tests for preference * j.j. dik

TESTS FOR PREFERENCE

UNIVERSITEIT VAN AMSTERDAM

TESTS FOR PREFERENCE

ACADEMISCH PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN
DOCTOR IN DE WISKUNDE EN NATUURWETENSCHAPPEN
AAN DE UNIVERSITEIT VAN AMSTERDAM,
OP GEZAG VAN DE RECTOR MAGNIFICUS,
DR. D.W. BRESTERS,
HOOGLERAAR IN DE FACULTEIT DER
WISKUNDE EN NATUURWETENSCHAPPEN,
IN HET OPENBAAR TE VERDEDIGEN IN DE
AULA DER UNIVERSITEIT
(TIJDELIJK IN DE LUTHERSE KERK,
INGANG SINGEL 411, HOEK SPUI)
OP WOENSDAG 1 APRIL 1981
DES NAMIDDAGS TE 15.00 UUR PRECIES

DOOR

JACOB JAN DIK

GEBOREN TE STAD DELDEN

1981
MATHEMATISCH CENTRUM, AMSTERDAM

PROMOTOR: PROF. DR. J. HEMELRIJK

COREFERENT: PROF. DR. J. OOSTERHOFF

Voor Irene

Acknowledgements

This thesis was written at the Institute of Applied Mathematics of the University of Amsterdam. I am most grateful to Prof. Dr. J. Hemelrijk, who drew my attention to the subject of this thesis, for many inspiring discussions which often clarified the issues considerably.

I am grateful to the Institute for providing excellent working conditions and for giving me ample computing facilities.

Dr. A.A. Balkema and Dr. G. Laman were consulted a number of times during the preparation of the manuscript. I thank them for their kind help.

Dr. P. Groeneboom provided me with a pre-preprint of the paper of Dr. Rothe of the University of Dortmund on Pitman efficiencies. I am grateful to Dr. G. Rothe for sending me the (slightly more general) preprint of his paper which will - to my knowledge - appear in the march number of the 1981 volume of the Annals of Statistics.

I am most grateful to Prof. Dr. J. Oosterhoff for his referential work.

I sincerely thank the typist (who definitely wants to remain anonymous) for the typing of the manuscript, Mr. T. Baanders for his typographical work and his cover design, Messrs. D. Zwarst, J. Schipper, J.W. v. d. Werf, J. Suiker and F.J.C. Swenneker for printing and binding. Finally I want to thank Irene Cannegieter, Hans van Niekerk and Paul Goedhart for their proofreading work.

Amsterdam, february 1981. Jaap Dik

CONTENTS

INTRODUCTION

This thesis contains the results of researches that we made into a practical statistical situation. Our aim has been to make the results actually accessible to practicing statisticians and therefore this work was written with this goal in mind. It gives for instance several methods to compute (approximate) critical values for the test-statistics that occur.

The practical situation concerns the detection of differences of preferences or aversions between individuals when the observations are the (repeated) choices they have made. Suppose for instance that n persons may choose from k brands of chocolate. All persons may have the same absolute preferences, possibly changing in time, for special brands, but it is the *difference* between the persons with respect to these preferences that we wish to detect. (The title of this thesis might thus have been "Tests for differences of preference").

The practical problem and the statistical solution of it are outlined in chapter 1, which gives the practicing statistician all the information he needs to be able to apply the test.

The basis of the solution of the problem is a vector of observable random variables, $\vec{\underline{t}}_*$, of which the asymptotic normality is established under certain conditions. (Section 4.2.). The class of quadratic forms in $\vec{\underline{t}}_*$

$$\mathcal{T} = \{\vec{\underline{t}}_*' Q \vec{\underline{t}}_* \mid Q \text{ non-negative definite}\}$$

is then considered as a possible class of, in practice, useful test-statistics. The use of quadratic forms is given extentive intuitive (section 2.1.) and theoretical (chapter 8) motivation.

Two problems arose in the determination of the asymptotic distribution of $\vec{\underline{t}}_*' Q \vec{\underline{t}}_*$. The first problem was the singularity of the dispersion-matrix of $\vec{\underline{t}}_*$ (also asymptotically) and the second problem was the (more or less) arbitrariness of the matrix Q. It could be expected that only for some special choices of Q the test-statistic would (asymptotically) have a χ^2-distribution. Both problems are solved by a theorem (theorem 3.2.1.) which gives the distribution of a non-negative definite quadratic form in normal variables, also in the case that the dispersion matrix of the normal variables is singular. Using this theorem, asymptotic distributions of $\underline{\vec{t}}_*'Q\underline{\vec{t}}_*$ are determined (both under the null-hypothesis and under alternatives)(section 4.3.).

