CHAPTER 2

Elementary theory of cumulants

2.1 Introduction

This chapter deals with the elementary theory of cumulants in the multivariate case as well as the univariate case. Little prior knowledge is assumed other than some familiarity with moments and the notion of mathematical expectation, at least in the univariate case. In what follows, all integrals and infinite sums are assumed to be convergent unless otherwise stated. If the random variable X has a density function $f_X(x)$ defined over $-\infty < x < \infty$, then the expectation of the function $g(X)$ is

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x) f_X(x) dx. \tag{2.1}$$

If the distribution of X is discrete, the integral is replaced by a sum over the discrete values. More generally, to take care of both the discrete case and the continuous case simultaneously, we may replace $f_X(x)dx$ in the above integral by $dF_X(x)$, where F_X is a probability measure. Such differences, however, need not concern us here. All that is required is a knowledge of some elementary properties of the integral and the expectation operator. In the particular case where $g(X)$ is the rth power of X, (2.1) is called the rth *moment* of X.

In the multivariate case, essentially the same definitions may be used except that the integral over R^1 is replaced by an integral over R^p. Not only do we have to consider the moments of each component of X but also the cross moments such as $E(X^1X^2)$ and $E(X^1X^1X^2)$. As always, in the multivariate case, superscripts denote components and not powers. The moments of a given component of X give information regarding the marginal distribution of that component. The cross moments are required to give information concerning the joint distribution of the components.

Cumulants are normally introduced as functions of the moments. It is entirely natural to inquire at the outset why it is preferable to work with cumulants rather than moments since the two are entirely equivalent. A single totally convincing answer to this query is difficult to find and, in a sense, Chapters 2 to 6 provide several answers. Simplicity seems to be the main criterion as the following brief list shows.

(i) Most statistical calculations using cumulants are simpler than the corresponding calculations using moments.
(ii) For independent random variables, the cumulants of a sum are the sums of the cumulants.
(iii) For independent random variables, the cross cumulants or mixed cumulants are zero.
(iv) Edgeworth series used for approximations to distributions are most conveniently expressed using cumulants.
(v) Where approximate normality is involved, higher-order cumulants can usually be neglected but not higher-order moments.

2.2 Generating functions

2.2.1 *Definitions*

As always, we begin with the random variable X whose components are $X^1, \ldots, X^p$. All arrays bearing superscripts refer to these components. Unless otherwise specified, the moments of X about the origin are assumed finite and are denoted by

$$\begin{aligned} \kappa^i = E(X^i), \quad \kappa^{ij} = E(X^i X^j), \\ \kappa^{ijk} = E(X^i X^j X^k) \end{aligned} \tag{2.2}$$

and so on. Moments about the mean, also called central moments, are rarely considered explicitly here, except as a special case of the above with $\kappa^i = 0$. The indices need not take on distinct values and there may, of course, be more than p indices, implying repetitions. Thus, for example, κ^{11} is the mean square of X^1 about the origin and κ^{222} is the mean cube of X^2, the second component of X. In this context, superscripts must not be confused with powers and powers should, where possible, be avoided. In this book, powers are avoided for the most part, the principal exceptions arising in Sections 2.5 and 2.6 where the connection with other notations is described and interpretations are given.

Consider now the infinite series

$$\begin{aligned} M_X(\xi) = 1 + \xi_i\kappa^i + \xi_i\xi_j\kappa^{ij}/2! + \xi_i\xi_j\xi_k\kappa^{ijk}/3! \\ + \xi_i\xi_j\xi_k\xi_l\kappa^{ijkl}/4! + \ldots, \end{aligned} \tag{2.3}$$

which we assume to be convergent for all $|\xi|$ sufficiently small. The sum may be written in the form

$$M_X(\xi) = E\{\exp(\xi_i X^i)\},$$

and the moments are just the partial derivatives of $M_X(\xi)$ evaluated at $\xi = 0$.

The cumulants are most easily defined via their generating function,

$$K_X(\xi) = \log M_X(\xi),$$

which has an expansion

$$\begin{aligned} K_X(\xi) = \xi_i\kappa^i + \xi_i\xi_j\kappa^{i,j}/2! + \xi_i\xi_j\xi_k\kappa^{i,j,k}/3! \\ + \xi_i\xi_j\xi_k\xi_l\kappa^{i,j,k,l}/4! + \ldots. \end{aligned} \tag{2.4}$$

This expansion implicitly defines all the cumulants, here denoted by κ^i, $\kappa^{i,j}$, $\kappa^{i,j,k}$ and so on, in terms of the corresponding moments. The major departure from standard statistical notation is that we have chosen to use the same letter for both moments and cumulants. Both are indexed by a set rather than by a vector, the only distinction arising from the commas, which are considered as separators for the cumulant indices. Thus the cumulants are indexed by a set of indices fully partitioned and the moments by the same set unpartitioned. One curious aspect of this convention is that the notation does not distinguish between moments with one index and cumulants with one index. This is convenient because the first moments and first cumulants are identical.

The infinite series expansion (2.3) for $M_X(\xi)$ may be divergent for all real $|\xi| > 0$ either because some of the higher-order moments are infinite or because the moments, though finite, increase sufficiently rapidly to force divergence (Exercise 2.2). In such cases, it is entirely legitimate to work with the finite series expansions up to any specified number of terms. This device can be justified by taking ξ to be purely imaginary, in which case, the integral $E\{\exp(\xi_i X^i)\}$ is convergent and its Taylor approximation for small imaginary ξ is just the truncated expansion. Similarly for

$K_X(\xi)$. Thus cumulants of any order are well defined if the corresponding moment and all lower-order marginal moments are finite: see (2.9) and Exercise 2.1.

One difficulty that arises in using moment or cumulant calculations to prove results of a general nature is that the infinite set of moments is, in general, not sufficient to determine the joint distribution uniquely. Feller (1971, Section VII.3) gives a pair of non-identical univariate density functions having identical moments of all orders. Non-uniqueness occurs only when the function $M_X(\xi)$ is not analytic at the origin. Thus, for a large class of problems, non-uniqueness can be avoided by including the condition that the series expansion for $M_X(\xi)$ be convergent for $|\xi| < \delta$ where $\delta > 0$ (Moran, 1968, Section 6.4; Billingsley, 1985, Exercise 30.5). In the univariate case, other conditions limiting the rate of increase of the even moments are given by Feller (1971, Sections VII.3, VII.6 and XV.4).

2.2.2 *Some examples*

Multivariate normal distribution: The multivariate normal density with mean vector κ^i and covariance matrix $\kappa^{i,j}$ may be written in the form

$$(2\pi)^{-p/2}|\kappa^{i,j}|^{-1/2}\exp\{-\tfrac{1}{2}(x^i-\kappa^i)(x^j-\kappa^j)\kappa_{i,j}\},$$

where $\kappa_{i,j}$ is the matrix inverse of $\kappa^{i,j}$ and x ranges over R^p. The moment generating function may be found by completing the square in the exponent and using the fact that the density integrates to 1. This gives

$$M_X(\xi) = \exp\{\xi_i\kappa^i + \tfrac{1}{2}\xi_i\xi_j\kappa^{i,j}\}$$

and

$$K_X(\xi) = \xi_i\kappa^i + \tfrac{1}{2}\xi_i\xi_j\kappa^{i,j}.$$

In other words, for the normal distribution, all cumulants of order three or more are zero and, as we might expect, the second cumulant is just the covariance array.

Multinomial distribution: For a second multivariate example, we take the multinomial distribution on k categories with index m and parameter $\pi = \pi_1, \ldots, \pi_k$. The joint distribution or probability function may be written

$$\operatorname{pr}(X^1 = x^1, \ldots, X^k = x^k) = \binom{m}{x^1, \ldots, x^k}\prod_{j=1}^{k}\pi_j^{x^j}$$

where $0 \le x^j \le m$ and $\sum_j x^j = m$. The moment generating function may be found directly using the multinomial theorem, giving

$$M_X(\xi) = \sum_x \binom{m}{x^1, \ldots, x^k}\prod_{j=1}^{k}\exp(\xi_j x^j)\pi_j^{x^j} = \Big(\sum \pi_j\exp(\xi_j)\Big)^m.$$

The cumulant generating function is

$$K_X(\xi) = m\log\Big(\sum \pi_j\exp(\xi_j)\Big).$$

Thus all cumulants are finite and have the form $m \times$ (function of π). The first four are given in Exercise 2.16.

Student's distribution: Our third example involves a univariate distribution whose moments of order three and higher are infinite. The t distribution on three degrees of freedom has density function

$$f_X(x) = \frac{2}{\pi\sqrt{3}(1 + x^2/3)^2} \qquad -\infty < x < \infty.$$

The moment generating function, $M_X(\xi) = E\{\exp(\xi X)\}$, diverges for all real $|\xi| > 0$ so that the function $M_X(\xi)$ is not defined or does not exist for real ξ. However, if we write $\xi = i\zeta$ where ζ is real, we find

$$M_X(i\zeta) = \int_{-\infty}^{\infty} \frac{2\exp(i\zeta x)}{\pi\sqrt{3}(1 + x^2/3)^2}\, dx.$$

The integrand has poles of order 2 at $x = \pm i\sqrt{3}$ but is analytic elsewhere in the complex plane. If $\zeta > 0$, the integral may be evaluated by deforming the contour into the positive complex half-plane, leaving the residue at $x = +i\sqrt{3}$. If $\zeta < 0$, it is necessary to deform in the other direction, leaving the residue at $x = -i\sqrt{3}$. This procedure gives

$$\begin{aligned} M_X(i\zeta) &= \exp(-\sqrt{3}|\zeta|)\{1 + \sqrt{3}|\zeta|\} \\ K_X(i\zeta) &= -\sqrt{3}|\zeta| + \log(1 + \sqrt{3}|\zeta|) \\ &= \begin{cases} -3\zeta^2/2 + \sqrt{3}\zeta^3 - \ldots & \text{if } \zeta > 0 \\ -3\zeta^2/2 - \sqrt{3}\zeta^3 - \ldots & \text{if } \zeta < 0 \end{cases} \end{aligned}$$

Thus $K_X(i\zeta)$ has a unique Taylor expansion only as far as the quadratic term. It follows that $E(X) = 0$, $\text{var}(X) = 3$ and that the higher-order cumulants are not defined.

