EDGEWORTH-CORRECTED SMALL-SAMPLE CONFIDENCE INTERVALS FOR RATIO PARAMETERS IN LINEAR REGRESSION

by

Kamanzi-wa-Binyavanga

Dissertation presented for the degree of Doctor of Philosophy in the Department of Statistics and Actuarial Science, Faculty of Economic and Management Sciences at the University of Stellenbosch.

March 2002

DECLARATION

I, the undersigned, hereby declare that the work contained in this dissertation is my original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Kamanzi-wa-Binyavanga
March 2002

ABSTRACT

In this thesis we construct a central confidence interval for a smooth scalar non-linear function of parameter vector β in a single general linear regression model $Y = X\beta + \varepsilon$. We do this by first developing an Edgeworth expansion for the distribution function of a standardised point estimator. The confidence interval is then constructed in the manner discussed. Simulation studies reported at the end of the thesis show the interval to perform well in many small-sample situations.

Central to the development of the Edgeworth expansion is our use of the index notation which, in statistics, has been popularised by McCullagh (1984, 1987).

The contributions made in this thesis are of two kinds. We revisit the complex McCullagh Index Notation, modify and extend it in certain respects as well as repackage it in the manner that is more accessible to other researchers.

On the new contributions, in addition to the introduction of a new small-sample confidence interval, we extend the theory of stochastic polynomials (SP) in three respects. A method, which we believe to be the simplest and most transparent to date, is proposed for deriving cumulants for these. Secondly, the theory of the cumulants of the SP is developed both in the context of Edgeworth expansion as well as in the regression setting. Thirdly, our new method enables us to propose a natural alternative to the method of Hall (1992a, 1992b) regarding skewness-reduction in Edgeworth expansions.

OPSOMMING

In hierdie proefskrif word daar aandag gegee aan die konstruksie van 'n sentrale vertrouensinterval vir 'n gladde skalare nie-lineêre funksie van die parametervektor β in 'n enkele algemene lineêre regressiemodel $Y = X\beta + \varepsilon$.. Dit behels eerstens die ontwikkeling van 'n Edgeworth uitbreiding vir die verdelingsfunksie van 'n gestandaardiseerde puntberamer. Die vertrouensinterval word dan op grond van hierdie uitbreiding gekonstrueer. Simulasiestudies wat aan die einde van die proefskrif gerapporteer word, toon dat die voorgestelde interval goed vertoon in verskeie klein-steekproef gevalle.

Die gebruik van indeksnotasie, wat in die statistiek deur McCullagh (1984, 1987) bekendgestel is, speel 'n sentrale rol in die ontwikkeling van die Edgeworth uitbreiding.

Die bydrae wat in hierdie proefskrif gemaak word, is van 'n tweërlei aard. Die ingewikkelde Indeksnotasie van McCullagh word ondersoek, aangepas en ten opsigte van sekere aspekte uitgebrei. Die notasie word ook aangebied in 'n vorm wat dit hopelik meer toeganklik sal maak vir ander navorsers.

Betreffende die bydrae wat gemaak word, word 'n nuwe klein-steekproef vertrouensinterval voorgestel, en word die teorie van stogastiese polinome (SP) ook in drie opsigte uitgebrei. 'n Metode word voorgestel om die kumulante van SP'e af te lei. Ons glo dat hierdie metode die duidelikste en eenvoudigste metode is wat tot dusver hiervoor voorgestel is. Tweedens word die teorie van die kumulante van SP'e ontwikkel binne die konteks van Edgeworth uitbreidings, sowel as die konteks van regressie. Derdens stel ons nuwe metode ons in staat om 'n natuurlike alternatief voor te stel vir die metode van Hall (1992a, 1992b) vir die vermindering van skeefheid in Edgeworth uitbreidings.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation for the support and encouragement that I received during the course of this work from the following: Professor J.S. Maritz (Promoter), Professor S.J. Steel (Co-Promoter) and Professor T. de Wet of the Department of Statistics and Actuarial Science, University of Stellenbosch. Very deep gratitude is also expressed to Professor Steel for generously getting involved in the simulation study that is carried out in Chapter seven. He programmed the software and executed the programmes on his RS 6000 IBM model 43 P computer. Dr L van der Merwe of the same department helped immensely during the editorial phase. This assistance is deeply acknowledged.

Furthermore, I sincerely wish to acknowledge the willingness of the Department to accept a thesis proposal which did not emanate from the research programmes of the individual members of the Department.

In the last two years, I have received full financial sponsorship from the United States Agency for International Development under bursary USTELP/STELLE/0371/2000, administered by the South African Institute of Race Relations. This support is kindly acknowledged.

Lastly, during the course of work of this kind, sacrifices are inevitably made. In this connection, I wish to thank and dedicate this work to my wife, Sylvia, my two daughters, Conchita and Kangyeyo, and my son, Brian, for their perseverance.

GLOSSARY OF NOTATION AND SYMBOLS

$E[\cdot]$	Expected value of $[\cdot]$
$\Pr\{\cdot\}$	probability of $\{\cdot\}$
$\lvert c \rvert$	absolute value of scalar c
$\lVert V \rVert$	Euclidean norm of vector V
$\rightarrow$	"tends to" operator
$\xrightarrow{d}$	convergence in distribution operator
$\xrightarrow{p}$	convergence in probability operator
$a_n = o(b_n)$	means $a_n/b_n \rightarrow 0$ as $n \rightarrow \infty$
$a_n = 0(b_n)$	means a_n/b_n is bounded as $n \rightarrow \infty$
$X_n = o_p(b_n)$	means for every $\varepsilon > 0$, $\lim_{n\rightarrow\infty} \Pr\{\lvert X_n/b_n \rvert \leq \varepsilon\} = 1$
$X_n = 0_p(b)$	means for every $\varepsilon > 0$ and every $\eta > 0$ there exists an integer $n_0 = n_0(\varepsilon, \eta)$ such that if $n \geq n_0$, then $\Pr\{\lvert X_n/b_n \rvert < \varepsilon\} \geq 1 - \eta$
$\hat{\beta}(n)$	least squares estimate of β based on regression sample of size n
$f_{r_1 r_2 \ldots r_v}$	partial derivative of $f\left(\hat{\beta}(n)\right)$ with respect to $\hat{\beta}_r(n)$, $\hat{\beta}_{r_2}(n)$, ..., $\hat{\beta}_{r_v}(n)$ evaluated at $\hat{\beta}(n) = \beta$
$N(a, b)$	normal distribution with mean a and variance b

GLOSSARY OF ABBREVIATIONS

c.g.f.	cumulant generating function
d.f.	distribution function
GNF	general nonlinear function $f(\beta)$
iid	independent and identically distributed
LHS	left-hand side
LSE	least squares estimate/estimator
mgf	moment generating function
p.d.	positive definite
p.d.f.	probability density function
r.v.	random variable/vector
RHS	right-hand side
s.t.	such that
w.r.t.	with respect to

CONTENTS

CHAPTER ONE

INTRODUCTION

1.1 EXAMPLES OF RATIO PARAMETERS IN LINEAR REGRESSION

In statistical practice ratio parameters in the linear regression setting arise in several diverse applications, and it is perhaps appropriate to begin with a few specific Examples illustrating the problem to be considered. The meaning of symbols, notations, and abbreviations when not provided in the text, should be checked in the glossary provided.

Example 1.1: RATIO β_0/β_1 IN STRAIGHT-LINE REGRESSION

A model that sometimes arises in animal population size estimation studies is the following:

$$E[Y] = \gamma(\theta - x) \tag{1.1}$$

where θ is the population size measure of interest. An application of (1.1) appears in De Lury (1947) in a study of a Lobster population; see Section 2.5 of Chapter two for more details.

Given the way that the two unknowns γ and θ appear in (1.1), one way to find θ is to seek a solution satisfying the condition

$$E[Y] = 0. \tag{1.2}$$

Rewriting (1.1) as

$$E[Y] = \beta_0 + \beta_1 x \tag{1.3}$$

we see that such solution is given by

$$\theta = -\beta_0/\beta_1 \tag{1.4}$$

Example 1.2 INVERSE ESTIMATION IN LINEAR REGRESSION

A common purpose of most determinations of regression equations is to use them to derive estimates. For other uses of regression see e.g. Montgomery and Peck (1982, pp. 5–6). In connection with the estimation applications, we often either seek the estimates of the dependent values, given values of all the others (*direct estimation/prediction*) or estimates of the independent, given the dependent variables (*inverse estimation/prediction*).

In the case of linear regression with one independent variable, the inverse prediction problem based on, say (1.1), involves the estimation of x_0 given by:

$$x_0 = \{y_0 - \beta_0\}/\beta_1, \qquad y_0 \text{ given} \tag{1.5}$$

The problem arises in calibration experiments and for this reason the term *calibration* is sometimes used in the place of inverse estimation/prediction. The reader is, however, referred to Williams (1969) for a more formal interpretation of calibration.

Example 1.3: THE TURNING POINT OF A PARABOLIC REGRESSION FUNCTION

For the parabolic regression model

$$E[Y] = \beta_1 + \beta_2 x + \beta_3 x^2 \tag{1.6}$$

it is known that the co-ordinates of extremal point, say $(x_m, E[Y]_m)$, are given by

$$\begin{aligned} x_m &= -\beta_2/2\beta_3 \\ E[Y]_m &= \beta_1 - \frac{\beta_2^2}{4\beta_3} \end{aligned} \tag{1.7}$$

Example 1.4: THE CHANGE-OVER VALUE OF A TWO-PHASE REGRESSION LINE

In the description of a nonlinear relationship between X and Y, a natural competitor to (1.6) is a model represented by two straight lines, one being appropriate when x takes

values below and the other when x takes values above a certain fixed but often unknown value corresponding to the intersection. Following Sprent (1961), the intersection phase in such regressions is referred to as the *change-over point* and the value of x at which this occurs is called the *changeover value.*

Let

$$\begin{array}{ll} E[Y] = \alpha_1 + \beta_1 x & x \leq \theta \\ E[Y] = \alpha_2 + \beta_2 x & x > \theta\,. \end{array} \tag{1.8}$$

Then the change-over value θ is given by

$$\theta = (\alpha_1 - \alpha_2) / (\beta_1 - \beta_2) \tag{1.9}$$

Situations in which two-phase regressions arise include applications of a treatment having an immediate stimulating or inhibiting effect, the onset of disease resulting in reduced growth rate and so on. It is also likely that two-phase regressions arise in economic and industrial production problems.

As is generally the case when selecting models for statistical examination, the analyst will postulate a two-phase model rather than some alternative such as (1.6), largely on intuitive grounds, guided by, say, experience and common sense.

Example 1.5: TWO-SIDED CONFIDENCE INTERVAL AS A TEST FOR PARALLELISM

As is well-known, confidence intervals for a parameter θ, say, are closely related to tests of the simple null hypothesis $H_0 : \theta = \theta_0$ against the two-sided alternative $H_a : \theta \neq \theta_0$. Consider two regression models

$$E[Y] = \alpha_j + \beta_j x \qquad j = 1, 2\,. \tag{1.10}$$

Then a confidence interval for β_1/β_2 is equivalent to testing hypotheses $H_0 : \beta_1 = \lambda\beta_2$ for a range of values of λ.

1.2 PROBLEM-SETTING

1.2.1 MODEL SPECIFICATION

We consider the following regression model:

$$Y_i = C_i^T \beta + \varepsilon_i\,, \qquad i = 1, 2, \ldots, n \tag{1.11}$$

where ε_i, hereafter called *model errors*, are random variables, β is a $p \times 1$ vector $(\beta_0, \beta_1, ..., \beta_{p-1})$ of unknown constants with $p < n$ and with $q = 1 - p$

$$\begin{aligned} C_i^T &= (1, c_{i1} - \bar{c}_{\cdot 1}, ..., c_{iq} - \bar{c}_{\cdot q}) \\ \bar{c}_i &= n^{-1} \textstyle\sum_{j=1}^{q} c_{ij}, \qquad i = 1, ..., n \end{aligned} \tag{1.12}$$

the c_{ij} being known constants which are *deliberately* chosen by the researcher/analyst. We will refer to the C_i as the *corrected covariate vectors* or simply the covariate vectors, while (1.11) is commonly known as the *deviations-from-the-sample-average* form for the general linear model. In the later part of this thesis, we will comment on the reason for choosing to work with this form of model rather than the one with the design points uncorrected for their means. Writing

$$\begin{aligned} Y &= (Y_1, Y_2, ..., Y_n)^T \\ X &= (C_1, C_2, ..., C_n)^T \\ \varepsilon &= (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)^T \end{aligned} \tag{1.13}$$

we have the following usual matrix form of (1.11)

$$Y = X\beta + \varepsilon\,. \tag{1.14}$$

Several additional specifications accompany the above model. They are as follows:

C.1.1 MODEL ERRORS

(a) Model errors ε_i are all iid;

(b) The common distribution of the errors, hereafter denoted by G, is not necessarily normal but must have zero mean. As is pointed out by Kendall & Stuart (1973, p. 80) this zero-mean assumption is less restrictive than it appears. The essential assumption is that the ε_i have *equal* means.

The zero mean requirement serves at least two purposes. Firstly, it makes the least squares estimate of β unbiased and secondly it, together with C.1.1(a) imply that

$$\begin{aligned} E\left[Y_i | C_i^T\right] &= E\left[C_i^T\beta + \varepsilon_i | C_i^T\right] \\ &= E\left[C_i^T\beta \,|\, C_i^T\right] + E\left[\varepsilon_i \,|\, C_i^T\right] && \text{(1.15a)} \\ &= C_i^T\beta\,. && \text{(1.15b)} \end{aligned}$$

In moving from (1.15a) to (1.15b) we use the fact that since the ε_i's are selected independently of C_i^T, the expectation $E\left[\varepsilon_i | C_i^T\right]$ equals the unconditional $E\left[\varepsilon_i\right]$. For the purposes of simulation, useful p.d.f.'s meeting the zero mean criterion include:

(i) the central student-t distribution

(ii) the standard Laplace p.d.f., Chapter 22 of Johnson and Kotz (1970)

(iii) the exponential distribution

$$\sigma^{-1}\exp[-\sigma^{-1}(\varepsilon+\sigma)], \quad \text{for } \varepsilon > -\sigma,\ \sigma > 0$$

and zero elsewhere, Johnson & Kotz (1970, Chapter 18).

Distribution (iii) is asymmetrical with the traditional coefficient of skewness of 2 for all values of σ.

C.1.2 MOMENTS LEMMA

We will require that for every υ belonging to $(0,\infty)$.

$$\sum_{r=1}^{\infty} \Pr\left\{|\varepsilon_i| \geq r^{\frac{1}{\upsilon}}\right\} < \infty\,. \tag{1.16}$$

This requirement ensures that the υ-th moment of ε_i exists as do all the lower moments, see e.g. Laha & Rohatgi (1979, p. 38). Hereafter

$$\mu_\upsilon = E\left[\{\varepsilon_i\}^\upsilon\right]\,. \tag{1.17}$$

C.1.3 VECTOR β

Parameter vector β is assumed to be a member of a parameter vector space $\Omega \subset \mathbb{R}^P$ of all the linear combinations of the columns of X. Its point estimator, denoted by $\hat{\beta}$ is any vector in Ω which minimizes the error sum of squares

$$SS(\beta) = (Y - X\beta)^T (Y - X\beta)\,. \tag{1.18}$$

As is well-known, if X^TX is nonsingular then

$$\hat{\beta} = \left(X^TX\right)^{-1} X^TY \tag{1.19}$$

and we will use the common name of least squares estimator (LSE) to refer to (1.19).

Given our relaxation of the usual normality assumption for the error distribution, the universal suitability of (1.19) may be in question. To guard against the undesirable effects of the departure from the normality assumption we shall assume the absence of outliers in the regression data.

C.1.4 DESIGN MATRIX

We require that the design matrix X be of full rank p. Furthermore we make the following assumption concerning the notational *sequence* $\{X_n\}$ of model matrices, namely

$$\lim_{n\to\infty} n^{-1} X_n^T X_n = \Sigma\,, \qquad \text{a p.d. matrix} \tag{1.20}$$

For an example where $X_n^T X_n$ is p.d. but limiting matrix Σ is not, see Judge, et al. (1985, p. 145).

X having full rank ensures that $X^T X$ is nonsingular, and we will see that condition (1.20) and Corollary 4.2.1 are the only conditions that we need to be able to construct an Edgeworth expansion from the cumulants of $\hat{\beta} - \beta$.

1.2.2 MODEL APPLICATION STUDIED AND ASSOCIATED CONDITIONS

We seek to construct an equal-tailed approximate confidence interval for a nonlinear function

$$\theta = f(\beta) \tag{1.21}$$

where this interval is suitable for small-to-moderate values of n. Besides nonlinearity, the following are additional conditions that we place on $f(\cdot)$.

C.1.5 $f(\cdot)$ belongs to Class C^∞ functions;

C.1.6 Let D be a convex subset of Ω. We require that every β belonging to D have a δ-neighbourhood

$$S\left(\hat{\beta}, \delta\right) = \left\{\hat{\beta} : ||\hat{\beta} - \beta|| < \delta,\ \delta > 0\right\} \tag{1.22}$$

in which for every v

$$\left|f_{r_1 r_2 \ldots r_v}\left(\hat{\beta}\right)\right| \leq c, \quad 0 < c < \infty. \tag{1.23}$$

Under C.1.5 and C.1.6 we then have that:

C.1.7 An infinite Taylor series of $f(\hat{\beta})$ at $\hat{\beta} = \beta$ converges to $f\left(\hat{\beta}\right)$ for every $\hat{\beta}$ in (1.22).

This Taylor series has the form

$$\begin{aligned} f\left(\hat{\beta}\right) &= \theta + \sum_r f_r Z_r + \frac{1}{2!}\sum_s \sum_r f_{rs} Z_r Z_s \\ &\quad + \frac{1}{3!}\sum_t \sum_s \sum_r f_{rst} Z_r Z_s Z_t \\ &\quad + \frac{1}{4!}\sum_u \sum_t \sum_s \sum_r f_{rstu} Z_r Z_s Z_t Z_u \\ &\quad + \\ &\quad \vdots \end{aligned} \tag{1.24}$$

where all summations run from zero to $p - 1$.

In the above Taylor expansion

$$L = \left(X^T X\right)^{-1} X^T \tag{1.25}$$

$$Z = \hat{\beta} - \beta \tag{1.26a}$$

$$= L\varepsilon \tag{1.26b}$$

At this point, some remarks concerning the above stochastic Taylor series are in order.

Remark 1.1

Our choice of $f\left(\hat{\beta}\right)$ as the point estimator of $f(\beta)$ is based on the principle of *functional equivalence*; see Barnard (1974, p. 5). According to this principle, if one's estimate of ψ is $\hat{\psi}$ and $h(\cdot)$ is a smooth function, then one's estimate of $h(\psi)$ is $h\left(\hat{\psi}\right)$. We will call such estimates *equivariant* estimates.

Remark 1.2

Identity (1.26b) is critical to our problem. It shows how the distributional properties of Z can be derived from the corresponding properties of the model-error vector. The identity also demonstrates the very complicated nature of the stochastic Taylor series

Remark 1.3

Notwithstanding the second feature of Remark 1.2, some welcome simplification does occur when $f(\cdot)$ is a simple ratio of the kind examined in Examples 1.1 to 1.5. This is because in this instance all the higher-order partial derivatives w.r.t. the numerator vanish, starting as early as the second derivative in the case of a ratio such as β_r/β_s, $r \neq s$.

It is this feature of the simple ratio parameters, together with the fact that this class of parameters arises frequently in regression, that have largely influenced our choice for the nonlinear aspect of our thesis title. However, as will be seen throughout, the theory developed in this thesis applies to any (scalar) nonlinear function of β that fulfils C.1.5, C.1.6, and C.1.7. Hereafter, when dimension p is arbitrary but, of course, finite, we will sometimes refer to $f(\beta)$ simply as a general nonlinear function or we will use the abbreviation GNF.

1.2.3 PROBLEM SOLUTION STRATEGY

(a) GENERAL OVERVIEW

As is well-known, an interval estimate is often more useful than just a point estimate. Taken together, the point estimate and interval estimate provide us the "best" guess for the unknown value, and how far in error that guess may reasonably be.

Let

$$\hat{\theta}(n) = f\left(\hat{\beta}\right), \quad \hat{\beta} = \hat{\beta}(n). \tag{1.27}$$

A common approach to confidence interval construction uses a pivotal quantity. A function, say W_n, of both sample data and an unknown parameter is said to be *exactly pivotal* if it has the same distribution for all values of the unknowns. It is said to be *asymptotically pivotal* if its limiting distribution is independent of the unknowns. In relation to our problem, one common candidate member of the family of such functions has the form

$$W_n = \left\{\hat{\theta}(n) - \theta\right\} / \upsilon_n \tag{1.28}$$

for some choice of the scaling factor υ_n^{-1}.

Confidence intervals for θ may be obtained by referring to the percentiles of the d.f. of (1.28) hereafter denoted by G_n, in the following way:

For $0 < \alpha < \frac{1}{2}$, let w_α and $w_{1-\alpha}$ respectively denote the 100α-th and $100(1-\alpha)$-th percentiles of G_n. If υ_n were a known constant and if G_n were independent of θ, one would obtain from the inequality

$$w_\alpha \leq W_n \leq w_{1-\alpha} \tag{1.29}$$

a unique pair of numbers, say $\hat{\theta}_L$ and $\hat{\theta}_U$, such that the probability, say $1-q$, of one being correct in stating that

$$\theta \in \left[\hat{\theta}_L,\ \hat{\theta}_U\right], \quad \text{for all } \theta \tag{1.30}$$

whenever we observe the sample

$$S_O = \left\{C_i^T,\ Y_i\right\}_{i=1}^n \tag{1.31}$$

is equal to a given value $1-2\alpha$ chosen in advance. We call $1-q$ the *coverage probability* and $1-2\alpha$ the *confidence coefficient or confidence level* of interval (1.30). When the coverage probability coincides with the confidence coefficient for all n, the associated confidence interval is said to be *exact*. When for some n there is a mismatch, the confidence interval is termed

approximate. Two conditions lead to (1.30) being exact. The first one is that both v_n and G_n must be known and the second condition is that G_n must be independent of θ for all n. When at least one of these factors does not hold, the percentile points w_α and $w_{1-\alpha}$ must be estimated. This estimation may cause a finite-sample confidence interval to be approximate, with the magnitude of the difference between $1-q$ and $1-2\alpha$ depending on four factors:

(a) the value of n;

(b) the nature of v_n;

(c) the method of estimating v_n and G_n, if unknown;

(d) the method used to extract the percentile points.

(b) STRATEGY FOLLOWED

We will be proceeding in the following way.

(i) CHOICE OF v_n

Since $f(\cdot)$ is nonlinear and since the model errors are not necessarily normally distributed, the first consideration in our task is the choice of v_n.

Suppose that there are conditions such that $n^{\frac{1}{2}}\left\{\hat{\theta}(n)-\theta\right\}$ is asymptotically $N(0,\kappa_{2,\infty})$ then we will choose

$$v_n = n^{-\frac{1}{2}}\kappa_{2,\infty}^{\frac{1}{2}} \tag{1.32}$$

and then replace $\kappa_{2,\infty}$ by its equivariant estimate at the last step of estimating w_α and $w_{1-\alpha}$.

(ii) APPROXIMATION OF G_n

The strategy that we will adopt in approximating G_n proceeds in two main steps as well as invokes a rule of thumb concerning the adequacy of relying on the first four cumulants in an exercise of this kind. First we represent G_n as an asymptotic series. To motivate this first step, we have that under certain conditions which will be seen in Chapter two

$$\lim_{n\to\infty} G_n(w) = \Phi(w) \tag{1.33}$$

where $\Phi(\cdot)$ denotes the standard normal d.f. Hence one may write

$$G_n(w) = \Phi(w) + R_n(w) \tag{1.34}$$

where $R_n(w)$ is small for large n. Then by further analysis of $R_n(w)$, we find a higher-order approximation than Φ which we use in our task as described in Chapter six.

In connection with the study of approximation of distributions and for the purposes intended in this thesis, it has been found during the early part of the last century that the most useful way to express $R_n(w)$ is in the *asymptotic series form*

$$R_n(w) = n^{-\frac{1}{2}}p_1(w)\phi(w) + n^{-1}p_2(w)\phi(w) + n^{-\frac{3}{2}}p_3(w)\phi(w) + \cdots \tag{1.35}$$

where $\phi(w)$ is the first derivative of $\Phi(\cdot)$ and functions $p_v(w)$ contain cumulants and cumulant products as the coefficients of Hermite polynomials in w. Following its introduction by Edgeworth (1905), the identity

$$G_n(w) = \Phi(w) + n^{-\frac{1}{2}}p_1(w)\phi(w) + n^{-1}p_2(w)\phi(w) + \cdots \tag{1.36}$$

is referred to as an *Edgeworth expansion for* G_n. Conceptually we think of (1.36) as starting *after the normal approximation*, so that $n^{-\frac{1}{2}}p_1(w)\phi(w)$ is the *first term* rather than the second term.

Again in connection with the study of approximation of distributions, it has been found (see e.g. Pearson (1963); Pearson & Tukey (1965)) that there is often a remarkable similarity *in shape* between distributions of different functional forms, whose p.d.f.'s are given *the first four moments* (and hence the first four cumulants). While from the strictly mathematical point of view the correspondence cannot be regarded as exact, Pearson & Hartley (1976, p. 75) observe that for many practical purposes the agreement may be satisfactory and in so far as there is agreement the choice of function can be based on *practical convenience.*

It turns out, see Chapter six, that the first term in (1.36) is due entirely to *skewness* while the second term is due to the *main effect of kurtosis* and the *secondary effect of skewness.* Consequently, usually not more than the first two terms will suffice to give a better approximation of G_n than the one provided by Φ.

The second consideration when approximating G_n using expansion (1.36) is the manner of estimating the unknown constants in the $p_v(w)$'s. As the expansion presently stands every $p_v(w)$ is independent of n so that each d.f. G_n in a sequence of such d.f.'s may be approximated by a partial sum of the RHS of (1.36) and the errors satisfy the condition

$$|R_n(w) - \textstyle\sum_{v=1}^{r} n^{-\frac{v}{2}}p_v(w)\phi(w)| \leq n^{-\frac{1}{2}(r+1)}C_r(w), \text{ say} \tag{1.37}$$

where $C_r(\cdot)$ is free of n. When property (1.37) holds, we say that expansion (1.36) is *valid to* r *terms. Therefore, when estimating* $p_v(\cdot)$ by $\hat{p}_v(\cdot)$, say, ideally this should be done in such a way as to recover the property. This aspect is, however, not pursued in the thesis.

An Edgeworth expansion in which all the unknown constants have been replaced by their corresponding estimators is said to be *empirical.*

1.3 WHEN DO SMALL SAMPLES ARISE?

While it is difficult to give an exhaustive account of regression situations which give rise to small samples, some common circumstances that may bring this about relate to two factors: data collection expense or effort; and statistical modelling imperatives.

1.3.1 DATA COLLECTION EFFORT

Sometimes a regression sample is small because the associated data collection is expensive or requires great effort or both. This is particularly the case in one kind of the so-called *Calibration Problem.* According to Williams (1969, pp. 17–18), the term calibration is applied to two distinct but related activities: first, the relating of a measurement made according to a non-standard technique to that which would be expected according to a standard technique, and second, the standardization or graduation of one technique or instrument against another, as for instance in the calibration of thermometers or other measuring instruments. Let us refer to these as *calibration problems of the first and second kind, respectively.*

The problem of calibration of the first kind occurs when a physical characteristic can be measured in two ways: first by a precise method requiring great skill and expense or by a second method involving an easily measured surrogate response. The analyst obtains n pairs of measurements $\{x_i, y_i\}$, $i = 1, ..., n$, the x_i data representing the precisely measured response and the y_i data being the surrogate or easily measured response.

On assuming that all the random variability may be assigned to the y_i data and if the straight line model is appropriate, the calibration problem of the first kind becomes the inverse estimation problem represented by Example 1.2.

It is likely then that in this kind of calibration the sample size may be small.

1.3.2 MODELLING IMPERATIVES

In some instances, the analyst finding that a regression model developed is appropriate in every other respect except its application to the full data, may choose to restrict the region of validity to only

a subset of the data, and this action may result in a small sample. An interesting example of this situation is an application of model (1.1) considered by De Lury (1947). Here the full data calls for a parabolic model yet the problem at hand is modelled by the author to employ a straight line model. Upon making the necessary restriction, the author remains with seventeen of the original forty-two (full)data points. More details of this problem are reviewed in Chapter two.

1.4 SUMMARY OF THIS CHAPTER AND THE REST OF THE THESIS

This chapter has been an introduction. In it, we have stated the problem to be tackled as well as the underlying accompanying conditions. Important Examples were given showing how the problem often arises in the real world. A sketch is outlined showing the strategy framework that guides the work done in this thesis. We point out that the theory is developed for any smooth nonlinear scalar function of β. However, because ratio parameters arise in practice often and because some aspects of our theory will be demonstrated using simple ratio parameters, the nonlinear aspect of the thesis title has been chosen to highlight this.

In Chapter two, we give an overview of the alternative approaches and methods that have appeared in the literature, concerning the estimation of smooth nonlinear functions of β.

In order to be able to develop an Edgeworth expansion for use in the construction of the confidence interval for θ, one must have, among others, explicit knowledge of the cumulants (or moments) of $\hat{\theta}(n) - \theta$. Given the complicated nature of this r.v., we have chosen to obtain these descriptive constants through the use of the special index notation of McCullagh (1984, 1987). However, it is our suspicion that familiarity with this notation is not widespread among statisticians. Furthermore, during the course of our learning the notation, we have found some aspects of it in need of further clarification or amendment. Consequently, we have devoted the whole of Chapter three to describing this notation.

The actual determination of cumulants and accompanying conditions starts in Chapter four and is completed in Chapter five. Chapter four determines the cumulants of random vector Z (recall that the latter denotes $\hat{\beta}(n) - \beta$) and also discusses certain important conditions. Chapter five is devoted solely to the cumulants of $\hat{\theta}(n) - \theta$. Introduced there is a simple method which, in conjunction with

McCullagh's work in this regard, goes a long way to bring the much-needed transparency into this otherwise difficult problem.

Chapter six constructs an Edgeworth expansion and the confidence interval that we seek.In the case of the former, no attempt is made to discuss the asymptotic convergence of the expansion obtained. In the practical applications considered, it is of little value to know the convergence properties of the expansions. What is sought, is to demonstrate, through simulation study, that a small number of terms suffice to give a good approximate confidence interval. If we show this to be the case, it should not be of much concern whether the Edgeworth series is convergent or not.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

In this chapter we review the different approaches and methods that have appeared in the literature concerning confidence interval estimation of $f(\beta)$. We proceed in two phases. In the first phase we review the methods that apply when $f(\beta)$ is a GNF and in the second phase we look at those methods that are tailored for some special forms of $f(\beta)$. In each phase we distinguish between exact and approximate methods.

2.2 METHODS FOR GNF

2.2.1 EXACT

For the general linear model (1.11) with normally distributed errors (and uncentred design points), Durand (1954) has discussed the use of the joint confidence region of the β_v's, an ellipsoid in the p-dimensional space, for the construction of confidence intervals for linear functions of the regression coefficients. He points out that the chosen confidence coefficient, corresponding to the ellipsoid, is a lower bound for the joint confidence of any set of intervals thus derived.

Mandel (1958) generalises this procedure by removing the restriction of linearity. The form of function is arbitrary but known.

2.2.2 APPROXIMATE

(a) THE DELTA METHOD

Asymptotic normality is a widely applied procedure for obtaining asymptotically pivotal quantities, see e.g. Barndorff-Nielsen & Cox (1989, p. 30). In particular, when the r.v. involved is a nonlinear function, asymptotic normality is often established by local linearisation, typically via the mean-value theorem. Although this procedure is commonly known as the delta method, in Chapters six we will have occasion to review other interpretations of the delta method in the context of Edgeworth expansions.

Now, the asymptotic normality of the statistic

$$T_n = n^{\frac{1}{2}} \left\{ \hat{\theta}(n) - \theta \right\} \tag{2.1}$$

is easily established by appealing to the proof of Corollary A.14.17 of Bickel & Doksum (1977, pp. 461–2). This proof rests entirely on the Slutsky's Theorem as well as the Mean Value theorem for smooth functions. Under conditions $\mu_2 > 0$ and (1.20), Lemma 3.1 of Miller (1974) states that

$$n^{\frac{1}{2}} Z \xrightarrow{d} N_p \left(\underset{\sim}{0},\ \mu_2 \sum^{-1} \right). \tag{2.2}$$

Then applying reasoning similar to that of Bickel and Doksum we have

$$\begin{aligned} T_n &\xrightarrow{d} N(0,\ \kappa_{2,\infty}) \\ \kappa_{2,\infty} &= \mu_2 \underset{\sim}{f'}^T \sum^{-1} \underset{\sim}{f'} \\ \underset{\sim}{f'} &= [f_0,\ f_1,\ \ldots,\ f_{p-1}]^T . \end{aligned} \tag{2.3}$$

Since $\underset{\sim}{f'}$ is not a null vector, the requirement of $\sum$ to be a p.d. matrix ensures that always $\kappa_{2,\infty} > 0$. Now let $\hat{\kappa}_{2,\infty}$ denote a weakly consistent estimator of $\kappa_{2,\infty}$. Then the intervals

$$\left[\hat{\theta}(n) - n^{-\frac{1}{2}} z_{1-\alpha} \hat{\kappa}_{2,\infty}^{\frac{1}{2}},\quad \hat{\theta}(n) - n^{-\frac{1}{2}} z_{\alpha} \hat{\kappa}_{2,\infty}^{\frac{1}{2}} \right] \tag{2.4}$$

$$z_\alpha = \Phi^{-1}(\alpha)$$

satisfy

$$\lim_{n\to\infty} \Pr \left\{ \hat{\theta}(n) - n^{-\frac{1}{2}} z_{1-\alpha} \hat{\kappa}_{2,\infty}^{\frac{1}{2}} \le \theta \le \hat{\theta}(n) - n^{-\frac{1}{2}} z_{\alpha} \hat{\kappa}_{2,\infty}^{\frac{1}{2}} \right\} = 1 - 2\alpha . \tag{2.5}$$

However, Weber and Welsh (1983) have found that the distribution of the standardized T_n can be very skewed when n is small and so one would not expect the symmetric interval (2.4) to give a reasonable coverage when n is small. Under the conditions used to construct an Edgeworth expansion, the asymmetry of the distribution of $T_n/\kappa_{2,\infty}^{\frac{1}{2}}$ is of order $0\left(n^{-\frac{1}{2}}\right)$ in magnitude, which may be significant when n is small. Our objective is to construct Edgeworth-based confidence

intervals which capture the leading term of the third cumulant of T_n, which is of order $0(n^{-\frac{1}{2}})$. Hereafter, we shall refer to (2.4) as the *standard* confidence interval.

(b) EXPONENTIALLY-TILTED EMPIRICAL METHOD

Recently, Tingley (1992) has constructed a small-sample confidence interval for $f(\beta)$ using repeated tests under exponential tilt of the empirical distribution of a test statistic as well as a closed form for the empirical distribution, to estimate the p-value for the test. The author calls them *exponentially tilted empirical* confidence intervals to distinguish them from *exponentially tilted bootstrap* intervals. The latter use bootstrap resampling instead of a closed form for the empirical distribution. Vector β is estimated by a Mallows estimate see e.g. Hampel *et al* (1986).

Let β^0 be the unknown value of parameter vector β and assume $H_0 : \theta = \theta_0$ where θ_0 is $f(\beta^0)$ and let $\tilde{\beta}$ denote the Mallows estimate of β.

The key step in Tingley's method is the linear approximation

$$\tilde{\theta} - \theta_0 = \bar{g} + 0_p\left(n^{-\frac{1}{2}}\right) \tag{2.6}$$

which ensures that the g_i are *independent* and

$$\bar{g} = n^{-1}\sum_1^n g_i = 0\,. \tag{2.7}$$

Let $\hat{F}$ denote the empirical distribution that puts mass $\frac{1}{n}$ at each observed g_i. Since these observations are independent the cumulant generating function of $\hat{F}$ is

$$M(t) = \ln\left\{n^{-1}\sum_1^n \exp\left(tg_i\right)\right\}\,. \tag{2.8}$$

Given this empirical estimate, the author argues that the null hypothesis

$$H_0 : \theta - \theta_0 = 0 \tag{2.9}$$

is equivalent to a test concerning λ in the one parameter exponential family with "density"

$$h(g,\,\lambda) = e^{\lambda g - M(\lambda)} \tag{2.10}$$

having cumulant generating function

$$M_*(t) = M\left(\lambda + t\right) - M\left(\lambda\right)\,. \tag{2.11}$$

The author further asserts that the equivalent test is

$$H_0' : \lambda - \lambda_0 = 0 \tag{2.12}$$

where λ_0 solves the equation

$$\frac{d}{d\lambda}M(\lambda) = \theta_0 - \tilde{\theta}\,. \tag{2.13}$$

The author then gives an algorithm that tries successive values of θ_0 in order to locate the endpoints of the confidence interval sought. In this task, he appeals to the Lugannani & Rice (1980) tail-area approximation which adds error $0_p\left(n^{-\frac{1}{2}}\right)$ to the calculation of the p-value.

(c) BOOTSTRAP

EFRON PERCENTILE METHOD

Approximate confidence intervals based on the bootstrap methodology were introduced in regression context by Efron & Gong (1983). This approach is computer-intensive: it largely replaces theoretical analysis with considerable amount of computations based on a single repeated resampling procedure chosen in advance.

An explicit description of the bootstrap percentile method for our estimation problem starts with the probability structure of the underlying linear model. This probability structure is specified by three components:

(i) the data set S_0 given by (1.31)

(ii) the regression model given by (1.11) and

(iii) our earlier assumption that error distribution G is completely *unknown.*

Since β and G are not known while ε is not observable, the first step of the bootstrap methodology is to *reconstitute* the above probability structure. With $\hat{\beta}$ in hand the method calculates *approximate errors*

$$e_i = Y_i - C_i^T\hat{\beta} \tag{2.14}$$

which commonly are known as *ordinary empirical residuals* and with these in hand, G is estimated by their empirical distribution

$$\hat{G}: \text{ putting probability } n^{-1} \text{on } e_i \tag{2.15}$$

for all i. By design $\beta_0 \neq 0$ so that covariate vectors C_i^T all have first component 1. This makes the sum of (2.14) over all i equal to zero, implying that the mean of G is zero.

To complete the probability structure reconstitution process, suppose we select from $(e_1, e_2, \ldots, e_n)$ a simple random sample also of size n, with replacement. Let this sample be

$$S_* = (e_1^*, e_2^*, \ldots, e_n^*)\ . \tag{2.16}$$

Thus sample (2.16) consists of members of the original residual set $\{e_i\}_1^n$, some not appearing at all, some appearing once, and some appearing more than once. With (2.16) in hand, one may then

replace the original response-generating mechanism (1.11) by

$$Y_i^* = C_i^T \hat{\beta} + e_i^* \quad i = 1, 2, ..., n \tag{2.17}$$

thereby generating the so-called *bootstrap responses* $\{Y_i^*\}_i^n$ and the data set:

$$S_0^* = \{C_i^T, Y_i^*\}_1^n \tag{2.18}$$

which is usually referred to as *a bootstrap resample of the original* S_0 *or simply a bootstrap sample.*

Also introduced in the same Efron & Gong paper is another method of generating a bootstrap regression sample. We shall, however, not go into the details of it since the method applies when the c_{ij} in (1.12) are random, i.e. not controlled by the analyst.

After obtaining a bootstrap sample, the next step in the bootstrap methodology is to calculate what is known as a *replicate of the original point estimate.* In the context of our problem, a bootstrap replicate of $\hat{\theta}(n)$ is given by $\hat{\theta}^*(n)$, say, where

$$\hat{\theta}^*(n) = f(\hat{\beta}^*) \tag{2.19}$$

with $\hat{\beta}^*$ calculated by putting bootstrap sample (2.18) in formula (1.19).

We repeat some large number B of times, the above procedure starting from simple random sampling with replacement to generate another (2.16) to the last step of calculating replicate (2.19). By definition, let:

$$\widehat{CDF}(t) = \text{Prob}_* \left\{\hat{\theta}^*(n) \leq t\right\} = \# \left\{\hat{\theta}^*(n) \leq t\right\} / B \tag{2.20}$$

where "Prob$_*$" indicates bootstrap probability as induced by the mechanism of (2.16) and the last expression, a cumulative relative frequency, is assumed to equal this probability as B tends to ∞.

For a given α between 0 and 0.5, define

$$\theta(\alpha) = \widehat{CDF}^{-1}(\alpha) \tag{2.21}$$

then the bootstrap *percentile method* assigns

$$\theta \in [\theta(\alpha), \quad \theta(1-\alpha)] \tag{2.22}$$

as the approximate $(1-2\alpha)$ central confidence interval for θ. Here then is the algorithm for construction of confidence interval for $f(\beta)$ by Efron's *simple percentile method* (SPM).

ALGORITHM 2.1

Step 1: Construct $\hat{G}$ as at (2.15).

Step 2: From $\hat{G}$, draw a *bootstrap sample of residuals* as at (2.16).

Step 3: Using approximate regression model (2.17) calculate the bootstrap responses Y_i^*, thereby

generating data set (2.18). Use these data to calculate

$$\hat{\beta}^* = \left(X^{*T}X^*\right)^{-1} X^{*T}Y^* \quad \text{and}$$
$$\hat{\theta}^*(n) = f\left(\hat{\beta}^*\right).$$

Step 4: Independently repeat Step 2 some large B times and each time calculate $\hat{\theta}^*(n)$ as in Step 3.

Step 5: Construct $\widehat{CDF}(t)$ as at (2.20) and obtain the interval for θ as at (2.22).