A usual method to deal with singularity is to define a transformation to a lower dimensional space in which the dispersion matrix of the (transformed) variables is non-singular. This leads mostly to complicated statistics and obscures the working of the tests. Using Rao's theory on g-inverses of matrices (RAO (1973)) it is shown that such a transformation is unneccessary (chapter 6). MADANSKY (1963) used the method of transformation to a lower dimension when he proposed a generalisation to Cochran's Q-test (COCHRAN (1950)). Both Madansky's and Cochran's test can be seen as a special case of the tests we investigated (chapter 6).

Consistency properties and power of the tests from $\mathcal{T}$ are considered in chapter 4. The asymptotic relative efficiencies of pairs of tests from $\mathcal{T}$, according to Pitman and Bahadur are established in chapter 5. Neither of these efficiency concepts gives a clear indication which Q to use when an overall type test is desired.

Therefore, again mostly motivated by intuitive arguments, a χ^2-type statistic is recommended for practical use (section 6.4.). The recommendations are supported by the results of simulation which we give in chapter 9. It is shown there also that the tests can effectively be directed towards a special alternative by a suitable choice of the matrix Q.

Finally, in order to find a good and simple approximation for the distribution under H_0, the expectation and variance of the test-statistics are established for some special choices of Q (chapter 7).

CHAPTER 1

PRACTICE: A RECOMMENDED STATISTIC

1.1. A TESTING PROBLEM: GENERAL REMARKS

In this study, we investigate the properties of a class of statistical tests for a certain testing problem. As a result of the investigations a member of this class can be recommended for practical use. In this first chapter, we give the possible user all the information he needs when he wants to apply the recommended test.

A more formal approach is started in chapter 2 and the problem is developed further in chapters 3-8. Finally some numerical results are given as illustration in chapter 9.

In this section we begin with the statement of the problem. Although the proposed testing procedure is applicable in many other situations, it is convenient to adopt the terminology of the following example. This not only makes the description of the situation easier, but it is also natural, because this investigation was motivated by this example.

In 1975 KNEEPKENS (1975) wrote a report "De voornaamste kop op de voorpagina's van een vijftal landelijke Nederlandse dagbladen in de eerste twee maanden van 1964 en 1974", (in Dutch), in which he compared 5 Dutch newspapers in 1964 and 1974. In this report, Kneepkens asks the question if there exists a statistical test for the following situation.

'Newspaper' Problem

Each day, every one of n newspapers chooses a subject for its 'front-page' article from a category of subjects. The different categories are elements of a given categorical system

(1.1.1) $\mathcal{C} = \{C_1,\ldots,C_k\}$.

On the basis of the observed choices made by the n newspapers on m different days, we want to find out if there are one or more newspapers differing, more or less persistently from the others with respect to their preferences for categories from $\mathcal{C}$.

Formulated as a testing problem, we would want to test the null-hypothesis that the newspapers do not differ among themselves with respect to their choices, against the alternative that at least one newspaper has a persistent preference or aversion for at least one subject-category.

Although this is still somewhat loosely formulated, it follows that we need to construct an overall-test not unlike Friedman's m rankings test. As in Friedman's case, the test that we shall construct is not likely to work well against *all* deviations from the null-hypothesis.

Now, let's be more specific. The mathematical model that we make for the 'newspaper'-problem will be based on the following assumptions

(1.1.2) the newspapers make their choices independently of each other;

(1.1.3) the choices which are made on different days are independent.

Let

(1.1.4) $(\underline{C}_i^{(1)}, \underline{C}_i^{(2)}, \ldots, \underline{C}_i^{(n)}) \in \mathcal{C}^n$

denote an observation on the i'th day, i.e.

(1.1.5) $\underline{C}_i^{(\nu)}$ is the category that is chosen on the i'th day by the ν'th newspaper.

