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\sum_{i=0}^{p} \frac{D_i}{t_2^{n-1}} \sum_{j\in S_i'} \frac{e^{-it_1(y_1-b_{1i})-it_2(y_2-b_{2i})}}{\prod_{\substack{g\in S_i' \\ g\neq j}} (i\ j\ g)} \cdot \frac{(b_{1i}-b_{1j})^{p_i'-2}}{t_1+\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2} \right] dt_1\, dt_2 \quad .$$

Note that by Lemma 3.2 it follows that the above integrand is analytic everywhere in the complex two-dimensional space.

Holding t_2 fixed, distort the contour of integration for t_1 to avoid all singularities and integrate with respect to t_1 term by term using the formula

$$\int_{\Gamma} \frac{e^{-iaz}}{z+b}\, dz = \begin{cases} -2\pi i\, e^{iab} & \text{if } a \geq 0 \\ 0 & \text{if } a < 0 \end{cases}$$

for the contour integrals. Thus

$f(y_1, y_2)$

$$= \frac{n!}{(2\pi)^2 i^n} K_{np} \int_{-\infty}^{\infty} \left\{ \left[\sum_{i=0}^{p} \frac{D_i}{t_2^{n-1}} \sum_{j \in S_i} \frac{e^{-it_2(y_2-b_{2i})} (b_{1i}-b_{1j})^{p_i'-2}}{\prod_{\substack{g \in S_i' \\ g \neq j}} (i\, j\, g)} \right. \right.$$

$$\left. \cdot \int_{\Gamma} \left[\frac{e^{-it_1(y_1-b_{1i})}}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}} \right) t_2} \right] dt_1 \right] dt_2$$

and since

$$\int_{\Gamma} \frac{e^{-it_1(y_1-b_{1i})}}{t_1 + \left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}} \right) t_2}\, dt_1 = \begin{cases} -2\pi i\, e^{i(y_1-b_{1i})\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}} \right) t_2} & \text{if } y_1-b_{1i} \geq 0 \\ 0 & \text{if } y_1-b_{1i} < 0 \end{cases}$$

and

$$e^{-it_2(y_2-b_{2i})} \; e^{i(y_1-b_{1i})\left(\frac{b_{2i}-b_{2j}}{b_{1i}-b_{1j}}\right)t_2}$$

$$= e^{-it_2\left[(y_2-b_{2i})(b_{1j}-b_{1i}) + (y_1-b_{1i})(b_{2i}-b_{2j})\right] \frac{1}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}}}$$

$$= e^{-it_2\left[\frac{(i\ j\ y)}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}}\right]}$$

the density function can be written

$$f(y_1,y_2) = \frac{n!\ (-1)}{2\pi i}^{n-1} K_{np} \int_{-\infty}^{\infty} \left[\sum_{i=0}^{p} \frac{\lambda_i D_i}{t_2^{\,n-1}} \sum_{j \in S_i'} \frac{(b_{1i}-b_{1j})^{p_i'-2}}{\prod_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g)} \cdot e^{-it_2\left[\frac{(i\ j\ y)}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}}\right]} \right] dt_2 \; .$$

Now distort the contour of integration for t_2 to avoid the singularity at $t_2 = 0$ and integrate with respect to t_2 term by term using the formula

$$\int_{\Gamma} \frac{e^{-iaz}}{z^{n}}\, dz = \begin{cases} \dfrac{(-1)^{n}\, 2\pi i^{n}\, a^{n-1}}{(n-1)!} & \text{if } a \geq 0 \\[2ex] 0 & \text{if } a < 0 \end{cases}$$

for the contour integrals. Thus

$$f(y_1, y_2) = \frac{n!\,(-1)}{2\pi i^{n-1}} K_{np} \sum_{i=0}^{p} \lambda_i D_i \sum_{j \in S_i'} \frac{(b_{1i} - b_{1j})^{p_i'-2}}{\prod_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g)} \cdot \int_{\Gamma} \frac{e^{-it_2 \left[\frac{(i\ j\ y)}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}} \right]}}{t_2^{n-1}}\, dt_2$$