2.3 Cumulants and moments

To establish the relationships connecting moments with cumulants, we write $M_X(\xi) = \exp\{K_X(\xi)\}$ and expand to find

$$\begin{aligned} 1 + \xi_i\kappa^i &+ \xi_i\xi_j(\kappa^{i,j}/2! + \kappa^i\kappa^j/2!) \\ &+ \xi_i\xi_j\xi_k\kappa^{i,j,k}/3! \quad + \xi_i\xi_j\xi_k\xi_l\kappa^{i,j,k,l}/4! + \ldots \\ &+ \xi_i\xi_j\xi_k\kappa^i\kappa^{j,k}/2! \quad + \xi_i\xi_j\xi_k\xi_l\{\kappa^i\kappa^{j,k,l}/6 + \kappa^{i,j}\kappa^{k,l}/8\} + \ldots \\ &+ \xi_i\xi_j\xi_k\kappa^i\kappa^j\kappa^k/3! + \xi_i\xi_j\xi_k\xi_l\kappa^i\kappa^j\kappa^{k,l}/4 + \ldots \\ &\qquad\qquad\qquad + \xi_i\xi_j\xi_k\xi_l\kappa^i\kappa^j\kappa^k\kappa^l/4! + \ldots \\ &\qquad\qquad\qquad\qquad + \ldots . \end{aligned} \tag{2.5}$$

After combining terms and using symmetry, we find the following expressions for moments in terms of cumulants:

$$\begin{aligned} \kappa^{ij} &= \kappa^{i,j} + \kappa^i\kappa^j \\ \kappa^{ijk} &= \kappa^{i,j,k} + (\kappa^i\kappa^{j,k} + \kappa^j\kappa^{i,k} + \kappa^k\kappa^{i,j}) + \kappa^i\kappa^j\kappa^k \\ &= \kappa^{i,j,k} + \kappa^i\kappa^{j,k}[3] + \kappa^i\kappa^j\kappa^k \\ \kappa^{ijkl} &= \kappa^{i,j,k,l} + \kappa^i\kappa^{j,k,l}[4] + \kappa^{i,j}\kappa^{k,l}[3] + \kappa^i\kappa^j\kappa^{k,l}[6] \\ &\quad + \kappa^i\kappa^j\kappa^k\kappa^l, \end{aligned} \tag{2.6}$$

where, for example,

$$\kappa^{i,j}\kappa^{k,l}[3] = \kappa^{i,j}\kappa^{k,l} + \kappa^{i,k}\kappa^{j,l} + \kappa^{i,l}\kappa^{j,k}$$

is the sum over the three 2^2 partitions of four indices. The bracket notation is simply a convenience to avoid listing explicitly all 15 partitions of four indices in the last equation (2.6). Only the five distinct types, each corresponding to a partition of the *number* 4, together with the number of partitions of each type, need be listed. The following is a complete list of the 15 partitions of four items, one column for each of the five types.

$ijkl$	$i\|jkl$	$ij\|kl$	$i\|j\|kl$	$i\|j\|k\|l$
	$j\|ikl$	$ik\|jl$	$i\|k\|jl$	
	$k\|ijl$	$il\|jk$	$i\|l\|jk$	
	$l\|ijk$		$j\|k\|il$	
			$j\|l\|ik$	
			$k\|l\|ij$	

In the case of fifth-order cumulants, there are 52 partitions of seven different types and our notation makes the listing of such partitions feasible for sets containing not more than eight or nine items. Such lists are given in Tables 1 and 2 of the Appendix, and these may be used to find the expressions for moments in terms of cumulants. We find, for example, from the partitions of five items, that

$$\begin{aligned}\kappa^{ijklm} &= \kappa^{i,j,k,l,m} + \kappa^i\kappa^{j,k,l,m}[5] + \kappa^{i,j}\kappa^{k,l,m}[10] \\ &\quad + \kappa^i\kappa^j\kappa^{k,l,m}[10] + \kappa^i\kappa^{j,k}\kappa^{l,m}[15] \\ &\quad + \kappa^i\kappa^j\kappa^k\kappa^{l,m}[10] + \kappa^i\kappa^j\kappa^k\kappa^l\kappa^m.\end{aligned}$$

If $\kappa^i = 0$, all partitions having a unit part (a block containing only one element) can be ignored. The formulae then simplify to

$$\begin{aligned}\kappa^{ij} &= \kappa^{i,j}, \quad \kappa^{ijk} = \kappa^{i,j,k} \\ \kappa^{ijkl} &= \kappa^{i,j,k,l} + \kappa^{i,j}\kappa^{k,l}[3] \\ \kappa^{ijklm} &= \kappa^{i,j,k,l,m} + \kappa^{i,j}\kappa^{k,l,m}[10] \\ \kappa^{ijklmn} &= \kappa^{i,j,k,l,m,n} + \kappa^{i,j}\kappa^{k,l,m,n}[15] + \kappa^{i,j,k}\kappa^{l,m,n}[10] \\ &\quad + \kappa^{i,j}\kappa^{k,l}\kappa^{m,n}[15].\end{aligned}$$

These are the formulae for the central moments in terms of cumulants.

The reverse formulae giving cumulants in terms of moments may be found either by formal inversion of (2.6) or by expansion of $\log M_X(\xi)$ and combining terms. The expressions obtained for the first four cumulants are

$$\begin{aligned}\kappa^{i,j} &= \kappa^{ij} - \kappa^i\kappa^j \\ \kappa^{i,j,k} &= \kappa^{ijk} - \kappa^i\kappa^{jk}[3] + 2\kappa^i\kappa^j\kappa^k \\ \kappa^{i,j,k,l} &= \kappa^{ijkl} - \kappa^i\kappa^{jkl}[4] - \kappa^{ij}\kappa^{kl}[3] + 2\kappa^i\kappa^j\kappa^{kl}[6] - 6\kappa^i\kappa^j\kappa^k\kappa^l.\end{aligned} \tag{2.7}$$

Again, the sum is over all partitions of the indices but this time, the coefficient $(-1)^{\nu-1}(\nu-1)!$ appears, where ν is the number of blocks of the partition. The higher-order formulae follow the same pattern and the list of partitions in Tables 1,2 of the Appendix may be used for cumulants up to order eight.

More generally, the relationships between moments and cumulants may be written as follows. Let $\Upsilon = \{\upsilon_1, \ldots, \upsilon_\nu\}$ be a partition of a set of p indices into ν non-empty blocks. (Υ and υ are the upper and lower cases of the Greek letter 'upsilon'). For example, if $p = 4$ and $\nu = 2$ we might have $\Upsilon = \{(i,j),(k,l)\}$, $\{(i,k),(j,l)\}$ or $\{(i),(j,k,l)\}$ or any one of the four other possible partitions into 2 blocks. The partition comprising just a single block is denoted by Υ_1 and the partition comprising p unit blocks is denoted by Υ_p. There is no standard statistical notation for these partitions and it is sometimes convenient to use the alternatives, 1 and 0, borrowed from lattice theory: see Section 3.6. The moment involving all p indices is written $\kappa(\Upsilon_1)$ and the corresponding cumulant $\kappa(\Upsilon_p)$. In the above example, $\Upsilon_1 = \{(i,j,k,l)\}$ is the partition into one block, $\kappa(\Upsilon_1) = \kappa^{ijkl}$ is the corresponding moment, $\Upsilon_4 = \{(i),(j),(k),(l)\}$ is the partition into 4 blocks and $\kappa(\Upsilon_4) = \kappa^{i,j,k,l}$ is the corresponding cumulant. Cumulants involving only those indices in block υ_j are written $\kappa(\upsilon_j)$ and the corresponding moment is written $\mu(\upsilon_j)$.

By examining the general term in expansion (2.5) for $M_X(\xi)$, it is not difficult to see that the expression for the moment $\kappa(\Upsilon_1)$ in terms of cumulants may be written

$$\kappa(\Upsilon_1) = \sum_{\Upsilon} \kappa(\upsilon_1)\ldots\kappa(\upsilon_\nu) \tag{2.8}$$

where the sum extends over all partitions of the indices. Equivalently, the above may be expressed as a double sum, first over ν and then over all partitions of the indices into ν blocks.

The corresponding expression for the cumulant $\kappa(\Upsilon_p)$ in terms of the moments is

$$\kappa(\Upsilon_p) = \sum_{\Upsilon} (-1)^{\nu-1}(\nu-1)!\mu(\upsilon_1)\ldots\mu(\upsilon_\nu). \tag{2.9}$$

Note that the block sizes do not enter into either of the above expressions.

In fact, we may take (2.9) as an alternative definition of cumulant, more directly applicable than the definition relying on generating functions. The advantage of (2.9) as a definition is that it makes explicit the claim made at the end of Section 2.2, that cumulants of order r are well defined when the corresponding rth-order moment and all lower-order marginal moments are finite. Note that, apart from the univariate case, $\kappa^{ijk} < \infty$ does not imply that $\kappa^{i,j,k} < \infty$: see Exercise 2.1.