ABSTRACT BOOTSTRAP RESAMPLING

Recently, Holm (1993) proposed another technique for constructing bootstrap confidence intervals in regression situations where the design matrix is controlled by the researcher. It may be summarized as follows. *Imagine* bootstrap samples from the set of true model errors rather than from empirical residuals. In the former case, neither the original sample nor the bootstrap samples generated are observable since they involve unknown parameters. Calculate theoretically *what would happen* if these bootstrap samples were used to construct a confidence interval for θ. Then it may happen, as will be shown shortly, that the final result involves *only the observable variables.* The only approximation in the technique would be the pure bootstrap error imposed by *hypothetically* using

$$\hat{G}_a : \text{probability } n^{-1} \text{ on each } \varepsilon_i \tag{2.23}$$

instead of the unknown G.

We review below the method in detail.

From (1.11) write

$$\varepsilon = Y - X\beta. \tag{2.24}$$

Imagine drawing a simple random sample

$$\varepsilon^* = \{\varepsilon_{r_i}\}_{i=1}^n \tag{2.25}$$

with replacement from (2.24). This is *abstract sampling* since the model errors are not observable. With (2.25) define for $i = 1, 2, \ldots, n$

$$Y_i^* = C_i^T\beta + \varepsilon_{r_i} = C_i^T\beta + Y_{r_i} - C_{r_i}^T\beta \tag{2.26}$$
$$= Y_{r_i} + \left(C_i^T - C_{r_i}^T\right)\beta. \tag{2.27}$$

If $\beta_0 \neq 0$ as we have assumed, then in (2.27) the vector $C_i^T - C_{r_i}^T$ would have all first zero elements, implying that all the information concerning β_0 is lost in the above definition. Holm's technique therefore requires a re-writing of (1.11) as, say

$$Y_i = \beta_0 + C_i'^T\beta' + \varepsilon \tag{2.28}$$

where

$$\beta' = (\beta_1, ..., \beta_{p-1})^T \tag{2.29}$$

$$C'_i = (c_{i1} - \bar{c}_{\cdot 1}, ..., c_{iq} - \bar{c}_{\cdot q})^T \tag{2.30}$$

or in matrix form

$$\begin{aligned} Y &= \beta_0 + C'\beta' + \varepsilon \\ C' &= (C'_1, C'_2, ..., C'_n)^T . \end{aligned} \tag{2.31}$$

We also need the condition:

C.2.1 $C'^T C$ is nonsingular.

To re-write (2.26) in matrix form using (2.28), let r_i retain the same meaning as in (2.25) and define C'_{r_i} as in (2.30), i.e. we replace i by r_i. Then:

$$Y^* = \beta_0 + C'\beta' + Y^a - \beta_0 - C'_a\beta' \tag{2.32}$$

$$= C'\beta' + Y^a - C'_a\beta' \tag{2.33}$$

where

$$C'_a = \{C'_{r_i}\}_{i=1}^n \tag{2.34}$$

$$Y^a = (Y_{r_1}, Y_{r_2}, ..., Y_{r_n})^T . \tag{2.35}$$

Now, under C.2.1, the LSE of β' is given by

$$\hat{\beta}' = D^{-1}C'^T Y , \qquad D = C'^T C' . \tag{2.36}$$

Then using (2.33) the bootstrap replicate of (2.36) is:

$$\hat{\beta}'_* = D^{-1}\{D\beta' + C'^T Y^a - C'^T C'_a \beta'\} . \tag{2.37}$$

For reasons given by Holm (1993) at p. 160, but omitted here for simplicity, of all the plausible values of β' in the reduced Ω'-space, we are interested in the critical value for which (2.36) equals (2.37). Let this value be β'_c. Then, provided the inverse in (2.38) exists

$$\beta'_c = \{D - C'^T C'_a\}^{-1}\left\{C'^T Y - C'^T Y^a\right\} \tag{2.38}$$

with all the quantities appearing in the RHS of (2.38) being observable. The most crucial part of this method therefore is the equality between the unobservable $\hat{\beta}'_*$ given by (2.37) and the observable $\hat{\beta}'$ given by (2.36). The confidence limits for θ may now be found completely automatically. For each one of B bootstrap samples (2.25) we have

$$\theta_c = f(\beta'_c) \tag{2.39}$$

is calculated via (2.38) and when a large number B of these have been obtained, the values θ'_c are arranged in ascending order. Let

$$f[1] \leq f[2] \leq \ldots \leq f[B] \tag{2.40}$$

be the ordered values of $f(\beta'_c)$ in B bootstrap repetitions. Then the desired interval is

$$f[B_\alpha] \leq \theta \leq f[B_{1-\alpha}] \tag{2.41}$$

where $f[B_\alpha]$ is the $100\times \frac{B_\alpha}{B}$-th (nearest integer) ordered value in (2.40).

The algorithm for implementing the above procedure is given below.

ALGORITHM 2.2

Step 1: Construct $\hat{G}_a$ as at (2.23).

Step 2: From $\hat{G}_a$, draw an *abstract bootstrap sample* as at (2.25). Effectively, this means choosing by simple random sampling from 1 to n the index set

$$\{r_1, r_2, \ldots, r_n\} .$$

Step 3: Determine C'_a and Y^a as at (2.34) and (2.35) respectively.

Step 4: Calculate β'_c and θ_c as at (2.38) and (2.39) respectively.

Step 5: Repeat step 2 a large number B times, each time computing θ_c through steps 3 and 4.

Step 6: Determine the confidence interval as at (2.41).

The computer programme to execute this algorithm is given in Appendix A.2

OTHER BOOTSTRAP METHODS

There exist other bootstrap methods but we shall not go into their details here. Only a very brief reference is made in regard to them.

One of these additional bootstrap methods employs a simple adjustment to residuals e_i in the Efron method. These residuals are multiplied with a factor $\left(1-\frac{p}{n}\right)^{-\frac{1}{2}}$, where p is the dimension of vector β, before taking the bootstrap samples. This corresponds to the factor needed to get an unbiased estimate of the variance from the maximum likelihood estimate under normal errors. Such a factor is used by Wu (1986) in connection with Efron's percentile method. Another factor is proposed by Urban Hjorth (1994, p187) in relation to Holm's method.

Before leaving the bootstrap methodology, the following remarks are in order:

Remark 2.1

Computer simulations employing Algorithm 2.1 or 2.2 would draw B bootstrap resamples from a single original X-data set and a single original Y-data set, if the intention is to calculate a *single* confidence interval. However, if more than one interval is required, e.g. the many confidence intervals required to calculate empirical coverage/miscoverage probability, then every distinct interval calculation requires a fresh Y-data set.

Remark 2.2

The inverse matrix in (2.38) does not exist if C'_a happens to equal C'. The probability of this event occurring is n^{-n}. The non-existence of the inverse, which might occur in the implied cases, must therefore be taken into account in the registration procedure for the bootstrap simulations.

Remark 2.3

In the context of our problem, the limited scope of applicability of Holm's method is self-evident. The restriction to design matrices which have no unit vector columns (corresponding to the intercept parameter) implies that the only functions that can be estimated by the method are those devoid of the intercept parameter. Thus for instance in our Example 1.3, the method can be used to estimate the abscissa x_m but not the ordinate $E[Y]_m$, when $\beta_0 \neq 0$.

(d) JACKKNIFE

An alternative to the above bootstrap methods are the Jackknife procedures described by Miller (1974), Weber & Welsh (1983) and Wu (1986). For details, the reader is referred to those papers.

2.3 AD HOC METHODS

In this section we review briefly some methods that have been developed for special cases of nonlinearity when the model errors are assumed to be iid normal r.v.'s.

2.3.1 FIELLER TECHNIQUE

For $r,\ s = 0,\ 1,\ ...,\ p-1$ let:

$$\begin{aligned} v'_{rs} &= cov\left(\hat{\beta}_r,\ \hat{\beta}_s\right) = (r,\ s) \text{ entry of } \mu_2 \left(X^T X\right)^{-1} \\ \hat{v}'_{rs} &= (r,\ s) \text{ entry of } \hat{\mu}_2 \left(X^T X\right)^{-1} \\ \hat{\mu}_2 &= (n-p)^{-1} \textstyle\sum_1^n e_i^2 \end{aligned} \tag{2.42}$$

and let $k_r(\theta)$ be functions not involving the parameters β_r. Fieller's (1954) method considers the equation

$$k\left(\beta,\ \theta\right) = \beta_0 k_0(\theta) + \beta_1 k_1(\theta) + ... + \beta_q k_q(\theta) = 0 \tag{2.43}$$

and, assuming that the model errors are iid normal r.v.'s, the technique proposes a confidence interval based on a consideration of the region of the $(\theta,\ t^2)$-plane lying above the curve

$$\left\{k\left(\hat{\beta},\ \theta\right)\right\}^2 = t^2 \sum\sum \hat{v}_{rs} k_r(\theta) k_s(\theta) \tag{2.44}$$

where

$$t = t_{\alpha,\, n-p} \tag{2.45}$$

is the 100α-th percentile of the student-t distribution with $n-p$ degrees of freedom.

We can illustrate the method by considering the estimation of a ratio of affine transformations. Let

$$\theta = \eta_1 \eta_2^{-1}, \qquad \begin{aligned} \eta_1 &= \underset{\sim}{a}^T \beta + b = \eta_1\left(\beta\right) \\ \eta_2 &= \underset{\sim}{a}'^T \beta + b' = \eta_2\left(\beta\right) \end{aligned} \tag{2.46}$$

where $\underset{\sim}{a},\ \underset{\sim}{a}'$ are $p \times 1$ vectors of *known* constants and $b,\ b'$ are *known* scalars. For example, the calibration function represented by (1.5), has this form with

$$\begin{aligned} \underset{\sim}{a}^T &= (-1,\ 0) & \underset{\sim}{a}'^T &= (0,\ 1) \\ b &= y_0 & b' &= 0\,. \end{aligned} \tag{2.47}$$

Furthermore, let

$$\hat{\eta}_i = \eta_i\left(\hat{\beta}\right)\,. \tag{2.48}$$

When the error distribution is normal, then

$$\hat{\eta}_1 - \theta\hat{\eta}_2 \sim N\left[0,\ \mu_2\left(v_{11} + \theta^2 v_{22} - 2\theta v_{12}\right)\right] \tag{2.49}$$

where

$$\begin{aligned} v_{11} &= \underset{\sim}{a}^T \left(X^T X\right)^{-1} \underset{\sim}{a} \\ v_{22} &= \underset{\sim}{a}'^T \left(X^T X\right)^{-1} \underset{\sim}{a}' \\ v_{12} &= \underset{\sim}{a}^T \left(X^T X\right)^{-1} \underset{\sim}{a}' \end{aligned} \tag{2.50}$$

and $\hat{\beta}$ is independent of $\hat{\mu}_2$. The latter is a well-known result, see e.g. Theorem 3.1, Seber (1977, p. 54). Therefore, the pivotal quantity

$$\{\hat{\eta}_1 - \theta\hat{\eta}_2\}/\,\hat{\mu}_2^{\frac{1}{2}}\left\{v_{11} + \theta^2 v_{22} - 2\theta v_{12}\right\}^{\frac{1}{2}} \tag{2.51}$$

has a student-t distribution with $n - p$ degrees of freedom. Hence with probability $1 - 2\alpha$

$$\{\hat{\eta}_1 - \theta\hat{\eta}_2\}^2 = \eta_2\{\hat{\theta} - \theta\}^2 \leq t^2\hat{\mu}_2\left\{v_{11} + \theta^2 v_{22} - 2\theta v_{12}\right\} \tag{2.52}$$

and the equality sign gives a quadratic equation in θ whose solution is

$$\theta_L,\ \theta_U = \{1 - g\}^{-1}\left\{\hat{\theta} - g v_{12}v_{22}^{-1} \pm t\hat{\mu}_2^{\frac{1}{2}}\hat{\eta}_2 d\right\} \tag{2.53}$$

where

$$d = v_{11} - 2\hat{\theta}v_{12} + \hat{\theta}^2 v_{22} - g\left\{v_{11} - v_{12}^2 v_{22}^{-1}\right\}^{\frac{1}{2}} \tag{2.54}$$

$$g = t^2\hat{\mu}_2 v_{22}\hat{\eta}_2^{-2} \tag{2.55}$$

As is well-known, there are some anomalies that are associated with this method, see e.g. Finney (1978, p. 82).

2.3.2 CALIBRATION CHART METHODS

For the calibration problem of the second kind under iid normal errors, Scheffé (1973) has proposed a procedure of interval estimation based on what the author calls the *Calibration Chart.* Details of the method are presented in thirty-seven pages of the paper that the author has dedicated to Jerzy Neyman.

2.3.3 CHANGE-OVER POINT INTERVALS

Consider the following two-phase simple regression model

$$\begin{aligned} Y_{1i} &= \alpha_1 + \beta_1 x_{1i} + \varepsilon_{1i} \qquad i = 1, 2, \ldots, n_1 \\ Y_{2i} &= \alpha_2 + \beta_2 x_{2i} + \varepsilon_{2i} \qquad i = 1, 2, \ldots, n_2\,. \end{aligned} \tag{2.56}$$

When the position of the change-over value θ is known to lie between x_{1n_1} and x_{21} then the former is given by (1.9). Then under the assumption of normal errors Fieller's theorem may be invoked. If, however, the position of θ is not known, the value of n_1 is an additional unknown parameter to be considered and the problem becomes more difficult. It is this latter situation that Hinkley (1971) has considered. The author's work includes a procedure for approximate interval estimation for θ for this case.

2.4 Example 1.1 REVISITED

Often studies concerning the *"availability"* (defined as the ratio of total catch to total effort) of fish in consecutive fishing seasons, involve keeping records in such a way as to show, for each day (or some other period of time), the *weight* of individual fish caught and also the amount of *"effort"* spent in capturing the fish. De Lury (1947) discusses a particularly interesting application of the simple linear regression models (1.1), (1.3) in which the ratio $-\beta_0\beta_1^{-1}$ arises as a measure of the absolute total population size (in weight) of fish resources.

Let

$\theta(t)$ represent the population size by weight at time t.

$r(t)$ denote the relative rate at which the population size changes at time t. Thus function $r(t)$ is intended to cover such sources of population increase/decrease as migration, growth, natural mortality, etc.

$\gamma(t)$ represent the relative rate at which the population size changes by sampling method.

Furthermore, assume that:

C.2.1 the population is *closed,* i.e. the effects of migration and natural mortality are negligible, and the time interval is restricted to exclude times of growth/maturity, etc..

C.2.2 the *proportion* $\gamma(t)$ of the population captured during the time interval t is given by

$$\gamma(t) = \gamma'(t)E(t) \tag{2.57}$$

where $\gamma'(t)$ is the proportion of the population captured during time interval t by *one* unit of effort and $E(t)$ is the total effort expended during time interval t. Conveniently De Lury defines the unit of effort as 1000 traps fished for one day and the unit of catch as 1000 pounds.

C.2.3 $\gamma'(t) = \gamma$, a constant.

It is shown below that, under the above three assumptions, model (1.1) holds with X, Y, and θ appropriately interpreted.

Now, to the first-order, the population change during interval $(t, t + \triangle t)$ is given by

$$\theta(t + \triangle t) - \theta(t) = \theta(t)\{r(t) - \gamma(t)\} \triangle t. \tag{2.58}$$

Then as $\triangle t \to 0$

$$\frac{d}{dt}\theta(t) = \theta(t)\{r(t) - \gamma(t)\} \tag{2.59}$$

$$\frac{d}{dt}\ell n\,\theta(t) = r(t) - \gamma(t)\,. \tag{2.60}$$

Under C.2.1, $r(t)$ is zero while C.2.2 and C.2.3 allow us to write

$$\frac{d}{dt}\ell n\,\theta(t) = -\gamma E'(t)\,. \tag{2.61}$$

On integrating (2.61) between 0 and t we get:

$$\ell n\,\theta(t) - \ell n\,\theta(0) = -\gamma \int_0^t E'(t)dt = -\gamma E(t)\,, \quad \text{say} \tag{2.62}$$

where $E(t)$ is the total effort expended during the interval $(0,\ t)$. More compactly, (2.62) has the form

$$\theta(t) = \theta(0)\exp[-\gamma E(t)]\,. \tag{2.63}$$

Differentiating with respect to E, we have

$$Y(t) = \gamma\theta(0)\exp\{-\gamma E\,[t]\} \tag{2.64}$$

where $-Y(t)$ represents $\dfrac{d\theta}{dE}$ and the latter can be identified with the negative of the catch per unit effort, in view of assumption C.2.1.

Finally, under the same assumption C.2.1, the *total catch* during the interval $(0,\ t)$ is given by

$$\theta(0) - \theta(t) = \theta(0) - \theta(0)\exp[-\gamma E(t)] = X(t), \quad \text{say.} \tag{2.65}$$

Then using the last equality in (2.65) to substitute for $\theta(0)\exp[-\gamma E(t)]$ in (2.64) we have

$$Y(t) = \gamma\{\theta(0) - X(t)\} \tag{2.66}$$

which may be compared with model (1.1).

We have taken keen interest in the De Lury model for at least two reasons. The first one is that in some countries, $\theta(0)$ may represent *a conservation parameter of major national interest.* Secondly, the way the author has applied the model to his Table 1 data illustrates one way in which a small sample may arise from the statistical modelling imperatives earlier referred to in the last chapter. That table gives a day-by-day record of the catch of lobsters, in pounds, and the effort, measured by the number of traps fished each day from May 2 to June 12. These records were made in 1944 in the Tignish area of Prince Edward Island, Canada.

When the values of $Y(t)$ are plotted against those of $X\,(t)$, the scatter plot exhibits a curved trend with a parabolic shape. However, after May 22, the values of $Y(t)$ decrease fairly regularly and it

seems not unreasonable to fit a straight line to the records from May 23 to June 12, and disregard the records outside this period.

2.5 CHAPTER SUMMARY

This chapter has reviewed the various approaches and methods that have appeared in literature concerning the problem of constructing confidence intervals for $f(\beta)$. The spectrum surveyed has included the exact and approximate methods both for the general $f(\beta)$ as well as for some special nonlinear forms. With respect to approximate methods we have reviewed both the small sample and the large-sample methods. In connection with the delta method, we have pointed out a Weber and Welsh (1983) finding that the distribution of $n^{\frac{1}{2}}\left\{\hat{\theta}(n)-\theta\right\}$ can be very skewed when n is small and so one should not expect a symmetrical interval such as (2.4) to be appropriate.

Since our main focus is on the small-sample situation we saw little reason to go into the many other large-sample methods that have appeared in literature. Included among these is the work of Miller (1974) which introduces the jackknife resampling to regression data structures and its follow up by Weber and Welsh (1983).

The last item reviewed in this chapter are the details of a particularly interesting nonlinear conservation parameter function arising in the context of some application of the simple-linear-regression-modelling to catch-effort data.

CHAPTER THREE

INDEX NOTATION

3.1 INTRODUCTION

In the monograph presenting a very monumental work on index-notation-based methods in statistics, McCullagh (1987) introduces the first chapter as follows: "It is a fact not widely acknowledged that, with appropriate choice of notation, many multivariate statistical calculations can be made simpler and more transparent than the corresponding univariate calculations. This simplicity is achieved through the systematic use of index notation and special arrays called tensors."

Among the several applications that the author has in mind is the problem of determining the cumulants of a non-linear function of a random vector. In statistical practice it sometimes happens, for example, that the cumulants of an r.v. can be determined directly but not its distribution. When this is the case and one needs the distribution, one approach to the indirect determination of the distribution is to represent the latter in terms of a *known* distribution as well as those cumulants. A commonly used tool in this connection is the expansion of the kind given by (1.36).

In a problem of this kind, if the r.v. is a multivariate non-linear function and if the function is more manageable when expressed as a Taylor series, then the question of judicious choice of notation becomes particularly important *at the cumulant-determination phase.* To see this, let

$$Q = a + U_1 + ... + U_v + ... \tag{3.1}$$

be our representation of such Taylor series, where a is fixed and U_1, ..., U_v, ... are random variables of increasing complexity as v increases. The cumulants of (3.1) may be determined from an identity in which (assuming Q to be scalar):

(a) one side is the mgf of Q i.e.the expectation of

$$\exp(\xi Q) \quad , \quad |\xi| < \infty \tag{3.2}$$

and

(b) the value of that mgf is subsequently expressed in the Maclaurin series

$$1 + b_1 + ... + b_v + ..., \qquad \text{say} \tag{3.3}$$

about $\xi = 0$

Now, on the face of it, the task as outlined above may appear to be elementary. Yet the algebraic complexity of the results sought by this straightforward approach, and the amount of work required to reach them, is evidenced by the fact that three papers on this subject have appeared in leading statistical journals. The papers are Leonov & Shiryaev (1959), James & Mayne (1962), and McCullagh (1984). The main hurdle is the manner of representing the algebra that is involved. Not only should the representation be compact but also the compactness must be sufficiently transparent.

With respect to the problem that is addressed in this thesis, there is another important reason for the *necessity* to choose notation carefully. The formulas for the end-points of our confidence intervals contain some *components* of the first four cumulants of $\hat{\theta}(n)$ and therefore some task lies ahead regarding the *identification* of these components. Moreover, what McCullagh (1984) calls *generalised cumulants* arise quite naturally in the second- and higher-order cumulants of $\hat{\theta}(n)$. In their primary form, these cumulants appear to lack the many good properties enjoyed by the more familiar *ordinary cumulants* see e.g. Brillinger (1981, p. 19). Without due care on the question of choice of notation, therefore, it may prove difficult to proceed in our task. Fortunately, however, it is demonstrated on p. 48 of James & Mayne (1962) that *any* generalised cumulant can be converted to a combination of ordinary joint cumulants, and the procedure to accomplish this has been much simplified by the *indexing convention* which McCullagh (1984) has proposed to distinguish multivariate cumulants from moments.

Given that our r.v. of interest is in the form given by (1.24), we have been persuaded by the above reasons to determine the cumulants of $\hat{\theta}(n)$ using McCullagh's index notation, hereafter abbreviated by MINO.

While proceeding with this task, we have found some of the original account and subsequent use of MINO in need of amendment, further clarification, and extension in a few but important instances. Moreover, it is our suspicion that familiarity with MINO among statisticians remains very limited. Consequently, it is the primary purpose of this chapter to take the reader of this thesis through every

aspect of MINO that is relevant to our problem and secondarily, to record our minor but useful additions that especially will greatly help the beginner.

3.2 MEANING OF INDEX NOTATION

Index notation is a set of *conventions* for (a) assigning indices to components of any finite-dimensional array or (b) summarising some algebraic operations over indices or both (a) and (b), the objective being either (a)′ to economise mathematical representation or (b)′ to simplify some mathematical operations or both (a)′ and (b)′.

Thus, whenever we use the term "index notation", we will be referring to *all* the conventions for the purposes indicated above, *taken together for the purposes at hand.*

3.3 BRIEF HISTORY OF USE IN STATISTICS

The work of Kaplan (1952) is perhaps the first to draw attention to the benefits of employing index notation to derive certain multivariate statistical results by treating them as mild generalisations of known univariate relations. However, the author adopts the name "tensor notation" rather than "index notation". He defines his choice on p. 319 as follows:

> "A vector random variable at its simplest is represented by a letter with a single undetermined subscript...If the vector notation is generalised by allowing more than one variable index, tensors are obtained, while ordinary variables or numbers with no variable index, are called scalars"

More about the focus of this paper is given in the next chapter.

A decade later, James & Mayne (1962) employ a modified form of the "tensor" notation suggested by Kaplan.

By far, the most extensive treatment of the role which indexing conventions can play in statistics appears in McCullagh (1984, 1987). In the 1984 paper, MINO is used to present yet another look at the cumulants of multivariate non-linear transformations, while the 1987 monograph gives more details of the 1984 paper, with some correction, as well as discusses at length many sophisticated statistical applications in which the special use of indices can play invaluable role.

Also relevant at this point in time, is the use of the classical Cartesian tensor notation by McCullagh & Pregibon (1987) to derive two sets of unbiased estimates of the third and fourth ordinary joint cumulants of (1.26).

Since McCullagh's monumental work, several books have, in varying extent, included some aspects of MINO in places. Among these are Barndorff-Nielsen & Cox (1989), Kass & Vos (1997), and Pace & Salvan (1997).

3.4 INDEX VS TENSOR NOTATION

Often writers use the terms "index notation" and "tensor notation" interchangeably, leaving the reader to wonder why this is practised. Perhaps the term "tensor" notation should be restricted to describing only the index notation that is used to represent and/or manipulate the transformation laws of Cartesian Tensors.

3.5 MAIN FEATURES OF MINO, WITH SOME MODIFICATIONS AND EXTENSION

3.5.1 ORDINARY AND COMPOUND INDEX SETS

Let

$$R = \{r_1, \ldots, r_v\} \tag{3.4a}$$

denote an arbitrary finite set of indices, *taken one at a time.* We will call (3.4a) an *ordinary* index set to distinguish it from, say

$$\{r, st\}\ , \quad \{rs, t, u\}\ , \quad \{rs, tuv\} \tag{3.4b}$$

which are examples of what we will call *compound* index sets. In (3.4a), v is called the *size* of R while in (3.4b) the sizes are 2 for $\{r, st\}$, 3 for $\{rs, t, u\}$, and 2 for $\{rs, tuv\}$.

Since readers typically find *indexed* indices difficult to follow, it is advisable to use them sparingly. Consequently, in those parts of the thesis where we work with index notation, we will as far as possible work with i, j, k, ℓ, m or r, s, t, u, v, w and expect the reader to make the necessary generalisation of type (3.4a).

3.5.2 PARTITIONS

A *partition* of an ordinary index set, say R, is a collection of *disjoint* subsets of R whose union is all of R. Let α denote the number of subsets in a partition. Then a typical partition of (3.4a) will be shown as follows

$$R_1 \mid ..... \mid R_\alpha \tag{3.5}$$

for any α in the interval $[1,\ v]$. The R_s, $s = 1, ..., \alpha$ are sometimes also called *blocks*. The following definition will assist in our subsequent exposition.

DEFINITION 3.1

The indices of a set such as (3.5) are said to be:

(i) unpartitioned if $\alpha = 1$

(ii) Type-I-partitioned if $\alpha = v$

(iii) Type-II-partitioned if $1 < \alpha < v$

To assist analysts, McCullagh (1987) has provided on pp. 254–6 a table of natural numbers which may be used to *identify* partitions of ordinary index sets of sizes up to $v = 6$. On using those natural numbers, we obtain the following three tables to which we will refer from time to time later.

TABLE 3.1: POSSIBLE PARTITIONS OF $R = \{r, s, t\}$

1	2	3
rst	$r\|st$ $s\|rt$ $t\|rs$	$r\|s\|t$

N.B.: When using partitions, it is essential that one regards the individual members of an index set as independent and therefore distinct. Thus in the table above when $\alpha = 2$ the partitions $r|st$, $s|rt$, and $t|rs$ are considered as formally different even in cases like $r = s = t = 1$ or $r = 1$, $s = t = 3$.

TABLE 3.2: POSSIBLE PARTITIONS OF $\{r, s, t, u\}$

1	2	3	4
	1st KIND		
$rstu$	$r\|stu$	$r\|s\|tu$	$r\|s\|t\|u$
	$s\|rtu$	$r\|t\|su$	
	$t\|rsu$	$r\|u\|st$	
	$u\|rst$	$s\|t\|ru$	
	2nd KIND	$s\|u\|rt$	
	$rs\|tu$	$t\|u\|rs$	
	$rt\|su$		
	$ru\|st$		

TABLE 3.3: POSSIBLE PARTITIONS OF $\{r, s, t, u, v\}$

1	2	3	4	5
	1st KIND	1st KIND		
$rstuv$	$r\|stuv$	$r\|s\|tuv$	$r\|s\|t\|uv$	$r\|s\|t\|u\|v$
	$s\|rtuv$	$r\|t\|suv$	$r\|s\|u\|tv$	
	$t\|rsuv$	$r\|u\|stv$	$r\|s\|v\|tu$	
	$u\|rstv$	$r\|v\|stu$	$r\|t\|u\|sv$	
	$v\|rstu$	$s\|t\|ruv$	$r\|t\|v\|su$	
		$s\|u\|rtv$	$r\|u\|v\|st$	
	2nd KIND	$s\|v\|rtu$	$s\|t\|u\|rv$	
		$t\|u\|rsv$	$s\|t\|v\|ru$	
	$rs\|tuv$	$t\|v\|rsu$	$s\|u\|v\|rt$	
	$rt\|suv$	$u\|v\|rst$	$t\|u\|v\|rs$	
	$ru\|stv$	2nd KIND		
	$rv\|stu$	$r\|st\|uv$		
	$st\|ruv$	$r\|su\|tv$		
	$su\|rtv$	$r\|sv\|tu$		
	$sv\|rtu$	$s\|rt\|uv$		
	$tu\|rsv$	$s\|ru\|tv$		
	$tv\|rsu$	$s\|rv\|tu$		
	$uv\|rst$	$t\|rs\|uv$		
		$t\|ru\|sv$		
		$t\|rv\|su$		
		$u\|rs\|tv$		
		$u\|rt\|sv$		
		$u\|rv\|st$		
		$v\|rs\|tu$		
		$v\|rt\|su$		
		$v\|ru\|st$		

3.5.3 INDEXING CONVENTIONS

We now consider the basic conventions of MINO, as modified and extended in places. The reader is referred to the chapter summary at the end of this chapter for a guide regarding the amendments and extensions.

We start with conventions for indexing coefficients, vectors, matrices, product moments, and multivariate cumulants. We end with conventions on summing over indices.

CONVENTION 3.1: COEFFICIENTS

In a given summand, the symbol(s) appearing with *subscript(s)* is(are) to be regarded as coefficient(s).

CONVENTION 3.2: VECTORS

Components of a vector are indexed using *superscripts.* Furthermore, a vector and its general component are represented by the *same* notation. Thus for example, vector Z in (1.26a) and (1.26b) may be represented as Z^r as would its r-th component.

CONVENTION 3.3: MATRICES

The (i, j) entry of matrix A is written as

$$a^i_j \tag{3.6}$$

instead of the usual notation a_{ij}. Note that to avoid ambiguity the superscript shall always refer to the row index and the subscript will correspond with the column index. Furthermore, as with vectors, (3.6) may represent matrix A as well.

CONVENTION 3.4: MULTIVARIATE MOMENTS AND CUMULANTS

In so far as the problem of determining cumulants/moments of a stochastic Taylor series is concerned, the key feature of MINO is McCullagh (1984)'s convention according to which both moments and cumulants use the *same letter symbol and the same set of indices.* The point of departure is as follows:

(a) product-moments are indexed by *unpartitioned* superscripts; while

(b) multivariate cumulants are indexed by *partitioned* superscripts, with commas replacing vertical lines as separators of blocks. In the context of our problem, the moments and cumulants of $\hat{\beta}(n) - \beta$ will be represented using the symbol π and indices $r, s, t, \ldots$ while the corresponding quantities for the error vector ε will be represented using μ and indices $i, j, k, \ldots$ Thus according to convention 3.4(a), we have, for example, that:

$$\begin{aligned} \mu^{ij} &= E\left[\varepsilon^i \varepsilon^j\right] \\ \mu^{ijk} &= E\left[\varepsilon^i \varepsilon^j \varepsilon^k\right] \\ \mu^{iii} &= E\left[\varepsilon^i \varepsilon^i \varepsilon^i\right] \\ \pi^{rstu} &= E\left[Z^r Z^s Z^t Z^u\right] . \end{aligned} \tag{3.7}$$

As for convention 3.4(b) the *type of partitioning* used distinguishes between three kinds of cumulants as follows:

(i) A cumulant notation as described above in which the superscripts are Type-I-partitioned specifies what is called an *ordinary cumulant.* The number of indices involved defines the cumulant's *order.* For example,

$$\pi^{r,s,t} \tag{3.8}$$

represents the third-order ordinary joint cumulant of components Z^r, Z^s, and Z^t.

(ii) A cumulant notation such as $\mu^{i,i,i,i}$ represents the fourth-order-cumulant of component ε^i. Since by assumption these components are identically distributed, we may let

$$\mu^{i,i,i,i} = \rho_4 \tag{3.9}$$

where ρ_υ is the notation that we will use for the υ-th order cumulant of the common error distribution.

(iii) Finally, a cumulant notation in which the superscripts are Type-2-partitioned specifies what is known as a *generalised cumulant.* The order of this cumulant is given by the number of partitions while the number of indices specify its *degree.* For example

$$\pi^{r,s,t,uv} \tag{3.10}$$

represents the fourth-order, fifth-degree generalised cumulant of components Z^r, Z^s, Z^t and product $Z^u Z^v$.

3.5.4 SUMMATION CONVENTIONS

Keeping track of all the *individual* terms that are involved in the manipulations that lead up to (3.3) requires careful consideration. To facilitate this, it is useful to distinguish between two *phases* of the summation operation. The first phase considers the summation of "similar" terms:

either (a) over all the *possible partitions of a given kind;*
or (b) over all the *possible "permutations" of the index set that is at hand.*

In (b) there is an important reason behind our placing quotation marks over the term *permutations*. As we know, the term, in general, refers to different *orderings* of *distinct* members of a discrete set. Moreover, ordinarily it is assumed that the identity of each member of the original set remains *unchanged* as the different orderings are formed. Thus, for example, in the ordinary sense:

(i) the different permutations of (e, f) are:

$$(e, f) \quad , \quad (f, e)$$

while

(ii) the different permutations of (e, f, g) are the following:

$$\begin{array}{lll} (e,f,g) & , \ (f,e,g) & , \ (g,e,f) \\ (e,g,f) & , \ (f,g,e) & , \ (g,f,e) \end{array}$$

What our quotation marks in (b) refer to is that the above-mentioned identity-feature is relaxed in the manner most easily explained by these examples below:

$$\begin{array}{lll} \text{(iii)} & (rs,t) & , \ (r,st) \\ \text{(iv)} & (rs,t,u) & , \ (r,st,u) \quad , \ (r,s,tu) \\ \text{(v)} & (rs,tu,v) & , \ (rs,t,uv) \quad , \ (r,st,uv) \end{array}$$

Not only are the identities of the members of each starting set lost in the subsequent ordering, but we also see some reduction in the number of premutations in the case of (iv) and (v) as compared to (ii).

We will refer to summation (a) as *Type-I First Phase Summation* and (b) will be called *Type-II First Phase Summation.*

The second summation phase consists of summing over *each index* that is involved. We shall refer to it as the *Second Phase Summation.*

MINO employs what we will call the "Rectangular Bracket Summation Convention" to both types of the first summation phase and the so-called "Einstein Summation Convention" to the Second Phase Summation.

We now describe these conventions in some detail, starting with the Einstein Convention.

CONVENTION 3.5: THE EINSTEIN SUMMATION CONVENTION

Whenever two or more indexed quantities are expressed as a product in which the *same index* appears both as a subscript and as a superscript, then summation is implied over the repeated index, which we will call a *summation index (or index of summation).* This convention is reportedly due to Einstein, see e.g. Lass (1950, p. 259). Any index which is present in the product in any other form is called a *free index.*

As an example, the Taylor expansion that is given by (1.24) in Chapter one, may be written according to the above convention as follows:

$$\hat{\theta}(n) = \theta + f_r Z^r + \frac{1}{2!} f_{r_s} Z^r Z^s + \frac{1}{3!} f_{r_{st}} Z^r Z^s Z^t + \ldots\ldots \tag{3.11}$$

$$Z^r = \ell_i^r \varepsilon^i . \tag{3.12}$$

In (3.12) the RHS may be interpreted either as the representation of the r-th component of $L\varepsilon$, which in the ordinary convention, is given by the sum $\sum_i \ell_{ri}\varepsilon_i$ or the vector $L\varepsilon$ itself with ℓ_i^r denoting L and vector ε represented by ε^i. Under the first interpretation we call index r a *free index* to distinguish it from *summation index "i".*

In view of (3.1), it is helpful to examine how, under the Einstein Convention, an arbitrary conventional sum $\sum_{i=1}^n a_i$ behaves when the exponentiation operation is carried out.

If we do not appeal to the classical multinomial expansion theorem, we have that:

$$\begin{aligned}
\left\{\sum_1^n a_i\right\}^2 &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \\
\left\{\sum_1^n a_i\right\}^3 &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_i a_j a_k \\
&\vdots \\
&\vdots \\
\left\{\sum_1^n a_i\right\}^\upsilon &= \sum_{i_\upsilon=1}^n \cdots \sum_{i_1=1}^n a_{i_1} a_{i_2} \ldots a_{i_\upsilon}
\end{aligned} \tag{3.13}$$

The above suggests that if we define

$$1_{i_1 i_2 \ldots i_\upsilon} \equiv 1 \tag{3.14}$$

for all $i_1, i_2, \ldots, i_\upsilon$, running from 1 to n we have

$$\sum_{i_\upsilon} \cdots \sum_{i_2} \sum_{i_1} a_{i_1} a_{i_2} \ldots a_{i_\upsilon} = 1_{i_1 i_2 \ldots i_\upsilon} a^{i_1} a^{i_2} \ldots a^{i_\upsilon} \tag{3.15}$$

in the Einstein Convention representation.

CONVENTION 3.6: THE RECTANGULAR BRACKET SUMMATION CONVENTION

(i) INTRODUCTION

McCullagh (1984, 1987) uses notation $[m]$, where m is some natural number, to summarise representation of certain summation operations. We will call this notation the *Rectangular Bracket Summation Convention* and abbreviate it by RBSC.

In the 1984 paper the author uses RBSC in three different ways which we will examine shortly, and in the 1987 monograph this practice is repeated in addition to the fourth sense of the use of the same notation in the first identity on p. 147. We will, however, not address this last case here, as it is not relevant to our problem. The other three are.

(ii) FIRST INTERPRETATION OF $[m]$

One sense in which the author has used $[m]$ was perhaps motivated by the author's wish to condense further Kaplan (1952)'s equation (1). We consider some illustrations using our symbols.

An expression such as $\pi^r\pi^{st}[m]$ can be made to represent compactly, a well-defined first phase summation. This can be done by referring to Table 3.1. We see there that setting $m = 3$ makes the following a well-defined identity:

$$\pi^r\pi^{st}[3] = \pi^r\pi^{st} + \pi^s\pi^{rt} + \pi^t\pi^{rs}\,. \tag{3.16}$$

In like manner, columns 2 and 3 of Table 3.2 allow us to do likewise and define, say

$$\pi^r\pi^{stu}[4], \quad \pi^{rs}\pi^{tu}[3], \quad \pi^r\pi^s\pi^{tu}[6] \tag{3.17}$$

as follows:

$$\begin{aligned}
\pi^r\pi^{stu}[4] &= \pi^r\pi^{stu} + \pi^s\pi^{rtu} + \pi^t\pi^{rsu} + \pi^u\pi^{rst} \\
\pi^{rs}\pi^{tu}[3] &= \pi^{rs}\pi^{tu} + \pi^{rt}\pi^{su} + \pi^{ru}\pi^{st} \\
\pi^r\pi^s\pi^{tu}[6] &= \pi^r\pi^s\pi^{tu} + \pi^r\pi^t\pi^{su} + \pi^r\pi^u\pi^{st} \\
&\quad + \pi^s\pi^t\pi^{ru} + \pi^s\pi^u\pi^{rt} + \pi^t\pi^u\pi^{rs}
\end{aligned} \tag{3.18}$$

In the illustrations just considered, we have applied the RBSC to the sums of moment products. To apply the same convention to the sums of cumulant products, we only need to recall convention 3.4(b) which distinguishes moments from cumulants. Thus, for example:

$$\pi^r\pi^{s,t,u}[4] = \pi^r\pi^{s,t,u} + \pi^s\pi^{r,t,u} + \pi^t\pi^{r,s,u} + \pi^u\pi^{r,s,t} \tag{3.19}$$

and

$$\pi^{r,s}\pi^{t,u}[3] = \pi^{r,s}\pi^{t,u} + \pi^{r,t}\pi^{s,u} + \pi^{r,u}\pi^{s,t}\,. \tag{3.20}$$

The message that is illustrated by (3.16) to (3.20) should by now be clear, namely that

$[m]$ notation as used above is unambiguous *if* m *refers to the number of partitions of the same kind.*

(iii) SECOND USE OF $[m]$

McCullagh (1984, 1987) uses notation $[m]$ in another sense *in relation to the generalised cumulants.* However, in neither publication is the reader warned of the difference. Barndorff-Nielsen and Cox (1989) appear to notice this new use when on p. 139 they state that "$[n]$ after a symbol will indicate a sum of n similar terms determined by suitable permutations". As we shall see shortly, "suitable permutations" in this case refers to an entirely different situation involving generalised cumulants. However, Barndorff-Nielsen and Cox (1989) also fail to distinguish their use of RBSC in, say, their equations (5.45) and (5.46) from the way the same notation is used in their formula (5.57).

From McCullagh's explanation of his method, it is evident that the terms $A_1A_1A_2[3]$ and $A_1A_2A_2[3]$ appearing in McCullagh (1984)'s equation (10), must have the following meanings:

$$A_1A_1A_2[3] = A_1A_1A_2 + A_1A_2A_1 + A_2A_1A_1 \tag{3.21}$$

$$A_1A_2A_2[3] = A_1A_2A_2 + A_2A_1A_2 + A_2A_2A_1\,. \tag{3.22}$$

Then as consequence, the terms

$$\kappa^{i,j,k\ell}[3]\,, \quad \kappa^{i,jk,\ell m}[3]\,, \quad \kappa^{i,j,k,\ell m}[4]$$

appearing in equation (3.10) of McCullagh (1987) have the following interpretations:

$$\kappa^{i,j,k\ell}[3] = \kappa^{i,j,k\ell} + \kappa^{i,jk,\ell} + \kappa^{ij,k,\ell} \tag{3.23}$$

$$\kappa^{i,jk,\ell m}[3] = \kappa^{i,jk,\ell m} + \kappa^{ij,k,\ell m} + \kappa^{ij,k\ell,m} \tag{3.24}$$

$$\kappa^{i,j,k,\ell m}[4] = \kappa^{i,j,k,\ell m} + \kappa^{i,j,k\ell,m} + \kappa^{i,jk,\ell,m} + \kappa^{ij,k,\ell,m}\,. \tag{3.25}$$

In Chapter five, we will propose a much simpler alternative method which will, among others, justify (3.21) to (3.25).