Introduce the following random variables

(1.1.6) $\underline{x}_{ij}^{(\nu)} = \begin{cases} 1 \text{ if the } \nu\text{'th newspaper chooses } C_j \text{ on the i'th day;} \\ 0 \text{ otherwise.} \end{cases}$

and probabilities *)

(1.1.7) $p_{ij}^{(\nu)} \overset{d}{=} P(\underline{x}_{ij}^{(\nu)} = 1) = P(\underline{C}_i^{(\nu)} = C_j)$,

---

*) "$\overset{d}{=}$" indicates a definition.

with

(1.1.8) $$p_{i+}^{(\nu)} = 1.$$

(When an index is replaced by a "+" sign, we mean that the indexed quantity has been summed over that index, i.e.

$$p_{i+}^{(\nu)} = \sum_{j=1}^{k} p_{ij}^{(\nu)} \quad).$$

We define

(1.1.9) $$\underline{a}_{ij} \stackrel{d}{=} \underline{x}_{ij}^{(+)},$$

i.e. $\underline{a}_{ij}$ is the number of times that the category C_j has been chosen on the i'th day, and

(1.1.10) $$\underline{h}_j^{(\nu)} \stackrel{d}{=} \underline{x}_{+j}^{(\nu)},$$

i.e. $\underline{h}_j^{(\nu)}$ is the number of times that the ν'th newspaper has chosen category C_j. Note that

(1.1.11) $$\underline{h}_j^{(+)} \equiv \underline{a}_{+j}.$$

Because every newspaper chooses one category at a time, we have

(1.1.12) $$\underline{x}_{i+}^{(\nu)} \equiv 1.$$

The assumptions (1.1.2) and (1.1.3) mean that in our model we must take

(1.1.13) $\underline{x}_{ij}^{(\nu)}$ and $\underline{x}_{hl}^{(\mu)}$ to be independent whenever $i \neq h$ or $\nu \neq \mu$.

This completes our basic mathematical model. We can now formulate the null-hypothesis in terms of this model

(1.1.14) $$H_0\colon p_{ij}^{(1)} = p_{ij}^{(2)} = \dots = p_{ij}^{(n)} \qquad \text{for all } i \text{ and } j.$$

Denote the common value of $p_{ij}^{(\nu)}$, under H_0, by p_{ij}. It is then clear that the model still contains the m×k unknown parameters p_{ij}.

All the unknown (nuisance) parameters p_{ij}, however, can be eliminated when we condition on the event

$$(1.1.15) \qquad \underline{A} = A$$

where

$$(1.1.16) \qquad \underline{A} = \begin{pmatrix} \underline{a}_{11} & \underline{a}_{12} & \cdots & \underline{a}_{1k} \\ \underline{a}_{21} & \underline{a}_{22} & \cdots & \underline{a}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{a}_{m1} & \underline{a}_{m2} & \cdots & \underline{a}_{mk} \end{pmatrix}$$

and A is a similarly defined $m \times k$ matrix which contains the observed values of $\underline{A}$.

When we consider the random variables $\underline{x}_{ij}^{(\nu)}$, but now conditioned on $\underline{A} = A$, i.e.

$$(1.1.17) \qquad \underline{t}_{ij}^{(\nu)} \overset{d}{=} (\underline{x}_{ij}^{(\nu)} | \underline{A} = A)$$

then the joint distribution of the $\underline{t}_{ij}^{(\nu)}$ contains, under H_0, no more unknown parameters. In fact, it is evident that, under H_0, given $\underline{A} = A$, all the ('generalised') permutations of

$$(1.1.18) \qquad \underbrace{C_1 \ldots C_1}_{a_{i1}\times} \; \underbrace{C_2 \ldots C_2}_{a_{i2}\times} \; \ldots \; \underbrace{C_k \ldots C_k}_{a_{ik}\times}$$

are equally likely to occur as outcomes of (1.1.4). That is, each generalised permutation has (conditional) probability

$$(1.1.19) \qquad \frac{a_{i1}!\; a_{i2}! \;\ldots\; a_{ik}!}{n!} \; .$$

Therefore we can use the observable random variables $\underline{t}_{ij}^{(\nu)}$ as building-stones for possible test-statistics.

Finally we define, analogous to (1.1.10)

$$(1.1.20) \qquad \underline{f}_j^{(\nu)} \overset{d}{=} \underline{t}_{+j}^{(\nu)} .$$

1.2. PRESENTATION OF THE DATA

Assume that the matrix A contains the observed values of $\underline{a}_{ij}$. Then under the condition of the event $\underline{A} = A$, we may represent the data as follows.