2.4 Linear and affine transformations

The objective in this section is to examine how the moment arrays κ^{ij}, $\kappa^{ijk}, \ldots$ and the cumulant arrays $\kappa^{i,j}$, $\kappa^{i,j,k}, \ldots$ change when we make a simple transformation from the original variables $X^1, \ldots, X^p$ to new variables $Y = Y^1, \ldots, Y^q$. If Y is a linear function of X, we may write

$$Y^r = a_i^r X^i,$$

where a_i^r is an array of constants. It is not difficult to see that the moments of Y are

$$a_i^r \kappa^i, \quad a_i^r a_j^s \kappa^{ij}, \quad a_i^r a_j^s a_k^t \kappa^{ijk}, \ldots$$

while the cumulants are

$$a_i^r \kappa^i, \quad a_i^r a_j^s \kappa^{i,j}, \quad a_i^r a_j^s a_k^t \kappa^{i,j,k}, \ldots. \tag{2.10}$$

In other words, under linear transformation, both moments and cumulants transform like contravariant tensors. Note however, that the matrix a_i^r need not have full rank.

Affine transformations involve a change of origin according to the equation

$$Y^r = a^r + a_i^r X^i.$$

The cumulants of Y, derived at the end of this section, are

$$a^r + a_i^r \kappa^i, \quad a_i^r a_j^s \kappa^{i,j}, \quad a_i^r a_j^s a_k^t \kappa^{i,j,k}, \quad a_i^r a_j^s a_k^t a_l^u \kappa^{i,j,k,l} \tag{2.11}$$

and so on. The change of origin affects only the mean vector or first cumulant. For this reason, cumulants are sometimes called semi-invariants. On the other hand, the moments of Y are

$$\begin{aligned}
&a^r + a_i^r \kappa^i, \\
&a^r a^s + a^r a_i^s \kappa^i[2] + a_i^r a_j^s \kappa^{ij}, \\
&a^r a^s a^t + a^r a^s a_i^t \kappa^i[3] + a^r a_i^s a_j^t \kappa^{ij}[3] + a_i^r a_j^s a_k^t \kappa^{ijk}
\end{aligned}$$

and so on, where $a^r a_i^s \kappa^i[2] = a^r a_i^s \kappa^i + a^s a_i^r \kappa^i$. Thus, unlike the cumulants, the moments do not transform in a pleasant way under affine transformation of coordinates.

The transformation law for cumulants is similar to the transformation law of Cartesian tensors, (Jeffreys, 1952), the only difference being the dependence of the first cumulant on the choice of origin. To avoid any ambiguity of terminology, we use the term *cumulant tensor* to describe any array of quantities that transforms according to (2.11) under affine transformation of X. In Section

3.3, we develop rules for the transformation of cumulant tensors under non-linear transformation of X. These rules are quite different from tensor transformation laws that arise in differential geometry or in theoretical physics.

To prove (2.11) we use the method of generating functions, giving

$$\begin{aligned} M_Y(\xi) &= E[\exp\{\xi_r(a^r + a_i^r X^i)\}] \\ &= \exp(\xi_r a^r) M_X(\xi_r a_i^r). \end{aligned}$$

In other words,

$$K_Y(\xi) = \xi_r a^r + K_X(\xi_r a_i^r)$$

from which expressions (2.10) and (2.11) follow directly.

2.5 Univariate cumulants and power notation

Much of the literature on cumulants concentrates on the univariate case, $p = 1$ and uses the condensed *power* notation κ_r for the rth cumulant of X^1, written in this section as X without indices. In this and in the following section, we move quite freely from power notation to index notation and back: this should cause no confusion on account of the different positions of indices in the two notations. Following Kendall & Stuart (1977, Chapter 3) we write

$$\mu_r' = E(X^r) \quad \text{and} \quad \mu_r = E(X - \mu_1')^r$$

where the superscript here denotes a power. Expressions (2.6) giving moments in terms of cumulants become

$$\begin{aligned} \mu_2' &= \kappa_2 + \kappa_1^2 \\ \mu_3' &= \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3 \\ \mu_4' &= \kappa_4 + 4\kappa_1\kappa_3 + 3\kappa_2^2 + 6\kappa_1^2\kappa_2 + \kappa_1^4 \end{aligned}$$

where superscripts again denote powers. The reverse formulae (2.7) become

$$\begin{aligned} \kappa_2 &= \mu_2' - (\mu_1')^2 \\ \kappa_3 &= \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3 \\ \kappa_4 &= \mu_4' - 4\mu_1'\mu_3' - 3(\mu_2')^2 + 12(\mu_1')^2\mu_2' - 6(\mu_1')^4. \end{aligned}$$

In this notation, the permutation factors, previously kept in [.], and the arithmetic factors $(-1)^{\nu-1}(\nu-1)!$ become combined, with the result that the essential simplicity of the formulae (2.7) disappears. For this reason, unless we are dealing with a single random variable or a set of independent and identically distributed random variables, it is usually best to use index notation, possibly reverting to power notation at the last step of the calculations.

An extended version of power notation is sometimes used for bivariate cumulants corresponding to the random variables X and Y. For example, if $\mu_{rs}' = E(X^r Y^s)$, the corresponding cumulant may be written κ_{rs}. Note that in this notation, $\kappa_{rs} \neq \kappa_{sr}$. In other words, with power notation, the cumulants are indexed by a vector whereas, with index notation, the cumulants are indexed by a set or, more generally, by the partitions of a set.

To establish the relationships between bivariate moments and bivariate cumulants in this notation, we simply convert the terms in (2.7) into power notation. For $r = 2$, $s = 1$ this gives

$$\kappa_{21} = \mu_{21}' - 2\mu_{10}'\mu_{11}' - \mu_{01}'\mu_{20}' + 2(\mu_{10}')^2\mu_{01}'$$

and, for $r = 2$, $s = 2$,

$$\begin{aligned} \kappa_{22} = \mu_{22}' &- 2\mu_{10}'\mu_{12}' - 2\mu_{01}'\mu_{21}' - \mu_{20}'\mu_{02}' - 2(\mu_{11}')^2 \\ &+ 2(\mu_{10}')^2\mu_{02}' + 2(\mu_{01}')^2\mu_{20}' + 8\mu_{01}'\mu_{10}'\mu_{11}' - 6(\mu_{10}')^2(\mu_{01}')^2. \end{aligned}$$

Additional, more impressive formulae of this type may be found in David, Kendall & Barton (1966, Tables 2.1.1 and 2.1.2). The formula for κ_{44}, for example, occupies 24 lines.

Simpler formulae are available in terms of central moments, but it is clear that the above notation conceals the simplicity of the formulae. For example, if μ'_{01} and μ'_{10} are both equal to zero, the formulae for κ_{44}, κ_{35}, κ_{26} and so on, can be found from the list of partitions of $1, 2, \ldots, 8$ in Table 2 of the Appendix. For this reason, power notation is best avoided at least in formal manipulations. It is, however, very useful and convenient in the univariate case and is also useful more generally for interpretation.

2.6 Interpretation of cumulants

Although the definition of cumulants given in Sections 2.2 and 2.3 covered both univariate and mixed cumulants, their interpretations are best considered separately. Roughly speaking, mixed cumulants have an interpretation in terms of dependence or independence: univariate cumulants have a simple interpretation in terms of the shape of the marginal distribution. Of course, mixed cumulants could be interpreted also in terms of the shape of the joint distribution but such interpretations are not given here. We deal first with univariate cumulants using the notation of Section 2.5 and work our way up to mixed cumulants involving four distinct variables.

The first cumulant of X, denoted by κ_1, is the mean value and the second cumulant, κ_2, is the variance. In rigid body mechanics, κ_1 is the x-coordinate of the centre of gravity and κ_2 is the moment of inertia about the axis $x = \kappa_1$ of a uniform laminar body of unit mass in the (x, y) plane, bounded by $0 \le y \le f(x)$ where $f(x)$ is the density of X.

The third cumulant of X is a measure of asymmetry in the sense that $\kappa_3 = E(X - \kappa_1)^3$ is zero if X is symmetrically distributed. Of course $\kappa_3 = 0$ does not, on its own, imply symmetry: to guarantee symmetry, we require all odd cumulants to vanish and the distribution to be determined by its moments. For an example of an asymmetrical distribution whose odd cumulants are zero, see Kendall & Stuart (1977, Exercise 3.26), which is based on the note by Churchill (1946). The usual measure of skewness is the standardized third cumulant, $\rho_3 = \kappa_3/\kappa_2^{3/2}$, which is unaffected by affine transformations $X \to a + bX$ with $b > 0$. If $b < 0$, $\rho_3 \to -\rho_3$. Third and higher-order standardized cumulants given by $\rho_r = \kappa_r/\kappa_2^{r/2}$, can be interpreted as summary measures of departure from normality in the sense that if X is normal, all cumulants of order three or more are zero. This aspect is developed in greater detail in Chapter 5 where Edgeworth expansions are introduced.

Suppose now we have two random variables X^1 and X^2. With index notation, the mixed cumulants are denoted by $\kappa^{1,2}$, $\kappa^{1,1,2}$, $\kappa^{1,2,2}$, $\kappa^{1,1,1,2}, \ldots$. The corresponding quantities in power notation are κ_{11}, κ_{21}, κ_{12}, κ_{31} and so on. To the extent that third and higher-order cumulants can be neglected, we find from the bivariate normal approximation that

$$\begin{aligned} E(X^2|X^1 = x^1) &\simeq \kappa_{01} + (\kappa_{11}/\kappa_{20})(x^1 - \kappa_{10}) \\ E(X^1|X^2 = x^2) &\simeq \kappa_{10} + (\kappa_{11}/\kappa_{02})(x^2 - \kappa_{01}) \end{aligned} \tag{2.12}$$

so that $\kappa_{11} > 0$ implies positive dependence in the sense of increasing conditional expectations. Refined versions of the above, taking third- and fourth-order cumulants into account, are given in Chapter 5.