In the sequel, we shall retain the first use of $[m]$ as described in subsection (ii) above. In the place of the second use of $[m]$ that is described in this subsection, we will use $[m]_*$. Thus for example, the left hand side of (3.23) would, instead, be expressed as $\kappa^{i,j,k\ell}[3]_*$.

In short, notation $[m]$ will be applied to Type-I First Phase Summations and $[m]_*$ to Type-II First Phase Summations when the index set at hand is *compound.*

(iv) INTERPRETATION OF $[m-1]$

From (3.20) we have that

$$\pi^{r,s}\pi^{t,u}[3] - \pi^{r,s}\pi^{t,u} = \pi^{r,t}\pi^{s,u} + \pi^{r,u}\pi^{s,t} \tag{3.26}$$

Apparently this has led McCullagh (1984) on p. 463 and McCullagh (1987) on p. 58 to define

$$\pi^{r,t}\pi^{s,u}[2] = \pi^{r,t}\pi^{s,u} + \pi^{r,u}\pi^{s,t}\,. \tag{3.27}$$

If one adopts this interpretation, then one should similarly define, say

$$\pi^{r,t}\pi^{s,u,v}[9] = \pi^{r,s}\pi^{t,u,v}[10] - \pi^{r,s}\pi^{t,u,v} \tag{3.28}$$

and so on.

In Section 4.3.2 of the next chapter, a rule will be provided to govern the handling of the summation operations in which RBSC is present. It is that rule that motivated our explicit distinction between "First" and "Second" phase summations. Indeed when that section is reached, we recommend that some aspects of it be regarded as an extension of this chapter.

3.6 CHAPTER SUMMARY

This chapter has extensively reviewed index notation in so far as we plan to use it to determine the cumulants of $\hat{\theta}(n)$. As stated early in the chapter, the primary purpose of the review has been to *prepare* the reader to follow our use of the notation in the above task.

From the review, we see that while a few authors have, in varying extent, contributed to drawing attention to the benefits of using index notation when handling some difficult multivariate statistical calculations where matrix notation is insufficient, the work of McCullagh (1984, 1987) constitutes by far the most important contribution to date, in this regard. Hence our naming the notation "MINO" in a tribute to the author.

During the review of MINO, we have found it worthwhile to offer some definitions here and there, aimed mainly to assist in better understanding of what MINO is about. In addition, we have discussed at length the use of the rectangular bracket notation and subsequently chose to represent Types-I and II First Phase Summations *differently*. We recommend this practice because in the past ignoring this has led to the same symbol being used to represent the two, totally different, summation operations, sometimes even within a single expression or equation. For more in this regard, see Section 4.3.2 in the next chapter in connection with the generalised cumulants.

CHAPTER FOUR

THE CUMULANTS OF $\hat{\beta}(\mathbf{n}) - \beta$

4.1 INTRODUCTION

Let us recall that we seek to use an Edgeworth expansion to construct some approximate confidence interval for θ. The particular Edgeworth expansion that will be used is a function of, among others, the cumulants of the standardised statistic based on the asymptotic result (2.3). With this in mind, it is our desire to design a theory for those cumulants in such a way that:

(a) we preserve the asymptotic normality due to that delta method (recall that the standard normal distribution is uniquely determined by its cumulants, see e.g. Billingsley (1979, p. 344));

(b) the *first two* terms of the *asymptotic expansions* of the first four of those cumulants can be captured in the first two terms of the Edgeworth expansion *and* in the manner that follows tradition.

From the Taylor expansion (3.11), it is clear that the second- and higher-order cumulants of $\hat{\theta}(n)$ are functions of all the possible subsets of size two or more that can be formed from the vector

$$(Z^r,\ Z^r Z^s,\ Z^r Z^s Z^t,\ .....) \tag{4.1}$$

Following McCullagh (1984), we shall refer to the latter type of cumulants as *generalised* cumulants. A key property of these cumulants was mentioned early in the last chapter, namely, that they are expressible in terms of only ordinary joint cumulants.

It is furthermore evident from (3.12) that the cumulants of $\hat{\theta}(n)$ are also dependent on the cumulants of model errors. Consequently, before we consider the cumulants of the Taylor expansion and of the

standardized statistic, which is taken up in the next chapter, we wish to address in this chapter the preliminary results that will be needed there. These results relate to:

(i) the cumulants of Z as functions of the cumulants of model errors;

(ii) a crucial condition around which we construct all the cumulant-theory for the standardized $\hat{\theta}(n)$ that we will need at the Edgeworth expansion construction stage; and

(iii) the key property of the generalised cumulants to which we have just referred.

We begin, in Section 4.2, with a review of an old, short method that makes available two sets of relations: one set giving joint ordinary cumulants as functions of moments and the other set giving reverse relations. It is through the relationships between the two sets that the key property of the generalised cumulants conversion comes about. For the ultimate purpose at hand, we will need the above relations only as far as the fifth order. Before ending the section, we take the opportunity to justify our choice to work with cumulants instead of moments.

In Section 4.3 we discuss in detail the key property referred to in (iii) above and in Section 4.4 we consider the cumulants of Z in terms of the cumulants of the model errors. From the latter, a condition is established around which all the subsequent theory concerning the cumulants that we will need is constructed. We will call it a *crucial condition.* Since the cumulants of the model errors have unknown values except, by assumption, the first one, the question of their estimation is addressed in the last Section, 4.5. The last section examines the crucial condition in relation to the simple linear and quadratic regression models, as these will be studied in simulations.

4.2 ORDINARY JOINT CUMULANTS

The shortest and probably the oldest method for arriving at the relations between ordinary joint cumulants and moments is due to Kaplan (1952). The shortness of this method derives from two sources: (1) the *symmetry* property of both the cross-moments and the ordinary joint cumulants and (2) a 1949 paper by F.N. David & M.G. Kendall concerning the tables of symmetric functions.[1]

This method proceeds as follows:

In order to deduce the relationships between the ordinary cumulants

$$\pi^{r,s},\ \pi^{r,s,t},\ \pi^{r,s,t,u},\ \ldots\ldots \tag{4.2}$$

[1] **Recall that a function is symmetric if it is invariant under all the permutations of its arguments**

and the moments

$$\pi^r,\ \pi^{rs},\ \pi^{rst},\ \pi^{rstu},\ \ldots\ldots \tag{4.3}$$

we may use the identities

$$\ln E\left[\exp(\xi_r Z^r)\right] = \xi_r\pi^r + \frac{1}{2!}\xi_r\xi_s\pi^{r,s} + \frac{1}{3!}\xi_r\xi_s\xi_t\pi^{r,s,t} + \ldots\ldots \tag{4.4}$$

$$E\left[\exp(\xi_r Z^r)\right] = 1 + Q_* \tag{4.5}$$

$$Q_* = \xi_r\pi^r + \frac{1}{2!}\xi_r\xi_s\pi^{rs} + \frac{1}{3!}\xi_r\xi_s\xi_t\pi^{rst} + \ldots\ldots \tag{4.6}$$

in a purely formal way, without paying attention to the questions of existence of moments or the convergence of the series involved.

On substituting the RHS of (4.5) into the LHS of (4.4), the result expands in the following well-known way:

$$Q_* - \frac{1}{2}Q_*^2 + \frac{1}{3}Q_*^3 - \frac{1}{4}Q_*^4 + \ldots\ldots \tag{4.7}$$

from which the υ-th member of (4.2), denoted by

$$\pi^{r_1,\ldots,r_\upsilon} \tag{4.8}$$

should be extracted as the *coefficient* of the product

$$\frac{1}{\upsilon!}\xi_{r_1},\ \ldots,\ \xi_{r_\upsilon} \tag{4.9}$$

in that expansion.

Kaplan's (1952) short method of carrying out this extraction rests on the symmetry property of the cross moments with respect to the indices, i.e. their values remain unchanged under all the possible permutations of the indices. Since each index in (4.6), assumed distinct, must occur exactly once in each term of (4.7), the above symmetry is a sufficient and apparently also necessary condition for the coefficient of (4.9) *not* to be dependent upon the indices. This then permits one to deduce *uniquely* the formulas relating (4.8) to moments from their *univariate* counterparts, using tables of symmetric functions provided by David & Kendall (1949). This way, and on using the univariate formulas as given, for example, in Kendall (1948), we have the following expressed in MINO as far as the fifth order:

$$\pi^{r,s} = \pi^{rs} - \pi^r\pi^s \tag{4.10}$$

$$\pi^{r,s,t} = \pi^{rst} - \pi^r\pi^{st}[3] + 2\pi^r\pi^s\pi^t \tag{4.11}$$

$$\pi^{r,s,t,u} = \pi^{rstu} - \pi^r\pi^{stu}[4] - \pi^{rs}\pi^{tu}[3] + 2\pi^r\pi^s\pi^{tu}[6] - 6\pi^r\pi^s\pi^t\pi^u \tag{4.12}$$

$$\begin{aligned}\pi^{r,s,t,u,v} = {} & \pi^{rstuv} - \pi^r\pi^{stuv}[5] - \pi^{rs}\pi^{tuv}[10] + 2\pi^r\pi^s\pi^{tuv}[10] \\ & + 2\pi^r\pi^{st}\pi^{uv}[15] - 6\pi^r\pi^s\pi^t\pi^{uv}[10] + 24\pi^r\pi^s\pi^t\pi^u\pi^v.\end{aligned} \tag{4.13}$$

The general form of the coefficient occurring in each formula above is given by Section 10 of David & Kendall (1949) as being equal to

$$(-1)^{\rho-1}(\rho-1)! \tag{4.14}$$

where ρ is the number of the π's that are present in the product; see also formula (3.39) on p. 72 of Kendall & Stuart (1977).

Now, under the zero mean assumption for the model errors, we have

$$\pi^r = 0 \tag{4.15}$$

and this simplifies formulas (4.10) to (4.13) as follows

$$\pi^{r,s} = \pi^{rs} \tag{4.16}$$

$$\pi^{r,s,t} = \pi^{rst} \tag{4.17}$$

$$\pi^{r,s,t,u} = \pi^{rstu} - \pi^{rs}\pi^{tu}[3] \tag{4.18}$$

$$\pi^{r,s,t,u,v} = \pi^{rstuv} - \pi^{rs}\pi^{tuv}[10]\,. \tag{4.19}$$

Both sets of formulas, i.e. (4.10) to (4.13) and (4.16) to (4.19) will be needed. To facilitate their distinction we will refer to the first set as the *full version* and the second set as the *reduced version*.

Also needed are the full-version reverse formulas for the corresponding cross moments. Since each member of (4.2) is symmetric in the indices, exactly the same reasoning due to Kaplan gives the following:

$$\pi^{rs} = \pi^{r,s} + \pi^r\pi^s \tag{4.20}$$

$$\pi^{rst} = \pi^{r,s,t} + \pi^r\pi^{s,t}[3] + \pi^r\pi^s\pi^t \tag{4.21}$$

$$\begin{aligned}\pi^{rstu} &= \pi^{r,s,t,u} + \pi^r\pi^{s,t,u}[4] + \pi^{r,s}\pi^{t,u}[3] \\ &\quad + \pi^r\pi^s\pi^{t,u}[6] + \pi^r\pi^s\pi^t\pi^u\end{aligned} \tag{4.22}$$

$$\begin{aligned}\pi^{rstuv} &= \pi^{r,s,t,u,v} + \pi^r\pi^{s,t,u,v}[5] + \pi^{r,s}\pi^{t,u,v}[10] \\ &\quad + \pi^r\pi^s\pi^{t,u,v}[10] + \pi^r\pi^{s,t}\pi^{u,v}[15] \\ &\quad + \pi^r\pi^s\pi^t\pi^{u,v}[10] + \pi^r\pi^s\pi^t\pi^u\pi^v\,.\end{aligned} \tag{4.23}$$

The coefficient of $+1$ throughout each one of (4.20) to (4.23) is in accordance with Section 3 of David & Kendall (1949).

Before ending this section we wish to point out three factors that have motivated our choosing to work with cumulants rather than moments. Firstly, under the assumption for independent model errors, all the mixed ordinary cumulants of vector ε vanish. Secondly, the affine transformation law, which will play a very key role as we shall see, produces simpler results for second- and higher-order cumulants when compared to the corresponding results for moments. Lastly, an Edgeworth expansion is derived

in a straightforward way when cumulant generating functions are used, and the expansion obtained under this approach contains cumulants in the place of moments.

4.3 GENERALISED CUMULANTS

4.3.1 INTRODUCTION

Let the index set given by (3.5) be Type-II-partitioned. In keeping with our convention (3.4) of MINO, the α-th-order generalised cumulant of Z shall be denoted by

$$\pi^{R_1,\ldots,R_\alpha} \tag{4.24}$$

with the commas replacing the vertical lines as the separators of the index blocks that comprise the index set. Generalised cumulants arise quite naturally in the determination of the second- and higher-order cumulants of a multivariate stochastic Taylor series. Taking the case of (3.11), for example, the emergence of these cumulants is seen easily, upon regarding the series as an affine transformation of vector (4.1) and applying the basic structure of the results of the transformation law for cumulants of such function, see Section 4.4 and Chapter five.

4.3.2 RELATIONSHIP WITH ORDINARY JOINT CUMULANTS

As demonstrated by equations (8), (9) and (10) of James & Mayne (1962), any generalised cumulant is expressible in terms of the ordinary joint cumulants. To achieve this, we propose two simple rules that, among other considerations, should be followed.

Rule R.4.1

To convert a generalised cumulant of *a given order*, one *should* use a *full-version* formula for an ordinary joint cumulant *of the same order.*

In applying this rule, note that in the context of generalised cumulants, *order* refers to the *number of blocks* whereas in relation to ordinary joint cumulants, *order* refers to *the number of indices.*

For convenience, the ordinary cumulant formula that is used for such conversion shall be referred to simply as *the conversion formula.*

Rule R.4.2

Any rectangular bracket notation either present in the conversion formula or associated with the generalised cumulant due for conversion, must be eliminated *before the rest of the conversion steps can proceed.*

We consider some examples. Suppose that we wish to convert $\pi^{r,st}$ or $\pi^{r,stu}$. According to rule R.4.1 we would use *full-version* conversion formula (4.10). This is how it is done: first, we have

$$\pi^{r,st} = \pi^{rst} - \pi^{r}\pi^{st} \tag{4.25a}$$

$$= \pi^{rst} \quad \text{since} \quad \pi^{r} = 0 \tag{4.25b}$$

while in the case of $\pi^{r,stu}$

$$\pi^{r,stu} = \pi^{rstu} - \pi^{r}\pi^{stu} \tag{4.26a}$$

$$= \pi^{rstu} \tag{4.26b}$$

Then applying reduced formulas (4.17) and (4.18) respectively to π^{rst} and π^{rstu} we get

$$\pi^{r,st} = \pi^{r,s,t} \tag{4.27}$$

and

$$\pi^{r,stu} = \pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3] \tag{4.28}$$

where the second term in (4.28) results from the use of (4.16). The danger of using conversion formulas which are in reduced form at the *initial step* may be illustrated by the following example. Suppose that it is

$$\pi^{rs,tu} \tag{4.29}$$

that we wish to convert. If we use the reduced formula (4.16) we get

$$\pi^{rs,tu} = \pi^{rstu} \tag{4.30}$$

while if we use (4.10) we obtain

$$\pi^{rs,tu} = \pi^{rstu} - \pi^{rs}\pi^{tu}. \tag{4.31}$$

Clearly it is (4.31) that is correct.

To see the revelance of rule R.4.2 we take another conversion example. Suppose that we wish to convert the function

$$\pi^{r,s,tu}[3]_{*} \tag{4.32}$$

into a function of only the ordinary cumulants. In accordance with rule R.4.1, we first eliminate $[\cdot]_*$ as follows:

$$\pi^{r,s,tu}[3]_* = \pi^{r,s,tu} + \pi^{r,st,u} + \pi^{rs,t,u}\,. \tag{4.33}$$

Then, according to rule R.4.2, we convert *each* one of the summands in (4.33) using the full conversion formula for $\pi^{r,s,t}$, which is given by (4.11). That formula, however, contains a rectangular bracket notation [3]. Were we to ignore rule R.4.2 and simply proceed as follows:

$$\pi^{r,s,tu} = \pi^{rstu} - \pi^{r}\pi^{stu}[3] + 2\pi^{r}\pi^{s}\pi^{tu} \tag{4.34}$$

not only do we end up with the term

$$\pi^{r}\pi^{stu}[3] \tag{4.35}$$

which is difficult to interpret but also (4.34) reduces to

$$\pi^{r,s,tu} = \pi^{rstu} \tag{4.36}$$

since $\pi^{r} = \pi^{s} = 0$. But (4.36) is incorrect. To see this we apply rule R.4.2 to (4.11) first and obtain

$$\pi^{r,s,t} = \pi^{rst} - \left\{\pi^{r}\pi^{st} + \pi^{s}\pi^{rt} + \pi^{t}\pi^{rs}\right\} + 2\pi^{r}\pi^{s}\pi^{t}\,. \tag{4.37}$$

Then on using (4.37) we have, instead of (4.36), that:

$$\pi^{r,s,tu} = \pi^{rstu} - \pi^{rs}\pi^{tu} \tag{4.38}$$

which is the correct relation that would yield the desired formula in terms of ordinary cumulants. The reader may complete this exercise.

Finally, we will adopt the convention that an expression such as

$$f_{rs}f_{tu}\pi^{r,s}\pi^{t,u}[3] \tag{4.39}$$

shall mean

$$f_{rs}f_{tu}\pi^{r,s}\pi^{t,u} + f_{rt}f_{su}\pi^{r,t}\pi^{s,u} + f_{ru}f_{st}\pi^{r,u}\pi^{s,t}\,. \tag{4.40}$$

We end the examination of generalised cumulants by providing four tables which will serve as our future reference in Chapter five. One table gives some selected expansions of generalised cumulants in terms of ordinary cumulants. The other three present some expansions based on the notation

$$[m]_* \tag{4.41}$$

as these will be used in some of the formulas to come.

The tables are displayed in groups of the same cumulant-order, to make it easy for the reader wishing to do a fairly quick verification.

TABLE 4.1: SOME SELECTED GENERALISED CUMULANTS EXPRESSED IN TERMS OF ORDINARY JOINT CUMULANTS

GROUP A $\alpha = 2$

$$\pi^{r,st} = \pi^{rs,t} = \pi^{r,s,t}$$

$$\pi^{r,stu} = \pi^{rst,u} = \pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3]$$

$$\pi^{r,stuv} = \pi^{rstu,v} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10]$$

$$\pi^{rs,tu} = \pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3] - \pi^{r,s}\pi^{t,u}$$

$$\pi^{rs,tuv} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \pi^{r,s}\pi^{t,u,v}$$

$$\pi^{rst,uv} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \pi^{r,s,t}\pi^{u,v}$$

GROUP B $\alpha = 3$

$$\pi^{r,s,tu} = \pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3] - \pi^{r,s}\pi^{t,u}$$

$$\pi^{r,st,u} = \pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3] - \pi^{r,u}\pi^{s,t}$$

$$\pi^{rs,t,u} = \pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3] - \pi^{rs}\pi^{t,u}$$

$$\pi^{r,s,tuv} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \pi^{r,s}\pi^{t,u,v}$$

$$\pi^{r,stu,v} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \pi^{r,v}\pi^{s,t,u}$$

$$\pi^{rst,u,v} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \pi^{r,s,t}\pi^{u,v}$$

$$\pi^{r,st,uv} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \{\pi^{s,t}\pi^{r,u,v} + \pi^{u,v}\pi^{r,s,t}\}$$

$$\pi^{rs,t,uv} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \{\pi^{r,s}\pi^{t,u,v} + \pi^{u,v}\pi^{r,s,t}\}$$

$$\pi^{rs,tu,v} = \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] - \{\pi^{r,s}\pi^{t,u,v} + \pi^{t,u}\pi^{r,s,v}\}$$

GROUP C $\alpha = 4$

$$\begin{aligned}
\pi^{r,s,t,uv} &= \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] \\
&\quad - \{\pi^{u,v}\pi^{r,s,t} + \pi^{r,s}\pi^{t,u,v} + \pi^{r,t}\pi^{s,u,v} + \pi^{r,u,v}\pi^{s,t}\} \\
\pi^{r,s,tu,v} &= \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] \\
&\quad - \{\pi^{t,u}\pi^{r,s,v} + \pi^{r,s}\pi^{t,u,v} + \pi^{r,t,u}\pi^{s,v} + \pi^{r,v}\pi^{s,t,u}\} \\
\pi^{r,st,u,v} &= \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] \\
&\quad - \{\pi^{s,t}\pi^{r,u,v} + \pi^{r,s,t}\pi^{u,v} + \pi^{r,u}\pi^{s,t,v} + \pi^{r,v}\pi^{s,t,u}\} \\
\pi^{rs,t,u,v} &= \pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10] \\
&\quad - \{\pi^{r,s}\pi^{t,u,v} + \pi^{r,s,t}\pi^{u,v} + \pi^{r,s,u}\pi^{t,v} + \pi^{r,s,v}\pi^{t,u}\}
\end{aligned}$$

TABLE 4.2: SELECTED USE OF $[m]_*$

GROUP D $\alpha = 2$

$$\begin{aligned}
\pi^{r,st}[2]_* &= \pi^{r,st} + \pi^{rs,t} \\
\pi^{r,stu}[2]_* &= \pi^{r,stu} + \pi^{rst,u} \\
\pi^{r,stuv}[2]_* &= \pi^{r,stuv} + \pi^{rstu,v} \\
\pi^{rs,tuv}[2]_* &= \pi^{rs,tuv} + \pi^{rst,uv}
\end{aligned}$$

GROUP E $\alpha = 3$

$$\begin{aligned}
\pi^{r,s,tu}[3]_* &= \pi^{r,s,tu} + \pi^{r,st,u} + \pi^{rs,t,u} \\
\pi^{r,s,tuv}[3]_* &= \pi^{r,s,tuv} + \pi^{r,stu,v} + \pi^{rst,u,v} \\
\pi^{r,st,uv}[3]_* &= \pi^{r,st,uv} + \pi^{rs,t,uv} + \pi^{rs,tu,v}
\end{aligned}$$

GROUP F $\alpha = 4$

$$\pi^{r,s,t,uv}[4]_* = \pi^{r,s,t,uv} + \pi^{r,s,tu,v} + \pi^{r,st,u,v} + \pi^{rs,t,u,v}$$

TABLE 4.3: THE FULL EXPANSIONS OF THE SUMS APPEARING IN TABLE 4.2

GROUP G

$$\pi^{r,st}[2]_* = 2\pi^{r,s,t}$$

$$\pi^{r,stu}[2]_* = 2\left\{\pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3]\right\}$$

$$\pi^{r,stuv}[2]_* = 2\left\{\pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10]\right\}$$

$$\begin{aligned}\pi^{rs,tuv}[2]_* &= 2\left\{\pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10]\right\} \\ &\quad - \left\{\pi^{r,s}\pi^{t,u,v} + \pi^{u,v}\pi^{r,s,t}\right\}\end{aligned}$$

GROUP H

$$\begin{aligned}\pi^{r,s,tu}[3]_* &= 3\left\{\pi^{r,s,t,u} + \pi^{r,s}\pi^{t,u}[3]\right\} \\ &\quad - \left\{2\pi^{r,s}\pi^{t,u} + \pi^{r,u}\pi^{s,t}\right\}\end{aligned}$$

$$\begin{aligned}\pi^{r,s,tuv}[3]_* &= 3\left\{\pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10]\right\} \\ &\quad - \left\{\pi^{r,s}\pi^{t,u,v} + \pi^{r,v}\pi^{s,t,u} + \pi^{u,v}\pi^{r,s,t}\right\}\end{aligned}$$

$$\begin{aligned}\pi^{r,st,uv}[3]_* &= 3\left\{\pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10]\right\} \\ &\quad -2\left\{\pi^{r,s}\pi^{t,u,v} + \pi^{u,v}\pi^{r,s,t}\right\} \\ &\quad - \left\{\pi^{s,t}\pi^{r,u,v} + \pi^{t,u}\pi^{r,s,v}\right\}\end{aligned}$$

GROUP I

$$\begin{aligned}\pi^{r,s,t,uv}[4]_* &= 4\left\{\pi^{r,s,t,u,v} + \pi^{r,s}\pi^{t,u,v}[10]\right\} \\ &\quad -3\left\{\pi^{r,s}\pi^{t,u,v} + \pi^{u,v}\pi^{r,s,t}\right\} \\ &\quad -2\left\{\pi^{r,v}\pi^{s,t,u} + \pi^{s,t}\pi^{r,u,v} + \pi^{t,u}\pi^{r,s,v}\right\} \\ &\quad - \left\{\pi^{r,t}\pi^{s,u,v} + \pi^{r,u}\pi^{s,t,v} + \pi^{s,v}\pi^{r,t,u} + \pi^{t,v}\pi^{r,s,u}\right\}\end{aligned}$$

TABLE 4.4: FURTHER USES OF $[m]_*$ ILLUSTRATED

$$f_r f_s f_{tu}[3]_* = f_r f_s f_{tu} + f_r f_{st} f_u + f_{rs} f_t f_u$$

$$f_r f_s f_{tuv}[3]_* = f_r f_s f_{tuv} + f_r f_{stu} f_v + f_{rst} f_u f_v$$

$$f_r f_{st} f_{uv}[3]_* = f_r f_{st} f_{uv} + f_{rs} f_t f_{uv} + f_{rs} f_{tu} f_v$$

$$f_r f_{st} \pi^{r,st}[2]_* = f_r f_{st} \pi^{r,st} + f_{rs} f_t \pi^{rs,t}$$

$$f_r f_{stu} \pi^{r,stu}[2]_* = f_r f_{stu} \pi^{r,stu} + f_{rst} f_u \pi^{rst,u}$$

$$f_r f_{stuv} \pi^{r,stuv}[2]_* = f_r f_{stuv} \pi^{r,stuv} + f_{rstu} f_v \pi^{rstu,v}$$

$$f_{rs} f_{tuv} \pi^{rs,tuv}[2]_* = f_{rs} f_{tuv} \pi^{rs,tuv} + f_{rst} f_{uv} \pi^{rst,uv}$$

$$f_r f_s f_{tu} \pi^{r,s,tu}[3]_* = f_r f_s f_{tu} \pi^{r,s,tu} f_r f_{st} f_u \pi^{r,st,u} + f_{rs} f_t f_u \pi^{rs,t,u}$$

$$f_r f_s f_{tuv} \pi^{r,s,tuv}[3]_* = f_r f_s f_{tuv} \pi^{r,s,tuv} + f_r f_{stu} f_v \pi^{r,stu,v} + f_{rst} f_u f_v \pi^{rst,u,v}$$

$$f_r f_{st} f_{uv} \pi^{r,st,uv}[3]_* = f_r f_{st} f_{uv} \pi^{r,st,uv} + f_{rs} f_t f_{uv} \pi^{rs,t,uv} + f_{rs} f_{tu} f_v \pi^{rs,tu,v}$$

$$f_r f_s f_t f_{uv} \pi^{r,s,t,uv}[4]_* = f_r f_s f_t f_{uv} \pi^{r,s,t,uv} + f_r f_s f_{tu} f_v \pi^{r,s,tu,v} + f_r f_{st} f_u f_v \pi^{r,st,u,v} + f_{rs} f_t f_u f_v \pi^{rs,t,u,v}$$

4.4 EFFECT OF MODEL ERRORS

4.4.1 INTRODUCTION

In this section we consider the dependence of the cumulants of Z upon the cumulants of model errors. While proceeding in this task, we also will take the opportunity to lay some groundwork for what we will propose in the next chapter to be perhaps the simplest and most direct approach to date in respect of calculating the cumulants of a stochastic Taylor series such as (1.24). Essentially, we will view such r.v. as an *affine* transformation and then simply *invoke* the *basic structure* of the cumulants of such transformations. Lacking a more appropriate name, we will call this approach a *variant* of McCullagh's (1984) *operator-based* technique, see the third section of the next chapter.

To facilitate the laying of the groundwork, let us, instead of (3.12), write

$$Z^r = \ell_i^r \varepsilon^i + c\,, \qquad c = 0\,. \tag{4.42}$$

Since (1.24) is an affine transformation of (4.1), when it comes to the determination of the cumulants of $\hat{\theta}(n)$, which is done in the next chapter, we will feel free to re-use the basic structure of the cumulants of (4.42) as stated in Theorem 4.1 below.

4.4.2 ORDINARY CUMULANTS OF AFFINE TRANSFORMATIONS

THEOREM 4.1

Let $\mu^{i_1,\ldots,i_\upsilon}$ denote the υ-th ordinary joint cumulant of the error vector ε. Then the cumulants of (4.42) are given by:

$$\pi^r = c + \ell_i^r \mu^i \tag{4.43}$$

$$\left.\begin{aligned} \pi^{r,s} &= \ell_i^r \ell_j^s \mu^{i,j} \\ &\vdots \\ \pi^{r,s,t,u,v} &= \ell_i^r \ell_j^s \ell_k^t \ell_\ell^u \ell_m^v \mu^{i,j,k,\ell,m} \\ &\vdots \end{aligned}\right\} \tag{4.44}$$

PROOF

The proof involves straightforward application of the exponentiation identities (3.13) as follows:

$$E\left[\exp\left\{\xi_r Z^r\right\}\right] = e^{c\xi_r} E\left[\exp\left\{\xi_r \ell_i^r \varepsilon^i\right\}\right] \tag{4.45}$$

$$= e^{c\xi_r} \begin{pmatrix} 1 \\ +\xi_r \ell_i^r \mu^i \\ +\frac{1}{2!}\xi_r \xi_s \ell_i^r \ell_j^s \mu^{ij} \\ +\frac{1}{3!}\xi_r \xi_s \xi_t \ell_i^r \ell_j^s \ell_k^t \mu^{ijk} \\ + \\ \vdots \end{pmatrix} \tag{4.46}$$

Let

$$\lambda^{r_1 r_2 \ldots r_\upsilon} = \ell_{i_1}^{r_1} \ell_{i_2}^{r_2} \ldots \ell_{i_\upsilon}^{r_\upsilon} \mu^{i, i_2 \ldots i_\upsilon} \tag{4.47}$$

$$Q_{**} = \xi_r \lambda^r + \frac{1}{2!}\xi_r \xi_s \lambda^{rs} + \frac{1}{3!}\xi_r \xi_s \xi_t \lambda^{rst} + \ldots \tag{4.48}$$

Then

$$\ln E\left[\exp \xi_r Z^r\right] = c\xi_r + Q_{**} - \frac{1}{2}Q_{**}^2 + \frac{1}{3}Q_{**}^3 - \frac{1}{4}Q_{**}^4 + \cdots \tag{4.49}$$

from which the coefficient of ξ_r is evidently (4.43). Furthermore, applying Kaplan's method to obtain the coefficient of, say $\xi_r \xi_s \xi_t$, we have that

$$\pi^{r,s,t} = \lambda^{rst} - \lambda^r \lambda^{st}[3] + 2\lambda^r \lambda^s \lambda^t \tag{4.50}$$

$$= \ell_i^r \ell_j^s \ell_k^t \mu^{ijk} - \left\{\ell_i^r \ell_j^s \ell_k^t \mu^i \mu^{jk} + \ell_j^s \ell_i^r \ell_k^t \mu^j \mu^{ik} + \ell_k^t \ell_i^r \ell_j^s \mu^k \mu^{ij}\right\} + 2\ell_i^r \ell_j^s \ell_k^t \mu^i \mu^j \mu^k . \tag{4.51}$$

Hence

$$\pi^{r,s,t} = \ell_i^r \ell_j^s \ell_k^t \mu^{i,j,k} \tag{4.52}$$

where

$$\mu^{i,j,k} = \mu^{ijk} - \mu^i \mu^{jk}[3] + 2\mu^i \mu^j \mu^k \tag{4.53}$$

$$\mu^i \mu^{jk}[3] = \mu^i \mu^{jk} + \mu^j \mu^{ik} + \mu^k \mu^{ij} . \tag{4.54}$$

All the other members of (4.44) are obtained similarly.

Remark 4.1

Formulae for the cumulants of affine transformations are also given on p. 33 of McCullagh (1987). The author points out that they follow *directly* from the cumulant generating function.

Remark 4.2

Corresponding expressions for the product moments are complicated by the presence of c. This is because such expressions are obtained after expressing *both* sides of the identity:

$$\exp\left\{\xi_r\pi^r + \frac{1}{2!}\xi_r\xi_s\pi^{r,s} + \cdots\right\} = E\left[\exp\left\{c\xi_r + \xi_r\ell_i^r\varepsilon^i\right\}\right] \tag{4.55}$$

in the Maclaurin series.

Now, because the errors are iid, the members of (4.44) simplify further.

THEOREM 4.2

Under the iid assumption of model errors, we have

$$\pi^{r,s} = \rho_2\sum_{i=1}^{n}\ell_{ri}\ell_{si} \tag{4.56a}$$

$$\pi^{r,s,t} = \rho_3\sum_{i=1}^{n}\ell_{ri}\ell_{si}\ell_{ti} \tag{4.56b}$$

$$\pi^{r,s,t,u} = \rho_4\sum_{i=1}^{n}\ell_{ri}\ell_{si}\ell_{ti}\ell_{ui} \tag{4.56c}$$

$$\pi^{r,s,t,u,v} = \rho_5\sum_{i=1}^{n}\ell_{ri}\ell_{si}\ell_{ti}\ell_{ui}\ell_{vi} \tag{4.56d}$$

where ρ_υ is the υ-th cumulant of the common error distribution. In particular

$$\begin{aligned}\rho_2 &= \mu_2\\ \rho_3 &= \mu_3\\ \rho_4 &= \mu_4 - 3\mu_2^2\\ \rho_5 &= \mu_5 - 10\mu_3\mu_2\end{aligned} \tag{4.57}$$

μ_υ being defined by (1.17).

PROOF

$$\begin{aligned}\pi^{r,s} &= \ell_i^r\ell_j^s\mu^{i,j}\\ &= \ell_i^r\ell_i^s\mu^{i,i} && \text{independence}\\ &= \rho_2\ell_i^r\ell_i^s 1^{i,i} && \text{identical distributions}\end{aligned}$$

where

$$1^{i,i} = 1\,.$$

All the other members of (4.56) may be proved likewise.

The following corollary provides *the only other* condition, besides (1.20), from which our Edgeworth expansion will be constructed. We referred to it earlier as the *crucial condition.*

COROLLARY 4.2.1

A necessary condition for

$$\pi^{r_1,\ldots,r_\upsilon} = 0\left(n^{-\upsilon+1}\right) \tag{4.58}$$

is that

$$\ell_i^{r_1}.....\ell_i^{r_\upsilon} 1^{i,\ldots,i} = 0\left(n^{-\upsilon+1}\right) \tag{4.59}$$

where

$$1^{i,\ldots,i} = 1\,.$$

From the definition of the L-matrix given by (1.25), it is evident that to comply with (4.59) one needs to impose the necessary conditions on the entries of the C-matrix taking account of the nature of the inverse of X^TX. This matter is addressed in Section 4.6 in the context of two special cases.

4.5 ESTIMATION OF ρ_υ

4.5.1 INTRODUCTION

The estimation of the cumulants of model errors in the general linear regression model has been considered by Pukelsheim (1980), McCullagh (1987), McCullagh & Pregibon (1987) and perhaps others. The formulas of Pukelsheim are of little computational use and so we will not have more to say about that work.

Section 4.7 of McCullagh (1987) extensively discusses the estimation problem in the context of k-statistics, that is, the unbiased estimates of the cumulants. The discussion, however, proceeds largely in terms of the *multivariate* linear model and using Cartesian tensors. Anyone not well-acquainted with Cartesian tensors, therefore, would have some difficulty in following the account. Nevertheless, at the end of the section simple unbiased estimates are given for the model-error cumulants up to the fourth order, in the case of a single general linear model. These estimates also appear in McCullagh & Pregibon (1987) who also provide other estimates based on the ℓ-statistics. For the meaning of the latter see Section 12.22 of Kendall & Stuart (1977)

4.5.2 THE SIMPLE UNBIASED ESTIMATES

Let $\hat{\rho}_v$ denote the k-statistic for ρ_v and let

$$M = I - X\left(X^T X\right)^{-1} X^T . \tag{4.60}$$

Then, provided that m_3 and $\triangle$ below are not equal to zero:

$$\hat{\rho}_2 = S_2/m_2 \qquad \text{(well-known)} \tag{4.61}$$

$$\hat{\rho}_3 = S_3/m_3 \tag{4.62}$$

$$\hat{\rho}_4 = \triangle^{-1}\left\{m_2\left(m_2+2\right)S_4 - 3m_{22}S_2^2\right\} \tag{4.63}$$

where

$$S_v = \sum_{i=1}^{n} e_i^v \tag{4.64}$$

$$m_{22} = \sum_1 \sum_1^n \sum_1 m_{ri}^2 m_{r_j}^2 \tag{4.65a}$$

$$= \sum_{r=1}^{n} m_{rr}^2 \tag{4.65b}$$

$$m_v = \sum_1 \sum_1^n m_{ri}^v \tag{4.66}$$

$$m_2 = \sum_1 \sum_1^n m_{ri}^2 = n - p \tag{4.67}$$

$$\triangle = m_2 m_4\left(m_2 - 1\right) + 3\left(m_2 m_4 - m_{22}^2\right) \tag{4.68}$$

with both (4.65b) and (4.67) deriving from the fact that matrix M is not only idempotent but also symmetric. Similar formulas for $\hat{\rho}_3$ and $\hat{\rho}_4$ are given on p. 130 of McCullagh (1987) as well as in Sections 3 and 4 of McCullagh & Pregibon (1987). However, it is only in McCullagh (1987) that the details of their justification are discussed. The discussion proceeds in the context of multivariate heteroscedastic linear regression and those formulas are given as the special case of a single homoscedastic linear model. The generality of the regression models considered is at the level of, say, Bunke & Bunke (1986). This, coupled with the use of Cartesian tensor mathematics in the main discussion may perhaps prove difficult for some readers to follow in places. We give below a comparatively lighter derivation of (4.62) and (4.63).

Formula (4.62) is easily obtained by noting from (4.57) that ρ_3 is identical to μ_3. If we express the r-th residual as follows:

$$e^r = m_i^r \varepsilon^i . \tag{4.69}$$

Then upon using (3.13) and $1_r = 1$ we have

$$\sum_1^n \left\{e^r\right\}^3 = 1_r m_i^r m_j^r m_k^r \varepsilon^i \varepsilon^j \varepsilon^k \tag{4.70}$$

which, on taking expectation, gives

$$\begin{aligned} E\left[\sum_1^n \{e^r\}^3\right] &= 1_r m_i^r m_j^r m_k^r \mu^{ijk} \\ &= 1_r m_i^r m_j^r m_k^r \mu^{i,j,k} \\ &= 1_r m_i^r m_i^r m_i \mu^{i,i,i} \qquad \text{independence} \\ &= \rho_3 1_r m_i^r m_i^r m_i^r 1^i \qquad \text{identical distributions} \\ &= \rho_3 \sum_i \sum_r^n m_{ri}^3 . \end{aligned} \tag{4.71}$$

To obtain (4.63) we continue as above, this time with

$$\begin{aligned} E\left[\sum_1^n \{e^r\}^4\right] &= 1_r m_i^r m_j^r m_k^r m_\ell^r \mu^{ijk\ell} \\ &= 1_r m_i^r m_j^r m_k^r m_\ell^r \left\{\mu^{i,j,k,\ell} + \mu^{i,j}\mu^{k,\ell}[3]\right\} \\ &= 1_r m_i^r m_i^r m_i^r m_i^r \mu^{i,i,i,i} \\ &\quad + m_i^r m_j^r m_k^r m_k^r \left\{\mu^{i,j}\mu^{k,\ell} + \mu^{i,k}\mu^{j,\ell} + \mu^{i,\ell} + \mu^{j,k}\right\} \qquad (4.72) \\ &= m_4\rho_4 + 3m_{22}\rho_2^2 . \qquad (4.73) \end{aligned}$$

The non-zero cumulant products in (4.72) are

$$\mu^{i,i}\mu^{k,k} \quad , \quad \mu^{i,i}\mu^{j,j} \quad , \quad \mu^{i,i}\mu^{j,j} .$$

Now, result (4.73) suggests that if there is another expectation of some function of the residuals whose value is linear in ρ_4 and ρ_2^2 and if this identity is independent of (4.73) then the problem of the unbiased estimation of ρ_4 (and ρ_2^2) is solved. It therefore is our wish to find this identity.

The obvious candidate is the expectation of the square of

$$S_2 = m_2\hat{\rho} = 1_r m_i^r m_j^r \varepsilon^i \varepsilon^j . \tag{4.74}$$

This expectation is, of course, not entirely unrelated to the LHS of (4.73) but for the purposes at hand conditions for the linear independence of the identities in ρ_4 and ρ_2^2 can be imposed once the identities are known.

Now, rather than proceed directly, it is more convenient to exploit the power of MINO to evaluate first the variance of S_2^2 and then apply the well-known relationship.

Now

$$\begin{aligned} var\, S_2^2 &= cov\left(1_r m_i^r m_j^r \varepsilon^i \varepsilon^j,\ 1_s m_k^s m_\ell^s \varepsilon^k \varepsilon^\ell\right) \\ &= 1_r 1_s m_i^r m_j^r m_k^s m_\ell^s \mu^{ij,k\ell} \\ &= 1_r 1_s m_i^r m_j^r m_k^s m_\ell^s \mu^{i,j,k,\ell} \\ &\quad +1_r 1_s m_i^r m_j^r m_k^s m_\ell^s \left\{\mu^{i,j}\mu^{k,\ell}[3] - \mu^{i,j}\mu^{k,\ell}\right\} \\ &= 1_r 1_s m_i^r m_i^r m_i^s m_i^s \mu^{i,i,i,i} \\ &\quad +1_r 1_s m_i^r m_j^r m_k^s m_\ell^s \left\{\mu^{i,k}\mu^{j,\ell} + \mu^{i,\ell}\mu^{j,k}\right\} \qquad (4.75) \\ &= m_{22}\rho_4 + 2m_2\rho_2^2 . \qquad (4.76) \end{aligned}$$

This time in (4.75), the non-zero cumulant products are

$$\mu^{i,i}\mu^{j,j} \qquad \mu^{i,i}\mu^{j,j}$$

but it is now the indices r, s that entirely bring about coefficients m_{22} and m_2 in (4.76).