Table 1.2.1. Presentation of the data.

	newspaper					category				
	$\nu =$ 1	2	. . .	n		$j =$ 1	2	. . .	k	
1	$\underline{c}_1^{(1)}$	$\underline{c}_1^{(2)}$	. . .	$\underline{c}_1^{(n)}$		a_{11}	a_{12}	. . .	a_{1k}	n
2	$\underline{c}_2^{(1)}$	$\underline{c}_2^{(2)}$	. . .	$\underline{c}_2^{(n)}$		a_{21}	a_{22}	. . .	a_{2k}	n
.	.	.	.	.		.	.	.	.	.
m	$\underline{c}_m^{(1)}$	$\underline{c}_m^{(2)}$	. . .	$\underline{c}_m^{(n)}$		a_{m1}	a_{m2}	. . .	a_{mk}	n
						a_{+1}	a_{+2}	. . .	a_{+k}	mn
1	$\underline{f}_1^{(1)}$	$\underline{f}_1^{(2)}$	. . .	$\underline{f}_1^{(n)}$	a_{+1}					
2	$\underline{f}_2^{(1)}$	$\underline{f}_2^{(2)}$	. . .	$\underline{f}_2^{(n)}$	a_{+2}					
.	.	.	.	.	.					
k	$\underline{f}_k^{(1)}$	$\underline{f}_k^{(2)}$	. . .	$\underline{f}_k^{(n)}$	a_{+k}					
	m	m	. . .	m	nm					

EXAMPLE 1.2.1. In this example we present the data of 6 Dutch newspapers in 1964.

1. De Telegraaf
2. De Volkskrant
3. Het Parool
4. Algemeen Handelsblad
5. N.R.C.
6. De Waarheid.

We consider the following categories

C_1. Dutch economy
C_2. Dutch politics
C_3. remaining Dutch news
C_4. America
C_5. Africa & Middle East
C_6. remaining foreign news.

The source of the following table of data is KNEEPKENS (1975).

<u>Table 1.2.2.</u> Example of data of 6 Dutch newspapers.

	newspaper							category						
date	1	2	3	4	5	6		1	2	3	4	5	6	
2-1-'64	C_5	C_1	C_4	C_5	C_4	C_1		2	0	0	2	2	0	6
7-1-'64	C_6	C_6	C_6	C_6	C_1	C_1		2	0	0	0	0	4	6
11-1-'64	C_4	C_4	C_4	C_4	C_4	C_4		0	0	0	6	0	0	6
17-1-'64	C_6	C_1	C_5	C_5	C_6	C_1		2	0	0	0	2	2	6
23-1-'64	C_1	C_5	C_6	C_6	C_6	C_3		1	0	1	0	1	3	6
29-1-'64	C_5	C_5	C_6	C_5	C_6	C_5		0	0	0	0	4	2	6
4-2-'64	C_3	C_3	C_3	C_3	C_1	C_1		2	0	4	0	0	0	6
10-2-'64	C_3	C_3	C_3	C_3	C_3	C_3		0	0	6	0	0	0	6
15-2-'64	C_3	C_5	C_5	C_5	C_6	C_1		1	0	1	0	3	1	6
21-2-'64	C_2	C_6	C_6	C_6	C_6	C_2		0	2	0	0	0	4	6
27-2-'64	C_2	C_2	C_6	C_2	C_2	C_1		1	4	0	0	0	1	6
category								11	6	12	8	12	17	66
1	1	2	0	0	2	6	11							
2	2	1	0	1	1	1	6							
3	3	2	2	2	1	2	12							
4	1	1	2	1	2	1	8							
5	2	3	2	4	0	1	12							
6	2	2	5	3	5	0	17							
	11	11	11	11	11	11	66							

Inspection of the data of example 1.2.1. leads to the following remarks.

i. Between two subsequent days of observation, there are each time four days on which no observation was made (not counting sundays). This is done to satisfy as good as possible the assumption (1.1.3).

ii. On 11-1-'64 and 10-2-'64 all the newspapers chose a subject from the same category (C_4 and C_3 resp.). Because these observations cannot contribute to the detection of deviations from H_0, they are "useless", and they should play no role in our procedure. In section 2.5. we shall show that we may delete such observations when we use one of the tests that we have developed for this problem.