The simplest interpretations of bivariate cumulants are given in terms of independence. If X^1 and X^2 are independent, then all mixed cumulants involving X^1 and X^2 alone are zero. Thus, $\kappa^{1,2} = \kappa^{1,1,2} = \kappa^{1,2,2} = \ldots = 0$ or, more concisely using power notation, $\kappa_{rs} = 0$ for all $r, s \ge 1$. Provided that the moments determine the joint distribution, the converse is also true, namely that if $\kappa_{rs} = 0$ for all $r, s \ge 1$, then X^1 and X^2 are independent. The suggestion here is that if $\kappa_{rs} = 0$ for $r, s = 1, \ldots, t$, say, then X^1 and X^2 are approximately independent in some sense. However it is difficult to make this claim rigorous except in the asymptotic sense of Chapter 5 where all higher-order cumulants are negligible.

Consider now the case of three random variables whose mixed cumulants may be denoted by κ_{111}, κ_{211}, κ_{121}, κ_{112}, κ_{311}, κ_{221}, and so on by an obvious extension of the notation of Section 2.5. It is not difficult to see that if X^1 is independent of (X^2, X^3), or if X^2 is independent of (X^1, X^3), or if X^3 is independent of (X^1, X^2) then $\kappa_{111} = 0$ and, in fact, more generally, $\kappa_{rst} = 0$ for all $r, s, t \geq 1$. These independence relationships are most succinctly expressed using generalized power notation and it is for this reason that we switch here from one notation to the other.

More generally, if we have any number of random variables that can be partitioned into two independent blocks, then all mixed cumulants involving indices from both blocks are zero. Note that if X^1 and X^2 are independent, it does *not* follow from the above that, say, $\kappa^{1,2,3} = 0$. For example, if $X^3 = X^1X^2$, then it follows from (3.2) that $\kappa^{1,2,3} = \kappa_{20}\kappa_{02}$ and this is strictly positive unless X^1 or X^2 is degenerate. For a more interesting example, see the bivariate exponential recurrence process $\{X_j, Y_j\}$, $j = \ldots, -1, 0, 1, \ldots$ described in Exercise 2.35, in which each X_j is independent of the marginal process $\{Y_j\}$ and conversely, Y_j is independent of the marginal process $\{X_j\}$, but the two processes are dependent.

To see how the converse works, we note that $\kappa_{rs0} = 0$ for all $r, s \geq 1$ implies independence of X^1 and X^2, but only if the joint distribution is determined by its moments. Similarly, $\kappa_{r0s} = 0$ implies independence of X^1 and X^3. This alone does not imply that X^1 is independent of the pair (X^2, X^3): for this we require, in addition to the above, that $\kappa_{rst} = 0$ for all $r, s, t \geq 1$.

2.7 The central limit theorem

2.7.1 *Sums of independent random variables*

Suppose that $X_1, \ldots, X_n$ are n independent vector-valued random variables where X_r has components $X_r^1, \ldots, X_r^p$. We do not assume that the observations are identically distributed and so the cumulants of X_r are denoted by $\kappa_r^i, \kappa_r^{i,j}, \kappa_r^{i,j,k}$ and so on. One of the most important properties of cumulants is that the joint cumulants of $X_\bullet = X_1 + \cdots + X_n$ are just the sums of the corresponding cumulants of the individual variables. Thus we may write $\kappa_\bullet^i$, $\kappa_\bullet^{i,j}$, $\kappa_\bullet^{i,j,k}$ for the joint cumulants of $X_\bullet$ where, for example, $\kappa_\bullet^{i,j} = \sum_{r=1}^n \kappa_r^{i,j}$. This property is not shared by moments and there is, therefore, some risk of confusion if we were to write $\kappa_\bullet^{ij} = E(X_\bullet^i X_\bullet^j)$ because

$$\kappa_\bullet^{ij} = \kappa_\bullet^{i,j} + \kappa_\bullet^i \kappa_\bullet^j \neq \sum_r \kappa_r^{ij}.$$

For this reason, it is best to avoid the notation $\kappa_\bullet^{ij}$.

To demonstrate the additive property of cumulants we use the method of generating functions and write

$$M_{X_\bullet}(\xi) = E[\exp\{\xi_i(X_1^i + \cdots + X_n^i)\}].$$

By independence we have,

$$M_{X_\bullet}(\xi) = M_{X_1}(\xi) \cdots M_{X_n}(\xi)$$

and thus

$$K_{X_\bullet}(\xi) = K_{X_1}(\xi) + \cdots + K_{X_n}(\xi).$$

The required result follows on extraction of the appropriate coefficients of ξ.

In the particular case where the observations are identically distributed, we have that $\kappa_\bullet^i = n\kappa^i$, $\kappa_\bullet^{i,j} = n\kappa^{i,j}$ and so on. If the observations are not identically distributed, it is sometimes convenient to define average cumulants by writing

$$\kappa_\bullet^i = n\bar{\kappa}^i, \quad \kappa_\bullet^{i,j} = n\bar{\kappa}^{i,j}, \quad \kappa_\bullet^{i,j,k} = n\bar{\kappa}^{i,j,k}$$

and so on where, for example, $\bar{\kappa}^{i,j}$ is the average covariance of components i and j.

2.7.2 *Standardized sums*

With the notation of the previous section, we write

$$Y^i = n^{-1/2}\{X^i_{\bullet} - n\bar{\kappa}^i\}.$$

The cumulants of $X_{\bullet}$ are $n\bar{\kappa}^i, n\bar{\kappa}^{i,j}, n\bar{\kappa}^{i,j,k}$ and so on. It follows from Section 2.4 that the cumulants of Y are

$$0, \quad \bar{\kappa}^{i,j}, \quad n^{-1/2}\bar{\kappa}^{i,j,k}, \quad n^{-1}\bar{\kappa}^{i,j,k,l}, \ldots$$

the successive factors decreasing in powers of $n^{-1/2}$. Of course, the average cumulants are themselves implicitly functions of n and without further, admittedly mild, assumptions, there is no guarantee that $n^{-1/2}\bar{\kappa}^{i,j,k}$ or $n^{-1}\bar{\kappa}^{i,j,k,l}$ will be negligible for large n. We avoid such difficulties in the most direct way, simply by assuming that $\bar{\kappa}^{i,j}, \bar{\kappa}^{i,j,k}, \ldots$ have finite limits as $n \to \infty$ and that the limiting covariance matrix, $\bar{\kappa}^{i,j}$ is positive definite. In many problems, these assumptions are entirely reasonable but they do require checking. See Exercise 2.10 for a simple instance of failure of these assumptions.

The cumulant generating function of Y is

$$K_Y(\xi) = \xi_i\xi_j\bar{\kappa}^{i,j}/2! + n^{-1/2}\xi_i\xi_j\xi_k\bar{\kappa}^{i,j,k}/3! + \ldots. \tag{2.13}$$

Under the assumptions just given, and for complex ξ, the remainder in this series after r terms is $O(n^{-r/2})$. Now, $\xi_i\xi_j\bar{\kappa}^{i,j}/2$ is the cumulant generating function of a normal random variable with zero mean and covariance matrix $\bar{\kappa}^{i,j}$. Since convergence of the cumulant generating function implies convergence in distribution, subject to continuity of the limiting function at the origin (Moran, 1968, Section 6.2), we have just proved a simple version of the central limit theorem for independent but non-identically distributed random variables. In fact, it is not necessary here to use the generating function directly. Convergence of the moments implies convergence in distribution provided that the limiting moments uniquely determine the distribution, as they do in this case. For a more accurate approximation to the density of Y, we may invert (2.13) formally, leading to an asymptotic expansion in powers of $n^{-1/2}$. This expansion is known as the Edgeworth series after F.Y. Edgeworth (1845–1926). Note however, that although the error in (2.13) after two terms is $O(n^{-1})$, the error in probability calculations based on integrating the formal inverse of (2.13) need not be $O(n^{-1})$. In discrete problems, the error is typically $O(n^{-1/2})$.

Stronger forms of the central limit theorem that apply under conditions substantially weaker than those assumed here, are available in the literature. In particular, versions are available in which finiteness of the higher-order cumulants is not a requirement. Such theorems, on occasion, have statistical applications but they sometimes suffer from the disadvantage that the error term for finite n may be large and difficult to quantify, even in an asymptotic sense. Often the error is $o(1)$ as opposed to $O(n^{-1/2})$ under the kind of assumptions made here. Other forms of the central limit theorem are available in which certain specific types of dependence are permitted. For example, in applications related to time series, it is often reasonable to assume that observations sufficiently separated in time must be nearly independent. With additional mild assumptions, this ensures that the asymptotic cumulants of derived statistics are of the required order of magnitude in n and the central limit result follows.

2.8 Derived scalars

Suppose we are interested in the distribution of the statistic

$$T^2 = (X^i - \kappa_0^i)(X^j - \kappa_0^i)\kappa_{i,j}$$

where $\kappa_{i,j}$ is the matrix inverse or generalized inverse of $\kappa^{i,j}$. The Mahalanobis statistic, T^2, is a natural choice that arises if we are testing the hypothesis $H_0 : \kappa^i = \kappa_0^i$, where the higher-order cumulants are assumed known. One reason for considering this particular statistic is that it is invariant under affine transformation of X. Its distribution must therefore depend on scalars derived from the cumulants of X that are invariant under affine nonsingular transformations

$$X^r \to a^r + a_i^r X^i$$

where a_i^r is a $p\times p$ matrix of rank p. These scalars are the multivariate generalizations of $\rho_3^2, \rho_4, \rho_6, \rho_5^2$ and so on, in the univariate case: see Section 2.6.