Then, from the well-known relation between the variance and the second moment about zero, we have

$$E\left[S_2^2\right] = m_{22}\rho_4 + m_2\left(m_2 + 2\right)\rho_2^2 \qquad (4.77)$$

and so, provided that the elements of M do not render the relevant matrix singular, solving (4.73) and (4.77) simultaneously completes the proof of (4.63), in addition to giving

$$\widehat{\rho_2^2} = \triangle^{-1}\left\{m_4 S_2^2 - m_{22} S_4\right\} \qquad (4.78)$$

as the unbiased estimate of ρ_2^2.

4.6 COROLLARY 4.2.1 REVISITED

Since we shall carry out simulation studies in Chapter eight, based on the estimation problems cited in Examples 1.1 and 1.3, it is perhaps appropriate to conclude this chapter with an examination of the conditions on the design matrices under which (4.59) is satisfied in each of these two regression settings. Below we do this for cumulants up to fourth-order for each model context. All summations unspecified run from 1 to n.

4.6.1 THE SIMPLE LINEAR MODEL

Here, with x_i used to represent the deviation $c_i - \bar{c}$, we have:

$$\begin{aligned} L &= \begin{bmatrix} \ell_{01} & \ell_{02} \ldots\ldots \ell_{0n} \\ \ell_{11} & \ell_{12} \ldots\ldots \ell_{1n} \end{bmatrix} \\ &= \frac{1}{n\Sigma x_i^2}\begin{bmatrix} \Sigma x_i^2 & 0 \\ 0 & n \end{bmatrix}\begin{bmatrix} 1 & 1 & \ldots\ldots & 1 \\ x_1 & x_2 & \ldots\ldots & x_n \end{bmatrix} \end{aligned} \qquad (4.79)$$

so that for $j = 1, 2, ..., n$

$$\ell_{0j} = n^{-1} \tag{4.80}$$

$$\ell_{1j} = \triangle^{-1} x_j \,; \qquad \triangle = \Sigma x_i^2 \,. \tag{4.81}$$

Then (4.56a) gives the following:

$$\pi^{0,0} = \mu_2 \left\{ \ell_{01}^2 + \ell_{02}^2 + ... + \ell_{0n}^2 \right\} = n^{-1} \mu_2 \tag{4.82}$$

$$\pi^{1,1} = \mu_2 \left\{ \ell_{11}^2 + \ell_{12}^2 + ... + \ell_{1n}^2 \right\} = \triangle^{-1} \mu_2 \tag{4.83}$$

$$\pi^{0,1} = \mu_2 \left\{ \ell_{01}\ell_{11} + \ell_{02}\ell_{12} + ... + \ell_{0n}\ell_{1n} \right\} = \pi^{1,0} = 0 \,. \tag{4.84}$$

From (4.56b):

$$\pi^{0,0,0} = \mu_3 \left\{ \ell_{01}^3 + \ell_{01}^3 + ... + \ell_{0n}^3 \right\} = n^{-2} \mu_3 \tag{4.85}$$

$$\pi^{1,1,1} = \mu_3 \left\{ \ell_{11}^3 + \ell_{12}^3 + ... + \ell_{1n}^3 \right\} = \triangle^{-3} \mu_3 \Sigma x_j^3 \tag{4.86}$$

$$\begin{aligned} \pi^{0,0,1} &= \mu_3 \left\{ \ell_{01}^2 \ell_{11} + \ell_{02}^2 \ell_{12} + ... + \ell_{0n}^2 \ell_{1n} \right\} \\ &= 0 = \pi^{0,1,0} = \pi^{1,0,0} \end{aligned} \tag{4.87}$$

$$\begin{aligned} \pi^{0,1,1} &= \mu_3 \left\{ \ell_{01}\ell_{11}^2 + \ell_{02}\ell_{12}^2 + ... + \ell_{0n}\ell_{1n}^2 \right\} \\ &= (n\triangle)^{-1} \mu_3 = \pi^{1,0,1} = \pi^{1,1,0} \,. \end{aligned} \tag{4.88}$$

Finally, from (4.56c)

$$\pi^{0,0,0,0} = \rho_4 \left\{ \ell_{01}^4 + \ell_{02}^4 + ... + \ell_{0n}^4 \right\} = n^{-3} \rho_4 \tag{4.89}$$

$$\pi^{1,1,1,1} = \rho_4 \left\{ \ell_{11}^4 + \ell_{12}^4 + ... + \ell_{1n}^4 \right\} = \triangle^{-4} \rho_4 \Sigma x_j^4 \tag{4.90}$$

$$\begin{aligned} \pi^{0,0,0,1} &= \rho_4 \left\{ \ell_{01}^3 \ell_{11} + \ell_{02}^3 \ell_{12} + ... + \ell_{0n}^3 \ell_{1n} \right\} \\ &= 0 = \pi^{0,0,1,0} = \pi^{0,1,0,0} = \pi^{1,0,0,0} \end{aligned} \tag{4.91}$$

$$\begin{aligned} \pi^{0,0,1,1} &= \rho_4 \left\{ \ell_{01}^2 \ell_{11}^2 + \ell_{02}^2 \ell_{12}^2 + ... + \ell_{0n}^2 \ell_{1n}^2 \right\} \\ &= n^{-2} \triangle^{-1} \rho_4 = \pi^{0,1,0,1} = \pi^{1,0,1,0} = \pi^{1,1,0,0} \end{aligned} \tag{4.92}$$

$$\begin{aligned} \pi^{0,1,1,1} &= \rho_4 \left\{ \ell_{01}\ell_{11}^3 + \ell_{02}\ell_{12}^3 + ... + \ell_{01}\ell_{1n}^3 \right\} \\ &= n^{-1} \triangle^{-3} \rho_4 \Sigma x_j^3 = \pi^{1,0,1,1} = \pi^{1,1,0,1} = \pi^{1,1,1,0} \,. \end{aligned} \tag{4.93}$$

Hence if we choose condition

$$\Sigma x_j^v = 0(n) \,, \qquad v = 2, 3, ... \tag{4.94}$$

we satisfy (4.59).

4.6.2 THE QUADRATIC MODEL

For the quadratic regression model (1.6), the form of the inverse of matrix X^TX may be determined using, say, the rule of Neter & Wasserman (1974) on pp. 194–5, since in this case

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} . \tag{4.95}$$

However, for simplicity we will limit our examination to the case of estimating the x-co-ordinate of the turning point. As (1.7) shows, this parameter is independent of the intercept term, so that we can proceed using model (1.6) with $\beta_1 = 0$. For this case

$$\begin{aligned} L &= (\ell_{rs})_{2\times n} \qquad r = 2,\ 3 \\ &\qquad\qquad\quad s = 1,\ 2,\ \ldots,\ n \\ &= \triangle_*^{-1} \begin{bmatrix} \Sigma x_i^4 & -\Sigma x_i^3 \\ -\Sigma x_i^3 & \Sigma x_i^2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \end{bmatrix} \end{aligned} \tag{4.96}$$

with

$$\triangle_* = \Sigma x_i^2 \Sigma x_i^4 - \Sigma x_i^3 \Sigma x_i^3 . \tag{4.97}$$

Hence

$$\begin{aligned} \ell_{2s} &= \triangle_*^{-1} \{x_s \Sigma x_i^4 - x_s^2 \Sigma x_i^3\} \\ \ell_{3s} &= \triangle_*^{-1} \{x_s^2 \Sigma x_i^2 - x_s \Sigma x_i^3\} . \end{aligned} \tag{4.98}$$

Again all the summations unspecified run from 1 to n. Now, with (4.98) we have the following:

$$\ell_{2s}^2 = \triangle_*^{-2} \left\{ x_s^2 \left(\Sigma x_i^4\right)^2 - 2x_s^3 \Sigma x_i^3 \Sigma x_i^4 + x_s^4 \left(\Sigma x_i^3\right)^2 \right\} \tag{4.99}$$

$$\ell_{3s}^2 = \triangle_*^{-2} \left\{ x_s^2 \left(\Sigma x_i^3\right)^2 - 2x_s^3 \Sigma x_i^2 \Sigma x_i^3 + x_s^4 \left(\Sigma x_i^2\right)^2 \right\} \tag{4.100}$$

$$\ell_{2s}^3 = \triangle_*^{-3} \left\{ x_s^3 \left(\Sigma x_i^4\right)^3 - 3x_s^4 \Sigma x_i^3 \left(\Sigma x_i^4\right)^2 + 3x_s^5 \Sigma x_i^4 \left(\Sigma x_i^3\right)^2 - x_s^6 \left(\Sigma x_i^3\right)^3 \right\} \tag{4.101}$$

$$\ell_{3s}^3 = \triangle_*^{-3} \left\{ -x_s^3 \left(\Sigma x_i^3\right)^3 + 3x_s^4 \Sigma x_i^2 \left(\Sigma x_i^3\right)^2 - 3x_s^5 \Sigma x_i^3 \left(\Sigma x_i^2\right)^2 + x_s^6 \left(\Sigma x_i^2\right)^3 \right\} \tag{4.102}$$

$$\ell_{2s}^4 = \triangle_*^{-4} \left[\begin{array}{l} x_s^4 \left(\Sigma x_i^4\right)^4 - 4x_s^5 \Sigma x_i^3 \left(\Sigma x_i^4\right)^3 + 6x_s^6 \left(\Sigma x_i^3\right)^2 \left(\Sigma x_i^4\right)^2 \\ -4x_s^7 \Sigma x_i^4 \left(\Sigma x_i^3\right)^3 + x_s^8 \left(\Sigma_i^3\right)^4 \end{array} \right] \tag{4.103}$$

$$\ell_{3s}^4 = \triangle_*^{-4} \left[\begin{array}{l} x_s^4 \left(\Sigma x_i^3\right)^4 - 4x_s \Sigma x_i^2 \left(\Sigma x_i^3\right)^3 + 6x_s^6 \left(\Sigma x_i^2\right)^2 \left(\Sigma x_i^3\right)^2 \\ -4x_s^7 \Sigma x_i^3 \left(\Sigma x_i^2\right)^3 + x_s^8 \left(\Sigma_i^2\right)^4 \end{array} \right] \tag{4.104}$$

$$\ell_{2s}\ell_{3s} = \triangle_*^{-2} \left\{ -x_s^2 \Sigma x_i^3 \Sigma x_i^4 + x_s^3 \left\{ \Sigma x_i^3 \Sigma x_i^3 + \Sigma x_i^2 \Sigma x_i^4 \right\} - x_s^4 \Sigma x_i^2 \Sigma x_i^3 \right\} \tag{4.105}$$

$$\ell_{2s}^2 \ell_{3s} = \triangle_*^{-3} \left[\begin{array}{l} -x_s^3 \Sigma x_i^3 \left(\Sigma x_i^4\right)^2 + x_s^4 \left\{ \Sigma x_i^2 \left(\Sigma x_i^4\right)^2 + 2\Sigma x_i^4 \left(\Sigma x_i^3\right)^2 \right\} \\ -x_s^5 \left\{ 2\Sigma x_i^2 \Sigma x_i^3 \Sigma x_i^4 + \left(\Sigma x_i^3\right)^3 \right\} + x_s^6 \Sigma x_i^2 \left(\Sigma x_i^3\right)^2 \end{array} \right] \tag{4.106}$$

$$\ell_{2s}\ell_{3s}^2 = \triangle_*^{-3} \left[\begin{array}{l} x_s^3 \Sigma x_i^4 \left(\Sigma x_i^3\right)^2 - x_s^4 \left\{ 2\Sigma x_i^2 \Sigma x_i^3 \Sigma x_i^4 + \left(\Sigma x_j^3\right)^3 \right\} \\ +x_s^5 \left\{ 2\Sigma x_i^2 \left(\Sigma x_i^3\right)^2 + \Sigma x_i^4 \left(\Sigma x_i^2\right)^2 \right\} - x_s^6 \Sigma x_i^3 \left(\Sigma x_i^2\right)^2 \end{array} \right] \tag{4.107}$$

$$\ell_{2s}^2 \ell_{3s}^2 = \triangle_*^{-4} \left[\begin{array}{l} x_s^4 \left(\Sigma x_i^3\right)^2 \left(\Sigma x_i^4\right)^2 \\ -2x_s^5 \left\{ \Sigma x_i^2 \Sigma x_i^2 \Sigma x_i^3 \left(\Sigma x_i^4\right)^2 + \Sigma x_i^4 \left(\Sigma x_i^3\right)^3 \right\} \\ +x_s^6 \left\{ \left(\Sigma x_i^2\right)^2 \left(\Sigma x_i^4\right)^2 + 4\Sigma x_i^2 \Sigma x_i^4 \left(\Sigma x_i^3\right)^2 + \left(\Sigma x_i^3\right)^4 \right\} \\ -2x_s^7 \left\{ \Sigma x_i^3 \Sigma x_i^4 \left(\Sigma x_i^2\right)^2 + \Sigma x_i^2 \left(\Sigma x_i^3\right)^3 \right\} \\ +x_s^8 \left(\Sigma x_i^2\right)^2 \left(\Sigma x_i^3\right)^2 \end{array} \right] \tag{4.108}$$

$$\ell_{2s}^3 \ell_{3s} = \triangle_*^{-4} \left[\begin{array}{l} -x_s^4 \Sigma x_i^3 \left(\Sigma x_i^4\right)^3 \\ +x_s^5 \left\{ \Sigma x_i^2 \left(\Sigma x_i^4\right)^3 + 3\left(\Sigma x_i^3\right)^2 \left(\Sigma x_i^4\right)^2 \right\} \\ -3x_s^6 \left\{ \Sigma x_i^2 \Sigma x_i^3 \left(\Sigma x_i^4\right)^2 + \Sigma x_i^4 \left(\Sigma x_i^3\right)^3 \right\} \\ +x_s^7 \left\{ 3\Sigma x_i^2 \Sigma x_i^4 \left(\Sigma x_i^3\right)^2 + \left(\Sigma x_i^3\right)^4 \right\} \\ -x_s^8 \Sigma x_i^2 \left(\Sigma x_i^3\right)^3 . \end{array} \right] \tag{4.109}$$

Finally,

$$\ell_{2s}\ell_{3s}^3 = \triangle_*^{-4} \left[\begin{array}{l} -x_s^4 \Sigma x_i^4 \left(\Sigma x_i^3\right)^3 \\ +x_s^5 \left\{ 3\Sigma x_i^2 \Sigma x_i^4 \left(\Sigma x_i^3\right)^2 + \left(\Sigma x_i^3\right)^4 \right\} \\ -3x_s^6 \left\{ \Sigma x_i^3 \Sigma x_i^4 \left(\Sigma x_i^2\right)^2 + \Sigma x_i^2 \left(\Sigma x_i^3\right)^3 \right\} \\ +x_s^7 \left\{ \Sigma_i^2 \Sigma x_i^4 \left(\Sigma x_i^2\right)^3 + \left(\Sigma x_i^2\right)^2 \left(\Sigma x_i^3\right)^2 \right\} \\ -x_s^8 \Sigma x_i^3 \left(\Sigma x_i^2\right)^3 . \end{array} \right] \tag{4.110}$$

Then, putting (4.99) to (4.110) in (4.56a), (4.56b) and (4.56c) in the same way as in the case of the simple linear model, we find on summing over index s that condition (4.59) will be satisfied *also by* (4.94) provided in this case, that:

$$\triangle_* = 0\left(n^2\right) \tag{4.111}$$

4.7 CHAPTER SUMMARY

There are four main features of this chapter. Firstly, we have discussed in detail the property of generalised cumulants which relates them to moments and ordinary cumulants. During this discussion, some elementary but important operational cautionary notes were highlighted. We have not seen them in the previous accounts that we have read but as we have pointed out several times before, generalised cumulants arise naturally in the second- and higher-order cumulants of multivariate stochastic Taylor series and so manipulations involving the latter cumulants inevitably would have to involve the former.

Secondly, we have examined the effect of the model errors on the cumulants of $\hat{\beta}(n)-\beta$. In particular, quite apart from the fact that the results in Theorem 4.1 will, on their own, be used in simulations, in the next chapter we will propose a simple, direct alternative method for calculating the cumulants of a stochastic Taylor series, which is based entirely on the basic structure of those results. Consequently, the proof that we have given for Theorem 4.1 shall also serve as the sole justification of that method. Also in connection with the cumulants of $\hat{\beta}(n) - \beta$, we have identified, through Theorem 4.2, a condition on the elements of matrix L which will enable us to capture the first two terms of the first four cumulants of the standardised $\hat{\theta}(n)$ in the first two terms of our Edgeworth series and in the usual way. The condition is stated in Corollary 4.2.1.

Thirdly, we have reviewed the estimation of the cumulants of the model errors using k-statistics. We point out that this is extensively done in Section 4.7 of McCullagh (1987) using Cartesian tensor mathematics. For the reader who may find discomfort with the level of that discussion, a comparatively lighter account for a single homoscedastic general linear regression situation has been provided.

Fourthly and lastly, we have gone through an elaborate determination of the conditions on the C-matrix which lead to the fulfilment of (4.59) in the two special cases that we will consider in simulations.

CHAPTER FIVE

THE CUMULANTS OF $\kappa_{2,\infty}^{-\frac{1}{2}} T_n$

5.1 INTRODUCTION

From Leonov and Shiryaev (1959), it appears that interest in the methods of calculating cumulants of nonlinear random transformations dates back to the early 1950's in Russia, in relation to nonlinear transformation of random processes.

The work of Leonov and Shiryaev itself focuses on the cumulants of an arbitrary number of nonlinear functions of an arbitrary number of r.v.'s, the cumulants of the r.v.'s being known. However, the authors' main results are expressed in rather unwieldy formulas, based on some combinatorial method. Another combinatorial method, due to James & Mayne (1962), is comparatively more transparent but its main results are expressed using index notation that begs more simplicity. Nowhere in the James and Mayne paper do the authors indicate awareness of the work of Leonov and Shiryaev. However, in view of limitations just indicated, we will not pursue further the above two methods.

More recently, McCullagh (1984) has proposed a third method using a much simplified index notation. We review this method in the next section in the context of our problem. In Section 5.3 we propose a variant which takes an even simpler form. Section 5.4 examines certain cumulant conditions which the Edgeworth expansion in Chapter six will need. Also in Chapter six, we will have occasion to deal with an Edgeworth expansion of a *univariate* polynomial transformation in relation to a method due to Hall (1992a, 1992b). We will offer an alternative based on the theory of cumulants for such transformations. This theory is developed in Section 5.5. The last section

addresses the actual computation of an important component of skewness correction which we will need in simulations.

5.2 A REVIEW OF McCULLAGH'S METHOD

In the context of our problem, the method of McCullagh (1984) proceeds by first re-expressing the Taylor-expansion (3.11) in the following *operator notation*:

$$\hat{\theta}(n) = (A_0 + A_1 + A_2 + ...)\, Z \tag{5.1a}$$

$$= PZ, \quad \text{say} \tag{5.1b}$$

where

$$\begin{aligned} A_0 Z &= \theta \\ A_1 Z &= f_r Z^r \\ A_2 Z &= \frac{1}{2!} f_{rs} Z^r Z^s \end{aligned} \tag{5.2}$$

and so on. In a similar manner, the action of operator A acting *singly* on π, where π is a structure representing the cumulants of Z, is *defined* in the way that is illustrated by the following examples:

$$\begin{aligned} 1\pi &= 0 \\ A_0\pi &= \theta \\ A_1\pi &= f_r \pi^r \\ A_2\pi &= \frac{1}{2!} f_{rs} \pi^{rs} \\ A_3\pi &= \frac{1}{3!} f_{rs} \pi^{rst} \end{aligned} \tag{5.3}$$

and so on. Also by definition, compound operators introduce commas as follows:

$$\begin{aligned} A_1 A_1 \pi &= f_r f_s \pi^{r,\,s} \\ A_2 A_1 \pi &= \frac{1}{2!} f_{rs} f_t \pi^{rs,t} \\ A_2 A_3 A_1 \pi &= \frac{1}{3!} f_{rs} f_{tuv} f_w \pi^{rs,\,tuv,\,w} \end{aligned} \tag{5.4}$$

and any compound term involving A_0 acting on π gives zero. The operators A are not commutative.

With the above definitions, the method expresses the cumulant generating function of Z, say $\psi^*(\xi)$ as follows:

$$\psi^*(\xi) = \exp(\xi P)\,\pi\,. \tag{5.5}$$

Then expanding the exponent in (5.5) about $\xi = 0$ gives

$$\begin{aligned}\psi^*(\xi) \;=\; & \xi\,(A_0 + A_1 + A_2 \ldots)\,\pi \\ & + \frac{1}{2!}\xi^2\,(A_0 + A_1 + A_2 + \ldots)^2\,\pi \\ & + \frac{1}{3!}\xi^3\,(A_0 + A_1 + A_2 + \ldots)^3\,\pi \\ & + \\ & \vdots\end{aligned} \tag{5.6}$$

Let κ_υ denote the υ-th cumulant of (5.1a). Then (5.6) gives

$$\kappa_1 = (A_0 + A_1 + A_2 + A_3 + \ldots)\,\pi \tag{5.7}$$

and for $i \neq j \neq k \neq \ell = 1, 2, 3, \ldots$

$$\kappa_2 = (A_iA_i + A_iA_j\,[2]_* + \ldots)\,\pi \tag{5.8}$$

$$\kappa_3 = (A_iA_iA_i + A_iA_iA_j\,[3]_* + A_iA_jA_k\,[6]_* + \ldots)\,\pi \tag{5.9}$$

$$\begin{aligned}\kappa_4 \;=\; & (A_iA_iA_iA_i + A_iA_iA_iA_j\,[4]_* + A_iA_iA_jA_j\,[6]_* \\ & + A_iA_iA_jA_k\,[12]_* + A_iA_jA_kA_\ell\,[24]_* + \ldots)\,\pi\end{aligned} \tag{5.10}$$

where, for example

$$A_iA_j\,[2] \;=\; A_iA_j + A_jA_i \tag{5.11a}$$

$$A_iA_iA_j\,[3] \;=\; A_iA_iA_j + A_iA_jA_i + A_jA_iA_i \tag{5.11b}$$

$A_iA_iA_jA_j\,[6]_*$ is a sum of similar terms over

$$\begin{array}{cccc|cccc} i & i & j & j & j & j & i & i \\ i & j & i & j & j & i & j & i \\ i & j & j & i & j & i & i & j \end{array} \tag{5.12a}$$

while $A_iA_iA_jA_k\,[12]_*$ is a sum of similar terms over:

i	i	j	k	i	i	k	j
i	j	i	k	i	k	i	j
i	j	k	i	i	k	j	i
j	i	i	k	j	k	i	i
j	i	k	i			–	
k	i	i	j	k	j	i	i
k	i	j	i			–	

(5.12b)

The author points out that there are a number of ways in which (5.5) may be derived. The author's preference is to proceed directly by observing that the mgf of Z is given by

$$\psi^*(\xi) = \exp(\xi P) \times \pi \tag{5.13}$$

where the operator "$\times$" is defined as in (5.3) and (5.4), except that the commas are omitted in (5.4) and $1 \times \pi = 1$. The expansion of (5.13) then gives results (5.6) to (5.10), with minor modifications such as the removal of commas and addition of 1. The effect of taking logarithms is to delete the leading constant 1, and to insert commas appropriately.

5.3 A PROPOSED VARIANT

Let us re-write the Taylor expansion (3.11) as follows:

$$\hat{\theta}(n) = \theta + 1_i U^i \tag{5.14}$$

where

$$\begin{aligned} 1_i &= 1 \\ U^1 &= f_r Z^r \\ U^2 &= \frac{1}{2!} f_{rs} Z^r Z^s \\ U^3 &= \frac{1}{3!} f_{rst} Z^r Z^s Z^t \end{aligned} \tag{5.15}$$

and so on. Then an alternative to the use of (5.5) and the operator A in the manner described in the last section, is to invoke the basic structure in (4.43) and (4.44) twice. First, we invoke it to extract

the cumulants of $\hat{\theta}(n)$ as follows:

$$\begin{aligned} \kappa_1 &= \theta + 1_i K^i \\ \kappa_2 &= 1_i 1_j K^{i,j} \\ \kappa_3 &= 1_i 1_j 1_k K^{i,j,k} \qquad \kappa_4 = 1_i 1_j 1_k 1_\ell K^{i,j,k,\ell} \end{aligned} \tag{5.16}$$

and so on, where

$$\begin{aligned} K^i &= E[U^i] \\ K^{i,j} &= \text{cov}\left(U^i, U^j\right) \\ K^{i,j,k} &= \text{cumulant of } U^i, U^j \text{ and } U^k \\ &\vdots \end{aligned} \tag{5.17}$$

Then we invoke (4.44) once again to evaluate the relevant RHS of (5.17) to obtain explicit forms for

$$K^{i,j},\ K^{i,j,k},\ K^{i,j,k,\ell}, \cdots . \tag{5.18}$$

In this way we find that the first four cumulants of $\hat{\theta}(n)$ are derived below as follows:

$$\begin{aligned} \kappa_1 &= \theta + K^1 + K^2 + K^3 + ... \\ &= \theta + f_r \pi^r + \frac{1}{2!} f_{rs} \pi^{rs} + \frac{1}{3!} f_{rst} \pi^{rst} + \cdots \\ &= \theta + \frac{1}{2} f_{rs} \pi^{r,s} + \frac{1}{6} f_{rst} \pi^{r,s,t} + \cdots \end{aligned} \tag{5.19}$$

$$\begin{aligned} \kappa_2 &= K^{1,1} + K^{1,2} + K^{1,3} + \cdots \\ &\quad + K^{2,1} + K^{2,2} + K^{2,3} + \cdots \\ &\quad + K^{3,1} + K^{3,2} + K^{3,3} + \cdots \\ &\quad + ... \\ &= f_r f_s \pi^{r,s} + \frac{1}{2} f_r f_{st} \pi^{r,st} [2]_* + \frac{1}{6} f_r f_{stu} \pi^{r,stu} [2]_* + \frac{1}{4} f_{rs} f_{tu} \pi^{rs,tu} + \cdots \end{aligned} \tag{5.20}$$

where the rectangular bracket notation is used in the sense shown in Table 4.4. In like manner

$$\begin{aligned} 1_i 1_j 1_k K^{i,j,k} &= K^{1,1,1} + K^{1,1,2} [3]_* + K^{1,1,3} [3]_* + \cdots \\ &\quad + K^{1,2,2} [3]_* + K^{1,2,3} [6]_* + \cdots \\ &\quad + K^{2,2,2} + K^{2,2,3} [3]_* + K^{2,2,4} [3]_* + \cdots \\ &\quad + \cdots \end{aligned} \tag{4.21}$$

where for example

$$K^{1,1,2} [3]_* = K^{1,1,2} + K^{1,2,1} + K^{2,1,1} . \tag{5.22}$$

Hence

$$\begin{aligned}\kappa_3 &= f_r f_s f_t \pi^{r,s,t} + \frac{1}{2} f_r f_s f_{tu} \pi^{r,s,tu} [3]_* \\ &+ \frac{1}{6} f_r f_s f_{tuv} \pi^{r,s,tuv} [3]_* + \frac{1}{4} f_r f_{st} f_{uv} \pi^{r,st,uv} [3]_* \\ &+ \frac{1}{12} f_r f_{st} f_{uvw} \pi^{r,st,uvw} [3]_* + \frac{1}{8} f_{rs} f_{tu} f_{vw} \pi^{rs,tu,vw} \\ &+ \cdots . \end{aligned} \tag{5.23}$$

Finally

$$\begin{aligned}\kappa_4 &= 1_i 1_j 1_k 1_\ell K^{i,j,k,\ell} \\ &= K^{1,1,1,1} \\ &\quad + K^{1,1,1,2} [4]_* + K^{1,1,1,3} [4]_* + \ldots\ldots \\ &\quad + K^{1,1,2,2} [6]_* + K^{1,1,2,3} [9]_* + \ldots\ldots \\ &\quad + \\ &\quad \vdots \end{aligned} \tag{5.24}$$

$$\begin{aligned} &= f_r f_s f_t f_u \pi^{r,s,t,u,} \\ &\quad + \frac{1}{2} f_r f_s f_t f_{uv} \pi^{r,s,t,uv} [4]_* \\ &\quad + \frac{1}{6} f_r f_s f_t f_{uvw} \pi^{r,s,t,uvw} [4]_* \\ &\quad + \frac{1}{4} f_r f_s f_{tu} f_{vw} \pi^{r,s,tu,vw} [6]_* \\ &\quad + \\ &\quad \vdots \end{aligned} \tag{5.25}$$

where $f_r f_s f_t f_{uv} \pi^{r,s,t,uv} [4]_*$ is given by Table 4.4 while $f_r f_s f_{tu} f_{vw} \pi^{r,s,tu,vw} [6]_*$ represents a sum of similar terms over the following possible index structure

$$\begin{array}{cccc|cccc} r, & s, & tu, & vw & rs, & tu, & v, & w \\ r, & st, & u, & vw & rs, & t, & uv, & w \\ r, & s, & tu, & vw & rs, & t, & u, & vw \end{array} \tag{5.26}$$

Remark 5.1

Since formulas (5.19), (5.20), (5.23) and (5.25) derive entirely from the direct application of (4.43) and (4.44), it is the proof of the latter that justifies those formulas.

Remark 5.2

In the next section we will see that, so far as our main objective is concerned, there are *mandatory* terms that must be retained in each of the first four cumulant formulas. These terms are of the same order in n^{-v+1}, $v = 2,\ 3,\ 4$ and 5 and are determined by the application of result (4.59) of Corollary 4.2.1 as well as other cumulant conditions which we will see in that section.

Remark 5.3

In seeking to isolate those mandatory terms mentioned in the last remark, it is essential that all the generalised cumulants that are *known* to contain terms of the order sought be first expanded in terms of ordinary cumulants.

Remark 5.4

Beside being very transparent and simple, our method allows one to see each cumulant expansion as consisting of separate expansions. This way the full systematic sequencing of all the terms involved in a given cumulant expansion is seen more clearly. Thus for example, we can see that in the second line of equation (3.10) of McCullagh (1987,p. 64) which, incidentally corrects the coefficients in equation (11) of McCullagh (1984), the missing term

$$a_i^r a_{jk\ell}^s \kappa^{i,j,k,\ell}\,[2] \tag{5.27}$$

should appear as well, preferably before

$$a_{ij}^r a_{k\ell}^s \kappa^{ij,k\ell} \tag{5.28}$$

since the latter, i.e. (5.28), does appear. The significance of this omission will be examined in the next section.

5.4 ESSENTIAL CONDITIONS

Recall that the delta method described in Section 2.3.1 provides the asymptotic normal theory for the statistic

$$W_n = \kappa_{2,\infty}^{-\frac{1}{2}} T_n \tag{5.29}$$

Let $\kappa_{v,n}$ denote the v-th cumulant of this statistic, that is:

$$\kappa_{v,n} = \frac{n^{\frac{v}{2}}}{\kappa_{2,\infty}^{\frac{v}{2}}} \times \text{ cumulant of } \hat{\theta}(n) - \theta\,. \tag{5.30}$$

In the next chapter, our derivation of the Edgeworth series for the d.f. of W_n will proceed under the following conditions on (5.30)

$$\kappa_{1,n} = 0 + n^{-\frac{1}{2}}\lambda_1 + 0\left(n^{-\frac{3}{2}}\right) \tag{5.31}$$

$$\kappa_{2,n} = 1 + n^{-1}\lambda_2 + 0\left(n^{-2}\right) \tag{5.32}$$

$$\kappa_{3,n} = 0 + n^{-\frac{1}{2}}\lambda_3 + 0\left(n^{-\frac{3}{2}}\right) \tag{5.33}$$

$$\kappa_{r,n} = 0 + n^{-\frac{1}{2}(r-2)}\lambda_r + 0\left(n^{-\frac{r}{2}}\right) \qquad r = 4,\ 5,\ \ldots \tag{5.34}$$

where

$$\lambda_v = 0\,(1) \qquad v = 1,\ 2,\ \ldots \tag{5.35}$$

The *primary* purpose of (5.31) to (5.35) is to ensure the fulfilment of the two objectives stated early in the last chapter, namely, (1) to preserve the asymptotic normality of the delta method, and (2) to ensure that the first two terms of the first four cumulants of our statistic can be captured in the first two terms of the Edgeworth expansion and in the traditional way.

Now, under the limit condition (1.20) in Chapter one we have that:

$$\kappa_{2,\infty} = 0\,(1)\ . \tag{5.36}$$

Therefore, conditions (5.31) to (5.34) imply that

$$\begin{aligned} \kappa_1 &= a_{1,n} + 0\left(n^{-2}\right) \\ \kappa_2 &= a_{2,n} + 0\left(n^{-3}\right) \\ \kappa_3 &= a_{3,n} + 0\left(n^{-2}\right) \\ \kappa_r &= a_{r,n} + 0\left(n^{-r}\right) \qquad r = 4,\ 5,\ \ldots \end{aligned} \tag{5.37}$$

where, for each v, $a_{v,n}$ represents the lower-order *mandatory* terms mentioned in Remark 5.2. They must be retained for the purpose of calculating the λ_v. In determining these terms we should, as pointed out in Remark 5.3, take into account the fact that some members of $a_{v,n}$ may come from expanding some generalised cumulants in terms of ordinary cumulants. Also crucial in this regard is result (4.59) of Corollary 4.2.1. Accordingly then, our method proceeds further as follows. From

(5.19) we see that there is no generalised cumulant to consider. Hence:

$$a_{1,n} = \frac{1}{2} f_{rs} \pi^{r,s} \qquad \lambda_1 = n \kappa_{2,\infty}^{-\frac{1}{2}} a_{1,n} \,. \tag{5.38}$$

However, from (5.20) we see that upon expanding separately each of the following

$$\frac{1}{2} f_r f_{st} \pi^{r,st} [2]_* \tag{5.39}$$

$$\frac{1}{6} f_r f_{stu} \pi^{r,stu} [2]_* \tag{5.40}$$

$$\frac{1}{4} f_{rs} f_{tu} \pi^{rs,tu} \,, \tag{5.41}$$

using the appropriate tables at the end of the last chapter, we have

$$\begin{aligned} a_{2,n} \;=\; & f_r f_s \pi^{r,s} \\ & + f_r f_{st} \pi^{r,s,t} \\ & + \frac{1}{6} \left(f_r f_{stu} + f_{rst} f_u \right) \pi^{r,s} \pi^{t,u} [3] \\ & + \frac{1}{4} f_{rs} f_{tu} \left(\pi^{r,s} \pi^{t,u} [3] - \pi^{r,s} \pi^{t,u} \right) \,. \end{aligned} \tag{5.42}$$

Then from (5.32), (5.37), and (5.42), we see that we may set

$$\kappa_{2,\infty} \;=\; n f_r f_s \pi^{r,s} \tag{5.43}$$

$$\lambda_2 \;=\; n \kappa_{2,\infty}^{-1} a_{2,n}^* \tag{5.44}$$

where $a_{2,n}^*$ represents the sum of the three terms in (5.42) which are subsequent to $f_r f_s \pi^{r,s}$.

The significance of Remarks 5.3 and 5.4 can now be brought to the fore. With respect to Remark 5.3 we see that had we not expanded all those generalised cumulants which are *known* to contain terms that are of order n^{-2}, we would not have obtained identity (5.42). With respect to Remark 5.4, we see that the third term in (5.42) represents the contribution of the term

$$\frac{1}{6} f_r f_{stu} \pi^{r,stu} [2]_* \,. \tag{5.45}$$

This term corresponds to term (5.27) which is missing from equation (3.10) of McCullagh (1987, p. 64). Continuing in like manner, we have from (5.23)

$$\begin{aligned} a_{3,n} &= f_r f_s f_t \pi^{r,s,t} \\ &\quad + \frac{1}{2} f_r f_s f_{tu} \left\{ \pi^{r,s} \pi^{t,u}[3] - \pi^{r,s} \pi^{t,u} \right\} \\ &\quad + \frac{1}{2} f_r f_{st} f_u \left\{ \pi^{r,s} \pi^{t,u}[3] - \pi^{r,u} \pi^{s,t} \right\} \\ &\quad + \frac{1}{2} f_{rs} f_t f_u \left\{ \pi^{r,s} \pi^{t,u}[3] - \pi^{r,s} \pi^{t,u} \right\} \end{aligned} \tag{5.46}$$

$$\begin{aligned} &= f_r f_s f_t \pi^{r,s,t} \\ &\quad + \frac{1}{2} f_r f_s f_{tu} \pi^{r,t} \pi^{s,u} + \frac{1}{2} f_r f_s f_{tu} \pi^{r,u} \pi^{s,t} \\ &\quad + \frac{1}{2} f_r f_{st} f_u \pi^{r,s} \pi^{t,u} + \frac{1}{2} f_r f_{st} f_u \pi^{r,t} \pi^{s,u} \\ &\quad + \frac{1}{2} f_{rs} f_t f_u \pi^{r,t} \pi^{s,u} + \frac{1}{2} f_{rs} f_t f_u \pi^{r,u} \pi^{s,t} . \end{aligned} \tag{5.47}$$

Since $a_{3,n}$ will feature prominently in the formula for the end-points of our confidence interval, see Chapter six, we have accommodated its calculation in our computer programme for determining these end-points. For that purpose we will use (5.47) rather than (5.46). Also, explicit detailed expansion for (5.47) is given in Section 5.6 for the special case considered in simulations.

From (5.33) and (5.37) we have that

$$\lambda_3 = n^2 \kappa_{2,\infty}^{-\frac{3}{2}} a_{3,n}. \tag{5.48}$$

Finally, from (5.34) and (5.37)

$$\lambda_4 = n^3 \kappa_{2,\infty}^{-2} a_{4,n} \tag{5.49}$$

where, from (5.25) and Table 4.1 we have:

$$\begin{aligned} a_{4,n} &= f_r f_s f_t f_u \pi^{r,s,t,u} \\ &\quad + \frac{1}{2} f_r f_s f_t f_{uv} \left\{ \pi^{r,s} \pi^{t,u,v}[10] - \left(\pi^{u,v} \pi^{r,s,t} + \pi^{r,s} \pi^{t,u,v} + \pi^{r,t} \pi^{s,u,v} + \pi^{s,t} \pi^{r,u,v} \right) \right\} \\ &\quad + \frac{1}{2} f_r f_s f_{tu} f_v \left\{ \pi^{r,s} \pi^{t,u,v}[10] - \left(\pi^{t,u} \pi^{r,s,v} + \pi^{r,s} \pi^{t,u,v} + \pi^{s,v} \pi^{r,t,u} + \pi^{r,v} \pi^{s,t,u} \right) \right\} \\ &\quad + \frac{1}{2} f_r f_{st} f_u f_v \left\{ \pi^{r,s} \pi^{t,u,v}[10] - \left(\pi^{s,t} \pi^{r,u,v} + \pi^{u,v} \pi^{r,s,t} + \pi^{r,u} \pi^{s,t,v} + \pi^{r,v} \pi^{s,t,u} \right) \right\} \\ &\quad + \frac{1}{2} f_{rs} f_t f_u f_v \left\{ \pi^{r,s} \pi^{t,u,v}[10] - \left(\pi^{r,s} \pi^{t,u,v} + \pi^{u,v} \pi^{r,s,t} + \pi^{t,v} \pi^{r,s,u} + \pi^{t,u} \pi^{r,s,v} \right) \right\} \end{aligned} \tag{5.50}$$

5.5 EFFECT OF BIAS-CORRECTION AND POLYNOMIAL TRANSFORMATION

5.5.1 BIAS-CORRECTION

We will see in the next chapter that there are advantages of working with the statistic

$$W_n' = W_n - n^{-\frac{1}{2}}\lambda_1 \tag{5.51}$$

rather than W_n where, as can be seen from (5.31), W_n' makes a correction for the main effect of bias in $\hat{\theta}(n)$. Let $\kappa_{v,n}'$ denote the v-th cumulant of (5.51), then

$$\kappa_{v,n}' = \begin{cases} 0 + 0\left(n^{-\frac{3}{2}}\right) & \text{for } v = 1 \qquad (5.52a) \\ \kappa_{v,n} & \text{for } v = 2,\ 3, \qquad (5.52b) \end{cases}$$

5.5.2 POLYNOMIAL TRANSFORMATION OF W_n'

The correction for skewness which will be employed in our confidence intervals shall be based on the transformation technique of Hall (1992a, 1992b). Central to this transformation are three qualities: firstly it should be strictly monotonically increasing, in our case, in the statistic W_n'; secondly the transformation should be easily invertible, and thirdly it should be skewness-reducing in the sense that the first term of the Edgeworth expansion for the distribution of the transformed statistic, which is entirely due to skewness, is eliminated. In handling the skewness-reducing property, Hall appeals to the delta method for Edgeworth expansions, see next chapter for details of this method. We will propose another approach which appears to be more instructive. It is based on the direct computation of the cumulants of the transformed statistic. In preparation thereof, let us do the computations of those cumulants in this section.

Suppose that the transformation of W_n' which satisfies the three qualities is represented as follows:

$$g\left(W_n'\right) = c_0 + c_1 W_n' + c_2 W_n' W_n' + c_3 W_n' W_n' W_n' + \ldots \tag{5.53}$$

Function (5.53) may take the form either directly or via Taylor expansion. We have in mind the kind of special functions considered by Hall. These are, the third-order polynomial and

$$g\left(W_n'\right) = d_0 + \frac{1}{d_1}\left\{e^{d_1 W_n'} - 1\right\} \tag{5.54}$$

with

$$c_0 = \frac{1}{6}n^{-\frac{1}{2}}b \qquad d_0 = n^{-\frac{1}{2}}c_0 \tag{5.55}$$

$$c_1 = 1 \qquad d_1 = 2c_2 \tag{5.56}$$

$$c_2 = -c_0 \qquad c_3 = \tfrac{1}{3}c_2^2 \tag{5.57}$$

and for $\upsilon \geq 4$

$$c_\upsilon = \begin{cases} 0 & \text{for third-degree polynomial} \\ \frac{1}{\upsilon!}(d_1)^{\upsilon-1} & \text{for transformation (5.54)} \end{cases} \tag{5.58}$$

and where b is a skewness-related parameter, as will be seen in the next chapter.