$$= \frac{n!\,(-1)^{n}}{(n-2)!} K_{np} \sum_{i=0}^{p} \lambda_i D_i \sum_{j \in S_i'} \frac{\lambda_{ij} (b_{1i} - b_{1j})^{p_i'-2}}{\prod_{\substack{g \in S_i' \\ g \neq j}} (i\ j\ g)} \left[\frac{(i\ j\ y)}{\begin{vmatrix} 1 & b_{1i} \\ 1 & b_{1j} \end{vmatrix}} \right]^{n-2}$$

since

$$\int_{\Gamma} \frac{e^{-it_2\left[\frac{(i\ j\ y)}{\begin{vmatrix}1 & b_{1i}\\ 1 & b_{1j}\end{vmatrix}}\right]}}{t_2^{\,n-1}}\, dt_2 = \begin{cases} \dfrac{(-1)^{n-1}\, 2\pi i^{\,n-1}\left[\dfrac{(i\ j\ y)}{\begin{vmatrix}1 & b_{1i}\\ 1 & b_{1j}\end{vmatrix}}\right]^{n-2}}{(n-2)!} & \text{if } \dfrac{(i\ j\ y)}{\begin{vmatrix}1 & b_{1i}\\ 1 & b_{1j}\end{vmatrix}} \geq 0 \\[2ex] 0 & \text{if } \dfrac{(i\ j\ y)}{\begin{vmatrix}1 & b_{1i}\\ 1 & b_{1j}\end{vmatrix}} < 0 \end{cases}$$

$$= \frac{\lambda_{ij}\,(-1)^{n-1}\, 2\pi i^{\,n-1}\left[\dfrac{(i\ j\ y)}{\begin{vmatrix}1 & b_{1i}\\ 1 & b_{1j}\end{vmatrix}}\right]^{n-2}}{(n-2)!}\,.$$

Finally

$$f(y_1,y_2) = \frac{n!(-1)^n}{(n-2)!} K_{0p} \sum_{i=0}^{p} \lambda_i D_i \sum_{j\in S_i'} \frac{\lambda_{ij}(b_{1i}-b_{1j})^{p_i'-2}(i\ j\ y)^{n-2}}{(-1)^{n-2}(b_{1i}-b_{1j})^{n-2} \prod_{\substack{g\in S_i'\\ g\neq j}} (i\ j\ g)}$$

and formula (3.3.3) is obtained by a simple combination of some of the terms in the above line. ▮

3.4 : Application to the distribution of the sums of ordered intervals

To demonstrate an application of the results obtained in this chapter the paper by Barton and David (1955), which was used as an illustratión in the previous chapter, is considered. In their paper they wanted to derive the joint density function of G_{tv} and G_{rs} $(v < r)$ where

$$G_{tv} = \sum_{i=t}^{v} U_{(i)} \quad \text{and} \quad G_{rs} = \sum_{i=r}^{s} U_{(i)} \quad .$$

They concluded that the following expression seems to be the most convenient form of the joint density function.

$$f(g_{tv}, g_{rs}, z_1, \ldots, z_{v-t})$$

$$= K \sum_{k=1}^{t-1} \sum_{m=1}^{r-v-1} \sum_{j=0}^{s-r-1} \binom{t-1}{k} \binom{r-v-1}{m} \binom{s-r}{j+1}$$

$$\cdot \frac{(-1)^{m+v+s-r-j-1} (j+1)^{s-r} \Delta kmj}{(s-r+j)(n+1-s)-k(j+1)}$$

where

$$z_i = u_{(t+i-1)} \qquad (i=1,2,\ldots,v-t)$$

$$K = \frac{(n+1)!\, n!}{(n+1-s)^{s-r-1} (s-v+t-2)!\, (t-1)!\, (r-v-1)!\, (s-r-1)!\, (n+1-s)!}$$