After deletion of these observations we have

Table 1.2.3. Data of example 1.2.1. after deletion of "useless" observations.

	newspaper							category						
date	1	2	3	4	5	6		1	2	3	4	5	6	
2-1-'64	C_5	C_1	C_4	C_5	C_4	C_1		2	0	0	2	2	0	6
7-1-'64	C_6	C_6	C_6	C_6	C_1	C_1		2	0	0	0	0	4	6
17-1-'64	C_6	C_1	C_5	C_5	C_6	C_1		2	0	0	0	2	2	6
23-1-'64	C_1	C_5	C_6	C_6	C_6	C_3		1	0	1	0	1	3	6
29-1-'64	C_5	C_5	C_6	C_5	C_6	C_5		0	0	0	0	4	2	6
4-2-'64	C_3	C_3	C_3	C_3	C_1	C_1		2	0	4	0	0	0	6
15-2-'64	C_3	C_5	C_5	C_5	C_6	C_1		1	0	1	0	3	1	6
21-2-'64	C_2	C_6	C_6	C_6	C_6	C_2		0	2	0	0	0	4	6
27-2-'64	C_2	C_2	C_6	C_2	C_2	C_1		1	4	0	0	0	1	6
category								11	6	6	2	12	17	54
1	1	2	0	0	2	6	11							
2	2	1	0	1	1	1	6							
3	2	1	1	1	0	1	6							
4	0	0	1	0	1	0	2							
5	2	3	2	4	0	1	12							
6	2	2	5	3	5	0	17							
	9	9	9	9	9	9	54							

1.3. THE 'CONDITIONAL' SITUATION

Sometimes an experimental setup leads directly to the situation which we have in section 1.1. after the conditioning on the event $\underline{A} = A$. We mean that a researcher may determine the elements of A in advance and perform an experiment in which the outcomes are of the type (1.1.18), which have each, under a 'H_0', the same probability.

Without changing the notation, the random variable $\underline{t}_{ij}^{(\nu)}$ would now mean

(1.3.1) $$\underline{t}_{ij}^{(\nu)} = \begin{cases} 1 & \text{if in the outcome on the i'th day } C_j \text{ is in the } \nu\text{'th place;} \\ 0 & \text{otherwise.} \end{cases}$$

The random variables $\underline{f}_j^{(\nu)}$ may be defined as in (1.1.20).

In fact, once we have conditioned on the event $\underline{A} = A$, it is not possible to discriminate between the two kinds of experiments and the two testing problems anymore, apart from the fact that the alternatives we are interested in may be chosen differently. This 'conditional' situation will be the starting point of the theory in chapter 2.

We give an example of this situation.

EXAMPLE 1.3.1. Suppose that a foreman distributes each morning n jobs among n workers. Among the n jobs are a_{ij} of the kind C_j, on the i'th day. We would now want to test the hypothesis that the foreman distributes the jobs at random, for instance against the alternative that some worker gets jobs assigned to him that are persistently of the same kind.

1.4. THE PROPOSED TEST

For the testing problems described in sections 1.1. and 1.2. we propose the following test-statistic

(1.4.1) $$\underline{v} \stackrel{d}{=} \sum_{j=1}^{k} \sum_{\nu=1}^{n} \frac{(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})^2}{\frac{a_{+j}}{n}} .$$

If, under the experimental situation of section 1.1. some category has not been chosen, we have $a_{+j} = 0$ and $(\underline{f}_j^{(\nu)} - a_{+j}/n)^2 \equiv 0$. In those cases we define $(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})^2 / \frac{a_{+j}}{n} \equiv 0$.

It is easily shown that, under H_0, $E\underline{f}_j^{(\nu)} = \frac{a_{+j}}{n}$, so our test-statistic has the form of the traditional chi-squared statistic.

In chapter 4 and 6 we shall show that the asymptotic distribution of $\frac{n-1}{n} \underline{v}|H_0$ is in a special case a χ^2-distribution with (n-1)(k-1) degrees of freedom, and in general the distribution of a linear combination of independent χ^2-variables. In both cases the following approximation is an improvement on these asymptotic distributions.