Now, let $\kappa^*_{\upsilon,n}$ denote the υ-th cumulant of (5.53). By elementary operations we have

$$\kappa^*_{1,n} = c_0 + c_2\kappa'_{2,n} + c_3\kappa'_{3,n} + 0\left(n^{-\frac{3}{2}}\right) \tag{5.59}$$

$$= 0 + 0\left(n^{-\frac{3}{2}}\right) \tag{5.60}$$

with (5.59) giving (5.60) upon applying the relevant results from (5.52a) downward. To obtain $\kappa^*_{2,n}$ we re-use the technique that gave us (5.20). For this purpose, assume once again that our index set (3.5) is Type-II-partitioned. Then in keeping with MINO we may define

$$\kappa_n^{s_1,\ldots,s_\alpha} = \text{generalised cumulant of } \underbrace{W'_n \ldots W'_n}_{s_1\text{-times}}, \ldots, \underbrace{W'_n \ldots W'_n}_{s_\alpha\text{-times}} \tag{5.61}$$

where s_i denotes the size of partition R_i. Thus for example:

$$\kappa_n^{1,2} = \operatorname{cov}(W'_n,\ W'_nW'_n)$$

$$\kappa_n^{2,1} = \operatorname{cov}(W'_nW'_n,\ W'_n)$$

$$\kappa^{1,1,2} = \operatorname{cum}(W'_n,\ W'_n,\ W'_nW'_n)$$

$$\kappa^{3,1,2} = \operatorname{cum}(W'_nW'_nW'_n,\ W'_n,\ W'_nW'_n)$$

and so on. Only the first two examples are relevant to $\kappa^*_{2,n}$, the other two apply to $\kappa^*_{3,n}$ We retain the original meaning of $\kappa'_{\upsilon,n}$ since that makes it identically equal to the "joint" cumulant of the υ-vector

$$W'_n(1, \ldots, 1)\ . \tag{5.62}$$

With these definitions, we have the following cumulant expansions for $\kappa^*_{2,n}$, $\kappa^*_{3,n}$, and $\kappa^*_{4,n}$ in which the terms up to order n^{-1} are displayed explicitly in the final expression.

We start by re-using the technique applied to (5.20) as follows:

$$\begin{aligned}\kappa^{*}_{2,n} &= c_1c_1\kappa'_{2,n} + c_2c_2\kappa^{2,2}_{n} + c_3c_3\kappa^{3,3}_{n} + \ldots \\ &\quad + c_1c_2\kappa^{1,2}_{n}[2]_{*} + c_1c_3\kappa^{1,3}_{n}[2]_{*}\ldots \\ &\quad + c_2c_3\kappa^{2,3}_{n}[2]_{*} + \ldots \\ &\quad + \\ &\quad \vdots \end{aligned} \tag{5.63}$$

Then we use the basic structure of formula (4.10) to expand the relevant generalised cumulants in (5.63). By relevant cumulants we mean those whose corresponding coefficients c_ic_j are of order up to n^{-1}. Thus:

$$\begin{aligned}\kappa^{1,2}_{n} &= E\left[(W'n)^3\right] - E\left[W'_n\right]E\left[(W'_n)^2\right] \\ &= \kappa_{3,n} + 0\left(n^{-\frac{3}{2}}\right) = \kappa^{2,1}_{n}\end{aligned} \tag{5.64}$$

$$\begin{aligned}\kappa^{1,3}_{n} &= E\left[(W'_n)^4\right] - E\left[W'_n\right]E\left[(W'_n)^3\right] \\ &= \kappa_{4,n} + 3\kappa^{2}_{2,n} + 0\left(n^{-\frac{3}{2}}\right) = \kappa^{3,1}_{n}\end{aligned} \tag{5.65}$$

$$\begin{aligned}\kappa^{2,2}_{n} &= E\left[(W'_n)^4\right] - \left\{E\left[(W'_n)^2\right]\right\}^2 \\ &= \kappa_{4,n} + 2\kappa^{2}_{2,n} + 0\left(n^{-\frac{3}{2}}\right).\end{aligned} \tag{5.66}$$

Then putting these back into (5.63) we obtain

$$\kappa^{*}_{2,n} = 1 + n^{-1}\left(\lambda_2 - \frac{1}{3}b\lambda_3 + \frac{1}{6}b^2\right) + 0\left(n^{-2}\right). \tag{5.67}$$

To obtain the expansion for $\kappa^{*}_{3,n}$ we perform likewise starting with the re-application of the procedure that gave us (5.21). We get:

$$\begin{aligned}\kappa^{*}_{3,n} &= c_1c_1c_1\kappa'_{3,n} + c_2c_2c_2\kappa^{2,2,2}_{n} + \ldots \\ &\quad + c_1c_1c_2\kappa^{1,1,2}_{n}[3]_{*} + c_1c_1c_3\kappa^{1.1.3}_{n}[3]_{*} + \ldots \\ &\quad + c_1c_2c_2\kappa^{1,2,2}_{n}[3]_{*} + c_1c_2c_3\kappa^{1,2,3}_{n}[6]_{*}\ldots \\ &\quad + \\ &\quad \vdots \end{aligned} \tag{5.68}$$

Then we use the basic structure of (4.11) to expand those generalised cumulants in (5.68) whose coefficients are of order up to n^{-1}. Included here are the following:

$$\begin{aligned}\kappa^{1,1,2} &= E\left[(W'n)^4\right] \\ &\quad -\left\{2E\left[W_n'\right]E\left[(W_n')\right]+\left\{E\left[(W_n')^2\right]\right\}^2\right\} \\ &\quad +2\left\{E\left[W_n'\right]\right\}^2 E\left[(W_n')^2\right] \qquad (5.69)\\ &= \kappa_{4,n}+2\kappa_{2,n}^2+0\left(n^{-\frac{3}{2}}\right). \qquad (5.70)\end{aligned}$$

The term $2\kappa_{2,n}^2$ in the above result plays a particularly interesting role in the method which we will propose as an alternative to the explanation of Hall (1992a, 1992b) for the skewness-reducing property of the transformations which the author has introduced.

Next we expand $\kappa^{1,1,3}$ similarly. We get:

$$\begin{aligned}\kappa^{1,1,3} &= E\left[(W_n')^5\right] \\ &\quad -\left\{2E\left[W_n'\right]E\left[(W_n')^4\right]+E\left[(W_n')^3\right]E\left[(W_n')^2\right]\right\} \\ &\quad +2\left\{E\left[W_n'\right]\right\}^2 E\left[(W_n')^3\right] \qquad (5.71)\\ &= 9\kappa_{3,n}\kappa_{2,n}+0\left(n^{-\frac{3}{2}}\right) \qquad (5.72)\\ &= 9\kappa_{3,n}+0\left(n^{-\frac{3}{2}}\right). \qquad (5.73)\end{aligned}$$

The last relevant cumulant in (5.68) is:

$$\begin{aligned}\kappa^{1,2,2} &= E\left[(W_n')^5\right] \\ &\quad -\left\{E\left[W_n'\right]E\left[(W_n')^4\right]+2E\left[(W_n')^2\right]E\left[(W_n')^3\right]\right\} \\ &\quad +2E\left[W_n'\right]\left\{W\left[(W_n')^2\right]\right\}^2 \qquad (5.74)\\ &= 8\kappa_{3,n}\kappa_{2,n}+0\left(n^{-\frac{3}{2}}\right) \qquad (5.75)\\ &= 8\kappa_{3,n}+0\left(n^{-\frac{3}{2}}\right). \qquad (5.76)\end{aligned}$$

Then upon putting these expansions into (5.68) we get:

$$\kappa_{3,n}^{*}=n^{-\frac{1}{2}}\left(\lambda_3-b\right)+0\left(n^{-\frac{3}{2}}\right). \qquad (5.77)$$

This result is particularly interesting. But before we indicate why it is so, let us complete the present task by computing $\kappa^*_{4,n}$. By repeating the same reasoning on the expansion

$$\begin{aligned}\kappa^*_{4,n} &= c_1c_1c_1c_1\kappa'_{4,n} + c_2c_2c_2c_2\kappa_n^{2,2,2,2} + \ldots \\ &\quad + c_1c_1c_1c_2\kappa_n^{1,1,1,2}[4]_* + c_1c_1c_1c_3\kappa_n^{1,1,1,3}[4]_* + \ldots \\ &\quad + \\ &\quad \vdots \end{aligned} \tag{5.78}$$

using the basic structure of (4.12), we get

$$\kappa^*_{4,n} = 0 + n^{-1}\left(\lambda_4 - 4b\lambda_3 + \frac{5}{9}b^2\right) + 0\left(n^{-2}\right). \tag{5.79}$$

We may now address the significance of result (5.77) in the overall context of the intended application of the results obtained in this section. We will see in the next chapter that our promised alternative to Hall's delta method will rest on two factors. The first one is that it will turn out that

$$b = \lambda_3 \tag{5.80}$$

so that the third-order cumulant of $g(W'_n)$, like the mean, has the same form as (5.52a).

The second factor is that, subject to (5.80), the basic form of the expansions for the cumulants of $g(W'_n)$ has just been shown to mimic the cumulants of W'_n. For the reader who is tempted to recall Theorem 1 of James (1958) in this connection, we point out that it does not strictly apply to (5.53) as the coefficients there depend on n. From the similarity of cumulant structure, once the Edgeworth series for the distribution of W'_n has been constructed, the counterpart for the distribution of $g(W'_n)$ becomes obvious.

5.6 THE λ_3-CALCULATION

In the next chapter, it will be seen that the end-points of our confidence interval depend, among others, upon the estimates of the parameter functions λ_1 and λ_3. From (5.38) it is seen that because $a_{1,n}$ is a simple function, so is λ_1. On the other hand, from (5.47) we see that $a_{3,n}$ is not quite a simple function so that λ_3, which is given by (5.48), is likewise not so simple. To illustrate the latter fact, we provide below an explicit expansion of (5.47) in the special case of the ratio parameter representing the abscissa of the turning point for the parabolic model in Example 1.3. The second identity in each summation represents the simplification which derives from two sources: the fact that f belongs to the class C^2 so that there is equality of cross derivatives and, secondly, the fact that ordinary cumulants are symmetric in their arguments (indices). Thus, for this case we have the following:

$$f_r f_s f_{tu} \pi^{r,t} \pi^{s,u} = \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + f_2 f_2 f_{23} \pi^{2,3} \pi^{2,2} + f_2 f_2 f_{32} \pi^{2,2} \pi^{2,3} + f_2 f_2 f_{33} \pi^{2,3} \pi^{2,3} \\ + f_2 f_3 f_{22} \pi^{2,2} \pi^{3,2} + f_2 f_3 f_{23} \pi^{2,3} \pi^{3,2} + f_2 f_3 f_{32} \pi^{2,2} \pi^{3,3} + f_2 f_3 f_{33} \pi^{2,3} \pi^{3,3} \\ + f_3 f_2 f_{22} \pi^{3,2} \pi^{2,2} + f_3 f_2 f_{23} \pi^{3,3} \pi^{2,2} + f_3 f_2 f_{32} \pi^{3,2} \pi^{2,3} + f_3 f_2 f_{33} \pi^{3,3} \pi^{2,3} \\ + f_3 f_3 f_{22} \pi^{3,2} \pi^{3,2} + f_3 f_3 f_{23} \pi^{3,3} \pi^{3,2} + f_3 f_3 f_{32} \pi^{3,2} \pi^{3,3} + f_3 f_3 f_{33} \pi^{3,3} \pi^{3,3} \end{bmatrix}$$

$$= \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + 2 f_2 f_3 f_{22} \pi^{2,2} \pi^{2,3} + f_3 f_3 f_{22} \pi^{2,3} \pi^{2,3} \\ + f_3 f_3 f_{33} \pi^{3,3} \pi^{3,3} + 2 f_2 f_3 f_{33} \pi^{3,3} \pi^{2,3} + f_2 f_2 f_{33} \pi^{2,3} \pi^{2,3} \\ + 2 f_2 f_2 f_{23} \pi^{2,2} \pi^{2,3} + 2 f_3 f_3 f_{23} \pi^{3,3} \pi^{2,3} \\ + 2 f_2 f_3 f_{23} \pi^{2,2} \pi^{3,3} + 2 f_2 f_3 f_{23} \pi^{2,3} \pi^{2,3} \end{bmatrix} \tag{5.81}$$

$$f_r f_s f_{tu} \pi^{r,u} \pi^{s,t} = \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + f_2 f_2 f_{23} \pi^{2,3} \pi^{2,2} + f_2 f_2 f_{32} \pi^{2,2} \pi^{2,3} + f_2 f_2 f_{33} \pi^{2,3} \pi^{2,3} \\ + f_2 f_3 f_{22} \pi^{2,2} \pi^{3,2} + f_2 f_3 f_{23} \pi^{2,3} \pi^{3,2} + f_2 f_3 f_{32} \pi^{2,2} \pi^{3,3} + f_2 f_3 f_{33} \pi^{2,3} \pi^{3,3} \\ + f_3 f_2 f_{22} \pi^{3,2} \pi^{2,2} + f_3 f_2 f_{23} \pi^{3,3} \pi^{2,2} + f_3 f_2 f_{32} \pi^{3,2} \pi^{2,3} + f_3 f_2 f_{33} \pi^{3,3} \pi^{2,3} \\ + f_3 f_3 f_{22} \pi^{3,2} \pi^{3,2} + f_3 f_3 f_{23} \pi^{3,3} \pi^{3,2} + f_3 f_3 f_{32} \pi^{3,2} \pi^{3,3} + f_3 f_3 f_{33} \pi^{3,3} \pi^{3,3} \end{bmatrix}$$

$$= \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + 2 f_2 f_3 f_{22} \pi^{2,2} \pi^{2,3} + f_3 f_3 f_{22} \pi^{2,3} \pi^{2,3} \\ + f_3 f_3 f_{33} \pi^{3,3} \pi^{3,3} + 2 f_2 f_3 f_{33} \pi^{3,3} \pi^{2,3} + f_2 f_2 f_{33} \pi^{2,3} \pi^{2,3} \\ + 2 f_2 f_2 f_{23} \pi^{2,2} \pi^{2,3} + 2 f_3 f_3 f_{23} \pi^{3,3} \pi^{2,3} \\ + 2 f_2 f_3 f_{23} \pi^{2,2} \pi^{3,3} + 2 f_2 f_3 f_{23} \pi^{2,3} \pi^{2,3} \end{bmatrix} \tag{5.82}$$

$$f_r f_{st} f_u \pi^{r,s} \pi^{t,u} = \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + f_2 f_{22} f_3 \pi^{2,2} \pi^{2,3} + f_2 f_{23} f_2 \pi^{2,2} \pi^{3,2} + f_2 f_{23} f_3 \pi^{2,2} \pi^{3,3} \\ + f_2 f_{32} f_2 \pi^{2,3} \pi^{2,2} + f_2 f_{32} f_3 \pi^{2,3} \pi^{2,3} + f_2 f_{33} f_2 \pi^{2,3} \pi^{3,2} + f_2 f_{33} f_3 \pi^{2,3} \pi^{3,3} \\ + f_3 f_{22} f_2 \pi^{3,2} \pi^{2,2} + f_3 f_{22} f_3 \pi^{3,2} \pi^{2,3} + f_3 f_{23} f_2 \pi^{3,2} \pi^{3,2} + f_3 f_{23} f_3 \pi^{3,2} \pi^{3,3} \\ + f_3 f_{32} f_2 \pi^{3,3} \pi^{2,2} + f_3 f_{32} f_3 \pi^{3,3} \pi^{2,3} + f_3 f_{33} f_2 \pi^{3,3} \pi^{3,2} + f_3 f_{33} f_3 \pi^{3,3} \pi^{3,3} \end{bmatrix}$$

$$= \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + 2 f_2 f_3 f_{22} \pi^{2,2} \pi^{2,3} + f_3 f_3 f_{22} \pi^{2,3} \pi^{2,3} \\ + f_3 f_3 f_{33} \pi^{3,3} \pi^{3,3} + 2 f_2 f_3 f_{33} \pi^{3,3} \pi^{2,3} + f_2 f_2 f_{33} \pi^{2,3} \pi^{2,3} \\ + 2 f_2 f_2 f_{23} \pi^{2,2} \pi^{2,3} + 2 f_3 f_3 f_{23} \pi^{3,3} \pi^{2,3} \\ + 2 f_2 f_3 f_{23} \pi^{2,2} \pi^{3,3} + 2 f_2 f_3 f_{23} \pi^{2,3} \pi^{2,3} \end{bmatrix} \tag{5.83}$$

$$f_r f_{st} f_u \pi^{r,t} \pi^{s,u} = \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + f_2 f_{22} f_3 \pi^{2,2} \pi^{2,3} + f_2 f_{23} f_2 \pi^{2,3} \pi^{2,2} + f_2 f_{23} f_3 \pi^{2,3} \pi^{2,3} \\ + f_2 f_{32} f_2 \pi^{2,2} \pi^{3,2} + f_2 f_{32} f_3 \pi^{2,2} \pi^{3,3} + f_2 f_{33} f_2 \pi^{2,3} \pi^{3,2} + f_2 f_{33} f_3 \pi^{2,3} \pi^{3,3} \\ + f_3 f_{22} f_2 \pi^{3,2} \pi^{2,2} + f_3 f_{22} f_3 \pi^{3,2} \pi^{2,3} + f_3 f_{23} f_2 \pi^{3,3} \pi^{2,2} + f_3 f_{23} f_3 \pi^{3,3} \pi^{2,3} \\ + f_3 f_{32} f_2 \pi^{3,2} \pi^{3,2} + f_3 f_{32} f_3 \pi^{3,2} \pi^{3,3} + f_3 f_{33} f_2 \pi^{3,3} \pi^{3,2} + f_3 f_{33} f_3 \pi^{3,3} \pi^{3,3} \end{bmatrix}$$

$$= \begin{bmatrix} f_2 f_2 f_{22} \pi^{2,2} \pi^{2,2} + 2 f_2 f_3 f_{22} \pi^{2,2} \pi^{2,3} + f_3 f_3 f_{22} \pi^{2,3} \pi^{2,3} \\ + f_3 f_3 f_{33} \pi^{3,3} \pi^{3,3} + 2 f_2 f_3 f_{33} \pi^{3,3} \pi^{2,3} + f_2 f_2 f_{33} \pi^{2,3} \pi^{2,3} \\ + 2 f_2 f_2 f_{23} \pi^{2,2} \pi^{2,3} + 2 f_3 f_3 f_{23} \pi^{3,3} \pi^{2,3} \\ + 2 f_2 f_3 f_{23} \pi^{2,2} \pi^{3,3} + 2 f_2 f_3 f_{23} \pi^{2,3} \pi^{2,3} \end{bmatrix} \tag{5.84}$$

$$
\begin{aligned}
f_{rs}f_tf_u\pi^{r,t}\pi^{s,u} &= \begin{bmatrix} f_{22}f_2f_3\pi^{2,2}\pi^{2,3} + f_{22}f_3f_2\pi^{2,3}\pi^{2,2} + f_{22}f_3f_2\pi^{2,3}\pi^{2,2} + f_{22}f_3f_3\pi^{2,3}\pi^{2,3} \\ +f_{23}f_2f_2\pi^{2,2}\pi^{3,2} + f_{23}f_2f_3\pi^{2,2}\pi^{3,3} + f_{23}f_3f_2\pi^{2,3}\pi^{3,2} + f_{23}f_3f_3\pi^{2,3}\pi^{3,3} \\ +f_{32}f_2f_2\pi^{3,2}\pi^{2,2} + f_{32}f_2f_3\pi^{3,2}\pi^{2,3} + f_{32}f_3f_2\pi^{3,3}\pi^{2,2} + f_{32}f_3f_3\pi^{3,3}\pi^{2,3} \\ +f_{33}f_2f_2\pi^{3,2}\pi^{3,2} + f_{33}f_2f_3\pi^{3,2}\pi^{3,3} + f_{33}f_3f_2\pi^{3,3}\pi^{3,2} + f_{33}f_3f_3\pi^{3,3}\pi^{3,3} \end{bmatrix} \\
&= \begin{bmatrix} f_2f_2f_{22}\pi^{2,2}\pi^{2,2} + 2f_2f_3f_{22}\pi^{2,2}\pi^{2,3} + f_3f_3f_{22}\pi^{2,3}\pi^{2,3} \\ +f_3f_3f_{33}\pi^{3,3}\pi^{3,3} + 2f_2f_3f_{33}\pi^{3,3}\pi^{2,3} + f_2f_2f_{33}\pi^{2,3}\pi^{2,3} \\ +2f_2f_2f_{23}\pi^{2,2}\pi^{2,3} + 2f_3f_3f_{23}\pi^{3,3}\pi^{2,3} \\ +2f_2f_3f_{23}\pi^{2,2}\pi^{3,3} + 2f_2f_3f_{23}\pi^{2,3}\pi^{2,3} \end{bmatrix} \qquad (5.85)
\end{aligned}
$$

$$
\begin{aligned}
f_{rs}f_tf_u\pi^{r,u}\pi^{s,t} &= \begin{bmatrix} f_2f_2f_{22}\pi^{2,2}\pi^{2,2} + f_{22}f_2f_3\pi^{2,3}\pi^{2,2} + f_{22}f_3f_2\pi^{2,2}\pi^{2,3} + f_{22}f_3f_3\pi^{2,3}\pi^{2,3} \\ +f_{23}f_2f_2\pi^{2,2}\pi^{3,2} + f_{23}f_2f_3\pi^{2,3}\pi^{3,2} + f_{23}f_3f_2\pi^{2,2}\pi^{3,3} + f_{23}f_3f_3\pi^{2,3}\pi^{3,3} \\ +f_{32}f_2f_2\pi^{3,2}\pi^{2,2} + f_{32}f_2f_3\pi^{3,3}\pi^{2,2} + f_{32}f_2\pi^{3,2}\pi^{2,3} + f_{32}f_3f_3\pi^{3,3}\pi^{2,3} \\ +f_{33}f_2f_2\pi^{3,2}\pi^{3,2} + f_{33}f_2f_3\pi^{3,3}\pi^{3,2} + f_{33}f_3f_2\pi^{3,2}\pi^{3,3} + f_{33}f_3f_3\pi^{3,3}\pi^{3,3} \end{bmatrix} \\
&= \begin{bmatrix} f_2f_2f_{22}\pi^{2,2}\pi^{2,2} + 2f_2f_3f_{22}\pi^{2,2}\pi^{2,3} + f_3f_3f_{22}\pi^{2,2}\pi^{3,3} \\ + f_3f_3f_{33}\pi^{3,3}\pi^{3,3} + 2f_2f_3f_{33}\pi^{3,3}\pi^{2,3} + f_2f_2f_{33}\pi^{2,3}\pi^{2,3} \\ +2f_2f_2f_{23}\pi^{2,2}\pi^{2,3} + 2f_3f_3f_{23}\pi^{3,3}\pi^{2,3} \\ +2f_2f_3f_{23}\pi^{2,2}\pi^{3,3} + 2f_2f_3f_{23}\pi^{2,3}\pi^{2,3} \end{bmatrix} \qquad (5.86)
\end{aligned}
$$

Let

$$
\begin{aligned}
Sum_1 &= f_2f_2f_{22}\pi^{2,2}\pi^{2,2} + f_3f_3f_{33}\pi^{3,3}\pi^{3,3} + 2f_2f_3\pi^{2,3}\{f_{22}\pi^{2,2} + f_{33}\pi^{3,3}\} \\
Sum_2 &= \pi^{2,3}\pi^{2,3}\{f_3f_3f_{22} + f_2f_2f_{33}\} \\
Sum_3 &= 2f_{23}\pi^{2,3}\{f_2f_2\pi^{2,2} + f_3f_3\pi^{3,3}\} \\
Sum_4 &= 2f_2f_3f_{23}\{\pi^{2,2}\pi^{3,3} + \pi^{2,3}\pi^{2,3}\}
\end{aligned} \qquad (5.87)
$$

Then

$$
a_{3,n} = f_rf_sf_t\pi^{r,s,t} + 3\{Sum_1 + Sum_2 + Sum_3 + Sum_4\} \qquad (5.88)
$$

In the case at hand, further simplification is evident from the fact that all the second and higher-order derivatives w.r.t $\hat{\beta}_2(n)$ are zero.

5.7 CHAPTER SUMMARY

This chapter began with a brief mention of the methods that have been proposed over the years regarding the determination of the cumulants of multivariate nonlinear functions of random variables, from the cumulants of the random variables. We went into detail in the case of McCullagh's (1984) method because this method employs a much simplified form of index notation. We then proposed a variant to this method, whose simplicity and transparency derive entirely from the fact that the proposed method is merely a direct application of the basic structure of the results of Theorem 4.1. The additional attraction of this method is that one is able to see immediately the manner in which the terms follow one another in each cumulant expansion. Helped by this, we have suggested a correction to formula (3.10) of McCullagh (1987) and explained the significance of the correction.

To enable us, in the next chapter, to convert a Gram-Charlier series into an Edgeworth series which will serve our ultimate goal, we have proposed certain conditions which, together with result (4.59) of Corollary 4.2.1, *identify* the particular terms that must be explicitly retained for that very purpose. The theory of the cumulants of multivariate nonlinear random variables is thus extended.

Finally, by calculating the cumulants of (5.53) we have set the stage to propose an alternative to an important aspect of Hall's (1992a, 1992b) method. As may be seen in Section 5.5.2, the manner of deriving those cumulants has entirely relied on the *flexibility* of the method which we have proposed in Section 5.3.

CHAPTER SIX

EDGEWORTH-CORRECTED CONFIDENCE INTERVALS

6.1 INTRODUCTION

In this chapter we consider two aspects: first, the construction of an Edgeworth expansion for the d.f. of transformation (5.51) and second, we construct from this expansion, an approximate confidence interval for θ, which contains a $n^{-\frac{1}{2}}$-order correction for the primary effects of skewness in the d.f. of W_n'.

With respect to the first task, we could construct the expansion directly. However, as this would mask seeing the way λ_1 enters into the picture, we choose to construct the Edgeworth expansion for the d.f. of W_n and then from it, we deduce its counterpart for the d.f. of (5.51).

Furthermore, given the purposes for which the two Edgeworth expansions are constructed, we shall make no attempt to discuss their asymptotic convergence.

6.2 THE EXPANSIONS

Both the general form of identity (1.36) as well as the specific functional forms of $p_1(w)$, $p_2(w)$, etc., may be straightforwardly derived under the following three conditions, namely:

C.6.1 W_n is asymptotically $N(0, 1)$ as $n \to \infty$;

C.6.2 All the cumulants of W_n are finite;

C.6.3 G_n is uniquely determined by the cumulants in C.6.2.

Following Moran (1968, pp. 259–262), a sufficient condition for C.6.3 is that the characteristic function of W_n be convergent in the interval, say

$$|\xi| < \eta\,, \quad \eta > 0\,. \tag{6.1}$$

That is, in this interval both conditions below should hold:

$$\begin{aligned} &\lim_{a\to\infty} \int_0^a e^{i\xi w} dG_n(w) < \infty \\ &\qquad\qquad\qquad\qquad i^2 = -1\,. \\ &\lim_{a\to\infty} \int_{-a}^0 e^{i\xi w} dG_n(w) < \infty \end{aligned} \tag{6.2}$$

Let $\Psi_n(\xi)$ and $M(\xi)$ respectively denote the characteristic functions of G_n and Φ respectively. From their very definitions we have

$$\ell n \Psi_n(\xi) = \kappa_{1,n}(i\xi) + \kappa_{2,n}\frac{(i\xi)^2}{2!} + \kappa_{3,n}\frac{(i\xi)^3}{3!} + \ldots \tag{6.3}$$

$$\ell n M(\xi) = 0 \cdot i\xi + 1 \cdot \frac{(i\xi)^2}{2!} + 0 \cdot \frac{(i\xi)^3}{3!} + \ldots \tag{6.4}$$

and recall that the normal distribution is uniquely determined by its cumulants. Subtracting (6.4) from (6.3) and then exponentiating the result, followed by converting the exponential function into a MacLaurin series about $\xi = 0$ we get

$$\Psi_n(\xi) = M(\xi)\left[1 + \sum_{r=1}^{\infty} \rho_{r,n}\frac{(i\xi)^r}{r!}\right] \tag{6.5}$$

where the first six coefficients $\rho_{r,n}$ are given below:

$$\rho_{1,n} = \kappa_{1,n} \tag{6.6}$$

$$\rho_{2,n} = (\kappa_{2,n} - 1) + \kappa_{1,n}^2 \tag{6.7}$$

$$\rho_{3,n} = \kappa_{3,n} + 3(\kappa_{2,n} - 1)\kappa_{1,n} + \kappa_{1,n}^3 \tag{6.8}$$

$$\rho_{4,n} = \kappa_{4,n} + 4\kappa_{3,n}\kappa_{1,n} + 3(\kappa_{2,n} - 1)^2 + 6(\kappa_{2,n} - 1)\kappa_{1,n}^2 + \kappa_{1,n}^4 \tag{6.9}$$

$$\begin{aligned} \rho_{5,n} = {} & \kappa_{5,n} + 5\kappa_{4,n}\kappa_{1,n} + 10\kappa_{3,n}(\kappa_{2,n} - 1) + 10\kappa_{3,n}\kappa_{1,n}^2 \\ & + 15(\kappa_{2,n} - 1)^2\kappa_{1,n} + 10(\kappa_{2,n} - 1)\kappa_{1,n}^3 + \kappa_{1,n}^5 \end{aligned} \tag{6.10}$$

$$\begin{aligned} \rho_{6,n} = {} & \kappa_{6,n} + 6\kappa_{5,n}\kappa_{1,n} + 15\kappa_{4,n}(\kappa_{2,n} - 1) + 15\kappa_{4,n}\kappa_{1,n}^2 \\ & + 10\kappa_{3,n}^2 + 60\kappa_{3,n}(\kappa_{2,n} - 1)\kappa_{1,n} + 20\kappa_{3,n}\kappa_{1,n}^3 \\ & + 15(\kappa_{2,n} - 1)^3 + 45(\kappa_{2,n} - 1)^2\kappa_{1,n}^2 + 15(\kappa_{2,n} - 1)\kappa_{1,n}^4 \\ & + \kappa_{1,n}^6\,. \end{aligned} \tag{6.11}$$

Observe the way $\kappa_{2,n}$ appears in the formulas. Shortly, we will see that, under our conditions (5.31) to (5.35), we require full knowledge of $\rho_{r,n}$ up to $r = 6$ in order to be able to fulfil the

primary objective of those conditions, namely to allow us to capture the $n^{-\frac{1}{2}}$ and n^{-1} order terms in $\kappa_{1,n}, \ldots, \kappa_{4,n}$ in the first two terms of the Edgeworth expansion.

Now, replacing $\Psi_n(\xi)$ and $M(\xi)$ by their equivalents, identity (6.5) becomes:

$$\int_{-\infty}^{\infty} e^{i\xi w} dG_n(w) = e^{-\frac{1}{2}\xi^2} + \sum_{r=1}^{\infty} \frac{1}{r!} \rho_{r,n} (i\xi)^r e^{-\frac{1}{2}\xi^2} \tag{6.12}$$

$$= \int_{-\infty}^{\infty} e^{i\xi w} d\Phi(w) + \sum_{r=1}^{\infty} \frac{1}{r!} \rho_{r,n} \int_{-\infty}^{\infty} e^{i\xi w} (-1)^r d\Phi^{(r)}(w) \tag{6.13}$$

with (6.13) coming about in view of the well-known property

$$(i\xi)^r e^{-\frac{1}{2}\xi^2} = \int_{-\infty}^{\infty} e^{i\xi w} (-1)^r d\Phi^{(r)}(w) \tag{6.14}$$

$\Phi^{(r)}$ being the r-th derivative of Φ, see e.g. Cramér (1946, p. 100).

Then, applying the Inversion Theorem to (6.13) term by term, we obtain

$$G_n(w) = \Phi(w) + \sum_{r=1}^{\infty} (-1)^r \frac{1}{r!} \rho_{r,n} \Phi^{(r)}(w) \tag{6.15}$$

$$= \Phi(w) - \phi(w) \sum_{r=1}^{\infty} \frac{1}{r!} \rho_{r,n} H_{r-1}(w) \tag{6.16}$$

where H_r are the r-th degree Hermite polynomials, the first six of which are listed below.

$$\begin{aligned} H_0(w) &= 1 & H_3(w) &= w^3 - 3w \\ H_1(w) &= w & H_4(w) &= w^4 - 6w^2 + 3 \\ H_2(w) &= w^2 - 1 & H_5(w) &= w^5 - 10w^3 + 15w \end{aligned} \tag{6.17}$$

Expansion (6.16) is known as the *Gram-Charlier series* (of Type A) for G_n. Under conditions (5.31) to (5.35) the order of successive terms of this series does not decrease steadily with increasing n. Hence (6.16) is not asymptotic in the sense of (1.37). To make it so, we collect together the terms that are of equal order in the n. The resulting series has the form (1.36) with

$$p_1(w) = -\left\{ \lambda_1 + \frac{1}{6} \lambda_3 H_2(w) \right\} \tag{6.18}$$

$$p_2(w) = -\left[\frac{1}{2} \left\{ \lambda_2 + \lambda_1^2 \right\} H_1(w) - \frac{1}{24} \left\{ \lambda_4 + 4\lambda_3 \lambda_1 \right\} H_3(w) + \frac{1}{72} \lambda_3^2 H_5(w) \right]. \tag{6.19}$$

It is at this point that the reason for our going as far as $r = 6$ in specifying the formulas for $\rho_{r,n}$ becomes evident. For we now can see that as one checks out, using (5.31) to (5.35), the terms in each $\rho_{r,n}$ which should be retained in the first two terms of the Edgeworth expansion, one finds that the last one is the term $10\kappa_{3,n}^2$ in the formula for $\rho_{6,n}$, as one moves from $r = 1$ upwards.

Result (6.18) suggests that we may convert to the traditional form of the first term of an Edgeworth series by correcting W_n for the main effect of bias as we have done in (5.51). Let G_n' denote the

d.f. of W_n'. Then, clearly

$$G_n'(w) = \Phi(w) + n^{-\frac{1}{2}} p_1^*(w)\phi(w) + n^{-1} p_2(w)\phi(w) + \ldots \tag{6.20}$$

where

$$p_1^*(w) = -\frac{1}{6}\lambda_3 \left(w^2 - 1\right). \tag{6.21}$$

6.3 TRUNCATION

In deciding upon which terms of (6.20) we should retain, we adopt as our main guide the question of which aspect of $G_n'(w)$ is violated in the tails when n is small but nevertheless the standard normal approximation is used to construct confidence intervals. In a simulation study involving the estimation problem cited in Example 1.3, Weber & Welsh (1983) have found that when n is small the distribution of the standardised $\hat{\theta}(n)$ can be very skewed under the three error distributions considered: the normal, the central student-t with three degrees of freedom, and the exponential distribution. This finding is summarised in the authors' Table 4.4 on p. 435.

On the basis that the first term of (6.20) contains the dominant term for the primary effects of skewness and has a simple, albeit quadratic form, we choose the approximation

$$\begin{aligned} G_n'(w) &\approx \Phi(w) + \left\{ -\frac{1}{6} n^{-\frac{1}{2}} \lambda_3 w^2 + \frac{1}{6} n^{-\frac{1}{2}} \lambda_3 \right\} \phi(w) \\ &= G_{n,0}'(w), \, say \end{aligned} \tag{6.22}$$

6.4 BASIC FAULTS

The above approximation, however, violates two essential requirements of a d.f. and this may happen in the tail area of $G_{n,0}'$. The first possible violation is the $0 - 1$ range for every probability set function. To see how this may occur, let us recall the following well-known expansion:

$$1 - \Phi(w) = \frac{\phi(w)}{w} \left\{ 1 - \frac{1}{w^2} + \frac{3}{w^4} - \frac{3.5}{w^6} + \ldots \right\} \tag{6.23}$$

which derives from repeated integration by parts. In the tails where w tends to $\pm\infty$ we can ignore the terms of order $0(w^{-2})$ and replace $G_{n,0}'$ with, say

$$\hat{G}_{n,0}'(w) = 1 - \phi(w) \left\{ \frac{1}{w} + \left(-\frac{1}{6} n^{-\frac{1}{2}} \lambda_3 \right) w^2 + \frac{1}{6} n^{-\frac{1}{2}} \lambda_3 \right\} \tag{6.24}$$

from which the possibility of $0-1$ range violation is evident. However, since we do not intend to use the approximation to generate probability values, this deficiency is not crucial in our particular problem.

Of particular concern to us is the fact that (6.22) is not *strictly* monotonic in w, *over the entire range* $-\infty$ *to* ∞. Now, from the viewpoint of representing a d.f., in addition to satisfying the $0-1$ range, the values of $G'_{n,0}(w)$ need to be at least non-decreasing as w increases. However, for purposes of confidence interval construction the second requirement has to be sharpened to include strict monotonicity. With the latter we can then associate every *distinct* percentile value w with a *unique* confidence level.

6.5 REMEDIAL MEASURES

6.5.1 GENERAL

An obvious approach to handling the above shortcomings is simply to identify the feasible region, in this case, in the (w, λ_3)-plane. For example, Berndt (1957) does just this by graphing. From the discussion of Kendall & Stuart (1977, Section 6.3), however, it would appear this approach is useful only in the cases of moderate skewness. On the other hand, an examination of Table 4.4 of Weber & Welsh (1983) demonstrates that skewness can be severe when n is small.

A more appealing route to take is the elementary transformation proposed by Hall (1992a, 1992b). In its basic form, the technique chooses a transformation, say $g(\cdot)$, which has the following three properties, stated below using our past convention and present notation:

P.6.1 g is strictly monotonic

P.6.2 It is also easily invertible

P.6.3 $\Pr\{g(W'_n) \leq w\} = \Phi(w) + 0(n^{-1})$

Since the term eliminated, of order $n^{\frac{1}{2}}$, is entirely due to λ_3 we shall refer to P.6.3 as the *skewness-reducing* property.

We need an additional property which the author does not stress, namely

P.6.4 The monotonicity of g is of increasing kind.

With P.6.4, we have

$$\Pr\{g(W_n') \leq w\} = G_n'\left(g^{-1}(w)\right) . \tag{6.25}$$

Let us now consider separately the choice of g and the skewness-reducing property.

6.5.2 CHOICE OF TRANSFORMATION

Two approaches have been suggested by Hall regarding the transformation function having the above properties. The first one is to consider a third degree polynomial. This choice is motivated by two factors: firstly, the fact that polynomials of even degree are never strictly monotonic while those of odd degrees $5, 7, 9, \ldots$ are generally not solvable explicitly and, secondly, it happens that the coefficient of $n^{-\frac{1}{2}}\phi(\cdot)$ in Edgeworth expansion is typically a second-degree polynomial which can be adjusted accordingly to give the desired third-degree polynomial.

In view of (6.22), let the adjusted cubic transformation that we seek be

$$g_3(W_n') = c_0' + W_n' + c_2'\{W_n'\}^2 + c_3'\{W_n'\}^3 \tag{6.26}$$

where

$$c_0' = \frac{1}{6}n^{-\frac{1}{2}}\lambda_3 , \qquad c_2' = -c_0 \tag{6.27}$$

while c_3' is yet to be determined. In order for (6.26) to satisfy P.6.4

$$\frac{dg_3}{dW_n'} > 0 \tag{6.28}$$

must hold *at every* W_n', which is satisfied when

$$c_3' = \frac{1}{3}\{c_2'\}^2 \tag{6.29}$$

in which case

$$\frac{dg_3}{dW_n'} = \{1 + c_2'W_n'\}^2 . \tag{6.30}$$

Then on integrating (6.30) and eliminating the constant of integration by setting $W_n = 0$ in both the integral and (6.26) we get

$$g_3(W_n') = \frac{1}{3c_2'}\left[\{1 + c_2'W_n'\}^3 + 3c_0'c_2' - 1\right] \tag{6.31}$$

with the unique solution of $g(W_n') = w$ being

$$g_3^{-1}(w) = \frac{1}{c_2'}\left[\{1 + 3c_2'(w - c_0)\}^{\frac{1}{3}} - 1\right] . \tag{6.32}$$

Rather than adjusting the quadratic function in the $n^{-\frac{1}{2}}$ term of an Edgeworth expansion, Hall suggests the use of such functions as

$$g(W_n') = d_0' + \frac{1}{d_1'}\left\{e^{d_1'W_n'} - 1\right\} \tag{6.33}$$

with

$$d'_0 = n^{-\frac{1}{2}} c'_0 \,, \qquad d'_1 = 2c'_2$$

whose inverse is given by

$$g^{-1}(w) = \frac{1}{d'_1} \ell n \left\{ d'_1 \left(w - d_0 \right) + 1 \right\} \tag{6.34}$$

provided the argument of the logarithm is positive.

6.5.3 SKEWNESS-REDUCING PROPERTY

In this section we describe an approach that is suggested by Hall to demonstrate skewness-reducing property P.6.3, as well as propose another method.

(a) HALL's APPROACH

Central to the approach used by Hall is the so-called *Delta Method for Edgeworth expansions,* hereafter conveniently abbreviated by DMEE. This method comes in at least three versions which we describe below in the context of our problem.