and

$$\Delta kmj = \left[H\left(1-g_{rs}\frac{n-r+m+2}{s-r+1} - g_{tv}(r-v-m) + (r-v-m-1)\sum_{i=2}^{v-t} z_i + (r-v-m-k-1)z_1\right)\right]^{n-v+t-2}$$

$$-\left[H\left(1-g_{rs}\frac{n-s+j+2}{s-r+1} - (1+A_j)g_{tv} + A_j\sum_{i=2}^{v-t} z_i + (A_j-k)z_1\right)\right]^{n-v+t-2}$$

with

$$A_j = \frac{(r-v-1)(j+1) - (n+1-s)(s-r+j)}{j+1} .$$

To use the procedure of this chapter it is necessary to follow in a manner similar to that of section 2.4. Expressing G_{tv} and G_{rs} in terms of the coverages one has

$$G_{tv} = (v+1-t)\sum_{i=1}^{t}\frac{U_i}{n+2-i} + \sum_{i=t+1}^{v}\frac{v+1-i}{n+2-i}U_i$$

and

$$G_{rs} = (s+1-r)\sum_{i=1}^{r}\frac{U_i}{n+2-i} + \sum_{i=r+1}^{s}\frac{s+1-i}{n+2-i}U_i .$$

Considering the coefficients of corresponding coverages as coordinates of a point in the two-dimensional space the following types of points are obtained under the assumption that $v < r$.

Type I $\quad (a_{1i},a_{2i}) = (0,0) \quad (i=0,s+1,\ldots,n)$

Type II $\quad (a_{1i},a_{2i}) = \left(\frac{v+1-t}{n+2-i}, \frac{s+1-r}{n+2-i}\right) \quad (i=1,2,\ldots,t-1)$

Type III $\quad (a_{1i},a_{2i}) = \left(\frac{v+1-t}{n+2-t}, \frac{s+1-r}{n+2-t}\right) \quad i=t$

Type IV $\quad (a_{1i},a_{2i}) = \left(\frac{v+1-i}{n+2-i}, \frac{s+1-r}{n+2-i}\right) \quad (i=t+1,t+2,\ldots,v)$

Type V $\quad (a_{1i},a_{2i}) = \left(0, \frac{s+1-r}{n+1-v}\right) \quad i=v+1$

Type VI $\quad (a_{1i},a_{2i}) = \left(0, \frac{s+1-r}{n+2-i}\right) \quad (i=v+2,v+3,\ldots,r)$

Type VII $\quad (a_{1i},a_{2i}) = \left(0,, \frac{s+1-i}{n+2-i}\right) \quad (i=r+1,r+2,\ldots,s)$

To apply formula (3.3.3) certain information regarding the points is required. It can easily be shown that:

(1) Type II points are distinct from one another and all belong to one colinear set. The magnitude of both the coordinates of these points vary directly with the subscript $i \in (1,2,\ldots,t-1)$.

(2) Type IV points are distinct from one another and all belong to one colinear set. The magnitude of the first coordinate of these points varies indirectly with the

subscript $i \in (t+1,t+2,\ldots,v)$ whereas that of the second coordinate varies directly.

(3) Type VI points are distinct from one another and similarly for type VII points. The first coordinate of each of these types of points is zero. The magnitude of the second coordinate of the type VI point varies directly with the subscript $i \in (v+2,v+3,\ldots,r)$ whereas that of the type VII point varies indirectly with the subscript $i \in (r+1,r+2,\ldots,s)$. The magnitude of the second coordinate of the point (a_{1r},a_{2r}) is the largest. These two types of points constitute a vertical colinear set.

(4) Type I,III, and V points each belong to exactly two of the colinear sets given in (1), (2), and (3). The type I points belong to the colinear sets of (1) and (3) and the type II and V points belong to the colinear sets of (1), (2) and (2), (3) respectively. All other points in the three colinear sets of (1), (2), and (3) do not belong to the other two.