To obtain the multivariate generalization of $\rho_3^2 = \kappa_3^2/\kappa_2^3$, we first write $\kappa^{i,j,k}\kappa^{l,m,n}$ as the generalization of κ_3^2. Division by κ_2 generalizes to multiplication by $\kappa_{r,s}$. Thus we require a scalar derived from

$$\kappa^{i,j,k}\kappa^{l,m,n}\kappa_{r,s}\kappa_{t,u}\kappa_{v,w}$$

by contraction, i.e. by equating pairs of indices and summing. This operation can be done in just two distinct ways giving two non-negative scalars

$$\bar\rho_{13}^2 = \rho_{13}^2/p = \kappa^{i,j,k}\kappa^{l,m,n}\kappa_{i,j}\kappa_{k,l}\kappa_{m,n}/p, \tag{2.14}$$

$$\bar\rho_{23}^2 = \rho_{23}^2/p = \kappa^{i,j,k}\kappa^{l,m,n}\kappa_{i,l}\kappa_{j,m}\kappa_{k,n}/p. \tag{2.15}$$

Similarly, to generalize $\rho_4 = \kappa_4/\kappa_2^2$, we obtain just the single expression

$$\bar\rho_4 = \rho_4/p = \kappa^{i,j,k,l}\kappa_{i,j}\kappa_{k,l}/p. \tag{2.16}$$

In the univariate case, (2.14) and (2.15) reduce to the same quantity and $\bar\rho_4$ satisfies the familiar inequality $\bar\rho_4 \geq \bar\rho_3^2 - 2$. The multivariate generalization of this inequality applies most directly to $\bar\rho_{13}^2$, giving

$$\bar\rho_4 \geq \bar\rho_{13}^2 - 2.$$

Equality is achieved if and only if the joint distribution of the Xs is concentrated on some conic in p-space in which the coefficients of the quadratic term are $\kappa_{i,j}$. The support of the distribution may be degenerate at a finite number of points but it is assumed here that it is not contained in any lower dimensional subspace. Otherwise, the covariance matrix would be rank deficient and p would be replaced by the rank of the subspace. In the univariate case, equality is achieved if and only if the distribution is concentrated on two points: see Exercise 2.12.

The corresponding inequality for $\bar\rho_{23}^2$ is

$$\bar\rho_4 \geq \bar\rho_{23}^2 - p - 1.$$

See Exercise 2.14. This limit is attained if and only if the joint distribution is concentrated on $p+1$ points not contained in any linear subspace of R^p. The inequality for $\bar\rho_{23}^2$ is obtained by taking the trace of the residual covariance matrix of the products after linear regression on the linear terms. The trace vanishes only if this matrix is identically zero. Thus, achievement of the bound for $\bar\rho_{23}^2$ implies achievement of the bound for $\bar\rho_{13}^2$ and also that the higher-order cumulants are determined by those up to order four. See Section 3.8.

The simplest example that illustrates the difference between the two skewness scalars is the multinomial distribution with index m and parameter vector $\pi_1, \ldots, \pi_k$. The joint cumulants are given in Exercise 2.16 and the covariance matrix $\kappa^{i,j} = m\{\pi_i\delta_{ij} - \pi_i\pi_j\}$, has rank $p = k - 1$. The simplest generalized inverse is $\kappa_{i,j} = \{m\pi_i\}^{-1}$ for $i = j$ and zero otherwise. Substitution of the expressions in Exercise 2.16 for the third and fourth cumulants into (2.14) – (2.16) gives

$$\begin{aligned} m(k-1)\bar{\rho}_{13}^2 &= \sum\nolimits_j \pi_j^{-1} - k^2 \\ m(k-1)\bar{\rho}_{23}^2 &= \sum\nolimits_j (1-\pi_j)(1-2\pi_j)/\pi_j \\ m(k-1)\bar{\rho}_4 &= \sum\nolimits_j \pi_j^{-1} - k^2 - 2(k-1). \end{aligned}$$

Thus $\bar{\rho}_{13}^2$ is zero for the uniform multinomial distribution even though $\kappa^{i,j,k}$ is not identically zero. On the other hand, $\bar{\rho}_{23}^2$ is zero only if $\kappa^{i,j,k}$ is identically zero and, in the case of the multinomial distribution, this cannot occur unless $k = 2$ and $\pi = \frac{1}{2}$. For additional interpretations of the differences between these two scalars: see Exercise 2.15.

The above invariants are the three scalars most commonly encountered in theoretical work such as the expansion of the log likelihood ratio statistic or computing the variance of T^2. However, they are not the only invariant functions of the first four cumulants. A trivial example is $\kappa^{i,j}\kappa_{i,j} = p$, or more generally, the rank of $\kappa^{i,j}$. Also, if we were to generalize $\rho_4^2 = \kappa_4^2/\kappa_2^4$ by considering quadratic expressions in $\kappa^{i,j,k,l}$, there are two possibilities in addition to $(\bar{\rho}_4)^2$. These are

$$\begin{aligned} \bar{\rho}_{14}^2 &= \kappa^{i,j,k,l}\kappa^{r,s,t,u}\kappa_{i,j}\kappa_{k,r}\kappa_{l,s}\kappa_{t,u}/p, \\ \bar{\rho}_{24}^2 &= \kappa^{i,j,k,l}\kappa^{r,s,t,u}\kappa_{i,r}\kappa_{j,s}\kappa_{k,t}\kappa_{l,u}/p. \end{aligned}$$

In addition to these, there are integer invariants of a qualitatively different kind, obtained by extending the notions of rank and signature to multi-way arrays. Kruskal (1977) defines the rank of a three-way asymmetrical array in a way that is consistent with the standard definition for matrices. In addition, the four-way array, $\kappa^{i,j,k,l}$ can be thought of as a symmetric $p^2 \times p^2$ matrix whose rank and signature are invariants. See Section 1.5.2 and Exercise 1.9. There may also be other invariants unconnected with the notions of rank or signature but none have appeared in the literature. However, it seems unnatural to consider $\kappa^{i,j,k,l}$ as a two-way array and, not surprisingly, the integer invariants just mentioned do not arise in the usual statistical calculations. For completeness, however, it would be good to know the complete list of all invariants that can be formed from, say, the first four cumulants.

To see that (2.14)–(2.16) are indeed invariant under affine transformation, we note that

$$\kappa^{i,j,k} \to a_i^r a_j^s a_k^t \kappa^{i,j,k} \quad \text{and} \quad \kappa_{i,j} \to b_r^i b_s^j \kappa_{i,j},$$

where b_r^i is the matrix inverse of a_i^r. Direct substitution followed by cancellation reveals the invariance. This invariance property of scalars derived by contraction of tensors is an elementary consequence of the tensor transformation property. Provided that we work exclusively with tensors, it is not necessary to check that scalars derived in this way are invariant.

2.9 Conditional cumulants

Suppose we are given the conditional joint cumulants of the random variables $X^1, \ldots, X^p$ conditional on some event A. How do we combine the conditional cumulants to obtain the unconditional joint cumulants? In the case of moments, the answer is easy because

$$E(X^1X^2\cdots) = E_A E(X^1X^2\cdots|A). \tag{2.17}$$

In other words, the unconditional moments are just the average of the conditional moments. However, it is not difficult to show, for example, that the covariance of X^i and X^j satisfies

$$\begin{aligned}\kappa^{i,j} =& E_A\{\text{cov}(X^i, X^j|A)\} \\ &+ \text{cov}_A\{E(X^i|A), E(X^j|A)\}.\end{aligned} \tag{2.18}$$

To see how these expressions generalize to cumulants of arbitrary order, we denote the conditional cumulants by λ^i, $\lambda^{i,j}$, $\lambda^{i,j,k}$ and we use the identity connecting the moment generating functions

$$M_X(\xi) = E_A M_{X|A}(\xi).$$

Expansion of this identity and comparison of coefficients gives

$$\begin{aligned}\kappa^i &= E_A\{\lambda^i\} \\ \kappa^{i,j} + \kappa^i\kappa^j &= E_A\{\lambda^{i,j} + \lambda^i\lambda^j\} \\ \kappa^{i,j,k} + \kappa^i\kappa^{j,k}[3] + \kappa^i\kappa^j\kappa^k &= E_A\{\lambda^{i,j,k} + \lambda^i\lambda^{j,k}[3] + \lambda^i\lambda^j\lambda^k\}.\end{aligned}$$

Expression (2.18) for the unconditional covariance follows from the second expression above. On using this result in the third expression, we find

$$\kappa^{i,j,k} = E(\lambda^{i,j,k}) + \kappa_2(\lambda^i, \lambda^{j,k})[3] + \kappa_3(\lambda^i, \lambda^j, \lambda^k).$$

The generalization is easy to see though, for notational reasons, a little awkward to prove, namely

$$\kappa(\Upsilon_p) = \sum_{\Upsilon} \kappa_\nu\{\lambda(\upsilon_1), \ldots, \lambda(\upsilon_\nu)\}$$

with summation over all partitions of the p indices. In this expression, $\lambda(\upsilon_j)$ is the conditional mixed cumulant of the random variables whose indices are in υ_j and $\kappa_\nu\{\lambda(\upsilon_1), \ldots, \lambda(\upsilon_\nu)\}$ is the νth order cumulant of the ν random variables listed as arguments. For details of a proof see the papers by Brillinger (1969) or Speed (1983).