(i) FIRST VERSION

In accordance with Section 3 of Hill & Davis (1968), the fact that we have the formal Edgeworth expansion (6.20) means that we can express the percentile points w as an asymptotic series

$$w = z + n^{-\frac{1}{2}} k_1(z) + n^{-1} k_2(z) + \cdots \tag{6.35}$$

where z is the percentile of Φ and the functions $k_1, k_2, \ldots$ are given by the authors' equation (22). Now, even if the values of $k_1, k_2, \ldots$ were not given, we can still make use of (6.35) in other ways. For example if we take

$$w = z + n^{-\frac{1}{2}} k_1(z) \tag{6.36}$$

by ignoring the other terms, we have, on Taylor-expanding about $w = z$

$$G'_n(w) = \Phi(z) + n^{-\frac{1}{2}} \left\{ k_1(z) + p_1^*(z) \right\} \phi(z) + 0\left(n^{-1}\right) . \tag{6.37}$$

In other words, setting $k_1(z) = -p_1^*(z)$ eliminates the term in (6.37) of order $n^{-\frac{1}{2}}$. An application of this is that

$$\Pr\left\{ W'_n \leq z - n^{-\frac{1}{2}} p_1^*(z) \right\} = \Phi(z) + 0\left(n^{-1}\right) . \tag{6.38}$$

In fact, taken at a more general level, this version of DMEE is an application of Lagrange's inversion rule for series functions, see e.g. Whittaker & Watson (1927, p. 133). See also Barndorff-Nielsen & Cox (1989, Section 3.5) for a sketch as well as its application by the authors on p. 117 to obtain a Cornish-Fisher inversion. Observe that the argument in

(6.38) is the percentile of *another distribution*, in this case Φ, and not the percentile of G'_n.

(ii) SECOND VERSION

If S_n and S'_n are two asymptotically normal statistics that satisfy

$$S'_n = S_n + 0_p\left(n^{-\frac{\upsilon}{2}}\right) \tag{6.39}$$

for an integer $\upsilon \geq 1$, then

$$\Pr\{S'_n \leq x\} = \Pr\{S_n \leq x\} + 0\left(n^{-\frac{\upsilon}{2}}\right). \tag{6.40}$$

For a proof, see Hall (1992b, Section 2.7).

(iii) THIRD VERSION

Another form of DMEE refers to the construction of Edgeworth expansion for a distribution of a transformed r.v. from cumulants/moments which are formally calculated from a Taylor series expansion of the transformed r.v. An example of this is our approach in this thesis. See also Skovgaard (1981) as well as a brief mention in Section 6.8 of Barndorff-Nielsen & Cox (1989).

We have gone through the trouble of distinguishing between the above common forms of DMEE in the hope that we can be more specific when describing Hall's approach since it employs the first two versions, as well as make clear the point of departure of our method which will shortly be proposed. It will be seen to belong to the third category of DMEE.

Now, following pp. 122–3 of Hall (1992b), let[1]

$$g_2(W'_n) = g_3(W'_n) - c'_3\{W'_n\}^3 \tag{6.41a}$$

$$= n^{-\frac{1}{2}}a\lambda_3\{W'_n\}^2 + W'_n - n^{-\frac{1}{2}}a\lambda_3. \tag{6.41b}$$

Then the event $g_2(W'_n) \leq w$ is equivalent to the event

$$\{W'_n\}^2 + n^{\frac{1}{2}}(a\lambda_3)^{-1}W'_n \leq n^{\frac{1}{2}}(a\lambda_3)^{-1}\left(w + n^{-\frac{1}{2}}a\lambda_3\right) \tag{6.42}$$

which, on completing the square, is in turn equivalent to the event

$$|W'_n| \leq n^{\frac{1}{2}}(2a\lambda_3)^{-1}\left[\left\{1 + 4n^{-\frac{1}{2}}a\lambda_3\left(w + n^{-\frac{1}{2}}a\lambda_3\right)\right\}^{\frac{1}{2}} - 1\right] \tag{6.43}$$

$$= w - n^{-\frac{1}{2}}p_1^*(w) + 0\left(n^{-1}\right) \tag{6.44}$$

where (6.44) is obtained by applying the binomial expansion rule to the RHS of (6.43). Since, as far as the first first version DMEE is concerned, the change of sign that is implied by $|W'_n| \leq w$ does not affect the order of the terms in the Edgeworth expansion, we have, on

[1] A correction needs to be made on p. 122 of Hall (1992b) where $a = \frac{1}{3}$ should be $a = -\frac{1}{6}$ as can be seen in our (6.22) as well as in Hall (1992a).

application of the latter to (6.44)

$$\Pr\left\{g_2\left(W_n'\right) \leq w\right\} = \Phi\left(z\right) + 0\left(n^{-1}\right) \tag{6.45}$$

as in (6.38). Then, because $g_3\left(W_n'\right)$ happens to be $g_2\left(W_n'\right)$ adjusted by a term of order n^{-1}, we have, on application of the second version of DMEE

$$\Pr\left\{g_3\left(W_n'\right) \leq w\right\} = \Phi(z) + 0\left(n^{-1}\right) . \tag{6.46}$$

Our explanation above differs slightly from Hall's in two respects.

In the first instance, we have made clear that there are two distinct forms of DMEE at play in Hall's (1992b) account on pp. 122–3. Secondly, the argument in our final expansions is the percentile of the *asymptotic* distribution rather than the percentile of the *exact* distribution.

It is not difficult to extend the same reasoning to the case of transformation (6.33). For, upon conversion to a Maclaurin series, we have

$$g\left(W_n'\right) = d_0' + \sum_{\upsilon=1}^{\infty} \frac{1}{\upsilon!} \left\{d_1\right\}^{\upsilon-1} \left\{W_n'\right\}^{\upsilon} \tag{6.47a}$$

$$= g_3\left(W_n'\right) + 0_p\left(n^{-\frac{3}{2}}\right) . \tag{6.47b}$$

(b) A PROPOSED OTHER METHOD

A natural alternative approach to the above is to demonstrate (6.46) using the third version of DMEE. Replacing (5.53) with (6.26), we see that identity (5.80) holds, which means that $\kappa_{3,n}^*$ has the same form as (5.60). In other words, were we to follow the same steps as in Section 6.2, the Edgeworth expansion for the d.f. of (6.26) would have no term in $n^{-\frac{1}{2}}$. In view of (5.67) and (5.79), the first term in that expansion, would be of order n^{-1}.

Now, because $g_3\left(W_n'\right)$ is in its original form and not in a Taylor series, the above does not strictly fit the description of our third version DMEE convention. However, we opt to retain the same name for our method so as to cover the general cases where the transformation in question requires to be converted into Taylor series. Hall's skewness-reducing transformation raises a potentially interesting open question regarding its extension to embrace the reduction of primary effects of kurtosis as well as the secondary effects of skewness present in the second term of the Edgeworth expansion.

6.6 CONFIDENCE INTERVAL

Let z_α denote the 100α-th percentile of the standard normal distribution. We want to construct $(1-\alpha)\,100\%$ approximate confidence interval for θ by transforming $z_{\frac{\alpha}{2}}$ onto the W_n'-scale using a third-degree polynomial of the kind (6.31). We start with

$$\Pr\left\{-z_{\frac{\alpha}{2}} \leq g_3\left(W_n'\right) \leq z_{\frac{\alpha}{2}}\right\} = \Pr\left\{g_3^{-1}\left(-z_{\frac{\alpha}{2}}\right) \leq W_n' \leq g_3^{-1}\left(z_{\frac{\alpha}{2}}\right)\right\} \tag{6.48}$$

Then the confidence interval that we desire is given by

$$\left[\underline{\theta}_n,\ \bar{\theta}_n\right] \tag{6.49}$$

where

$$\underline{\theta}_n = \hat{\theta}(n) + \left\{g_*^{-1}\left(z_{\frac{\alpha}{2}}\right) - n^{-\frac{1}{2}}\hat{\lambda}_1\right\} n^{-\frac{1}{2}}\hat{\kappa}_{2,\infty}^{\frac{1}{2}} \tag{6.50}$$

$$\bar{\theta}_n = \hat{\theta}(n) + \left\{g_*^{-1}\left(-z_{\frac{\alpha}{2}}\right) - n^{-\frac{1}{2}}\hat{\lambda}_1\right\} n^{-\frac{1}{2}}\hat{\kappa}_{2,\infty}^{\frac{1}{2}} \tag{6.51}$$

$$\begin{aligned} g_*^{-1}(z) &= \frac{6n^{\frac{1}{2}}}{\hat{\lambda}_3}\left[\left\{1 + \frac{1}{12}\hat{\lambda}_3 n^{-\frac{1}{2}}\left(\hat{\lambda}_3 n^{-\frac{1}{2}} - 6z\right)\right\}^{\frac{1}{3}} - 1\right] \\ &= -g_3^{-1}(z) \end{aligned} \tag{6.52}$$

and $\hat{\lambda}_1$, $\hat{\lambda}_3$, $\hat{\kappa}_{2,\infty}$ are weakly consistent equivariant estimates i.e. they are weakly consistent estimates based on the principle of functional equivalence. In this connection, it should be pointed out that $\hat{\beta}(n)$ is weakly consistent under condition (1.20), see e.g. Wu (1981, p.502). However, the matter of weak consistency estimators for λ_1 and $\lambda_2$need further investigation.

6.7 OTHER INTERVAL ESTIMATES

There are other Edgeworth-expansion-based transformations of $z_{\frac{\alpha}{2}}$ that have appeared in literature. Chief among these include the method due to Hill & Davis (1968), the techniques of Johnson (1978), as well as those of Abramovitch & Singh (1985). It is perhaps possible to adapt them to regression situations. However, all of these transformations are not strictly monotonic.

6.8 CHAPTER SUMMARY

The main activity of this chapter has been the completion of the theory leading to the statement of the confidence interval that is given by (6.49) to (6.52). During the proceedings, the alternative explanation for Hall's (1992a) skewness-reducing transformations which was started in Chapter five has also been completed. In the next and final chapter, we demonstrate, through simulation studies, both the internal coherence of our theory as well as the performance of our interval estimate.

CHAPTER SEVEN

NUMERICAL RESULTS AND CONCLUSION

7.1 INTRODUCTION

This last chapter examines Monte Carlo behaviour of the empirical average widths and miscoverage probabilities that are associated with three kinds of confidence intervals : our Edgeworth interval (6.49); the classical standard interval (2.4) and the Holm (1993)-method-based bootstrap interval (2.41).

Note that knowledge of the miscoverage probabilities of the lower and upper ends of the intervals implies knowledge of the coverage probability but not vice-versa and also that focusing on the former has the added advantage of allowing one to see any uneven miscoverage by the end-points that may arise.

Under consideration is the estimation of the abscissa of the turning point of the quadratic model (1.6). Four error distributions are considered : the normal distribution; the exponential distribution; the central student-t distribution with three degrees of freedom, hereafter abbreviated as student-t_3 and, the Laplace distribution.

7.2 MOTIVATION FOR MODEL CHOICE

Several factors have influenced our choosing the quadratic model. The first one is that this particular model as well as the associated estimation problem have appeared several times in literature , see

e.g. Weber & Welsh (1983, p. 443), Wu (1986, pp 1287 and 1288), and Urban Hjorth (1994, p. 193). The last author has illustrated the Holm's abstract bootstrap method using the following data generated using the following: $\beta_1 = 10, \beta_2 = -1, \beta_3 = 0.1, x_i = i$ for $i = 1, 2, \ldots, 20$, and an exponential distribution[1], with $\sigma = 4$. The author obtains the following 90% confidence interval for the x-coordinate of the turning point, namely

$$[-0.62\ ,\ 6.87] \tag{7.1}$$

This result has motivated us to use the same model and data to check the internal coherence of the entire theory that we have developed for our confidence interval. On doing this, we have obtained the following 90% confidence interval .

$$[3.22\ ,\ 7.46] \tag{7.2}$$

The point estimate is 4.67. Interval (7.2) was obtained both manually as well as by the use of a subroutine of the main programme that is appended as Appendix A.1.

The third and last reason which motivates our choice is that the ratio parameter given by (1.7) is, for all purposes under consideration, of the same form as the ratio function given by (1.4). Thus when we work with the former, we also in effect cover the case for (1.4). Indeed this is what motivated our verification of the crucial condition in Sections 4.6.1 and 4.6.2 for these two cases.

Finally, the decision to work with the deviations-from-the-mean form of regression models was not taken arbitrarily . In the early stages of this thesis, it became evident that the crucial condition (4.59) is not easily verified for the case of the simple linear model with uncentred design points. The reader is invited to check this in the case of, for example, $\pi^{0,0,1}$, $\pi^{0,1,1}$, $\pi^{0,0,0,1}$, and $\pi^{0,0,1,1}$.

Moreover, it also turns out that the deviations-from-the-mean form of the quadratic model avoids the pitfalls that have been pointed out by Bradley & Srivastava (1979) in connection with correlation in polynomial regression.

7.3 NATURE OF SIMULATIONS AND MODE OF REPORTING

The Monte Carlo simulations, which have been performed using an IBM computer RS6000 Model 43P, are characterised by the following : Ten thousand simulations have been used for each of the

[1] see Section 1 of Chapter 1

Edgeworth and standard confidence intervals. In the case of the bootstrap intervals, one thousand simulations have been run, each simulation being based on B = 1000 abstract bootstrap resamples (Holm's method). The sample sizes considered are 10,20, ..., 100. Three confidence levels, namely 0.99, 0.95, 0.90 were chosen.

The design points have been spaced in two ways : equally and unequally. In the equally-spaced case, we have used $x_i = i$ for $i = 1, 2 \ldots\ldots\ldots n_k$, $k = 1, 2, \ldots, 10$, $n_1 = 10, n_2 = 20, \ldots, n_{10} = 100$. In the case of unequally spaced design points, we have taken the Weber & Welsh (1983 , p.434) data points and simply added ten as follows :

1, 3, 5, 6, 6, 7, 8, 8.5, 9, 10

11, 13, 15, 16, 16, 17, 18., 18.5, 19, 20

:

:

:

:

91, 93, 95, 96, 96, 97, 98, 98.5, 99, 100

The results that are associated with the equally spaced x - levels are presented in sixteen tables, numbered 7.3 to 7.18 while the corresponding results for the unequally-spaced case are provided in Tables 7.19 to 7.34. To make clear the distinction between these two sets of results, we have reported them in clearly-marked Sections A and B. Below we give the meanings of abbreviations that are used in the miscoverage tables as well as a summary of the relevant moments of the distributions considered.

TABLE 7.1: Meanings of abbreviations used

L-ERROR=miscoverage rate for the lower end-point

U-ERROR=miscoverage rate for the upper end-point

TABLE 7.2: First three moments of distributions considered

Distribution	μ_1	μ_2	μ_3
Normal	0	σ^2	0
Student-t_3	0	3	0
Laplace	0	2	0
Exponential	0	σ^2	$2\sigma^3$

We have chosen the same values of σ^2 that have been used by Weber & Welsh (1983), namely 0.09, 0.25 and 1.00.

The general format of the title of each table is as follows. We start the title with the type of error distribution from which the corresponding responses are sampled. This is followed by the type of interval under consideration. The last part of the title describes the output characteristic(s) under display. All the results are reported to three decimal places.

7.4 RESULTS

We will focus only on the small samples ($n = 10, 20, 30$) and moderate samples ($n = 40, 50$).

Since the results show no noticable differences in respect of the two design points spacing modes considered, we will limit our summary to the features observed in the equally-spaced case.

7.4.1 AVERAGE WIDTH

At $\alpha = 0.01$ and $n = 10$ the bootstrap interval widths are seen to be unacceptably large for all the error distributions considered, though, in the cases relating to both the normal and exponential errors, this feature is observed only at $\sigma^2 = 1.00$. Also at $\sigma^2 = 1.00$ and $n = 10$ the same problem is observed in respect of the standard interval under exponential errors, at all the α levels. In no instance does our Edgeworth-corrected interval exhibit unacceptably large width.

Furthermore, in both the normal and exponential error cases, small values of σ^2, namely 0.09, 0.25 appear to make the intervals quite short as n increases from 10. This is seen to happen irrespective of the interval procedure used.

7.4.2 COVERAGE AND MISCOVERAGE

GENERAL

Table 7.11 shows that the unacceptably large widths of the standard interval mentioned above still fail to provide 100% empirical coverage at any of the α levels considered. However the empirical

coverage is virtually 100% for those bootstrap intervals with unacceptably large widths, see Tables 7.9, 7.12 and 7.18. Other instances exhibiting 100% empirical coverage in small to moderate samples are seen in Tables 7.10, 7.11 and 7.12 in respect of the exponential errors with $\sigma^2 = 0.09, 0.25$ for all three intervals.

Let us now turn to the main characteristics of the intervals under each error distribution.

NORMAL ERRORS : Tables 7.7 – 7.9

In small samples, the coverage of the standard interval is good at all σ and α considered. In the case of the Edgeworth interval, the coverage appears to improve as σ assumes smaller values than unity. The picture for the bootstrap intervals appears to depend on the level of confidence. Miscoverage is uneven for all three intervals but is most pronounced in respect of the Edgeworth intervals.

EXPONENTIAL ERRORS : Tables 7.10 – 7.12

The Edgeworth and Standard intervals display 100% empirical coverage for $\sigma^2 = 0.09, 0.25$ but at $\sigma^2 = 1.00$ their behaviour is reasonable. The bootstrap interval appears to suffer less from the problem.

STUDENT-t_3 ERRORS : Tables 7.13 – 7.15

The coverage of the Edgeworth interval is very poor at $n = 10$ at all α levels considered, but a sharp improvement at $n = 20$ is observed which gets even better at $n = 30$. Miscoverage is very uneven. The coverage of the standard interval has an edge. It, too, has quite uneven miscoverage. The coverage of the bootstrap intervals virtually mimmicks that of the standard intervals but the miscoverage is roughly even after $n = 20$.

LAPLACE ERRORS : Tables 7.16 – 7.18

The general picture is strikingly similar to that described above for student-t_3 errors.

SECTION A

Results for the case of equally spaced design points

TABLE 7.3: Normal distribution, average length of confidence intervals

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	Edgeworth	1.529	1.932	2.410	1.240	1.613	1.889	0.998	1.385	1.557
n=10	Standard	0.944	1.700	5.376	0.719	1.286	4.236	0.603	1.070	3.708
	Bootstrap	2.520	8.099	60.852	0.924	2.103	10.971	0.673	1.447	5.538
	Edgeworth	0.903	2.422	3.211	0.587	1.937	2.868	0.473	0.910	2.618
n=20	Standard	0.709	1.184	2.393	0.540	0.900	1.830	0.453	0.757	1.532
	Bootstrap	0.836	1.428	3.050	0.570	0.956	1.981	0.461	0.782	1.574
	Edgeworth	0.482	0.837	3.578	0.363	0.616	1.405	0.303	0.510	1.084
n=30	Standard	0.472	0.787	1.577	0.359	0.599	1.200	0.301	0.502	1.006
	Bootstrap	0.518	0.861	1.743	0.370	0.617	1.241	0.302	0.506	1.015
	Edgeworth	0.337	0.566	1.187	0.256	0.428	0.874	0.214	0.358	0.725
n=40	Standard	0.335	0.559	1.118	0.255	0.425	0.851	0.214	0.357	0.714
	Bootstrap	0.356	0.592	1.200	0.258	0.435	0.870	0.215	0.358	0.720
	Edgeworth	0.253	0.422	0.857	0.192	0.321	0.645	0.161	0.269	0.540
n=50	Standard	0.252	0.421	0.841	0.192	0.320	0.640	0.161	0.269	0.537
	Bootstrap	0.264	0.438	0.883	0.195	0.325	0.649	0.162	0.271	0.539
	Edgeworth	0.198	0.331	0.666	0.151	0.252	0.505	0.127	0.211	0.423
n=60	Standard	0.198	0.331	0.661	0.151	0.251	0.503	0.127	0.211	0.422
	Bootstrap	0.206	0.344	0.684	0.153	0.254	0.506	0.127	0.212	0.425
	Edgeworth	0.161	0.269	0.538	0.123	0.204	0.409	0.103	0.171	0.343
n=70	Standard	0.161	0.268	0.537	0.122	0.204	0.408	0.103	0.171	0.343
	Bootstrap	0.165	0.275	0.555	0.124	0.207	0.413	0.103	0.172	0.345
	Edgeworth	0.134	0.223	0.447	0.102	0.170	0.340	0.086	0.143	0.285
n=80	Standard	0.134	0.223	0.447	0.102	0.170	0.340	0.086	0.143	0.285
	Bootstrap	0.137	0.229	0.459	0.102	0.171	0.345	0.086	0.144	0.286
	Edgeworth	0.114	0.190	0.379	0.087	0.144	0.289	0.073	0.121	0.242
n=90	Standard	0.114	0.190	0.379	0.087	0.144	0.288	0.073	0.121	0.242
	Bootstrap	0.116	0.194	0.389	0.087	0.146	0.292	0.073	0.122	0.243
	Edgeworth	0.098	0.164	0.327	0.075	0.124	0.249	0.063	0.104	0.209
n=100	Standard	0.098	0.163	0.327	0.075	0.124	0.249	0.063	0.104	0.209
	Bootstrap	0.101	0.168	0.335	0.075	0.126	0.251	0.063	0.105	0.209

TABLE 7.4: Exponential distribution, average length of confidence intervals

		α								
		0.01			0.05			0.1		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	Edgeworth	1.657	1.744	3.215	1.433	1.244	2.531	1.189	1.021	2.029
n=10	Standard	0.917	1.556	960.264	0.698	1.185	284.329	0.586	0.993	278.422
	Bootstrap	0.598	1.988	49.859	0.245	0.724	8.572	0.177	0.537	4.765
	Edgeworth	1.227	2.488	2.517	1.071	1.765	1.862	0.938	0.872	1.545
n=20	Standard	0.708	1.181	2.395	0.539	0.899	1.827	0.452	0.754	1.532
	Bootstrap	0.250	0.709	3.131	0.167	0.463	1.943	0.134	0.377	1.584
	Edgeworth	0.861	1.544	3.359	0.744	1.298	1.370	0.627	0.645	1.075
n=30	Standard	0.472	0.787	1.577	0.359	0.599	1.200	0.301	0.502	1.007
	Bootstrap	0.157	0.426	1.765	0.109	0.302	1.245	0.089	0.243	0.989
	Edgeworth	0.648	1.122	2.459	0.549	0.929	1.107	0.288	0.438	0.793
n=40	Standard	0.335	0.559	1.118	0.255	0.425	0.851	0.214	0.357	0.714
	Bootstrap	0.107	0.300	1.210	0.077	0.217	0.854	0.064	0.175	0.705
	Edgeworth	0.513	0.874	1.842	0.420	0.694	0.846	0.194	0.314	0.599
n=50	Standard	0.252	0.421	0.841	0.192	0.320	0.640	0.161	0.269	0.537
	Bootstrap	0.079	0.223	0.893	0.057	0.160	0.648	0.048	0.133	0.537
	Edgeworth	0.420	0.710	1.465	0.314	0.375	0.638	0.145	0.239	0.467
n=60	Standard	0.198	0.330	0.661	0.151	0.251	0.503	0.127	0.211	0.422
	Bootstrap	0.062	0.173	0.690	0.046	0.125	0.511	0.038	0.106	0.419
	Edgeworth	0.352	0.592	1.205	0.163	0.263	0.498	0.115	0.190	0.375
n=70	Standard	0.161	0.268	0.537	0.122	0.204	0.408	0.103	0.171	0.343
	Bootstrap	0.050	0.138	0.556	0.037	0.103	0.412	0.030	0.085	0.339
	Edgeworth	0.300	0.502	1.013	0.126	0.207	0.402	0.094	0.156	0.309
n=80	Standard	0.134	0.223	0.447	0.102	0.170	0.340	0.086	0.143	0.285
	Bootstrap	0.041	0.116	0.464	0.031	0.085	0.340	0.026	0.071	0.286
	Edgeworth	0.258	0.431	0.862	0.103	0.170	0.334	0.079	0.131	0.260
n=90	Standard	0.114	0.190	0.379	0.087	0.144	0.288	0.073	0.121	0.242
	Bootstrap	0.035	0.097	0.391	0.026	0.073	0.290	0.022	0.060	0.241
	Edgeworth	0.223	0.371	0.735	0.086	0.143	0.283	0.067	0.112	0.222
n=100	Standard	0.098	0.163	0.327	0.075	0.124	0.249	0.063	0.104	0.209
	Bootstrap	0.030	0.084	0.335	0.023	0.062	0.250	0.019	0.052	0.207

TABLE 7.5: Student-t_3 distribution, average length of confidence intervals

		α		
		0.01	**0.05**	**0.10**
n=10	Edgeworth	2.564	1.930	1.569
	Standard	9.941	8.371	7.300
	Bootstrap	96.155	17.520	9.301
n=20	Edgeworth	3.920	3.491	3.195
	Standard	4.297	3.275	2.749
	Bootstrap	8.106	4.141	3.367
n=30	Edgeworth	4.848	4.211	3.611
	Standard	2.749	2.089	1.748
	Bootstrap	3.611	2.083	1.722
n=40	Edgeworth	2.762	1.628	1.301
	Standard	1.938	1.476	1.238
	Bootstrap	2.023	1.425	1.157
n=50	Edgeworth	1.547	1.140	0.945
	Standard	1.458	1.109	0.931
	Bootstrap	1.492	1.062	0.880
n=60	Edgeworth	1.171	0.880	0.736
	Standard	1.145	0.871	0.731
	Bootstrap	1.165	0.841	0.700
n=70	Edgeworth	0.939	0.711	0.595
	Standard	0.929	0.707	0.594
	Bootstrap	0.947	0.692	0.581
n=80	Edgeworth	0.777	0.590	0.495
	Standard	0.774	0.589	0.494
	Bootstrap	0.785	0.573	0.474
n=90	Edgeworth	0.658	0.500	0.420
	Standard	0.657	0.500	0.419
	Bootstrap	0.670	0.493	0.401
n=100	Edgeworth	0.567	0.431	0.362
	Standard	0.566	0.431	0.362
	Bootstrap	0.576	0.425	0.344

TABLE 7.6: Laplace distribution, average length of confidence intervals

		α		
		0.01	**0.05**	**0.10**
	Edgeworth	2.510	1.890	1.574
n=10	Standard	8.495	7.031	6.248
	Bootstrap	88.152	16.371	8.079
	Edgeworth	3.645	3.261	2.989
n=20	Standard	3.444	2.627	2.196
	Bootstrap	4.825	2.921	2.306
	Edgeworth	4.433	3.690	1.835
n=30	Standard	2.234	1.700	1.426
	Bootstrap	2.497	1.731	1.432
	Edgeworth	1.832	1.277	1.043
n=40	Standard	1.582	1.204	1.010
	Bootstrap	1.705	1.225	1.009
	Edgeworth	1.236	0.921	0.768
n=50	Standard	1.190	0.905	0.760
	Bootstrap	1.252	0.912	0.758
	Edgeworth	0.949	0.716	0.599
n=60	Standard	0.935	0.711	0.597
	Bootstrap	0.972	0.718	0.591
	Edgeworth	0.764	0.579	0.485
n=70	Standard	0.759	0.578	0.485
	Bootstrap	0.788	0.583	0.483
	Edgeworth	0.634	0.481	0.404
n=80	Standard	0.632	0.481	0.403
	Bootstrap	0.651	0.485	0.404
	Edgeworth	0.537	0.408	0.343
n=90	Standard	0.536	0.408	0.342
	Bootstrap	0.550	0.411	0.341
	Edgeworth	0.463	0.352	0.295
n=100	Standard	0.462	0.352	0.295
	Bootstrap	0.476	0.353	0.296

TABLE 7.7: Normal distribution, Edgeworth intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.018	0.002	0.047	0.056	0.032	0.060	0.087	0.075	0.081
n=10	U-error	0.000	0.004	0.119	0.000	0.015	0.198	0.000	0.027	0.254
	Coverage	0.982	0.994	0.834	0.944	0.953	0.742	0.913	0.898	0.665
	L-error	0.020	0.038	0.073	0.051	0.070	0.102	0.076	0.095	0.124
n=20	U-error	0.000	0.000	0.000	0.003	0.000	0.000	0.018	0.000	0.000
	Coverage	0.980	0.962	0.928	0.946	0.930	0.898	0.906	0.905	0.876
	L-error	0.010	0.015	0.028	0.035	0.042	0.059	0.061	0.069	0.080
n=30	U-error	0.001	0.001	0.000	0.014	0.008	0.000	0.040	0.031	0.011
	Coverage	0.989	0.985	0.972	0.951	0.950	0.941	0.899	0.901	0.910
	L-error	0.008	0.011	0.015	0.027	0.034	0.042	0.057	0.062	0.065
n=40	U-error	0.003	0.001	0.000	0.021	0.017	0.010	0.043	0.038	0.031
	Coverage	0.990	0.988	0.985	0.953	0.950	0.948	0.901	0.900	0.904
	L-error	0.006	0.007	0.010	0.031	0.029	0.032	0.057	0.056	0.060
n=50	U-error	0.003	0.004	0.001	0.021	0.023	0.015	0.047	0.045	0.038
	Coverage	0.991	0.989	0.989	0.948	0.948	0.953	0.896	0.899	0.902
	L-error	0.007	0.007	0.008	0.030	0.027	0.034	0.049	0.059	0.058
n=60	U-error	0.004	0.003	0.003	0.024	0.025	0.020	0.049	0.049	0.044
	Coverage	0.989	0.991	0.989	0.946	0.948	0.946	0.901	0.892	0.899
	L-error	0.006	0.006	0.005	0.026	0.030	0.027	0.050	0.049	0.052
n=70	U-error	0.004	0.004	0.004	0.024	0.022	0.025	0.050	0.050	0.048
	Coverage	0.989	0.990	0.992	0.950	0.948	0.949	0.900	0.901	0.900
	L-error	0.006	0.005	0.007	0.026	0.026	0.026	0.051	0.057	0.053
n=80	U-error	0.005	0.006	0.002	0.023	0.024	0.021	0.049	0.045	0.047
	Coverage	0.990	0.990	0.991	0.951	0.950	0.952	0.900	0.898	0.900
	L-error	0.006	0.006	0.006	0.026	0.025	0.023	0.048	0.053	0.050
n=90	U-error	0.005	0.004	0.005	0.026	0.024	0.022	0.048	0.053	0.045
	Coverage	0.989	0.990	0.990	0.948	0.951	0.954	0.905	0.894	0.905
	L-error	0.005	0.006	0.006	0.026	0.026	0.026	0.051	0.053	0.053
n=100	U-error	0.006	0.005	0.006	0.024	0.024	0.023	0.049	0.045	0.048
	Coverage	0.989	0.990	0.989	0.950	0.951	0.952	0.900	0.902	0.899

TABLE 7.8: Normal distribution, standard intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.007	0.009	0.006	0.032	0.033	0.025	0.056	0.062	0.053
n=10	U-error	0.000	0.000	0.000	0.011	0.003	0.000	0.032	0.015	0.001
	Coverage	0.993	0.991	0.994	0.958	0.964	0.975	0.912	0.923	0.946
	L-error	0.007	0.010	0.012	0.027	0.034	0.036	0.055	0.063	0.062
n=20	U-error	0.003	0.002	0.000	0.021	0.015	0.010	0.044	0.042	0.029
	Coverage	0.990	0.988	0.988	0.951	0.952	0.954	0.900	0.895	0.909
	L-error	0.006	0.006	0.009	0.027	0.028	0.033	0.051	0.054	0.058
n=30	U-error	0.004	0.004	0.003	0.022	0.022	0.019	0.049	0.045	0.043
	Coverage	0.990	0.990	0.989	0.951	0.951	0.948	0.901	0.901	0.899
	L-error	0.006	0.007	0.007	0.023	0.027	0.028	0.052	0.053	0.052
n=40	U-error	0.004	0.004	0.004	0.024	0.023	0.020	0.046	0.045	0.045
	Coverage	0.990	0.989	0.989	0.954	0.949	0.952	0.902	0.902	0.902
	L-error	0.005	0.005	0.006	0.029	0.026	0.026	0.055	0.053	0.053
n=50	U-error	0.005	0.006	0.005	0.023	0.026	0.021	0.049	0.049	0.048
	Coverage	0.991	0.989	0.989	0.948	0.949	0.953	0.897	0.898	0.900
	L-error	0.006	0.005	0.005	0.029	0.024	0.027	0.048	0.056	0.052
n=60	U-error	0.004	0.004	0.005	0.025	0.027	0.025	0.051	0.051	0.050
	Coverage	0.990	0.992	0.990	0.946	0.950	0.948	0.902	0.893	0.898
	L-error	0.006	0.006	0.004	0.025	0.028	0.024	0.049	0.048	0.049
n=70	U-error	0.005	0.005	0.005	0.025	0.024	0.028	0.051	0.051	0.053
	Coverage	0.989	0.989	0.991	0.950	0.948	0.948	0.900	0.902	0.898
	L-error	0.005	0.004	0.005	0.025	0.025	0.024	0.051	0.056	0.051
n=80	U-error	0.005	0.006	0.003	0.024	0.025	0.024	0.050	0.046	0.050
	Coverage	0.990	0.990	0.992	0.951	0.950	0.953	0.899	0.898	0.899
	L-error	0.006	0.006	0.005	0.026	0.025	0.022	0.047	0.052	0.048
n=90	U-error	0.006	0.004	0.006	0.027	0.024	0.024	0.048	0.054	0.047
	Coverage	0.989	0.991	0.990	0.948	0.951	0.955	0.905	0.894	0.905
	L-error	0.005	0.006	0.005	0.025	0.026	0.025	0.050	0.052	0.052
n=100	U-error	0.006	0.005	0.007	0.024	0.024	0.023	0.050	0.046	0.049
	Coverage	0.990	0.989	0.988	0.950	0.950	0.952	0.901	0.902	0.899

TABLE 7.9: Normal distribution, bootstrap intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.003	0.001	0.000	0.020	0.019	0.020	0.065	0.051	0.048
n=10	U-error	0.001	0.000	0.000	0.019	0.020	0.009	0.065	0.049	0.022
	Coverage	0.996	0.999	1.000	0.961	0.961	0.971	0.870	0.900	0.930
	L-error	0.004	0.005	0.002	0.032	0.029	0.027	0.056	0.053	0.071
n=20	U-error	0.005	0.002	0.002	0.018	0.033	0.034	0.046	0.054	0.050
	Coverage	0.991	0.993	0.996	0.950	0.938	0.939	0.898	0.893	0.879
	L-error	0.009	0.007	0.004	0.024	0.029	0.019	0.057	0.062	0.063
n=30	U-error	0.004	0.010	0.002	0.037	0.029	0.030	0.051	0.060	0.066
	Coverage	0.987	0.983	0.994	0.939	0.942	0.951	0.892	0.878	0.871
	L-error	0.005	0.004	0.003	0.026	0.032	0.035	0.068	0.061	0.049
n=40	U-error	0.005	0.003	0.006	0.028	0.028	0.025	0.052	0.054	0.068
	Coverage	0.990	0.993	0.991	0.946	0.940	0.940	0.880	0.885	0.883
	L-error	0.001	0.002	0.006	0.024	0.016	0.033	0.048	0.048	0.044
n=50	U-error	0.006	0.006	0.007	0.032	0.031	0.034	0.049	0.057	0.060
	Coverage	0.993	0.992	0.987	0.944	0.953	0.933	0.903	0.895	0.896
	L-error	0.002	0.013	0.008	0.018	0.020	0.024	0.042	0.053	0.045
n=60	U-error	0.008	0.008	0.006	0.035	0.026	0.033	0.062	0.059	0.043
	Coverage	0.990	0.979	0.986	0.947	0.954	0.943	0.896	0.888	0.912
	L-error	0.012	0.003	0.002	0.029	0.019	0.025	0.049	0.043	0.042
n=70	U-error	0.010	0.004	0.005	0.028	0.039	0.031	0.049	0.054	0.058
	Coverage	0.978	0.993	0.993	0.943	0.942	0.944	0.902	0.903	0.900
	L-error	0.006	0.004	0.003	0.026	0.015	0.020	0.054	0.042	0.050
n=80	U-error	0.002	0.004	0.008	0.026	0.027	0.025	0.056	0.058	0.048
	Coverage	0.992	0.992	0.989	0.948	0.958	0.955	0.890	0.900	0.902
	L-error	0.004	0.004	0.003	0.033	0.028	0.026	0.057	0.041	0.052
n=90	U-error	0.006	0.005	0.003	0.026	0.023	0.018	0.046	0.045	0.061
	Coverage	0.990	0.991	0.994	0.941	0.949	0.956	0.897	0.914	0.887
	L-error	0.003	0.007	0.007	0.023	0.024	0.026	0.045	0.050	0.057
n=100	U-error	0.009	0.004	0.009	0.028	0.022	0.021	0.053	0.046	0.045
	Coverage	0.988	0.989	0.984	0.949	0.954	0.953	0.902	0.904	0.898

TABLE 7.10: Exponential distribution, Edgeworth intervals, miscoverage and coverage rate

		α								
		0.01			0.05			0.1		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.000	0.001	0.047	0.000	0.002	0.066	0.000	0.006	0.084
n=10	U-error	0.000	0.000	0.059	0.000	0.000	0.100	0.000	0.000	0.140
	Coverage	1.000	0.999	0.894	1.000	0.998	0.835	1.000	0.994	0.776
	L-error	0.000	0.000	0.001	0.000	0.000	0.035	0.000	0.000	0.067
n=20	U-error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001
	Coverage	1.000	1.000	0.999	1.000	1.000	0.965	1.000	0.999	0.932
	L-error	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.028
n=30	U-error	0.000	0.000	0.008	0.000	0.001	0.028	0.000	0.001	0.050
	Coverage	1.000	1.000	0.992	1.000	0.999	0.966	1.000	0.999	0.922
	L-error	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.023
n=40	U-error	0.000	0.000	0.014	0.000	0.001	0.039	0.000	0.002	0.062
	Coverage	1.000	1.000	0.986	1.000	0.999	0.960	1.000	0.999	0.915
	L-error	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.020
n=50	U-error	0.000	0.000	0.015	0.000	0.000	0.045	0.000	0.001	0.067
	Coverage	1.000	1.000	0.985	1.000	1.000	0.954	1.000	0.999	0.913
	L-error	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.022
n=60	U-error	0.000	0.000	0.018	0.000	0.001	0.045	0.000	0.002	0.067
	Coverage	1.000	1.000	0.982	1.000	0.999	0.954	1.000	0.998	0.912
	L-error	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.019
n=70	U-error	0.000	0.000	0.018	0.000	0.000	0.042	0.000	0.002	0.068
	Coverage	1.000	1.000	0.982	1.000	1.000	0.956	1.000	0.998	0.914
	L-error	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.024
n=80	U-error	0.000	0.000	0.017	0.000	0.000	0.048	0.000	0.002	0.066
	Coverage	1.000	1.000	0.983	1.000	1.000	0.949	1.000	0.998	0.911
	L-error	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.023
n=90	U-error	0.000	0.000	0.015	0.000	0.000	0.043	0.000	0.001	0.064
	Coverage	1.000	1.000	0.985	1.000	1.000	0.953	1.000	0.999	0.913
	L-error	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.024
n=100	U-error	0.000	0.000	0.017	0.000	0.000	0.043	0.000	0.002	0.070
	Coverage	1.000	1.000	0.981	1.000	1.000	0.953	1.000	0.998	0.906

TABLE 7.11: Exponential distribution, standard intervals, miscoverage and coverage rates

		α = 0.01			α = 0.05			α = 0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
n=10	L-error	0.000	0.001	0.013	0.000	0.003	0.041	0.000	0.008	0.060
	U-error	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.005
	Coverage	1.000	0.999	0.987	1.000	0.997	0.959	1.000	0.992	0.936
n=20	L-error	0.000	0.000	0.024	0.000	0.002	0.045	0.000	0.004	0.071
	U-error	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.022
	Coverage	1.000	1.000	0.977	1.000	0.998	0.947	1.000	0.995	0.907
n=30	L-error	0.000	0.000	0.014	0.000	0.001	0.037	0.000	0.003	0.062
	U-error	0.000	0.000	0.002	0.000	0.000	0.014	0.000	0.001	0.036
	Coverage	1.000	1.000	0.985	1.000	0.999	0.948	1.000	0.996	0.903
n=40	L-error	0.000	0.000	0.010	0.000	0.000	0.034	0.000	0.002	0.061
	U-error	0.000	0.000	0.003	0.000	0.000	0.018	0.000	0.000	0.039
	Coverage	1.000	1.000	0.987	1.000	1.000	0.948	1.000	0.998	0.900
n=50	L-error	0.000	0.000	0.010	0.000	0.000	0.037	0.000	0.002	0.056
	U-error	0.000	0.000	0.003	0.000	0.000	0.021	0.000	0.000	0.040
	Coverage	1.000	1.000	0.987	1.000	1.000	0.942	1.000	0.998	0.903
n=60	L-error	0.000	0.000	0.010	0.000	0.000	0.031	0.000	0.002	0.055
	U-error	0.000	0.000	0.003	0.000	0.000	0.020	0.000	0.000	0.045
	Coverage	1.000	1.000	0.987	1.000	1.000	0.949	1.000	0.998	0.901
n=70	L-error	0.000	0.000	0.008	0.000	0.001	0.033	0.000	0.002	0.053
	U-error	0.000	0.000	0.003	0.000	0.000	0.019	0.000	0.001	0.044
	Coverage	1.000	1.000	0.989	1.000	0.999	0.948	1.000	0.998	0.903
n=80	L-error	0.000	0.000	0.008	0.000	0.001	0.031	0.000	0.002	0.053
	U-error	0.000	0.000	0.004	0.000	0.000	0.021	0.000	0.000	0.044
	Coverage	1.000	1.000	0.988	1.000	0.999	0.948	1.000	0.998	0.903
n=90	L-error	0.000	0.000	0.008	0.000	0.000	0.031	0.000	0.001	0.054
	U-error	0.000	0.000	0.004	0.000	0.000	0.021	0.000	0.000	0.043
	Coverage	1.000	1.000	0.988	1.000	1.000	0.948	1.000	0.999	0.904
n=100	L-error	0.000	0.000	0.008	0.000	0.000	0.028	0.000	0.001	0.050
	U-error	0.000	0.000	0.004	0.000	0.000	0.021	0.000	0.001	0.047
	Coverage	1.000	1.000	0.989	1.000	1.000	0.951	1.000	0.999	0.903