(5) Type I points form a set of n-s+1 coincident points at (0,0). All other points are distinct from one another except for the possibility of a type VI point (a_{1i},a_{2i}), $i \in (v+2,v+3,\ldots,r-1)$, being coincident with a type VII point (a_{1i},a_{2i}), $i \in (r+1,r+2,\ldots,e)$ where e is the largest integer such that the type VII

point (a_{1e},a_{2e}) has a larger ordinate then that of the type V point.

(6) All type II points have a smaller ordinate then any type IV point. The type III point has the largest abscissa among all points and the type V point has a smaller ordinate then all type VI points.

In any problem formula (3.3.3) can be used to find the joint density function of G_{tv} and G_{rs} $(v < r)$. The above remarks supply, in general, the information required to use formula (3.3.3) except for certain possible situations. That is, three possible situations exist which cannot, to any appreciable degree, be resolved. They are:

(i) the possibility of having a type II and IV point being vertical

(ii) the possibility of having a type VI and VII point being coincident

(iii) the possibility of having type II , IV, and VI points or type II, IV, and VII points being colinear.

As a result of these problems the form of the density function in Theorem 3.4 cannot, in this particular problem, be reduced to a much simpler form. However special cases can be cited in which all or some of the above coincident, vertical, and colinear problems are eliminated and the desired reduction is possible. Some of these cases are:

(1) If $t = 1$ then the vertical and colinear problems are eliminated.

(2) If $t = 1$ and $r = v+1$ or $r = v+2$ then all three problems are eliminated.

(3) If $r = v+1$ then the coincident and colinear problems are eliminated.

(4) If $\frac{k(n+1-s)}{s+1-r-k}$ is not an integer in $(1,2,\ldots,r-v-2)$ for some $k \in (1,2,\ldots,s-r)$ then the coincident problem is eliminated.

(5) If $\frac{k(n+1-v)}{v+1-t-k}$ is not an integer in $(1,2,\ldots,t-1)$ for some $k \in (1,2,\ldots,v-t)$ then the vertical problem is eliminated.

(6) If $\frac{k_1 k_2}{v+1-t-k_1}$ is not an integer in $(1,2,\ldots,t-1)$ for some $k_1 \in (1,2,\ldots,v-t)$ and some $k_2 \in (1,2,\ldots,r-v-1)$ and in addition if

$$\frac{k_1[(r+k_2)(s+1-r) - (v+1)(s+1-r-k_2) - (n+2)k_2]}{(v+1-t-k_1)(s+1-r-k_2)}$$

is also not such an integer for some k_1 in the range given above and for some $k_2 \in (1,2,\ldots,s-r)$ then the colinear problem is eliminated.

To illustrate a possible selection of the sets S_i and S_i' assume that the coincident, vertical, and colinear problems have been prevented by assuming (4), (5), and (6) above. Then under these conditions the number of distinct points is $n+1-(n-s) = s+1$. Denote these points by

$$(b_{1i},b_{2i}) \qquad (i=0,1,\ldots,s)$$

where

$$(b_{10},b_{20}) = (a_{1i},a_{2i}) = (0,0) \qquad (i=0,s+1,\ldots,n)$$

and

$$(b_{1i},b_{2i}) = (a_{1i},a_{2i}) \qquad (i=1,2,\ldots,s) .$$

Also note that $n_0 = n-s+1$ and $n_i = 1$ $(i=1,2,\ldots,s)$. The sets S_i $(i=0,1,\ldots,s)$ can be chosen to be

$$S_0 = (t,t+1,\ldots,v,v+1 \ ; \ t,1,\ldots,1,s-v \ ; \ v-t+2)$$

$$S_i = (t,t+1,\ldots,v,v+1,\ldots,s \ ; \ t,1,\ldots,1,1,\ldots,1 \ ; \ s-t+1)$$

$$(i=1,2,\ldots,t-1)$$

$$S_t = (0,v+1,v+2,\ldots,s \ ; \ t,v+1-t,1,\ldots,1 \ ; \ s-v+1)$$

$$S_i = (0,1,\ldots,t-1,v+1,v+2,\ldots,s \ ; \ 1,1,\ldots,1,v-t+1,1,\ldots,1 \ ; \ s-v+t)$$