In many circumstances, it is required to compute conditional cumulants from joint unconditional cumulants, the converse of the result just described. A little reflection soon shows that the converse problem is considerably more difficult and the best that can be expected are approximate conditional cumulants. Expansions of this type are given in Chapter 5.

However, if $M_{X,Y}(\xi,\zeta) = E\{\exp(\xi_i X^i + \zeta_r Y^r)\}$ is the joint moment generating function of X, Y, we can at least write down an expression for the conditional moment generating function $M_{X|Y}(\xi)$. Since

$$M_{X,Y}(\xi,\zeta) = \int \exp(\zeta_r Y^r) M_{X|Y}(\xi) f_Y(y)dy,$$

we may invert the integral transform to find

$$k\, M_{X|Y}(\xi) f_Y(y) = \int_{c-i\infty}^{c+i\infty} M_{X,Y}(\xi,\zeta)\exp(-\zeta_r y^r)d\zeta.$$

Division by $f_Y(y)$ gives the conditional moment generating function in the form

$$M_{X|Y}(\xi) = \frac{\int M_{X,Y}(\xi,\zeta)\exp(-\zeta_r y^r)d\zeta}{\int M_{X,Y}(0,\zeta)\exp(-\zeta_r y^r)d\zeta}. \tag{2.19}$$

This expression, due to Bartlett (1938), can be used for generating expansions or approximations for conditional moments: it is rarely used directly for exact calculations. See, however, Moran (1968, Section 6.14).

2.10 Bibliographic notes

Cumulants were first defined by Thiele in about 1889. Thiele's work is mostly in Danish and the most accessible English translation is of his book *Theory of Observations* reprinted in *The Annals of Mathematical Statistics* (1931), pp.165-308. In Chapter 6 of that book, Thiele defines the univariate cumulants and calls them *half-invariants* on account of their simple transformation properties. He also derives a version of the central limit theorem by showing that the higher-order cumulants of a standardized linear combination of random variables converge rapidly to zero provided that 'the coefficient of any single term is not so great ... that it throws all the other terms into the shade', a delightful statement closely approximating the spirit of the Lindeberg-Feller condition.

For an excellent readable account of the central limit theorem and the Lindeberg condition, see LeCam (1986).

Fisher (1929), in an astonishing tour de force, rediscovered cumulants, recognized their superiority over moments, developed the corresponding sample cumulants and cross-cumulants, gave the formulae for the cumulants and cross-cumulants of the sample cumulants and formulated combinatorial rules for computing the cumulants of such statistics. Both Thiele and Fisher used what we call 'power notation' for cumulants and cross-cumulants and their achievements are the more remarkable for that reason.

Formulae giving univariate moments in terms of cumulants and vice versa are listed in Kendall & Stuart (1977, Chapter 3): multivariate versions of these formulae are given in Chapter 13. See also David, Kendall & Barton (1966, Tables 2.1.1 and 2.1.2) and David & Barton (1962, Chapter 9). Similar formulae are given by Brillinger (1975, Chapter 2).

The derived scalars $\bar{\rho}_4$ and $\bar{\rho}_{23}^2$ were given by Mardia (1970) as summary measures of multivariate kurtosis and skewness. The additional skewness scalar, $\bar{\rho}_{13}^2$ is given by McCullagh & Cox (1986) who show how it arises in calculations involving likelihood ratio tests. See also Davis (1980) who shows how the scalar $\psi_p = p(\bar{\rho}_{13}^2 - \bar{\rho}_{23}^2)$ arises in calculations concerning the effect of non-normality on the distribution of Wilks's Λ.

2.11 Further results and exercises 2

2.1 Let X have density function given by

$$f_X(x) = \begin{cases} 2/x^3 & x \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Show that the mean of X is finite but that all higher-order moments are infinite. Find an expression for the density of $Y = 1/X$ and show that all moments and cumulants are finite. Let $\mu'_{rs} = E(X^r Y^s)$ be the joint moment of order $r + s$ (using the power notation of Section 2.5). Show that $\mu'_{21} = 2$ but that the corresponding cumulant, κ_{21}, is infinite.

2.2 Let X be a standard normal random variable and set $Y = \exp(X)$. Show that the rth moment of Y about the origin is $\mu'_r = \exp(r^2/2)$. Hence find expressions for the first four cumulants of Y. Show that the series expansions for $M_Y(\xi)$ and $K_Y(\xi)$ about $\xi = 0$ are divergent for all real $\xi > 0$ even though all cumulants are finite (Heyde, 1963).

2.3 Let X be a scalar random variable. Prove by induction on r that the derivatives of $M_X(\xi)$ and $K_X(\xi)$, if they exist, satisfy

$$M_X^{(r)}(\xi) = \sum_{j=1}^{r} \binom{r-1}{j-1} M_X^{(r-j)}(\xi) K_X^{(j)}(\xi) \qquad r \geq 1.$$

Hence show that

$$\mu'_r = \kappa_r + \sum_{j=1}^{r-1} \binom{r-1}{j-1} \kappa_j \mu'_{r-j}$$

and, for $r \geq 4$, that

$$\mu_r = \kappa_r + \sum_{j=2}^{r-2} \binom{r-1}{j-1} \kappa_j \mu_{r-j}$$

(Thiele, 1897, eqn. 22; Morris, 1982).

2.4 If $\mu(\Upsilon)$ and $\kappa(\Upsilon)$ denote the ordinary moment and the ordinary cumulant corresponding to the indices in $\Upsilon = \{i_1, \ldots, i_p\}$, show that

$$\mu(\Upsilon) = \kappa(\Upsilon) + \sum_{\{\upsilon_1, \upsilon_2\}} \kappa(\upsilon_1)\mu(\upsilon_2)$$

where $\{\upsilon_1, \upsilon_2\}$ is a partition of Υ into two non-empty blocks and the sum extends over all partitions such that $i_1 \in \upsilon_1$. What purpose does the condition $i_1 \in \upsilon_1$ serve? Show that this result generalizes the univariate identity in Exercise 2.3.

2.5 Show that the central and non-central moments satisfy

$$\mu'_r = \sum_{j=0}^{r} \binom{r}{j} \mu_{r-j} (\mu'_1)^j$$

$$\mu_r = \sum_{j=0}^{r} \binom{r}{j} \mu'_{r-j} (-\mu'_1)^j.$$

(Kendall and Stuart, 1977, p. 58).

2.6 The density function of Student's distribution on ν degrees of freedom is

$$\frac{(1 + t^2/\nu)^{-(\nu+1)/2}}{\nu^{1/2} B(\frac{1}{2}, \nu/2)} \qquad -\infty < t < \infty,$$

where $B(.,.)$ is the beta function. Show that the odd moments that exist are zero and that the even moments are

$$\mu_{2r} = \frac{1.3.\cdots(2r-1)\nu^r}{(\nu-2)\cdots(\nu-2r)} \qquad 2r < \nu.$$

Hence show that

$$\rho_4 = 6/(\nu - 4) \quad \text{for} \quad \nu > 4$$
$$\rho_6 = 240/\{(\nu-4)(\nu-6)\} \quad \text{for} \quad \nu > 6.$$

2.7 Prove directly using (2.7) that if X^i is independent of the pair (X^j, X^k), then $\kappa^{i,j,k} = 0$.

2.8 Show that if X^1 is independent of (X^2, X^3), then the cumulant κ_{rst} of order $r + s + t$ with $r, s, t \geq 1$ is equal to zero.

2.9 Derive expressions (2.7) for cumulants in terms of moments (i) directly from (2.6) and (ii) by expansion of $\log M_X(\xi)$.

2.10 Let $h_r(x)$ be the standardized Hermite polynomial of degree r satisfying the condition $\int h_r(x)h_s(x)\phi(x)dx = \delta_{rs}$ where $\phi(x)$ is the standard normal density. If $X_i = h_i(Z)$ where Z is a standard normal variable, show that $X_1, \ldots$ are uncorrelated but not independent. Show also that the second cumulant of $n^{1/2}\bar{X}$ is exactly one but that the third cumulant does not converge to zero. Construct a similar example in which the first three cumulants converge to those of the standard normal density, but where the central limit theorem does not apply.

2.11 Verify that $\bar{\rho}_4 = p^{-1}\kappa^{i,j,k,l}\kappa_{i,j}\kappa_{k,l}$ is invariant under affine non-singular transformation of X.

2.12 By considering the expression

$$\mathrm{var}\{\kappa_{i,j}X^iX^j - c_iX^i\}$$

with $c_i = \kappa_{i,r}\kappa_{s,t}\kappa^{r,s,t}$, show that $\bar{\rho}_4 \geq \bar{\rho}_{13}^2 - 2$, where $\bar{\rho}_4$ and $\bar{\rho}_{13}^2$ are defined by (2.14) and (2.16). In addition, by examining the expression

$$\int (a + a_ix^i + a_{ij}x^ix^j)^2 f_X(x)\,dx,$$

show that $\bar{\rho}_4 = \bar{\rho}_{13}^2 - 2$ if and only if the joint distribution of X is concentrated on a particular class of conic.

2.13 Show that $\bar{\rho}_{23}^2 \geq 0$ with equality only if $\kappa^{i,j,k} = 0$ identically.

2.14 Show that $\kappa^{ij,kl} - \kappa^{i,j,r}\kappa^{k,l,s}\kappa_{r,s}$, regarded as a $p^2 \times p^2$ symmetric matrix, is non-negative definite. By examining the trace of this matrix in the case where $\kappa^i = 0$, $\kappa^{i,j} = \delta^{ij}$, show that

$$\bar{\rho}_4 \geq \bar{\rho}_{23}^2 - p - 1,$$

with equality if and only if the joint distribution is concentrated on $p+1$ points not contained in any linear subspace of R^p. Deduce that $\bar{\rho}_4 = \bar{\rho}_{23}^2 - p - 1$ implies $\bar{\rho}_4 = \bar{\rho}_{13}^2 - 2$ and that $\bar{\rho}_4 > \bar{\rho}_{13}^2 - 2$ implies $\bar{\rho}_4 > \bar{\rho}_{23}^2 - p - 1$.