TABLE 7.12: Exponential distribution, bootstrap intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.003	0.002	0.003	0.033	0.024	0.027	0.052	0.055	0.056
n=10	U-error	0.001	0.000	0.000	0.015	0.013	0.015	0.036	0.054	0.026
	Coverage	0.996	0.998	0.997	0.952	0.963	0.958	0.912	0.891	0.918
	L-error	0.003	0.002	0.002	0.025	0.018	0.036	0.060	0.058	0.063
n=20	U-error	0.000	0.004	0.003	0.013	0.032	0.026	0.053	0.043	0.047
	Coverage	0.997	0.994	0.995	0.962	0.950	0.938	0.887	0.899	0.890
	L-error	0.004	0.003	0.004	0.024	0.022	0.025	0.050	0.066	0.052
n=30	U-error	0.002	0.002	0.006	0.025	0.027	0.026	0.062	0.052	0.052
	Coverage	0.994	0.995	0.990	0.951	0.951	0.949	0.888	0.882	0.896
	L-error	0.003	0.005	0.003	0.019	0.016	0.019	0.065	0.051	0.051
n=40	U-error	0.002	0.002	0.003	0.018	0.018	0.023	0.062	0.042	0.050
	Coverage	0.995	0.993	0.994	0.963	0.966	0.958	0.873	0.907	0.899
	L-error	0.004	0.005	0.007	0.029	0.026	0.030	0.064	0.049	0.073
n=50	U-error	0.005	0.004	0.008	0.023	0.029	0.025	0.046	0.045	0.048
	Coverage	0.991	0.991	0.985	0.948	0.945	0.945	0.890	0.906	0.879
	L-error	0.004	0.004	0.005	0.021	0.017	0.018	0.060	0.053	0.062
n=60	U-error	0.012	0.005	0.003	0.024	0.030	0.024	0.043	0.038	0.050
	Coverage	0.984	0.991	0.992	0.955	0.953	0.958	0.897	0.909	0.888
	L-error	0.004	0.003	0.005	0.025	0.028	0.031	0.042	0.050	0.048
n=70	U-error	0.005	0.002	0.004	0.014	0.024	0.042	0.047	0.038	0.050
	Coverage	0.991	0.995	0.991	0.961	0.948	0.927	0.911	0.912	0.902
	L-error	0.006	0.006	0.008	0.030	0.021	0.025	0.041	0.066	0.050
n=80	U-error	0.005	0.004	0.002	0.013	0.031	0.029	0.048	0.047	0.056
	Coverage	0.989	0.990	0.990	0.957	0.948	0.946	0.911	0.887	0.894
	L-error	0.002	0.001	0.004	0.023	0.030	0.026	0.053	0.052	0.051
n=90	U-error	0.002	0.006	0.003	0.010	0.020	0.022	0.044	0.049	0.051
	Coverage	0.996	0.993	0.993	0.967	0.950	0.952	0.903	0.899	0.898
	L-error	0.003	0.003	0.004	0.023	0.026	0.025	0.052	0.046	0.052
n=100	U-error	0.010	0.001	0.005	0.036	0.030	0.032	0.043	0.048	0.055
	Coverage	0.987	0.996	0.991	0.941	0.944	0.943	0.905	0.906	0.893

TABLE 7.13: Student t_3 distribution, Edgeworth intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.152	0.181	0.198
	U-error	0.164	0.251	0.290
	Coverage	0.684	0.569	0.512
n=20	L-error	0.088	0.116	0.136
	U-error	0.002	0.003	0.004
	Coverage	0.911	0.881	0.860
n=30	L-error	0.035	0.063	0.091
	U-error	0.000	0.000	0.000
	Coverage	0.965	0.937	0.909
n=40	L-error	0.021	0.043	0.066
	U-error	0.000	0.005	0.013
	Coverage	0.979	0.953	0.921
n=50	L-error	0.017	0.036	0.059
	U-error	0.002	0.011	0.029
	Coverage	0.981	0.953	0.912
n=60	L-error	0.012	0.034	0.053
	U-error	0.004	0.014	0.032
	Coverage	0.984	0.951	0.915
n=70	L-error	0.010	0.029	0.049
	U-error	0.005	0.018	0.037
	Coverage	0.985	0.953	0.914
n=80	L-error	0.009	0.032	0.052
	U-error	0.005	0.021	0.043
	Coverage	0.986	0.947	0.906
n=90	L-error	0.007	0.027	0.048
	U-error	0.005	0.022	0.037
	Coverage	0.988	0.951	0.915
n=100	L-error	0.007	0.023	0.049
	U-error	0.007	0.022	0.039
	Coverage	0.986	0.955	0.912

TABLE 7.14: Student t_3 distribution, standard intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.008	0.017	0.029
	U-error	0.000	0.000	0.002
	Coverage	0.992	0.983	0.969
n=20	L-error	0.019	0.039	0.062
	U-error	0.000	0.003	0.011
	Coverage	0.981	0.958	0.927
n=30	L-error	0.011	0.030	0.054
	U-error	0.004	0.017	0.030
	Coverage	0.985	0.953	0.916
n=40	L-error	0.009	0.025	0.047
	U-error	0.005	0.019	0.034
	Coverage	0.986	0.956	0.919
n=50	L-error	0.010	0.025	0.048
	U-error	0.006	0.021	0.040
	Coverage	0.985	0.954	0.912
n=60	L-error	0.008	0.027	0.045
	U-error	0.007	0.021	0.040
	Coverage	0.985	0.953	0.915
n=70	L-error	0.007	0.024	0.043
	U-error	0.008	0.023	0.043
	Coverage	0.985	0.954	0.914
n=80	L-error	0.007	0.029	0.047
	U-error	0.006	0.025	0.047
	Coverage	0.986	0.946	0.907
n=90	L-error	0.006	0.025	0.045
	U-error	0.007	0.025	0.040
	Coverage	0.988	0.950	0.915
n=100	L-error	0.007	0.020	0.047
	U-error	0.008	0.025	0.040
	Coverage	0.985	0.956	0.913

TABLE 7.15: Student t_3 distribution, bootstrap intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.000	0.015	0.045
	U-error	0.000	0.002	0.015
	Coverage	1.000	0.983	0.940
n=20	L-error	0.000	0.031	0.056
	U-error	0.003	0.024	0.057
	Coverage	0.997	0.945	0.887
n=30	L-error	0.002	0.018	0.035
	U-error	0.004	0.018	0.057
	Coverage	0.994	0.964	0.908
n=40	L-error	0.002	0.022	0.046
	U-error	0.004	0.028	0.049
	Coverage	0.994	0.950	0.905
n=50	L-error	0.005	0.026	0.058
	U-error	0.004	0.019	0.045
	Coverage	0.991	0.955	0.897
n=60	L-error	0.006	0.028	0.052
	U-error	0.002	0.028	0.057
	Coverage	0.992	0.944	0.891
n=70	L-error	0.008	0.014	0.040
	U-error	0.003	0.016	0.069
	Coverage	0.989	0.970	0.891
n=80	L-error	0.002	0.025	0.044
	U-error	0.002	0.027	0.052
	Coverage	0.996	0.948	0.904
n=90	L-error	0.005	0.029	0.063
	U-error	0.007	0.023	0.057
	Coverage	0.988	0.948	0.880
n=100	L-error	0.005	0.021	0.041
	U-error	0.003	0.026	0.064
	Coverage	0.992	0.953	0.895

TABLE 7.16: Laplace distribution, Edgeworth intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.115	0.139	0.154
	U-error	0.154	0.239	0.292
	Coverage	0.731	0.622	0.554
n=20	L-error	0.092	0.125	0.150
	U-error	0.000	0.000	0.000
	Coverage	0.908	0.875	0.849
n=30	L-error	0.035	0.068	0.093
	U-error	0.000	0.000	0.000
	Coverage	0.965	0.932	0.907
n=40	L-error	0.020	0.046	0.074
	U-error	0.000	0.006	0.023
	Coverage	0.980	0.949	0.904
n=50	L-error	0.011	0.039	0.060
	U-error	0.001	0.014	0.034
	Coverage	0.988	0.947	0.906
n=60	L-error	0.010	0.035	0.055
	U-error	0.003	0.017	0.042
	Coverage	0.988	0.948	0.903
n=70	L-error	0.008	0.031	0.055
	U-error	0.003	0.021	0.041
	Coverage	0.990	0.949	0.904
n=80	L-error	0.009	0.028	0.056
	U-error	0.004	0.020	0.048
	Coverage	0.988	0.952	0.897
n=90	L-error	0.008	0.028	0.049
	U-error	0.005	0.025	0.044
	Coverage	0.987	0.947	0.907
n=100	L-error	0.005	0.027	0.051
	U-error	0.005	0.024	0.046
	Coverage	0.990	0.950	0.903

TABLE 7.17: Laplace distribution, standard intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.005	0.020	0.041
	U-error	0.000	0.000	0.001
	Coverage	0.995	0.979	0.959
n=20	L-error	0.018	0.051	0.078
	U-error	0.000	0.006	0.022
	Coverage	0.983	0.943	0.900
n=30	L-error	0.011	0.036	0.061
	U-error	0.002	0.016	0.038
	Coverage	0.987	0.948	0.902
n=40	L-error	0.007	0.028	0.054
	U-error	0.003	0.023	0.046
	Coverage	0.990	0.949	0.900
n=50	L-error	0.006	0.028	0.050
	U-error	0.005	0.024	0.046
	Coverage	0.989	0.948	0.904
n=60	L-error	0.007	0.029	0.050
	U-error	0.005	0.023	0.049
	Coverage	0.988	0.949	0.901
n=70	L-error	0.005	0.025	0.051
	U-error	0.004	0.025	0.045
	Coverage	0.991	0.951	0.904
n=80	L-error	0.007	0.024	0.052
	U-error	0.006	0.024	0.052
	Coverage	0.988	0.952	0.897
n=90	L-error	0.007	0.026	0.047
	U-error	0.006	0.028	0.048
	Coverage	0.987	0.946	0.905
n=100	L-error	0.005	0.025	0.049
	U-error	0.005	0.026	0.047
	Coverage	0.990	0.950	0.904

TABLE 7.18: Laplace distribution, bootstrap intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.000	0.015	0.047
	U-error	0.000	0.003	0.014
	Coverage	1.000	0.982	0.939
n=20	L-error	0.003	0.029	0.063
	U-error	0.004	0.031	0.054
	Coverage	0.993	0.940	0.883
n=30	L-error	0.000	0.031	0.052
	U-error	0.002	0.020	0.045
	Coverage	0.998	0.949	0.903
n=40	L-error	0.005	0.029	0.048
	U-error	0.005	0.025	0.059
	Coverage	0.990	0.946	0.893
n=50	L-error	0.005	0.021	0.044
	U-error	0.009	0.023	0.046
	Coverage	0.986	0.956	0.910
n=60	L-error	0.005	0.029	0.055
	U-error	0.004	0.033	0.049
	Coverage	0.991	0.938	0.896
n=70	L-error	0.004	0.033	0.053
	U-error	0.002	0.021	0.059
	Coverage	0.994	0.946	0.888
n=80	L-error	0.003	0.023	0.054
	U-error	0.003	0.038	0.052
	Coverage	0.994	0.939	0.894
n=90	L-error	0.009	0.026	0.035
	U-error	0.004	0.023	0.061
	Coverage	0.987	0.951	0.904
n=100	L-error	0.008	0.027	0.037
	U-error	0.005	0.031	0.071
	Coverage	0.987	0.942	0.892

SECTION B

Results for the case of unequally spaced design points

TABLE 7.19: Normal distribution, average length of confidence intervals

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	Edgeworth	1.667	2.103	2.634	1.351	1.747	2.076	1.096	1.511	1.706
n=10	Standard	1.026	1.826	5.582	0.780	1.392	4.446	0.656	1.164	3.704
	Bootstrap	3.377	10.224	60.051	1.056	2.453	12.948	0.745	1.405	6.299
	Edgeworth	1.412	2.700	3.548	0.702	2.277	3.165	0.559	1.257	2.901
n=20	Standard	0.831	1.387	2.814	0.631	1.056	2.141	0.530	0.886	1.798
	Bootstrap	1.030	1.740	3.743	0.662	1.112	2.333	0.542	0.904	1.889
	Edgeworth	0.532	0.930	3.916	0.400	0.680	1.592	0.334	0.563	1.208
n=30	Standard	0.520	0.867	1.736	0.396	0.659	1.321	0.332	0.554	1.110
	Bootstrap	0.573	0.955	1.922	0.407	0.684	1.366	0.337	0.557	1.119
	Edgeworth	0.360	0.606	1.273	0.273	0.457	0.936	0.229	0.383	0.775
n=40	Standard	0.358	0.597	1.194	0.273	0.454	0.909	0.229	0.381	0.763
	Bootstrap	0.381	0.639	1.279	0.278	0.464	0.932	0.230	0.385	0.765
	Edgeworth	0.265	0.434	0.900	0.202	0.337	0.678	0.169	0.282	0.567
n=50	Standard	0.265	0.442	0.884	0.202	0.336	0.672	0.169	0.282	0.564
	Bootstrap	0.277	0.461	0.930	0.205	0.344	0.682	0.171	0.284	0.566
	Edgeworth	0.206	0.344	0.692	0.157	0.262	0.525	0.132	0.220	0.440
n=60	Standard	0.206	0.344	0.687	0.157	0.261	0.523	0.132	0.219	0.439
	Bootstrap	0.214	0.356	0.714	0.159	0.264	0.532	0.133	0.222	0.442
	Edgeworth	0.166	0.277	0.556	0.126	0.211	0.422	0.106	0.177	0.354
n=70	Standard	0.166	0.277	0.554	0.126	0.211	0.422	0.106	0.177	0.354
	Bootstrap	0.173	0.285	0.568	0.128	0.213	0.425	0.107	0.178	0.355
	Edgeworth	0.138	0.230	0.460	0.105	0.175	0.349	0.088	0.147	0.293
n=80	Standard	0.138	0.229	0.459	0.105	0.175	0.349	0.088	0.146	0.293
	Bootstrap	0.142	0.237	0.473	0.105	0.177	0.350	0.089	0.147	0.295
	Edgeworth	0.116	0.194	0.388	0.089	0.148	0.295	0.074	0.124	0.248
n=90	Standard	0.116	0.194	0.338	0.089	0.148	0.295	0.074	0.124	0.248
	Bootstrap	0.119	0.198	0.399	0.090	0.149	0.297	0.075	0.124	0.248
	Edgeworth	0.100	0.167	0.334	0.076	0.127	0.254	0.064	0.107	0.213
n=100	Standard	0.100	0.167	0.334	0.076	0.127	0.254	0.064	0.107	0.213
	Bootstrap	0.102	0.170	0.342	0.076	0.128	0.255	0.064	0.107	0.213

TABLE 7.20: Exponential distribution, average length of confidence intervals

		α = 0.01			α = 0.05			α = 0.10		
		σ^2			σ^2			σ^2		
		0.09	**0.25**	**1.00**	**0.09**	**0.25**	**1.00**	**0.09**	**0.25**	**1.00**
	Edgeworth	0.908	1.959	4.209	0.810	1.751	3.349	0.743	1.601	2.754
n=10	Standard	0.997	1.689	5.150	0.759	1.287	3.974	0.637	1.079	3.437
	Bootstrap	0.214	1.078	53.049	0.083	0.388	10.159	0.059	0.273	5.209
	Edgeworth	0.934	1.757	5.328	0.835	1.567	3.860	0.767	1.438	2.614
n=20	Standard	0.829	1.382	2.811	0.631	1.052	2.135	0.529	0.883	1.799
	Bootstrap	0.090	0.430	4.001	0.058	0.270	2.407	0.047	0.215	1.899
	Edgeworth	0.891	1.597	3.868	0.779	1.376	1.664	0.686	1.142	1.219
n=30	Standard	0.520	0.867	1.736	0.396	0.660	1.321	0.332	0.553	1.109
	Bootstrap	0.051	0.238	1.931	0.036	0.167	1.352	0.029	0.137	1.105
	Edgeworth	0.696	1.205	2.646	0.588	0.993	1.156	0.303	0.464	0.844
n=40	Standard	0.358	0.597	1.195	0.273	0.454	0.909	0.229	0.381	0.763
	Bootstrap	0.035	0.158	1.292	0.025	0.116	0.930	0.020	0.095	0.765
	Edgeworth	0.547	0.932	1.962	0.440	0.711	0.849	0.200	0.325	0.623
n=50	Standard	0.265	0.442	0.884	0.202	0.336	0.673	0.169	0.282	0.564
	Bootstrap	0.025	0.115	0.939	0.018	0.084	0.675	0.015	0.070	0.560
	Edgeworth	0.443	0.748	1.539	0.234	0.355	0.643	0.149	0.246	0.481
n=60	Standard	0.206	0.344	0.687	0.157	0.261	0.523	0.132	0.219	0.439
	Bootstrap	0.019	0.091	0.725	0.014	0.066	0.530	0.012	0.054	0.435
	Edgeworth	0.367	0.617	1.253	0.162	0.264	0.504	0.118	0.195	0.384
n=70	Standard	0.166	0.277	0.554	0.126	0.211	0.422	0.106	0.177	0.354
	Bootstrap	0.015	0.072	0.573	0.011	0.053	0.430	0.010	0.044	0.350
	Edgeworth	0.310	0.519	1.044	0.127	0.209	0.408	0.096	0.159	0.316
n=80	Standard	0.138	0.229	0.459	0.105	0.175	0.349	0.088	0.146	0.293
	Bootstrap	0.013	0.059	0.474	0.010	0.044	0.347	0.008	0.036	0.294
	Edgeworth	0.265	0.441	0.880	0.104	0.172	0.338	0.080	0.133	0.265
n=90	Standard	0.116	0.194	0.388	0.089	0.148	0.295	0.074	0.124	0.248
	Bootstrap	0.011	0.050	0.399	0.008	0.037	0.299	0.007	0.031	0.248
	Edgeworth	0.227	0.377	0.740	0.087	0.145	0.287	0.068	0.114	0.226
n=100	Standard	0.100	0.167	0.334	0.076	0.127	0.254	0.064	0.107	0.213
	Bootstrap	0.009	0.043	0.340	0.007	0.032	0.254	0.006	0.027	0.211

TABLE 7.21: Student t_3 distribution, average length of confidence intervals

		α		
		0.01	**0.05**	**0.10**
n=10	Edgeworth	2.794	2.101	1.748
	Standard	10.323	8.621	7.832
	Bootstrap	110.645	18.942	9.780
n=20	Edgeworth	4.335	3.840	3.498
	Standard	5.097	3.877	3.261
	Bootstrap	11.697	5.896	3.722
n=30	Edgeworth	5.186	4.530	3.969
	Standard	3.022	2.299	1.929
	Bootstrap	3.297	2.280	1.854
n=40	Edgeworth	3.422	1.754	1.395
	Standard	2.072	1.578	1.322
	Bootstrap	2.130	1.521	1.295
n=50	Edgeworth	1.630	1.199	0.993
	Standard	1.532	1.165	0.978
	Bootstrap	1.592	1.140	0.926
n=60	Edgeworth	1.217	0.916	0.765
	Standard	1.190	0.906	0.760
	Bootstrap	1.198	0.884	0.713
n=70	Edgeworth	0.969	0.734	0.614
	Standard	0.960	0.730	0.613
	Bootstrap	0.959	0.705	0.585
n=80	Edgeworth	0.799	0.606	0.508
	Standard	0.795	0.605	0.508
	Bootstrap	0.810	0.587	0.490
n=90	Edgeworth	0.674	0.512	0.430
	Standard	0.672	0.511	0.429
	Bootstrap	0.672	0.501	0.411
n=100	Edgeworth	0.579	0.432	0.369
	Standard	0.578	0.440	0.369
	Bootstrap	0.585	0.432	0.358

TABLE 7.22: Laplace distribution, average length of confidence intervals

		α		
		0.01	**0.05**	**0.10**
n=10	Edgeworth	2.746	2.101	1.700
	Standard	0.893	7.240	6.351
	Bootstrap	91.491	19.062	8.508
n=20	Edgeworth	4.022	3.595	3.317
	Standard	4.057	3.091	2.623
	Bootstrap	6.434	3.530	2.826
n=30	Edgeworth	4.760	4.025	2.310
	Standard	2.462	1.873	1.571
	Bootstrap	2.761	1.917	1.588
n=40	Edgeworth	1.985	1.371	1.117
	Standard	1.691	1.287	1.080
	Bootstrap	1.814	1.310	1.077
n=50	Edgeworth	1.300	0.968	0.807
	Standard	1.250	0.951	0.798
	Bootstrap	1.310	0.958	0.801
n=60	Edgeworth	0.986	0.745	0.623
	Standard	0.972	0.740	0.621
	Bootstrap	1.016	0.742	0.614
n=70	Edgeworth	0.789	0.598	0.501
	Standard	0.783	0.596	0.500
	Bootstrap	0.810	0.599	0.495
n=80	Edgeworth	0.651	0.495	0.415
	Standard	0.649	0.494	0.414
	Bootstrap	0.670	0.497	0.416
n=90	Edgeworth	0.550	0.418	0.351
	Standard	0.549	0.418	0.350
	Bootstrap	0.569	0.419	0.348
n=100	Edgeworth	0.473	0.359	0.302
	Standard	0.472	0.359	0.301
	Bootstrap	0.484	0.363	0.301

TABLE 7.23: Normal distribution, Edgeworth intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.020	0.002	0.043	0.058	0.037	0.056	0.087	0.073	0.070
n=10	U-error	0.000	0.004	0.115	0.000	0.016	0.196	0.000	0.027	0.255
	Coverage	0.980	0.994	0.842	0.942	0.947	0.748	0.913	0.900	0.675
	L-error	0.023	0.042	0.074	0.058	0.069	0.112	0.079	0.098	0.131
n=20	U-error	0.000	0.000	0.000	0.001	0.000	0.000	0.016	0.000	0.000
	Coverage	0.977	0.958	0.926	0.941	0.931	0.888	0.905	0.902	0.869
	L-error	0.010	0.016	0.030	0.036	0.042	0.061	0.060	0.067	0.085
n=30	U-error	0.002	0.000	0.000	0.017	0.009	0.000	0.037	0.028	0.009
	Coverage	0.989	0.984	0.970	0.947	0.950	0.939	0.903	0.905	0.906
	L-error	0.006	0.010	0.015	0.032	0.034	0.039	0.054	0.059	0.668
n=40	U-error	0.003	0.002	0.000	0.022	0.017	0.010	0.045	0.038	0.033
	Coverage	0.991	0.988	0.984	0.946	0.949	0.951	0.901	0.903	0.900
	L-error	0.007	0.008	0.011	0.029	0.028	0.034	0.052	0.054	0.061
n=50	U-error	0.003	0.002	0.001	0.021	0.019	0.015	0.047	0.044	0.043
	Coverage	0.989	0.990	0.988	0.950	0.953	0.952	0.901	0.902	0.896
	L-error	0.005	0.008	0.007	0.026	0.030	0.032	0.049	0.057	0.058
n=60	U-error	0.005	0.005	0.002	0.025	0.023	0.020	0.050	0.051	0.042
	Coverage	0.990	0.988	0.991	0.949	0.947	0.947	0.901	0.893	0.900
	L-error	0.005	0.006	0.007	0.030	0.027	0.029	0.051	0.050	0.054
n=70	U-error	0.005	0.005	0.004	0.022	0.023	0.019	0.050	0.045	0.046
	Coverage	0.990	0.989	0.989	0.948	0.950	0.952	0.900	0.905	0.900
	L-error	0.004	0.006	0.008	0.027	0.027	0.029	0.049	0.050	0.053
n=80	U-error	0.005	0.004	0.003	0.024	0.020	0.021	0.051	0.047	0.051
	Coverage	0.990	0.990	0.989	0.949	0.953	0.950	0.900	0.903	0.897
	L-error	0.006	0.005	0.006	0.027	0.027	0.025	0.054	0.054	0.049
n=90	U-error	0.004	0.005	0.005	0.025	0.024	0.023	0.050	0.049	0.043
	Coverage	0.990	0.990	0.989	0.947	0.950	0.955	0.896	0.897	0.907
	L-error	0.005	0.004	0.006	0.022	0.028	0.027	0.051	0.050	0.053
n=100	U-error	0.005	0.005	0.005	0.027	0.023	0.023	0.050	0.051	0.046
	Coverage	0.991	0.991	0.989	0.951	0.950	0.950	0.899	0.899	0.901

TABLE 7.24: Normal distribution, standard intervals, miscoverage and coverage rates

		α								
		0.01			**0.05**			**0.10**		
		σ^2			σ^2			σ^2		
		0.09	**0.25**	**1.00**	**0.09**	**0.25**	**1.00**	**0.09**	**0.25**	**1.00**
	L-error	0.006	0.008	0.003	0.033	0.037	0.028	0.060	0.061	0.055
n=10	U-error	0.001	0.000	0.000	0.010	0.002	0.000	0.030	0.014	0.001
	Coverage	0.993	0.992	0.997	0.956	0.961	0.972	0.910	0.926	0.944
	L-error	0.007	0.009	0.016	0.031	0.030	0.045	0.056	0.065	0.067
n=20	U-error	0.003	0.003	0.000	0.018	0.016	0.009	0.042	0.042	0.028
	Coverage	0.990	0.988	0.984	0.951	0.954	0.946	0.902	0.893	0.905
	L-error	0.006	0.007	0.009	0.028	0.028	0.033	0.051	0.054	0.055
n=30	U-error	0.004	0.004	0.002	0.025	0.023	0.020	0.045	0.042	0.043
	Coverage	0.990	0.990	0.989	0.948	0.950	0.947	0.904	0.905	0.902
	L-error	0.004	0.006	0.005	0.028	0.028	0.027	0.050	0.052	0.053
n=40	U-error	0.004	0.005	0.004	0.024	0.024	0.023	0.050	0.048	0.049
	Coverage	0.991	0.989	0.991	0.948	0.948	0.950	0.901	0.901	0.899
	L-error	0.006	0.005	0.005	0.026	0.024	0.026	0.049	0.051	0.053
n=50	U-error	0.005	0.004	0.003	0.023	0.023	0.022	0.049	0.049	0.051
	Coverage	0.989	0.991	0.992	0.950	0.953	0.953	0.901	0.901	0.896
	L-error	0.005	0.006	0.005	0.025	0.027	0.026	0.047	0.054	0.054
n=60	U-error	0.006	0.006	0.005	0.026	0.026	0.026	0.052	0.053	0.046
	Coverage	0.990	0.989	0.990	0.950	0.947	0.949	0.901	0.894	0.900
	L-error	0.005	0.006	0.005	0.029	0.026	0.025	0.050	0.048	0.050
n=70	U-error	0.005	0.006	0.005	0.022	0.024	0.022	0.051	0.047	0.050
	Coverage	0.990	0.988	0.990	0.949	0.950	0.953	0.899	0.906	0.900
	L-error	0.004	0.005	0.007	0.027	0.026	0.026	0.048	0.486	0.050
n=80	U-error	0.005	0.005	0.004	0.024	0.021	0.023	0.052	0.050	0.053
	Coverage	0.991	0.991	0.990	0.950	0.953	0.950	0.900	0.902	0.897
	L-error	0.005	0.005	0.005	0.026	0.026	0.024	0.052	0.054	0.047
n=90	U-error	0.005	0.006	0.005	0.026	0.025	0.025	0.051	0.050	0.046
	Coverage	0.990	0.990	0.990	0.948	0.950	0.952	0.897	0.897	0.907
	L-error	0.005	0.004	0.006	0.022	0.027	0.026	0.050	0.050	0.052
n=100	U-error	0.005	0.006	0.006	0.027	0.023	0.025	0.051	0.052	0.047
	Coverage	0.991	0.991	0.989	0.951	0.950	0.950	0.899	0.899	0.901

TABLE 7.25: Normal distribution, bootstrap intervals, miscoverage and coverage rates

		α								
		0.01			**0.05**			**0.10**		
		σ^2			σ^2			σ^2		
		0.09	**0.25**	**1.00**	**0.09**	**0.25**	**1.00**	**0.09**	**0.25**	**1.00**
	L-error	0.000	0.002	0.003	0.023	0.028	0.029	0.053	0.044	0.046
n=10	U-error	0.000	0.000	0.000	0.021	0.015	0.005	0.055	0.043	0.023
	Coverage	1.000	0.998	0.997	0.956	0.957	0.966	0.892	0.913	0.931
	L-error	0.001	0.002	0.004	0.024	0.023	0.021	0.070	0.051	0.049
n=20	U-error	0.005	0.003	0.005	0.026	0.025	0.036	0.054	0.049	0.052
	Coverage	0.994	0.995	0.991	0.950	0.952	0.943	0.876	0.900	0.899
	L-error	0.005	0.006	0.002	0.029	0.029	0.023	0.051	0.054	0.065
n=30	U-error	0.007	0.004	0.006	0.027	0.023	0.033	0.050	0.053	0.054
	Coverage	0.988	0.990	0.992	0.944	0.948	0.944	0.899	0.893	0.881
	L-error	0.004	0.004	0.005	0.027	0.025	0.025	0.042	0.044	0.051
n=40	U-error	0.010	0.003	0.008	0.021	0.033	0.027	0.046	0.055	0.055
	Coverage	0.986	0.998	0.987	0.952	0.942	0.948	0.912	0.901	0.894
	L-error	0.007	0.004	0.004	0.024	0.028	0.024	0.051	0.047	0.058
n=50	U-error	0.011	0.012	0.006	0.020	0.031	0.026	0.051	0.056	0.040
	Coverage	0.982	0.984	0.990	0.956	0.941	0.950	0.898	0.897	0.902
	L-error	0.007	0.004	0.005	0.028	0.034	0.025	0.053	0.045	0.043
n=60	U-error	0.004	0.009	0.005	0.032	0.025	0.026	0.050	0.054	0.054
	Coverage	0.989	0.987	0.990	0.940	0.941	0.949	0.897	0.901	0.903
	L-error	0.004	0.007	0.005	0.019	0.025	0.029	0.062	0.062	0.055
n=70	U-error	0.004	0.009	0.002	0.033	0.022	0.026	0.057	0.055	0.044
	Coverage	0.992	0.984	0.993	0.948	0.953	0.945	0.881	0.883	0.901
	L-error	0.005	0.002	0.004	0.029	0.022	0.018	0.048	0.043	0.041
n=80	U-error	0.008	0.005	0.003	0.038	0.024	0.035	0.050	0.054	0.052
	Coverage	0.987	0.993	0.993	0.933	0.954	0.947	0.902	0.903	0.907
	L-error	0.002	0.004	0.007	0.033	0.025	0.024	0.055	0.047	0.042
n=90	U-error	0.009	0.011	0.001	0.034	0.018	0.027	0.058	0.053	0.045
	Coverage	0.989	0.985	0.992	0.933	0.957	0.949	0.887	0.900	0.913
	L-error	0.007	0.006	0.004	0.033	0.023	0.024	0.042	0.034	0.043
n=100	U-error	0.007	0.004	0.006	0.036	0.028	0.025	0.048	0.053	0.048
	Coverage	0.986	0.990	0.990	0.931	0.949	0.951	0.910	0.913	0.909

TABLE 7.26: Exponential distribution, Edgeworth intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
n=10	L-error	0.000	0.000	0.028	0.000	0.000	0.035	0.000	0.000	0.043
	U-error	0.000	0.000	0.032	0.000	0.000	0.058	0.000	0.000	0.086
	Coverage	1.000	1.000	0.940	1.000	1.000	0.907	1.000	1.000	0.870
n=20	L-error	0.000	0.000	0.002	0.000	0.000	0.004	0.000	0.000	0.004
	U-error	0.000	0.000	0.000	0.000	0.001	0.004	0.000	0.002	0.018
	Coverage	1.000	1.000	0.998	1.000	0.999	0.992	1.000	0.998	0.978
n=30	L-error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023
	U-error	0.000	0.000	0.011	0.000	0.000	0.031	0.000	0.001	0.049
	Coverage	1.000	1.000	0.989	1.000	1.000	0.969	1.000	0.999	0.928
n=40	L-error	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.021
	U-error	0.000	0.000	0.015	0.000	0.001	0.039	0.000	0.002	0.059
	Coverage	1.000	1.000	0.985	1.000	0.999	0.960	1.000	0.998	0.920
n=50	L-error	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.019
	U-error	0.000	0.000	0.016	0.000	0.000	0.042	0.000	0.002	0.068
	Coverage	1.000	1.000	0.984	1.000	1.000	0.957	1.000	0.998	0.913
n=60	L-error	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.021
	U-error	0.000	0.000	0.016	0.000	0.000	0.041	0.000	0.001	0.069
	Coverage	1.000	1.000	0.984	1.000	1.000	0.956	1.000	0.999	0.910
n=70	L-error	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.021
	U-error	0.000	0.000	0.020	0.000	0.000	0.039	0.000	0.001	0.065
	Coverage	1.000	1.000	0.981	1.000	1.000	0.958	1.000	0.999	0.914
n=80	L-error	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.022
	U-error	0.000	0.000	0.017	0.000	0.001	0.042	0.000	0.001	0.068
	Coverage	1.000	1.000	0.983	1.000	1.000	0.955	1.000	0.999	0.910
n=90	L-error	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.024
	U-error	0.000	0.000	0.020	0.000	0.000	0.043	0.000	0.001	0.067
	Coverage	1.000	1.000	0.981	1.000	1.000	0.952	1.000	0.999	0.909
n=100	L-error	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.025
	U-error	0.000	0.000	0.015	0.000	0.000	0.039	0.000	0.001	0.072
	Coverage	1.000	1.000	0.985	1.000	1.000	0.956	1.000	0.998	0.903

TABLE 7.27: Exponential distribution, standard intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.000	0.002	0.019	0.000	0.004	0.047	0.000	0.010	0.067
n=10	U-error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
	Coverage	1.000	0.999	0.981	1.000	0.996	0.952	1.000	0.990	0.931
	L-error	0.000	0.001	0.028	0.000	0.003	0.056	0.000	0.008	0.076
n=20	U-error	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.014
	Coverage	1.000	0.999	0.972	1.000	0.997	0.941	1.000	0.992	0.911
	L-error	0.000	0.000	0.016	0.000	0.001	0.038	0.000	0.003	0.064
n=30	U-error	0.000	0.000	0.002	0.000	0.000	0.015	0.000	0.000	0.033
	Coverage	1.000	1.000	0.982	1.000	0.999	0.947	1.000	0.997	0.903
	L-error	0.000	0.000	0.009	0.000	0.001	0.032	0.000	0.003	0.056
n=40	U-error	0.000	0.000	0.003	0.000	0.000	0.018	0.000	0.000	0.038
	Coverage	1.000	1.000	0.988	1.000	0.999	0.950	1.000	0.997	0.906
	L-error	0.000	0.000	0.009	0.000	0.000	0.034	0.000	0.002	0.057
n=50	U-error	0.000	0.000	0.003	0.000	0.000	0.019	0.000	0.000	0.045
	Coverage	1.000	1.000	0.988	1.000	1.000	0.947	1.000	0.998	0.899
	L-error	0.000	0.000	0.009	0.000	0.001	0.031	0.000	0.002	0.054
n=60	U-error	0.000	0.000	0.002	0.000	0.000	0.019	0.000	0.000	0.043
	Coverage	1.000	1.000	0.989	1.000	1.000	0.950	1.000	0.998	0.903
	L-error	0.000	0.000	0.010	0.000	0.000	0.031	0.000	0.002	0.053
n=70	U-error	0.000	0.000	0.004	0.000	0.000	0.019	0.000	0.000	0.041
	Coverage	1.000	1.000	0.986	1.000	1.000	0.950	1.000	0.998	0.906
	L-error	0.000	0.000	0.008	0.000	0.000	0.030	0.000	0.001	0.052
n=80	U-error	0.000	0.000	0.003	0.000	0.000	0.020	0.000	0.000	0.044
	Coverage	1.000	1.000	0.990	1.000	1.000	0.950	1.000	0.999	0.904
	L-error	0.000	0.000	0.007	0.000	0.000	0.032	0.000	0.002	0.053
n=90	U-error	0.000	0.000	0.004	0.000	0.000	0.021	0.000	0.000	0.045
	Coverage	1.000	1.000	0.989	1.000	1.000	0.946	1.000	0.998	0.903
	L-error	0.000	0.000	0.008	0.000	0.000	0.031	0.000	0.001	0.053
n=100	U-error	0.000	0.000	0.003	0.000	0.000	0.021	0.000	0.001	0.049
	Coverage	1.000	1.000	0.989	1.000	1.000	0.948	1.000	0.998	0.898

TABLE 7.28: Exponential distribution, bootstrap intervals, miscoverage and coverage rates

		α								
		0.01			0.05			0.10		
		σ^2			σ^2			σ^2		
		0.09	0.25	1.00	0.09	0.25	1.00	0.09	0.25	1.00
	L-error	0.002	0.001	0.009	0.027	0.022	0.031	0.067	0.061	0.067
n=10	U-error	0.000	0.000	0.000	0.006	0.004	0.005	0.027	0.038	0.021
	Coverage	0.998	0.999	0.991	0.967	0.974	0.964	0.906	0.901	0.912
	L-error	0.003	0.005	0.005	0.027	0.020	0.025	0.061	0.056	0.048
n=20	U-error	0.003	0.002	0.001	0.020	0.022	0.021	0.049	0.046	0.053
	Coverage	0.994	0.993	0.994	0.953	0.958	0.954	0.890	0.898	0.899
	L-error	0.001	0.003	0.002	0.020	0.031	0.042	0.061	0.058	0.050
n=30	U-error	0.001	0.000	0.001	0.026	0.020	0.023	0.052	0.058	0.060
	Coverage	0.998	0.997	0.997	0.954	0.949	0.935	0.887	0.884	0.890
	L-error	0.004	0.007	0.001	0.023	0.029	0.021	0.050	0.062	0.053
n=40	U-error	0.008	0.004	0.001	0.021	0.028	0.030	0.042	0.067	0.048
	Coverage	0.988	0.989	0.998	0.956	0.943	0.949	0.908	0.871	0.899
	L-error	0.002	0.004	0.004	0.026	0.027	0.020	0.066	0.051	0.056
n=50	U-error	0.005	0.002	0.006	0.022	0.023	0.018	0.047	0.065	0.041
	Coverage	0.993	0.994	0.990	0.952	0.950	0.962	0.887	0.884	0.903
	L-error	0.007	0.003	0.007	0.031	0.023	0.024	0.053	0.057	0.046
n=60	U-error	0.002	0.001	0.005	0.018	0.028	0.026	0.056	0.055	0.052
	Coverage	0.991	0.996	0.988	0.951	0.949	0.950	0.891	0.888	0.902
	L-error	0.002	0.006	0.002	0.021	0.026	0.022	0.060	0.049	0.054
n=70	U-error	0.006	0.004	0.002	0.028	0.025	0.025	0.051	0.041	0.054
	Coverage	0.992	0.990	0.996	0.951	0.949	0.953	0.889	0.910	0.892
	L-error	0.003	0.002	0.003	0.027	0.030	0.026	0.045	0.052	0.053
n=80	U-error	0.002	0.004	0.003	0.023	0.031	0.028	0.041	0.048	0.046
	Coverage	0.995	0.994	0.994	0.950	0.939	0.946	0.914	0.900	0.901
	L-error	0.002	0.009	0.006	0.026	0.022	0.017	0.064	0.051	0.049
n=90	U-error	0.001	0.009	0.004	0.027	0.023	0.019	0.049	0.033	0.057
	Coverage	0.997	0.982	0.990	0.947	0.955	0.964	0.887	0.916	0.894
	L-error	0.003	0.004	0.004	0.030	0.023	0.033	0.057	0.072	0.056
n=100	U-error	0.005	0.005	0.003	0.022	0.021	0.027	0.048	0.047	0.051
	Coverage	0.992	0.991	0.993	0.948	0.956	0.940	0.895	0.881	0.893

TABLE 7.29: Student t_3 distribution, Edgeworth intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.142	0.172	0.192
	U-error	0.156	0.241	0.287
	Coverage	0.702	0.588	0.521
n=20	L-error	0.095	0.119	0.136
	U-error	0.002	0.006	0.009
	Coverage	0.903	0.875	0.855
n=30	L-error	0.038	0.659	0.086
	U-error	0.000	0.000	0.000
	Coverage	0.962	0.934	0.914
n=40	L-error	0.023	0.043	0.069
	U-error	0.000	0.004	0.015
	Coverage	0.977	0.952	0.916
n=50	L-error	0.013	0.037	0.058
	U-error	0.002	0.012	0.029
	Coverage	0.986	0.951	0.913
n=60	L-error	0.013	0.031	0.057
	U-error	0.003	0.017	0.031
	Coverage	0.984	0.952	0.912
n=70	L-error	0.010	0.032	0.052
	U-error	0.005	0.019	0.038
	Coverage	0.985	0.950	0.910
n=80	L-error	0.009	0.028	0.047
	U-error	0.006	0.018	0.037
	Coverage	0.986	0.954	0.916
n=90	L-error	0.010	0.024	0.046
	U-error	0.006	0.020	0.040
	Coverage	0.984	0.956	0.914
n=100	L-error	0.010	0.025	0.046
	U-error	0.006	0.021	0.043
	Coverage	0.984	0.954	0.911

TABLE 7.30: Student t_3 distribution, standard intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.006	0.015	0.031
	U-error	0.000	0.001	0.001
	Coverage	0.994	0.984	0.968
n=20	L-error	0.019	0.042	0.062
	U-error	0.000	0.001	0.009
	Coverage	0.981	0.957	0.929
n=30	L-error	0.013	0.031	0.052
	U-error	0.004	0.014	0.030
	Coverage	0.983	0.955	0.918
n=40	L-error	0.010	0.027	0.049
	U-error	0.005	0.021	0.038
	Coverage	0.985	0.952	0.914
n=50	L-error	0.007	0.026	0.048
	U-error	0.006	0.021	0.041
	Coverage	0.987	0.953	0.910
n=60	L-error	0.008	0.025	0.049
	U-error	0.006	0.023	0.039
	Coverage	0.986	0.952	0.912
n=70	L-error	0.008	0.026	0.047
	U-error	0.007	0.023	0.043
	Coverage	0.985	0.951	0.910
n=80	L-error	0.007	0.245	0.043
	U-error	0.007	0.022	0.041
	Coverage	0.986	0.954	0.917
n=90	L-error	0.008	0.217	0.044
	U-error	0.007	0.024	0.043
	Coverage	0.985	0.955	0.913
n=100	L-error	0.009	0.023	0.044
	U-error	0.007	0.023	0.045
	Coverage	0.985	0.955	0.911

TABLE 7.31: Student t_3 distribution, bootstrap intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.003	0.016	0.059
	U-error	0.000	0.002	0.012
	Coverage	0.997	0.982	0.929
n=20	L-error	0.000	0.026	0.040
	U-error	0.000	0.026	0.052
	Coverage	1.000	0.948	0.908
n=30	L-error	0.004	0.020	0.050
	U-error	0.002	0.025	0.054
	Coverage	0.994	0.955	0.896
n=40	L-error	0.002	0.024	0.047
	U-error	0.006	0.024	0.054
	Coverage	0.992	0.952	0.899
n=50	L-error	0.004	0.027	0.043
	U-error	0.004	0.024	0.069
	Coverage	0.992	0.949	0.888
n=60	L-error	0.003	0.027	0.051
	U-error	0.004	0.030	0.066
	Coverage	0.993	0.943	0.883
n=70	L-error	0.002	0.020	0.051
	U-error	0.006	0.017	0.049
	Coverage	0.992	0.963	0.900
n=80	L-error	0.001	0.024	0.049
	U-error	0.003	0.026	0.049
	Coverage	0.996	0.950	0.902
n=90	L-error	0.004	0.019	0.061
	U-error	0.006	0.031	0.043
	Coverage	0.990	0.950	0.896
n=100	L-error	0.004	0.020	0.049
	U-error	0.007	0.018	0.051
	Coverage	0.989	0.962	0.900

TABLE 7.32: Laplace distribution, Edgeworth intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.111	0.140	0.149
	U-error	0.146	0.224	0.290
	Coverage	0.763	0.636	0.561
n=20	L-error	0.101	0.124	0.138
	U-error	0.000	0.001	0.003
	Coverage	0.899	0.876	0.859
n=30	L-error	0.041	0.007	0.095
	U-error	0.000	0.000	0.000
	Coverage	0.959	0.934	0.905
n=40	L-error	0.021	0.047	0.077
	U-error	0.000	0.006	0.026
	Coverage	0.979	0.948	0.897
n=50	L-error	0.017	0.039	0.065
	U-error	0.001	0.012	0.033
	Coverage	0.982	0.950	0.902
n=60	L-error	0.011	0.034	0.056
	U-error	0.002	0.019	0.042
	Coverage	0.988	0.947	0.903
n=70	L-error	0.009	0.032	0.055
	U-error	0.003	0.020	0.046
	Coverage	0.988	0.948	0.899
n=80	L-error	0.007	0.032	0.053
	U-error	0.004	0.020	0.045
	Coverage	0.989	0.949	0.903
n=90	L-error	0.007	0.030	0.056
	U-error	0.004	0.024	0.043
	Coverage	0.988	0.946	0.901
n=100	L-error	0.006	0.030	0.056
	U-error	0.004	0.025	0.046
	Coverage	0.990	0.944	0.898

TABLE 7.33: Laplace distribution, standard intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.007	0.022	0.045
	U-error	0.000	0.000	0.000
	Coverage	0.993	0.978	0.955
n=20	L-error	0.020	0.046	0.068
	U-error	0.000	0.005	0.022
	Coverage	0.980	0.949	0.910
n=30	L-error	0.012	0.032	0.060
	U-error	0.005	0.016	0.040
	Coverage	0.983	0.952	0.900
n=40	L-error	0.008	0.028	0.057
	U-error	0.003	0.024	0.050
	Coverage	0.989	0.948	0.894
n=50	L-error	0.009	0.029	0.054
	U-error	0.005	0.023	0.046
	Coverage	0.986	0.948	0.900
n=60	L-error	0.007	0.027	0.048
	U-error	0.003	0.027	0.051
	Coverage	0.989	0.946	0.901
n=70	L-error	0.006	0.026	0.050
	U-error	0.006	0.024	0.052
	Coverage	0.988	0.949	0.898
n=80	L-error	0.005	0.028	0.050
	U-error	0.005	0.024	0.049
	Coverage	0.990	0.948	0.901
n=90	L-error	0.006	0.027	0.052
	U-error	0.006	0.027	0.046
	Coverage	0.989	0.946	0.901
n=100	L-error	0.005	0.029	0.054
	U-error	0.005	0.027	0.048
	Coverage	0.990	0.944	0.898

TABLE 7.34: Laplace distribution, bootstrap intervals, miscoverage and coverage rates

		α		
		0.01	**0.05**	**0.10**
n=10	L-error	0.002	0.021	0.055
	U-error	0.001	0.002	0.014
	Coverage	0.997	0.977	0.931
n=20	L-error	0.005	0.033	0.056
	U-error	0.004	0.022	0.045
	Coverage	0.991	0.945	0.899
n=30	L-error	0.002	0.019	0.051
	U-error	0.005	0.024	0.051
	Coverage	0.993	0.957	0.898
n=40	L-error	0.005	0.018	0.057
	U-error	0.004	0.041	0.050
	Coverage	0.991	0.941	0.893
n=50	L-error	0.003	0.025	0.048
	U-error	0.001	0.025	0.039
	Coverage	0.996	0.950	0.913
n=60	L-error	0.002	0.024	0.039
	U-error	0.008	0.021	0.053
	Coverage	0.990	0.955	0.908
n=70	L-error	0.004	0.024	0.034
	U-error	0.007	0.026	0.050
	Coverage	0.989	0.950	0.916
n=80	L-error	0.004	0.024	0.049
	U-error	0.002	0.032	0.056
	Coverage	0.994	0.944	0.895
n=90	L-error	0.003	0.025	0.062
	U-error	0.002	0.031	0.049
	Coverage	0.995	0.944	0.889
n=100	L-error	0.005	0.031	0.047
	U-error	0.005	0.023	0.046
	Coverage	0.990	0.946	0.907

7.5 CONCLUSION

This thesis should, in the main, be viewed as complementary to the work of McCullagh & Pregibon (1987) in the sense that index notation is applied to a regression problem which the traditional matrix notation is not able to handle. In this connection, we wish to feel that, along the way, we have made some modest contributions in several respects.