$$(i=t+1,t+2,\ldots,v)$$

$$S_{v+1} = (0,1,\ldots,t-1,t \ ; \ s-v,1,\ldots,1,v-t+1 \ ; \ t+1)$$

$$S_i = (1,\ldots,t-1,t,t+1,\ldots,v,0 \ ; \ 1,\ldots,1,1,1,\ldots,1,s-v \ ; \ v+1)$$

$$(i=v+2,v+3,\ldots,s)$$

and it follows that

$S'_0 = (t,t+1,\ldots,v \; ; \; t,1,\ldots,1 \; ; \; v-t+1)$

$S'_{v+1} = (1,\ldots,t-1,t \; ; \; 1,\ldots,1,v-t+1 \; ; \; t)$

$S'_i = (1,\ldots,t-1,t,t+1,\ldots,v \; ; \; 1,\ldots,1,1,1,\ldots,1 \; ; \; v)$

$(i=v+2,v+3,\ldots,s)$

$S'_i = S_i \quad (i=1,2,\ldots,v)$.

Consider the following simple example of the Barton and David problem. Let $n = 4$, $t = 1$, $v = 2$, $r = 3$, and $s = 4$. For convenience let $Y_1 = G_{tv} = G_{12}$ and $Y_2 = G_{rs} = G_{34}$. In this example the joint density function of

$$Y_1 = U_{(1)} + U_{(2)}$$

$$Y_2 = U_{(3)} + U_{(4)}$$

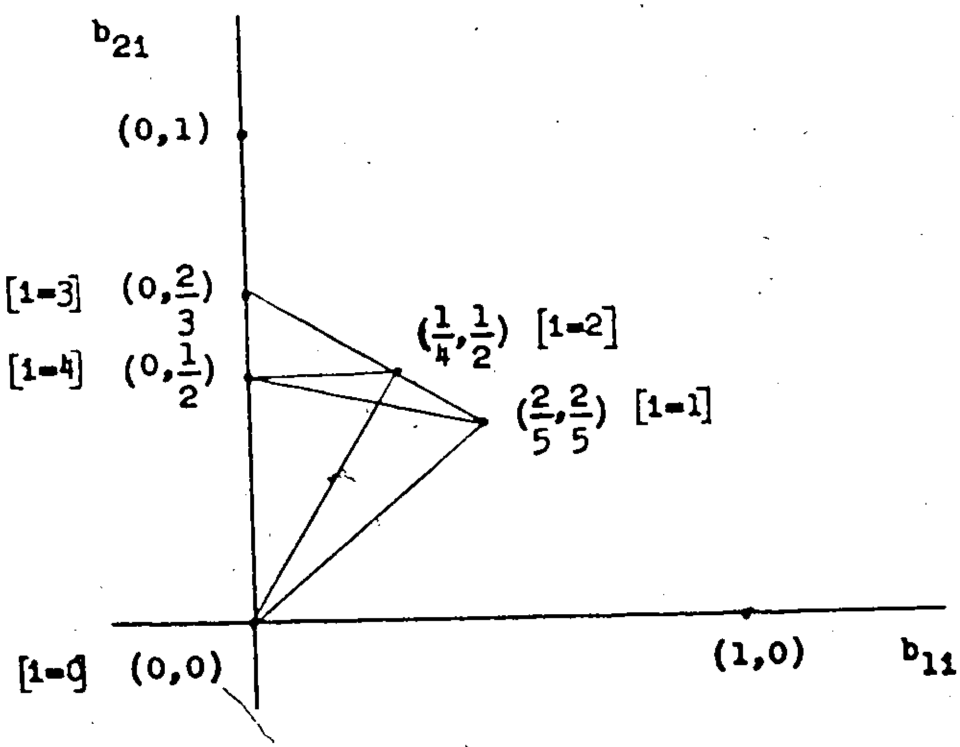
is considered.

The following points are involved, all of which are distinct.