2.15 Show that $\bar{\rho}_{13}^2 = 0$ if and only if every linear combination a_iX^i is uncorrelated with the quadratic form $\kappa_{i,j}X^iX^j$. (Take $\kappa^i = 0$.) Show also that $\bar{\rho}_{23}^2 = 0$ if and only if every linear combination a_iX^i is uncorrelated with every quadratic form $a_{ij}X^iX^j$.

2.16 Show that the multinomial distribution with index m and probability vector $\pi_1, \ldots, \pi_k$ has cumulant generating function

$$m\{k(\theta + \xi) - k(\theta)\},$$

where $k(\theta) = \log[\Sigma \exp(\theta_j)]$ and

$$\pi_i = \exp(\theta_i)/\sum \exp(\theta_j).$$

Hence show that the first four cumulants are

$$\begin{aligned}
\kappa^i &= m\pi_i \\
\kappa^{i,j} &= m\{\pi_i\delta_{ij} - \pi_i\pi_j\} \\
\kappa^{i,j,k} &= m\{\pi_i\delta_{ijk} - \pi_i\pi_j\delta_{ik}[3] + 2\pi_i\pi_j\pi_k\} \\
\kappa^{i,j,k,l} &= m\{\pi_i\delta_{ijkl} - \pi_i\pi_j(\delta_{ik}\delta_{jl}[3] + \delta_{jkl}[4]) + 2\pi_i\pi_j\pi_k\delta_{il}[6] \\
&\qquad - 6\pi_i\pi_j\pi_k\pi_l\},
\end{aligned}$$

where, for example, $\delta_{ijk} = 1$ if $i = j = k$ and zero otherwise, and no summation is implied where indices are repeated at the same level.

2.17 Evaluate explicitly the fourth cumulants of the multinomial distribution for the five distinct index patterns.

2.18 Show, for the multinomial distribution with index $m = 1$, that the moments are $\kappa^i = \pi_i$, $\kappa^{ij} = \pi_i \delta_{ij}$, $\kappa^{ijk} = \pi_i \delta_{ijk}$ and so on, where no summation is implied. Hence give an alternative derivation of the first four cumulants in Exercise 2.16.

2.19 For the multinomial distribution with $p = \mathrm{rank}(\kappa^{i,j}) = k - 1$, show that

$$
\begin{aligned}
(k-1)\bar{\rho}_{13}^2 &= m^{-1}\{\textstyle\sum \pi_j^{-1} - k^2\}, \\
(k-1)\bar{\rho}_{23}^2 &= m^{-1}\{\textstyle\sum \pi_j^{-1} - 3k + 2\} \\
(k-1)\bar{\rho}_{4} &= m^{-1}\{\textstyle\sum \pi_j^{-1} - k^2 - 2(k-1)\},
\end{aligned}
$$

and hence that

$$\bar{\rho}_4 = \bar{\rho}_{13}^2 - 2/m = \bar{\rho}_{23}^2 - k/m,$$

showing that the inequalities in Exercises 2.12 and 2.14 are sharp for $m = 1$. Show also that the minimum value of $\bar{\rho}_{23}^2$ for the multinomial distribution is $(k-2)/m$.

2.20 Hölder's inequality for a pair of random variables X and Y is

$$E|XY| \le \{E|X|^p\}^{1/p}\{E|Y|^q\}^{1/q}$$

where $p^{-1} + q^{-1} = 1$. Deduce from the above that

$$\{E|X_1 X_2 \cdots X_r|\}^r \le E|X_1|^r \cdots E|X_r|^r$$

for random variables $X_1, \ldots, X_r$. Hence prove that if the diagonal elements of cumulant tensors are finite then all other elements are finite.

2.21 Using (2.6) and (2.7), express $\kappa^{i,jkl} = \mathrm{cov}(X^i, X^j X^k X^l)$ in terms of ordinary moments and hence, in terms of ordinary cumulants.

2.22 Let $a = a^1, \ldots, a^p$ and $b = b^1, \ldots, b^p$, where $a^j \le b^j$, be the coordinates of two points in R^p and denote by (a, b) the Cartesian product in R^p of the intervals (a^j, b^j). Let $f(x) = f(x^1, \ldots, x^p)$ be a p-dimensional joint density function and define

$$F(a, b) = \int_{x \in (a,b)} f(x)\, dx$$

where $dx = dx^1 \cdots dx^p$. Express $F(a, b)$ in terms of the cumulative distribution function $F(x) \equiv F(-\infty, x)$ evaluated at points with coordinates $x^j = a^j$ or b^j. Comment briefly on the similarities with and differences from (2.9).

2.23 Let a^{ij} be the elements of a square matrix, not necessarily symmetrical, and let its inverse, a_{ij}, satisfy $a^{ij} a_{kj} = \delta^i_k = a^{ji} a_{jk}$, being careful regarding the positions of the indices. Show that the derivatives satisfy

$$
\begin{aligned}
\partial a_{ij} / \partial a^{rs} &= -a_{is} a_{rj} \\
\partial a^{ij} / \partial a_{rs} &= -a^{is} a^{rj}.
\end{aligned}
$$

2.24 Show that if a_i are the components of a vector of coefficients satisfying $a_i a_j \kappa^{i,j} = 0$, then

$$a_i \kappa^{i,j} = 0, \quad a_i \kappa^{i,j,k} = 0, \quad a_i \kappa^{i,j,k,l} = 0$$

and so on. Hence deduce that the choice of generalized inverse has no effect on the scalars derived in Section 2.8.

2.25 Let $X_1, \ldots, X_n$ be independent and identically distributed scalar random variables having cumulants κ_1, κ_2, $\kappa_3, \ldots$. Show that

$$\begin{aligned} 2\kappa_2 &= E(X_1 - X_2)^2 \\ 3\kappa_3 &= E(X_1 + \omega X_2 + \omega^2 X_3)^3 \quad (\omega = e^{2\pi i/3}) \\ 4\kappa_4 &= E(X_1 + \omega X_2 + \omega^2 X_3 + \omega^3 X_4)^4 \quad (\omega = e^{2\pi i/4}) \end{aligned}$$

where $i^2 = -1$. Hence, by writing $\omega = e^{2\pi i/r}$ and

$$r\kappa_r = \lim_{n\to\infty} n^{-1}[X_1 + \omega X_2 + \omega^2 X_3 + \cdots + \omega^{nr-1} X_{nr}]^r,$$

give an interpretation of cumulants as coefficients in the Fourier transform of the randomly ordered sequence $X_1, X_2, \ldots$. Express κ_r as a *symmetric* function of $X_1, X_2, \ldots$, (Good, 1975, 1977).

2.26 Show that if $\Upsilon = \{\upsilon_1, \ldots, \upsilon_\nu\}$ is a partition of the indices $j_1, \ldots, j_n$ and $\omega = e^{2\pi i/n}$ is a primitive nth root of unity, then

$$\sum_j \omega^{j_1+\cdots+j_n}\, \delta(\upsilon_1)\cdots\delta(\upsilon_\nu) = \begin{cases} 0 \text{ if } \Upsilon < 1 \\ n \text{ if } \Upsilon = 1 \end{cases}$$

where the sum extends over all positive integer vectors having components in the range $(1, n)$.

2.27 Let $X_1, \ldots, X_n$ be independent and identically distributed p-dimensional random vectors having cumulants κ^r, $\kappa^{r,s}$, $\kappa^{r,s,t}, \ldots$. Define the random vector $Z_{(n)}$by

$$Z^r_{(n)} = \sum_{j=1}^n X^r_j \exp(2\pi i j/n)$$

where $i^2 = -1$. Using the result in the previous exercise or otherwise, show that the nth-order moments of $Z_{(n)}$ are the same as the nth-order cumulants of X, i.e.

$$E(Z^{r_1}_{(n)} \cdots Z^{r_n}_{(n)}) = \kappa^{r_1,\ldots,r_n},$$

(Good, 1975, 1977). Hence give an interpretation of mixed cumulants as Fourier coefficients along the lines of the interpretation in Exercise 2.25.

2.28 Let $X_1, \ldots, X_n$ be independent χ^2_1 random variables. Show that the joint cumulant generating function is $-\frac{1}{2}\sum_i \log(1 - 2\xi_i)$. Show also that the joint cumulant generating function of $Y_1 = \sum X_j$ and $Y_2 = \sum \lambda_j X_j$ is $-\frac{1}{2}\sum_i \log(1 - 2\xi_1 - 2\lambda_i\xi_2)$. Hence show that the joint cumulants of (Y_1, Y_2) are given by

$$\kappa_{rs} = 2^{r+s-1}(r+s-1)! \sum \lambda_i^s$$

using the notation of Section 2.5.

2.29 Show that if the ratio $R = Y/X$ is independent of X, then the moments of the ratio are the ratio of the moments, i.e.

$$\mu'_r(R) = \mu'_r(Y)/\mu'_r(X).$$

2.30 In the notation of Exercise 2.28, let $R = Y_2/Y_1$. Show that the first four cumulants of this ratio are

$$\begin{aligned} \kappa_1(R) &= \kappa_1(\lambda) \\ \kappa_2(R) &= 2\kappa_2(\lambda)/(n+2) \\ \kappa_3(R) &= 8\kappa_3(\lambda)/\{(n+2)(n+4)\} \\ \kappa_4(R) &= 48\kappa_4(\lambda)/\{(n+2)(n+4)(n+6)\} \\ &\quad + 48n\kappa_2^2(\lambda)/\{(n+2)^2(n+4)(n+6)\} \end{aligned}$$

and explain what is meant by the notation $\kappa_r(\lambda)$.