Firstly, our review of MINO has included the clarification of some aspects of it which hitherto have sometimes been sources of much ambiguity and occasional statement of error. It is hoped that Chapter three and Section 3 of Chapter four now offer an improved MINO package which beginners would find useful. Secondly, we also would like to think that we have extended the theory of cumulants of nonlinear random functions by focussing on these in the context of regression as well as Edgeworth expansions. Thirdly, we have proposed a very simple and very transparent way of approaching the determination of the cumulants of multivariate polynomial expansions. Our procedure has also turned out to be flexible as evidenced by its easy adaptation when we determined the cumulants of univariate polynomials. Fourthly, the recent delta method of demonstrating skewness reduction by Hall has been given a natural alternative.

Lastly, we have introduced a new confidence interval for θ which has been shown by simulation studies to perform reasonably well in a variety of small-sample situations.

REFERENCES

ABRAMOVITCH, L. & SINGH, K. (1985). Edgeworth-corrected pivotal statistics and the bootstrap. *Ann. Statist.*, Vol **13**, No. 1, pp. 116–32.

BARNARD, G.A. (1974). Can we all agree on what we mean by estimation. *Utilitas Mathematica*, Vol. **6**, pp. 3–22.

BARNDORFF-NIELSEN, O.E. & COX, D.R. (1989). *Asymptotic techniques for use in statistics.* Chapman & Hall. London.

BERNDT, G. (1957). The regions of unimodality and positivity in the abbreviated Edgeworth and Gram-Charlier series. *J. Am. Stat. Ass.* **52**, pp. 253–6.

BICKEL, P.J. & DOKSUM, K.A. (1977). *Mathematical Statistics: Basic ideas and selected topics.* Holden-Day, Inc. San Francisco.

BILLINGSLEY, P. (1979). *Probability measure.* Wiley, N.Y.

BRADLEY, R.A. & SRIVASTAVA, S.S. (1979). *Correlation in polynomial regression.* Amer. Statist. Vol. **33**, No. 1, pp 11-14.

BRILLINGER, D.R. (1981). *Time Series: data analysis and Theory.* Expanded Edition. McGraw-Hill, Inc. N.Y.

BUNKE, H. & BUNKE, O. (1986). *Statistical inference in linear models,* Vol. **1**, Wiley, N.Y.

CRAMÉR, H. (1946). *Mathematical methods of statistics.* Princeton Univ. Press. Princeton.

DAVID, F.N.& KENDALL, M.G. (1949). Tables of symmetric functions, Part I. *Biometrika*, **36**, pp. 431–49.

DELURY, D.B. (1947). On the estimation of biological populations. *Biometrics*, **3**, pp. 145–67.

DURAND, D. (1954). Joint Confidence Regions for Multiple Regressions Coefficients. *J. Am. Stat. Assoc.*, **49**, pp. 130–46.

EDGEWORTH, F.Y. (1905). The law of error. Proc. *Cambridge Philos. Soc.*, **20**, pp. 36–65.

EFRON, B. & GONG, G. (1983). A leisurely look at the bootstrap, the Jackknife, and Cross-Validation. *Am. Statist.* Vol. **37**, 1, pp. 36–48.

FIELLER, E.C. (1954). Some problems in interval estimation. *J.R. Statist. Soc.*, B16, pp. 175–85.

FINNEY, D.J. (1978). *Statistical methods in biological assay.* 3rd ed. Charles Griffin & Co. London.

HALL, P. (1992a). On the removal of skewness by transformation. *J.R. Statist. Soc.* B, 54, No. 1, pp. 221–228.

HALL, P. (1992b). *The bootstrap and Edgeworth Expansion*. Springer-Verlag.

HAMPEL, F.R., RONCHETTI, E.M., ROUSSEEUW, P.J., & STAHEL, W.A. (1986). *Robust Statistics, the approach based on Influence Functions*. Wiley. N.Y.

HILL, G.W. & DAVIS, A.W. (1968). Generalised asymptotic expansions of Cornish-Fisher Type. *Ann. Math. Statist.*, Vol. **39**, No. 4, pp. 1264–73.

HINKLEY, D.V. (1971). Inference in two-phase regression. *J. Am. Stat. Assoc.*, Vol. **66**, No. 33, pp. 736–43.

HOLM, S. (1993). Abstract bootstrap confidence intervals in Linear Models. *Scand. J. Statist.*, **20**, pp. 157–170.

JAMES, G.S. & MAYNE, A.J. (1962). Cumulants of functions of random variables. *Sankhyā*, **24**, pp. 47–54.

JAMES, G.S. (1958). Cumulants of a transformed variate. *Biometrika*, **42**, pp. 529–31.

JOHNSON, N.J. (1978). Modified t-tests and confidence intervals for asymmetrical populations. *J. Amer. Statist. Assoc.*, **73**, pp. 536–544.

JOHNSON, N.L. & KOTZ, S. (1970). *Continuous univariate distributions – 1*. Wiley. N.Y.

JOHNSON, N.L. & KOTZ, S. (1970). *Continuous univariate distributions – 2*. Wiley. N.Y.

JUDGE, G.G., GRIFFITHS, W.E., HILL, R.C., LÜTKEPOHL, H. & LEE, T.C. (1985). *The Theory and Practice of Econometrics*. 2nd Ed. Wiley. N.Y.

KAPLAN, E.L. (1952). Tensor notation and the sampling cumulants of k-statistics. *Biometrika*, **39**, pp. 319–23.

KASS, R.E. & VOS, P.W. (1997). *Geometric foundations of asymptotic inference*. Wiley. N.Y.

KENDALL, M.G. (1948). *Advanced Theory of Statistics*. Vol. 1. 3^{rd} Ed. Charles Griffith and Company Ltd. London.

KENDALL, M. & STUART, A. (1973). *The Advanced Theory of Statistics*. Vol. 2. Charles Griffith and Company Ltd. London.

KENDALL, M. & STUART, A. (1977). *The Advanced Theory of Statistics*. Vol. 1. 4^{th} ed. Charles Griffith and Company Ltd. London.

LAHA, R.G. & ROHATGI, V.K. (1979). *Probability Theory*. Wiley. N.Y.

LASS, H. (1950). *Vector and tensor analysis*. McGraw-Hill. N.Y.

LEONOV, V.P. & SHIRYAEV, A.N. (1959). On a method of calculation of semi-invariants. *Theory Prob. Applic.*, **4**, pp. 319–29.

LUGANNANI, R. & RICE, S. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables. *Adv. Appl. Probab.*, **12**, pp. 475–490.

MANDEL, J. (1958). A note on confidence intervals in regression problems. *Ann. Math. Statist.*, **29**, pp. 903–7.

McCULLAGH, P. & PREGIBON, D. (1987). k-statistics and dispersion effects in regression. *Ann. Statist.*, **15**, pp. 202–219.

McCULLAGH, P. (1984). Tensor notation and cumulants of polynomials. *Biometrika*, **71**, 3, pp. 461–2.

McCULLAGH, P. (1987). *Tensor Methods in Statistics*. Chapman & Hall. London.

MILLER, R.G. (1974). An unbalanced jackknife. *Ann. Statist.* **2**, pp. 880–91.

MONTGOMERY, D.C. & PECK, E.A. (1982). *Introduction to linear regression analysis*. Wiley. N.Y.

MORAN, P.A.P. (1968). *An introduction to probability theory*. Clarendon Press. Oxford.

NETER, J. & WASSERMAN, W. (1974). *Applied linear statistical models*. Richard D. Irwin, Inc. Homewood. Illinois.

PACE, L. & SALVAN, A. (1997). *Principles of Statistical Inference: from a neo-Fisherian perspective*. World Scientific. Singapore.

PEARSON, E.S. & HARTLEY, H.O. (1976). Biometrika tables for statisticians. *Biometrika* Trust 1971.

PEARSON, E.S. & TUKEY, J.W. (1965). Approximate means and standard deviations based on distances between percentage points of frequency curves. *Biometrika*, **52**, 3 and 5, pp. 533–546.

PEARSON, E.S. (1963). Some problems arising in approximating probability distributions, using moments. *Biometrika*, **50**, 1 and 2, pp. 95–112.

PUKELSHEIM, F. (1980). Multilinear estimation of skewness and kurtosis in linear models. *Metrika*, **27**, pp. 103–113.

SCHEFFÉ, H. (1973). A statistical theory of calibration. *Ann. Statist.*, Vol. **1**, No. 1, pp. 1–37.

SEBER, G.A.F. (1977). *Linear regression analysis*. Wiley. N.Y.

SKOVGAARD, I.M. (1981). Transformation of an Edgeworth expansion by a sequence of Functions. *Scand. J. Statist.*, **8**, pp. 207–17.

SPRENT, P. (1961). Some hypotheses concerning two phase regression lines. *Biometrics*, **17**, pp. 634–49.

TINGLEY, M.A. (1992). Small-sample intervals for regression. *Canad. J. Statist.*, Vol. **20**, 3, pp. 271–80.

URBAN HJORTH, J.S.U. (1994). *Computer intensive statistical methods: validation, model selection, and bootstrap*. Chapman & Hall. London.

WEBER,N.C. & WELSH, A.H. (1983). Jackknifing the general linear model. *Austral. J. Statist.*, **25**(3), pp. 425–36.

WHITTAKER, E.T. & WATSON, G.N. (1927). *A course of modern analysis.* Cambridge Univ. Press. Cambridge.

WU, C.F. (1981). Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.*, Vol. **9,** No.3, pp. 501-15.

WU, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.*, Vol. **14**, pp. 1261–95.

WILLIAMS, E.J. (1969). Regression methods in calibration problems. *Bull. ISI*, Vol. **43**, pp. 17–28.

APPENDICES

APPENDIX A.1

```
C
C     THIS PROGRAM CONSIDERS THE EDGEWORTH BASED CONFIDENCE INTERVAL FOR
C     THE ABSCISSA OF THE TURNING POINT OF THE MODEL IN EXAMPLE 1.3,
C     AS WELL AS THE STANDARD CONFIDENCE INTERVAL. THE DESIGN-MATRIX IS READ
C     FROM A FILE, xdata.d, AND VALUES OF THE RESPONSE ARE GENERATED BY MEANS
C     OF SIMULATION. THE FOLLOWING QUANTITIES FOR BOTH INTERVALS ARE ESTIMATED:
C       1. THE LEFT MISCOVERAGE RATE FOR THE EDGEWORTH INTERVAL (ERRORLE)
C       2. THE RIGHT MISCOVERAGE RATE FOR THE EDGEWORTH INTERVAL (ERRORRE)
C       3. THE TOTAL MISCOVERAGE RATE FOR THE EDGEWORTH INTERVAL (AMISCOVERE)
C       4. THE COVERAGE PROBABILITY OF THE EDGEWORTH INTERVAL (COVERE)
C       5. THE EXPECTED LENGTH OF THE EDGEWORTH INTERVAL (ALENGTHE)
C       6. THE LEFT MISCOVERAGE RATE FOR THE STANDARD INTERVAL (ERRORLS)
C       7. THE RIGHT MISCOVERAGE RATE FOR THE STANDARD INTERVAL (ERRORRS)
C       8. THE TOTAL MISCOVERAGE RATE FOR THE STANDARD INTERVAL (AMISCOVERS)
C       9. THE COVERAGE PROBABILITY OF THE STANDARD INTERVAL (COVERS)
C      10. THE EXPECTED LENGTH OF THE STANDARD INTERVAL (ALENGTHS)
C
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
C
C     THE NEXT STATEMENT SPECIFIES THE NUMBER OF COLUMNS IN THE
C     DESIGN MATRIX (IP), THE SAMPLE SIZE (NN), 1 MINUS THE COVERAGE
C      (ALPHA), THE VARIANCE (SIGMA), AND THE NUMBER OF SIMULATION REPETITIONS
(MCH).
C
      PARAMETER (IP=3,NN=10,IPP1=IP+1,ALPHA=0.1D0,SIGMA=1.0D0,MCH=5000)
C
C     THE REQUIRED MATRICES AND VECTORS ARE DIMENSIONED.
C
      DIMENSION XX(NN,IP),Y(NN)
      DIMENSION Z(NN,IP),ZTZ(IP,IP),ZTZI(IP,IP)
      DIMENSION EL(IP,NN),EMOM(5)
      DIMENSION XTX(IP,IP),XTXB(IP,IP)
      DIMENSION BETA(IP),F1(10),F2(5,5),ERR(NN),T(NN)
      DIMENSION IVECNU(5)

      CHARACTER*70 FILEIN,FILEOUT1,FILEOUT2
      FILEIN='xdata.d'
      FILEOUT1='/home/sjst/KAMAN/SIMUL/NORMAL/normale.d'
      FILEOUT2='/home/sjst/KAMAN/SIMUL/NORMAL/normals.d'
      ZLOWER=DNORIN(ALPHA/2.0D0)
      ZUPPER=DNORIN(1.0D0-ALPHA/2.0D0)

C
C     xdata.d CONTAINS THE ORIGINAL X-MATRIX
C
C     READ THE ORIGINAL X-MATRIX FROM FILE
C
      OPEN(1,FILE=FILEIN,STATUS='OLD')
      DO 3 I=1,NN
      READ(1,*) (XX(I,J),J=1,IP)
3     CONTINUE
      CLOSE(1)
C
C     TRANSFORM THE X-MATRIX TO A Z-MATRIX, WHERE Z=X-XBAR
C
      S=0.0D0
      DO 5 I=1,NN
      S=S+XX(I,2)
5     CONTINUE
```

```
      S=S/NN
      XMEAN=S
      DO 7 I=1,NN
      Z(I,1)=1.0D0
      Z(I,2)=XX(I,2)-S
      Z(I,3)=Z(I,2)**2.0D0
7     CONTINUE
C
C     DO THE NECESSARY CALCULATIONS AND FIND THE L-MATRIX, WHERE
C     L=(Z-TRANSPOSE Z)-INVERSE TIMES Z-TRANSPOSE
C
      NRA=NN
      NCA=IP
      LDA=NN
      NBB=IP
      LDB=IP
      CALL DMXTXF(NRA,NCA,Z,LDA,NBB,ZTZ,LDB)
      CALL DLINDS(IP,ZTZ,IP,ZTZI,IP)
      DO 15 I=1,IP
      DO 14 J=1,NN
      S=0.0D0
      DO 12 K=1,IP
      S=S+ZTZI(I,K)*Z(J,K)
12    CONTINUE
      EL(I,J)=S
14    CONTINUE
15    CONTINUE
C
C     SET UP THE FIRST THREE CENTRAL MOMENTS OF THE DISTRIBUTION
C     OF THE ERROR-TERM
C
      EMOM(1)=0.0D0
      EMOM(2)=SIGMA
      EMOM(3)=0.0D0
C
C     INITIALISE SIMULATION COUNTERS
C
      COVERE=0.0D0
      ALENGTHE=0.0D0
      ERRORLE=0.0D0
      ERRORRE=0.0D0
      COVERS=0.0D0
      ALENGTHS=0.0D0
      ERRORLS=0.0D0
      ERRORRS=0.0D0
C
C     START THE SIMULATION LOOP
C
      DO 300 MC=1,MCH
      write(6,*) MC
C
C     USE SIMULATION TO GENERATE THE Y-VALUES FROM THE APPROPRIATE
C     DISTRIBUTION
C
      CALL DRNNOR(NN,ERR)
      DO 16 I=1,NN
      T(I)=I
16    CONTINUE
      DO 17 I=1,NN
      Y(I)=10.0D0-T(I)+0.1D0*T(I)*T(I)+ERR(I)
17    CONTINUE
```

```
C
C     FIND THE LEAST SQUARES ESTIMATES OF THE BETA-COEFFICIENTS,AND
C     CALCULATE THETA-HAT
C
      NOBS=NN
      LDX=NN
      INTCEP=0
      NIND=IP
      CALL DRLSE(NOBS,Y,NIND,Z,LDX,INTCEP,BETA,SST,SSE)
      THETAHAT=-BETA(2)/(2.0D0*BETA(3))+XMEAN
C
C     EXPRESSIONS FOR THE REQUIRED FIRST AND SECOND DERIVATIVES OF THE
C     THETA-HAT FUNCTION
C
      F1(2)=-0.5D0/BETA(3)
      F1(3)=0.5D0*BETA(2)/(BETA(3)**2.0D0)
      F2(2,2)=0.0D0
      F2(3,3)=-BETA(2)/(BETA(3)**3.0D0)
      F2(2,3)=0.5D0/(BETA(3)**2.0D0)
      F2(3,2)=F2(2,3)
C
C     CALCULATE SQRT(ASYMPTOTIC KAPPA-SUB-2)
C
      Q1=0.0D0
      DO 25 IS=2,3
      DO 24 IR=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IS
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL)
      Q1=Q1+F1(IR)*F1(IS)*BCUMUL
24    CONTINUE
25    CONTINUE
      Q1=DSQRT(NN*Q1)
C
C     NOW CALCULATE THE LOWER AND THE UPPER LIMIT OF THE
C     STANDARD CONFIDENCE INTERVAL
C
      SCONL=THETAHAT+ZLOWER*Q1/(DSQRT(1.0D0*NN))
      SCONU=THETAHAT+ZUPPER*Q1/(DSQRT(1.0D0*NN))
C
C     CALCULATE LAMBDA-1
C
      ALAM1=0.0D0
      DO 35 IS=2,3
      DO 34 IR=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IS
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL)
      ALAM1=ALAM1+F2(IR,IS)*BCUMUL
34    CONTINUE
35    CONTINUE
      ALAM1=NN*0.5D0*ALAM1/Q1
C
C     CALCULATE LAMBDA-3
C
C     FIRST CALCULATE a-sub-3,n
C
      S0=0.0D0
      DO 45 IR=2,3
```

```
      DO 44 IS=2,3
      DO 43 IT=2,3
      NU=3
      IVECNU(1)=IR
      IVECNU(2)=IS
      IVECNU(3)=IT
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL)
      S0=S0+F1(IR)*F1(IS)*F1(IT)*BCUMUL
43    CONTINUE
44    CONTINUE
45    CONTINUE

      S1=0.0D0
      DO 55 IR=2,3
      DO 54 IS=2,3
      DO 53 IT=2,3
      DO 52 IU=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IT
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL1)
      IVECNU(1)=IS
      IVECNU(2)=IU
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL2)
      S1=S1+F1(IR)*F1(IS)*F2(IT,IU)*BCUMUL1*BCUMUL2
52    CONTINUE
53    CONTINUE
54    CONTINUE
55    CONTINUE
      S1=0.5D0*S1

      S2=0.0D0
      DO 65 IR=2,3
      DO 64 IS=2,3
      DO 63 IT=2,3
      DO 62 IU=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IU
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL1)
      IVECNU(1)=IS
      IVECNU(2)=IT
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL2)
      S2=S2+F1(IR)*F1(IS)*F2(IT,IU)*BCUMUL1*BCUMUL2
62    CONTINUE
63    CONTINUE
64    CONTINUE
65    CONTINUE
      S2=0.5D0*S2

      S3=0.0D0
      DO 75 IR=2,3
      DO 74 IS=2,3
      DO 73 IT=2,3
      DO 72 IU=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IS
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL1)
      IVECNU(1)=IT
      IVECNU(2)=IU
```

```
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL2)
      S3=S3+F1(IR)*F1(IU)*F2(IS,IT)*BCUMUL1*BCUMUL2
72    CONTINUE
73    CONTINUE
74    CONTINUE
75    CONTINUE
      S3=0.5D0*S3

      S4=0.0D0
      DO 85 IR=2,3
      DO 84 IS=2,3
      DO 83 IT=2,3
      DO 82 IU=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IT
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL1)
      IVECNU(1)=IS
      IVECNU(2)=IU
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL2)
      S4=S4+F1(IR)*F1(IU)*F2(IS,IT)*BCUMUL1*BCUMUL2
82    CONTINUE
83    CONTINUE
84    CONTINUE
85    CONTINUE
      S4=0.5D0*S4

      S5=0.0D0
      DO 95 IR=2,3
      DO 94 IS=2,3
      DO 93 IT=2,3
      DO 92 IU=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IT
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL1)
      IVECNU(1)=IS
      IVECNU(2)=IU
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL2)
      S5=S5+F1(IT)*F1(IU)*F2(IR,IS)*BCUMUL1*BCUMUL2
92    CONTINUE
93    CONTINUE
94    CONTINUE
95    CONTINUE
      S5=0.5D0*S5

      S6=0.0D0
      DO 105 IR=2,3
      DO 104 IS=2,3
      DO 103 IT=2,3
      DO 102 IU=2,3
      NU=2
      IVECNU(1)=IR
      IVECNU(2)=IU
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL1)
      IVECNU(1)=IS
      IVECNU(2)=IT
      CALL CUMUL(NN,EL,EMOM,NU,IVECNU,BCUMUL2)
      S6=S6+F1(IT)*F1(IU)*F2(IR,IS)*BCUMUL1*BCUMUL2
102   CONTINUE
103   CONTINUE
```

```
104   CONTINUE
105   CONTINUE
      S6=0.5D0*S6
      A3N=S0+S1+S2+S3+S4+S5+S6
      ALAM3=NN*NN*A3N/(Q1*Q1*Q1)
C
C     CALCULATE THE LOWER AND THE UPPER CONFIDENCE LIMITS
C
      SQN=DSQRT(1.0D0*NN)
      FRACTION=1.0D0/3.0D0
      TERM1=(3.0D0*ALAM3*ALAM3/NN)-3.0D0*ALAM3*ZUPPER/SQN+1.0D0
      TERM2=(3.0D0*ALAM3*ALAM3/NN)-3.0D0*ALAM3*ZLOWER/SQN+1.0D0
      IF (TERM1.LT.0.0D0) TERM1=-((-TERM1)**FRACTION)
      IF (TERM2.LT.0.0D0) TERM2=-((-TERM2)**FRACTION)
      IF (TERM1.GT.0.0D0) TERM1=TERM1**FRACTION
      IF (TERM2.GT.0.0D0) TERM2=TERM2**FRACTION
      TERM3=Q1/DSQRT(1.0D0*NN)
      TERM4=ALAM1/DSQRT(1.0D0*NN)
      TERM5=DSQRT(1.0D0*NN)/ALAM3
C
C     NOW CALCULATE THE LOWER AND THE UPPER LIMIT OF THE
C     EDGEWORTH BASED CONFIDENCE INTERVAL
C
      CONL=THETAHAT-TERM3*(TERM4+TERM5*(1.0D0-TERM1))
      CONU=THETAHAT-TERM3*(TERM4+TERM5*(1.0D0-TERM2))
C
C     NOW CALCULATE THE LENGTH OF THE EDGEWORTH BASED CONFIDENCE
C     INTERVAL AND THE LENGTH OF THE STANDARD CONFIDENCE INTERVAL
C
      ALENGTHE=ALENGTHE+(CONU-CONL)
      ALENGTHS=ALENGTHS+(SCONU-SCONL)
C
C     CALCULATE THE RESPECTIVE COVERAGE PROBABILITIES
C
      IF (CONL.GT.5.0D0) ERRORLE=ERRORLE+1.0D0
      IF (CONU.LT.5.0D0) ERRORRE=ERRORRE+1.0D0
      IF (SCONL.GT.5.0D0) ERRORLS=ERRORLS+1.0D0
      IF (SCONU.LT.5.0D0) ERRORRS=ERRORRS+1.0D0
      IF ((CONL.LT.5.0D0).AND.(5.0D0.LT.CONU)) COVERE
     & =COVERE+1.0D0
      IF ((SCONL.LT.5.0D0).AND.(5.0D0.LT.SCONU)) COVERS
     & =COVERS+1.0D0

300   CONTINUE
      COVERE=COVERE/MCH
      ALENGTHE=ALENGTHE/MCH
      ERRORLE=ERRORLE/MCH
      ERRORRE=ERRORRE/MCH
      COVERS=COVERS/MCH
      ALENGTHS=ALENGTHS/MCH
      ERRORLS=ERRORLS/MCH
      ERRORRS=ERRORRS/MCH
      AMISCOVERE=ERRORLE+ERRORRE
      AMISCOVERS=ERRORLS+ERRORRS
C
C      OPEN(1,FILE=FILEOUT1,STATUS='NEW')
C      CLOSE(1)
C      OPEN(1,FILE=FILEOUT2,STATUS='NEW')
C      CLOSE(1)
C
C     WRITE THE RESULTS TO OUTPUT FILES
```

```
C
      OPEN(1,FILE=FILEOUT1,STATUS='OLD')
      OPEN(2,FILE=FILEOUT2,STATUS='OLD')
      WRITE(1,*)
      WRITE(2,*)
      WRITE(1,605) NN,SIGMA,ALPHA,'  EQUI-SPACED CASE','  FIRST G'
      WRITE(1,608) 'LEFT-ERROR','RIGHT-ERROR','ALL ERROR','COVERAGE'
     &,'EXP. LENGTH'
      WRITE(1,610) ERRORLE,ERRORRE,AMISCOVERE,COVERE,ALENGTHE
      WRITE(2,605) NN,SIGMA,ALPHA,'  EQUI-SPACED CASE','  FIRST G'
      WRITE(2,608) 'LEFT-ERROR','RIGHT-ERROR','ALL ERROR','COVERAGE'
     &,'EXP. LENGTH'
      WRITE(2,610) ERRORLS,ERRORRS,AMISCOVERS,COVERS,ALENGTHS
      CLOSE(1)
      CLOSE(2)
605   FORMAT(I5,2X,2(F12.6,2X),A20,A9)
608   FORMAT(A10,1X,A11,3X,A9,4X,A8,4X,A11)
610   FORMAT(4(F10.5,2X),2X,F10.5)

1000  STOP
      END



      SUBROUTINE CUMUL(N,EL,EMOM,NU,IVECNU,BCUMUL)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      PARAMETER (IP=3,NN=10,IPP1=IP+1,ALPHA=0.1D0,SIGMA=1.0D0,MCH=5000)
      DIMENSION EL(IP,NN),EMOM(5),IVECNU(5)
      IF (NU.EQ.2) THEN
          IR=IVECNU(1)
          IS=IVECNU(2)
          RHO2=EMOM(2)
          S=0.0D0
          DO 5 I=1,N
          S=S+EL(IR,I)*EL(IS,I)
5         CONTINUE
          BCUMUL=RHO2*S
      ENDIF
      IF (NU.EQ.3) THEN
          IR=IVECNU(1)
          IS=IVECNU(2)
          IT=IVECNU(3)
          RHO3=EMOM(3)
          S=0.0D0
          DO 15 I=1,N
          S=S+EL(IR,I)*EL(IS,I)*EL(IT,I)
15        CONTINUE
          BCUMUL=RHO3*S
      ENDIF
      IF (NU.EQ.4) THEN
          IR=IVECNU(1)
          IS=IVECNU(2)
          IT=IVECNU(3)
          IU=IVECNU(4)
          RHO4=EMOM(4)
          S=0.0D0
          DO 25 I=1,N
          S=S+EL(IR,I)*EL(IS,I)*EL(IT,I)*EL(IU,I)
25        CONTINUE
          BCUMUL=RHO4*S
      ENDIF
```

```
      IF (NU.EQ.5) THEN
          IR=IVECNU(1)
          IS=IVECNU(2)
          IT=IVECNU(3)
          IU=IVECNU(4)
          IV=IVECNU(5)
          RHO5=EMOM(5)
          S=0.0D0
          DO 35 I=1,N
          S=S+EL(IR,I)*EL(IS,I)*EL(IT,I)*EL(IU,I)*EL(IV,I)
35        CONTINUE
          BCUMUL=RHO5*S
      ENDIF
      RETURN
      END
```

APPENDIX A.2

```
C
C     THIS PROGRAM CONSIDERS THE BOOTSTRAP CONFIDENCE INTERVAL, DISCUSSED
C     BY HJORTH (1994), EXAMPLE 6.1, P.193, FOR THE ABSCISSA OF THE TURNING
C     POINT OF THE MODEL IN EXAMPLE 1.3. THE DESIGN-MATRIX IS READ
C     FROM A FILE, xdata.d, AND VALUES OF THE RESPONSE ARE GENERATED BY MEANS
C     OF SIMULATION. THE FOLLOWING QUANTITIES ARE ESTIMATED:
C       1. THE LEFT MISCOVERAGE RATE FOR THE BOOTSTRAP INTERVAL (ERRORLB)
C       2. THE RIGHT MISCOVERAGE RATE FOR THE BOOTSTRAP INTERVAL (ERRORRB)
C       3. THE TOTAL MISCOVERAGE RATE FOR THE BOOTSTRAP INTERVAL (AMISCOVERB)
C       4. THE COVERAGE PROBABILITY OF THE BOOTSTRAP INTERVAL (COVERB)
C       5. THE EXPECTED LENGTH OF THE BOOTSTRAP INTERVAL (ALENGTHB)
C
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
C
C     THE NEXT STATEMENT SPECIFIES THE NUMBER OF COLUMNS IN THE
C     DESIGN MATRIX (IP), THE SAMPLE SIZE (NN), 1 MINUS THE COVERAGE
C     (ALPHA), THE NUMBER OF BOOTSTRAP REPETITIONS (NBOOT), AND THE
C     NUMBER OF SIMULATION REPETITIONS (MCH).
C
      PARAMETER (IP=2,NN=100,IPP1=IP+1,ALPHA=0.1D0,NBOOT=1000,MCH=1000)
C
C     THE REQUIRED MATRICES AND VECTORS ARE DIMENSIONED.
C
      DIMENSION XX(NN,IP),Y(NN),ERR(NN),T(NN)
      DIMENSION XTX(IP,IP),XTXB(IP,IP),DIF1(IP,IP),DIF1I(IP,IP)
      DIMENSION BETA(IP),XTY(IP),XTYB(IP),DIF2(IP),BETAB(IP)
      DIMENSION XBOOT(NN,IP),YBOOT(NN),GVEC(NBOOT),GVECR(NBOOT)
      DIMENSION IPERM(NBOOT)

      CHARACTER*70 FILEIN,FILEOUT1,FILEOUT2
      FILEIN='xdata.d'
      FILEOUT1='/home/sjst/KAMAN/SIMUL/NORMAL/normalbootstrap.d'
      OPEN(1,FILE=FILEOUT1,STATUS='NEW')
      CLOSE(1)
      CALL ERSET(0,0,0)
C
C     READ THE ORIGINAL X-MATRIX FROM FILE
C
      OPEN(1,FILE=FILEIN,STATUS='OLD')
      DO 3 I=1,NN
      READ(1,*) DUMMY,(XX(I,J),J=1,IP)
3     CONTINUE
      CLOSE(1)

      DO 9 J=1,IP
      S=0.0D0
      DO 5 I=1,NN
      S=S+XX(I,J)
5     CONTINUE
      S=S/(1.0D0*NN)
      DO 7 I=1,NN
      XX(I,J)=XX(I,J)-S
7     CONTINUE
9     CONTINUE

      NRA=NN
      NCA=2
      LDA=N
      NBB=IP
```

```
      LDB=IP
      CALL DMXTXF(NRA,NCA,XX,LDA,NBB,XTX,LDB)
C
C     INITIALISE SIMULATION COUNTERS
C
      ERRORLB=0.0D0
      ERRORRB=0.0D0
      COVERB=0.0D0
      ALENGTHB=0.0D0
C
C     START THE SIMULATION LOOP
C
      DO 300 MC=1,MCH
      write(6,*) MC
C
C     USE SIMULATION TO GENERATE THE Y-VALUES FROM THE APPROPRIATE
C     DISTRIBUTION
C
      CALL DRNNOR(NN,ERR)
      DO 12 I=1,NN
      T(I)=I
12    CONTINUE
      DO 13 I=1,NN
      Y(I)=10.0D0-T(I)+0.1D0*T(I)*T(I)+DSQRT(SIGMA)*ERR(I)
13    CONTINUE

      NB=0
14    NB=NB+1
      IF (NB.GT.NBOOT) GOTO 50
      CALL BOOT(NN,XX,Y,XBOOT,YBOOT)
      NRA=NN
      NCA=IP
      LDA=N
      NRB=NN
      NCB=IP
      LDB=N
      NRC=IP
      NCC=IP
      LDC=IP
      CALL DMXTYF(NRA,NCA,XX,LDA,NRB,NCB,XBOOT,LDB,NRC,
     & NCC,XTXB,LDC)

      DO 20 I=1,IP
      DO 15 J=1,IP
      DIF1(I,J)=XTX(I,J)-XTXB(I,J)
15    CONTINUE
20    CONTINUE
      CALL DLINRG(IP,DIF1,IP,DIF1I,IP)
      ICODE=IERCD()
      IF (ICODE.NE.0) THEN
          NB=NB-1
          GOTO 14
      ENDIF
      IPATH=2
      CALL DMURRV(NRA,NCA,XX,LDA,NN,Y,IPATH,IP,XTY)
      IPATH=2
      CALL DMURRV(NRA,NCA,XX,LDA,NN,YBOOT,IPATH,IP,XTYB)
      DO 25 J=1,IP
      DIF2(J)=XTY(J)-XTYB(J)
25    CONTINUE
      DO 35 I=1,IP
```

```
      S=0.0D0
      DO 30 J=1,IP
      S=S+DIF1I(I,J)*DIF2(J)
30    CONTINUE
      BETAB(I)=S
35    CONTINUE
      GVEC(NB)=-BETAB(1)/(2.0D0*BETAB(2))
      GOTO 14

50    DO 210 I=1,NBOOT
      IPERM(I)=I
210   CONTINUE
      CALL DSVRGP(NBOOT,GVEC,GVECR,IPERM)
C
C     CALCULATE THE LOWER AND THE UPPER CONFIDENCE LIMITS
C
      IL=INT(((1.0D0-ALPHA)/2.0D0)*NBOOT)
      CONL=GVECR(IL)
      IU=INT((0.5D0+ALPHA/2.0D0)*NBOOT)
      CONU=GVECR(IU)
C
C     NOW CALCULATE THE LENGTH OF THE BOOTSTRAP CONFIDENCE INTERVAL
C
      ALENGTHB=ALENGTHB+CONU-CONL
C
C     CALCULATE THE COVERAGE PROBABILITY
C
      IF (CONL.GT.5.0D0) ERRORLB=ERRORLB+1.0D0
      IF (CONU.LT.5.0D0) ERRORRB=ERRORRB+1.0D0
      IF ((CONL.LT.5.0D0).AND.(5.0D0.LT.CONU)) COVERB
     & =COVERB+1.0D0
300   CONTINUE
      COVERB=COVERB/MCH
      ALENGTHB=ALENGTHB/MCH
      ERRORLB=ERRORLB/MCH
      ERRORRB=ERRORRB/MCH
      AMISCOVERB=ERRORLB+ERRORRB
C
C     WRITE THE RESULTS TO OUTPUT FILES
C
      OPEN(1,FILE=FILEOUT1,STATUS='OLD')
      WRITE(1,*)
      WRITE(1,*) '          EQUI-SPACED CASE    FIRST G'
      WRITE(1,602) 'VALUE OF N','SIGMA2','1-ALPHA'
      WRITE(1,605) NN,SIGMA,ALPHA
      WRITE(1,608) 'LEFT-ERROR','RIGHT-ERROR','ALL ERROR','COVERAGE'
     &,'EXP. LENGTH'
      WRITE(1,610) ERRORLB,ERRORRB,AMISCOVERB,COVERB,ALENGTHB
      CLOSE(1)

500   CONTINUE
510   CONTINUE
520   CONTINUE

602   FORMAT(A10,6X,A6,5X,A7)
605   FORMAT(5X,I5,2X,F10.4,2X,F10.4,A14,3X,A9)
608   FORMAT(A10,1X,A11,3X,A9,4X,A8,3X,A11)
610   FORMAT(4(F10.5,2X),2X,F10.5)

1000  STOP
      END
```

```
      SUBROUTINE BOOT(NP,X,Y,XBOOT,YBOOT)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      PARAMETER (IP=2,N=100,IPP1=IP+1,NBOOT=1000,MCH=1000)
      DIMENSION X(N,IP),XBOOT(N,IP),Y(N),YBOOT(N)
      DIMENSION IR(N)
      CALL RNUND(NP,NP,IR)
      DO 50 I=1,NP
      DO 40 J=1,IP
      XBOOT(I,J)=X(IR(I),J)
40    CONTINUE
      YBOOT(I)=Y(IR(I))
50    CONTINUE
      RETURN
      END
```