Type I	$(b_{10},b_{20}) = (0,0)$	$i = 0$
Type III	$(b_{11},b_{21}) = (\frac{2}{5},\frac{2}{5})$	$i = 1$
Type IV	$(b_{12},b_{22}) = (\frac{1}{4},\frac{1}{2})$	$i = 2$
Type V	$(b_{13},b_{23}) = (0,\frac{2}{3})$	$i = 3$
Type VII	$(b_{14},b_{24}) = (0,\frac{1}{2})$	$i = 4$

Note that $n_i = 1$ $(i=0,1,2,3,4)$ and that the coincident, vertical, and colinear problems do not exist in this example.

The point configuration is illustrated in the following diagram.



The colinear sets can be chosen to be

$S_0 = (1,2,3\ ;\ 1,1,2\ ;\ 3)$

$S_1 = (0,3,4\ ;\ 1,2,1\ ;\ 3)$

$S_2 = (0,3,4\ ;\ 1,2,1\ ;\ 3)$

$S_3 = (0,1\ ;\ 2,2\ ;\ 2)$

$S_4 = (1,2,0\ ;\ 1,1,2\ ;\ 3)$.

Also

$S_0' = (1,2\ ;\ 1,1\ ;\ 2)$

$S_3' = (1 \; ; \; 2 \; ; \; 1)$

$S_4' = (1,2 \; ; \; 1,1 \; ; \; 2)$

$S_i' = S_i \quad (i=1,2)$.

Noting that for the above values

$$\frac{n!}{(n-2)!} \prod_{p=0}^{s} \frac{1}{(n_p-1)!} \frac{\partial^{n-s}}{\partial_{b_{20}}^{n_0-1} \partial_{b_{21}}^{n_1-1} \cdots \partial_{b_{2s}}^{n_s-1}} = 12$$

to use formula (3.3.3) it remains to calculate

$$M_i = \frac{\lambda_i \prod_{h \in S_i'} (b_{1i}-b_{1h})^{s_{ih}}}{\prod_{\substack{g=0 \\ g \neq i \\ g \notin V_i}}^{s} (b_{1i}-b_{1g}) \prod_{q \in V_i} (b_{2i}-b_{2q})} \prod_{w \in S_i'} \frac{1}{(s_{iw}-1)!} \frac{\partial^{s_{iw}-1}}{\partial_{b_{2w}}^{s_{iw}-1}}$$

$$\cdot \sum_{j \in S_i'} \frac{\lambda_{ij}}{(b_{1i}-b_{1j})^{s-s_i'}} \frac{(i\ j\ y)^{n-2}}{\prod_{\substack{k \in S_i' \\ k \neq j}} (i\ j\ k)}$$

for $i=0,1,2,3,4$. In this regard one can show that:

$$M_0 = 30\,\lambda_0\,\lambda_{01}(y_1-y_2)^2 - 30\,\lambda_0\,\lambda_{02}(2y_1-y_2)^2$$

$$M_1 = -30\,\lambda_1\,\lambda_{10}(y_1-y_2)^2 + 30\,\lambda_1\,\lambda_{13}(4y_1-10y_1y_2+16y_2-15y_2^2-4)$$

$$+ 30\,\lambda_1\,\lambda_{14}(y_1+4y_2-2)^2$$

$$M_2 = 30\,\lambda_2\,\lambda_{20}(2y_1-y_2)^2 - 30\,\lambda_2\,\lambda_{23}(4y_1^2-4y_1y_2+16y_2-15y_2^2-4)$$

$$- 30\,\lambda_2\,\lambda_{24}(4y_2-2)^2$$

$$M_3 = 30\,\lambda_3\,\lambda_{31}y_1(4y_1+6y_2-4)$$

$$M_4 = -30\,\lambda_4\,\lambda_{41}(y_1+4y_2-2)^2 + 30\,\lambda_4\,\lambda_{42}(4y_2-2)^2 \quad .$$