2.31 Using the notation of Exercises 2.28 and 2.30, show that if the eigenvalues decrease exponentially fast, say $\lambda_j = \lambda^j$, with $|\lambda| < 1$, then nR has a non-degenerate limiting distribution for large n, with cumulants

$$\kappa_r(nR) \simeq 2^{r-1}(r-1)!\lambda^r/(1-\lambda^r).$$

Show that this result is false if λ is allowed to depend on n, say $\lambda_j = 1 - 1/j$.

2.32 Using (2.18), show that if X and Y are independent real-valued random variables, then the variance of the product is

$$\mathrm{var}(XY) = \kappa_{10}^2\kappa_{02} + \kappa_{01}^2\kappa_{20} + \kappa_{20}\kappa_{02}.$$

2.33 Let X_1 and X_2 be independent and identically distributed p-dimensional random variables with zero mean and identity covariance matrix. The spherical polar representation of X is written (R, θ) where θ has $p-1$ components and $R = |X|$ is the Euclidean norm of X. Show that

$$p\bar{\rho}_{13}^2 = E(R_1^3 R_2^3 \cos\theta_{12})$$
$$p\bar{\rho}_{23}^2 = E(R_1^3 R_2^3 \cos^3\theta_{12})$$

where $R_1 R_2 \cos\theta_{12} = X_1^i X_2^j \delta_{ij}$, so that θ_{12} is the random angle between X_1 and X_2. Hence give a geometrical interpretation of these two scalars in the special case where $p = 2$. Show also that

$$4p\bar{\rho}_{23}^2 - 3p\bar{\rho}_{13}^2 = E\{R_1^3 R_2^3 \cos(3\theta_{12})\}$$

and that this quantity is non-negative if $p \le 2$.

2.34 Let X be a scalar random variable and write

$$M_X^{(n)}(\xi) = 1 + \mu_1'\xi + \mu_2'\xi^2/2! + \cdots + \mu_n'\xi^n/n!$$

for the truncated moment generating function. The zeros of this function, $a_1^{-1}, \ldots, a_n^{-1}$, not necessarily real, are defined by

$$M_X^{(n)}(\xi) = (1-a_1\xi)(1-a_2\xi)\cdots(1-a_n\xi).$$

Show that the symmetric functions of the as

$$\langle rs\cdots u\rangle = \sum_{i<j<\cdots<l} a_i^r a_j^s \cdots a_k^u$$

are semi-invariants of X (unaffected by the transformation $X \to X + c$) if and only if the powers $r, s, \ldots, u$ that appear in the symmetric function are at least 2. Show also that the cumulants are given by the particular symmetric functions

$$\kappa_r = -(r-1)!\sum_i a_i^r = -(r-1)! < r > \qquad (r < n).$$

Express the semi-invariant $\langle 22\rangle$ in terms of κ_2 and κ_4. (MacMahon, 1884, 1886; Cayley, 1885).

2.35 Let $\{\epsilon_j\}$, $\{\epsilon'_j\}$ and $\{Z_j\}$, $j = \ldots, -1, 0, 1, \ldots$ be three doubly infinite, mutually independent sequences of independent unit exponential random variables. The bivariate sequence $\{X_j, Y_j\}$, $j = \ldots, -1, 0, 1, \ldots$, defined by

$$X_{j+1} = \begin{cases} X_j - Z_j & \text{if } X_j > Z_j \\ \epsilon_{j+1} & \text{otherwise} \end{cases}$$
$$Y_{j+1} = \begin{cases} Y_j - Z_j & \text{if } Y_j > Z_j \\ \epsilon'_{j+1} & \text{otherwise} \end{cases}$$

is known as a bivariate *exponential recurrence* process. Show that

(i) X_j and Y_j are unit exponential random variables.
(ii) $\text{cov}(X_j, X_{j+h}) = 2^{-h}$ where $h \geq 0$.
(iii) X_i is independent of Y_j.
(iv) X_i is independent of the sequence $\{Y_j\}$.

Hence deduce that all third-order mixed cumulants involving both Xs and Ys are zero. Show also that

$$\text{cum}(X_j, X_{j+1}, Y_j, Y_{j+1}) = 1/12$$

(McCullagh, 1984c).

2.36 In the two-dimensional case, show that the homogeneous cubic form $\kappa^{i,j,k} w_i w_j w_k$ can be written, using power notation, in the form

$$Q_3(w) = \kappa_{30} w_1^3 + \kappa_{03} w_2^3 + 3\kappa_{21} w_1^2 w_2 + 3\kappa_{12} w_1 w_2^2.$$

By transforming to polar coordinates, show that

$$Q_3(w) = r^3 \{\tau_1 \cos(\theta - \epsilon_1) + \tau_3 \cos(3\theta - 3\epsilon_3)\},$$

where

$$16\tau_1^2 = 9(\kappa_{30} + \kappa_{12})^2 + 9(\kappa_{03} + \kappa_{21})^2$$
$$16\tau_3^2 = (\kappa_{30} - 3\kappa_{12})^2 + (\kappa_{03} - 3\kappa_{21})^2.$$

Find similar expressions for ϵ_1 and ϵ_3 in terms of the κs.

2.37 By taking X to be a two-dimensional standardized random variable with zero mean and identity covariance matrix, interpret $Q_3(w)$, defined in the previous exercise, as a directional standardized skewness. [Take w to be a unit vector.] Show that, in the polar representation, $\epsilon_3 - \epsilon_1$ is invariant under rotation of X, but changes sign under reflection. Find an expression for this semi-invariant in terms of κ_{30}, κ_{03}, κ_{21} and κ_{12}. Discuss the statistical implications of the following conditions:

(i) $4\rho_{23}^2 - 3\rho_{13}^2 = 0$;
(ii) $\rho_{13}^2 = 0$;
(iii) $\epsilon_3 - \epsilon_1 = 0$.

2.38 In the two-dimensional case, show that the homogeneous quartic form $\kappa^{i,j,k,l} w_i w_j w_k w_l$ can be written, using power notation, in the form

$$Q_4(w) = \kappa_{40} w_1^4 + \kappa_{04} w_2^4 + 4\kappa_{31} w_1^3 w_2 + 4\kappa_{13} w_1 w_2^3 + 6\kappa_{22} w_1^2 w_2^2.$$

By transforming to polar coordinates, show that

$$Q_4(w) = r^4 \{\tau_0 + \tau_2 \cos(2\theta - 2\epsilon_2) + \tau_4 \cos(4\theta - 4\epsilon_4)\}.$$

Show that $8\tau_0 = 3\kappa_{40} + 3\kappa_{04} + 6\kappa_{22}$ is invariant under orthogonal transformation of X. Find similar expressions for τ_2, τ_4, ϵ_2 and ϵ_4 in terms of the κs.

2.39 By taking X to be a two-dimensional standardized random variable with zero mean and identity covariance matrix, interpret $Q_4(w)$, defined in the previous exercise, as a directional standardized kurtosis. Taking ϵ_1 as defined in Exercises 2.36 and 2.37, show, using the polar representation, that that $\epsilon_2 - \epsilon_1$ and $\epsilon_4 - \epsilon_1$ are both invariant under rotation of X, but change sign under reflection. Find expressions for these semi-invariants in terms of the κs. Interpret τ_0 as the mean directional kurtosis and express this as a function of $\bar{\rho}_4$. Discuss the statistical implications of the following conditions:

(i) $\tau_0 = 0$;
(ii) $\tau_2 = 0$;
(iii) $\tau_4 = 0$;
(iv) $\epsilon_4 - \epsilon_2 = 0$.

2.40 In the one-dimensional case, the functions of the cumulants of X that are invariant under affine transformation are

$$\frac{\kappa_{2r}}{\kappa_2^r}, \quad \frac{\kappa_{2r+1}^2}{\kappa_2^{2r+1}}, \quad r = 1, 2, \ldots$$

using power notation. All other invariants can be expressed as functions of this sequence. By extending the results described in the previous four exercises, describe the corresponding complete list of invariants and semi-invariants in the bivariate case.

2.41 Discuss the difficulties encountered in extending the above argument beyond the bivariate case.

2.42 *Spherically symmetric random variables*: A random variable X is said to be spherically symmetric if its distribution is unaffected by orthogonal transformation. Equivalently, in spherical polar coordinates, the radial vector is distributed independently of the angular displacement, which is uniformly distributed over the unit sphere. Show that the odd cumulants of such a random variable are zero and that the even cumulants must have the form

$$\kappa^{i,j} = \tau_2 \delta^{ij}, \quad \kappa^{i,j,k,l} = \tau_4 \delta^{ij}\delta^{kl}[3], \quad \kappa^{i,j,k,l,m,n} = \tau_6 \delta^{ij}\delta^{kl}\delta^{mn}[15]$$

and so on, for some set of coefficients τ_2, $\tau_4, \ldots$. Show that the standardized cumulants are

$$\bar{\rho}_4 = \tau_4(p+2)/\tau_2^2, \quad \bar{\rho}_6 = \tau_6(p+2)(p+4)/\tau_2^3$$

and hence that $\tau_4 \geq -2\tau_2^2/(p+2)$.

2.43 For the previous exercise, show that

$$\tau_6 \geq -\tau_2\tau_4 - \tau_2^3 + \frac{3(\tau_4 + \tau_2^2)^2}{(p+4)\tau_2}.$$