Thus

$$f(y_1,y_2) = 360\Big[\lambda_0\,\lambda_{01}(y_1-y_2)^2 - \lambda_0\,\lambda_{02}(2y_1-y_2)^2$$

$$- \lambda_1\,\lambda_{10}(y_1-y_2)^2 + \lambda_1\,\lambda_{13}(4y_1-10y_1y_2+16y_2-15y_2^2-4)$$

$$+ \lambda_1\,\lambda_{14}(y_1+4y_2-2)^2 + \lambda_2\,\lambda_{20}(2y_1-y_2)^2$$

$$- \lambda_2\,\lambda_{23}(4y_1^2-4y_1y_2+16y_2-15y_2^2-4) - \lambda_2\,\lambda_{24}(4y_2-2)^2$$

$$+ \lambda_3\,\lambda_{31}y_1(4y_1+6y_2-4) - \lambda_4\,\lambda_{41}(y_1+4y_2-2)^2$$

$$+ \lambda_4\,\lambda_{42}(4y_2-2)^2\Big] \quad .$$

The density function can also be put in an alternative form. Define $a = (b_{1a}, b_{2a})$ to be the intersection of the line through the points (0,0) and $(\frac{1}{4},\frac{1}{2})$ and the line through $(0,\frac{1}{2})$ and $(\frac{2}{5},\frac{2}{5})$. Then defining T_{ijk} to be the triangle with vertices (b_{1i},b_{2i}), (b_{1j},b_{2j}), and (b_{1k},b_{2k}) it can be shown that

$$f(y_1,y_2)=360\begin{cases}(y_1-y_2)^2 & \text{if } (y_1,y_2)\in T_{01a}\\ (4y_1-10y_1y_2+16y_2-15y_2^2-4) & \text{if } (y_1,y_2)\in T_{a12}\\ y_1(2y_2-3y_1) & \text{if } (y_1,y_2)\in T_{0a4}\\ (-4y_1^2-6y_1y_2+4y_1-16y_2^2+16y_2-4) & \text{if } (y_1,y_2)\in T_{4a2}\\ y_1(-4y_1-6y_2+4) & \text{if } (y_1,y_2)\in T_{423}\ .\end{cases}$$

Bibliography

1. Ali, Mir M., Distribution of linear combination of order statistics from rectangular population, Bulletin of the Institute of Statistical Research and Training, University of Dacca 3 (1969), 1-21.

2. Ali, Mir M. and Mead, E. R., On the distribution of several linear combinations of order statistics from the uniform distribution, Bulletin of the Institute of Statistical Research and Training, University of Dacca 3 (1969), 22-41.

3. Barton, D. E. and David, F. N., Sums of ordered intervals and distances, Mathematika 2 (1955), 150-159.

4. Barton, D. E. and David, F. N., Some notes on ordered random intervals, Journal of Royal Statistical Society, Series B, 18 (1956), 79-94.

5. Barton, D. E. and David, F. N., Tests for randomness of points on a line, Biometrika 43 (1956) 104-112.

6. Dempster, A. P. and Kleyle, R. M., Distributions determined by cutting a simplex with hyperplanes, Ann. Math. Statistics 39 (1968), 1473-1478.

7. Dwass. M., The distribution of linear combinations of random divisions of an interval, Trabajos De Estadistica 12 (1961), 10-17.

8. Mauldon, J. G., Random divisions of an interval, Proc. Cambridge Phil. Soc. 47 (1951), 331-339.

9. Mauldon, J. G., An inversion formula for a generalized transform, Mathematika 4 (1957), 146-155.

10. Mead, E. R., On the exact distribution of linear combinations of order statistics, Ph. D. Thesis, University of Western Ontario (1969).

11. Milne-Thomson, L. M., The calculus of finite differences, Macmillan and Co., Ltd., London, (1933).

12. Niven, B. S., The joint distribution of the sample mean and the sample range, Austral. Journal of Statistics 5 (1963), 127-132.

13. Sarhan, A. E. and Greenberg, B. G., Contributions to order statistics, Wiley, New York (1962).

14. Weisberg, Herbert, The distribution of linear combinations of order statistics from the uniform distribution, Ann. Math. Statistics 42 (1971), 704-709.