Approximate $c\underline{v}$ by a χ^2-distribution with η degrees of freedom, where c and η are determined such that, under H_0,

$$E c\underline{v} = cE\underline{v} = E\underline{\chi}^2[\eta] = \eta \tag{1.4.2}$$

and

$$\sigma^2(c\underline{v}) = c^2\sigma^2(\underline{v}) = \sigma^2(\underline{\chi}^2[\eta]) = 2\eta, \tag{1.4.3}$$

thus equating the first two moments of $c\underline{v}$ to those of $\chi^2[\eta]$. Hence

$$c = 2E\underline{v}/\sigma^2(\underline{v}), \tag{1.4.4}$$

$$\eta = 2(E\underline{v})^2/\sigma^2(\underline{v}). \tag{1.4.5}$$

$E\underline{v}$ and $\sigma^2(\underline{v})$ are given, under H_0, by

$$E\underline{v} = n\sum_{j=1}^{k}\frac{S_j}{E_j} \tag{1.4.6}$$

$$\sigma^2(\underline{v}) = \frac{2n^2}{n-1}\{\sum_{j=1}^{k}\frac{S_j^2 - T_j}{E_j} + \sum_{j\neq l}\sum\frac{S_{jl}^2 - T_{jl}}{E_jE_l}\}, \tag{1.4.7}$$

where

$$E_j = \frac{a_{+j}}{n} = \sum_{i=1}^{m}\frac{a_{ij}}{n}, \tag{1.4.8}$$

$$S_j = n^{-2}\sum_{i=1}^{m} a_{ij}(n-a_{ij}), \tag{1.4.9}$$

$$T_j = n^{-4}\sum_{i=1}^{m} a_{ij}^2(n-a_{ij})^2, \tag{1.4.10}$$

$$S_{jl} = n^{-2}\sum_{i=1}^{m} a_{ij}a_{il}, \tag{1.4.11}$$

$$T_{jl} = n^{-4}\sum_{i=1}^{m} a_{ij}^2a_{il}^2. \tag{1.4.12}$$

Critical values and tail-probabilities of the distribution of $\underline{v}$ may be approximated using this method.

For practical calculations of $\underline{v}$, we may use the following formula

$$(1.4.13)\qquad \underline{v} \equiv n \sum_{j=1}^{k} \frac{\sum_{\nu=1}^{n} \{\underline{f}_j^{(\nu)}\}^2}{a_{+j}} - nm .$$

Finally we apply the test to the data of example 1.2.1.

EXAMPLE 1.4.1. For the data of example 1.2.1., after the deletion of "useless" observations (see table 1.2.3.), our test-statistic takes the value $v = 33.19$. In this case we have $E(\underline{v}|H_0) = 20.11$ and $\sigma^2(\underline{v}|H_0) = 19.59$, so that $c = 2E/\sigma^2 = 2.0526$, $\eta = 41.27$ and $cv = 68.13$. The critical value of the $\chi^2[41.27]$ distribution for $\alpha = 0.05$ is equal to 57.26, so H_0 is rejected. The right tail-probability of 68.13 for the $\chi^2[41.27]$ distribution is equal to 0.0053.

An estimate of the right tail-probability of 33.19 in the exact distribution is equal to 0.007, indicating a close fit of the approximation. The estimate was obtained from 1000 simulated drawings from the exact distribution of $\underline{v}$.

CHAPTER 2

THEORY: PRELIMINARIES

2.1. THE PROBLEM

We consider a sequence $E_1,\ldots,E_m$ of m independent experiments. The possible outcomes of E_i $(i=1,\ldots,m)$ are the permutations of the n characters

(2.1.1) $$\underbrace{C_1\ldots C_1}_{a_{i1}\times}\ \underbrace{C_2\ldots C_2}_{a_{i2}\times}\ \ldots\ \underbrace{C_k\ldots C_k}_{a_{ik}\times}$$

where C_j $(j=1,\ldots,k)$ occurs a_{ij} times, with $0 \le a_{ij} < n$ and $\sum_{j=1}^{k} a_{ij} = n$. Because of the repetitions of the characters in the permutations, we shall call such a permutation a 'word'.

As indicated by the notation, the characters C_j $(j=1,\ldots,k)$ and the length n of the words will be the same for all experiments, but the numbers a_{ij} may differ from experiment to experiment. In asymptotic considerations we shall let $m \to \infty$ with n and k fixed.

The indices i,h will be used throughout this work to index the experiments, the indices j and l to index the characters and $\nu,\mu \in \{1,\ldots,n\}$ to indicate the ν'th and μ'th place in a word. So we shall always have

(2.1.2) $$i,h \in \{1,\ldots,m\};\ j,l \in \{1,\ldots,k\};\ \nu,\mu \in \{1,\ldots,n\}.$$

By this convention we can use these symbols without further explanation. Subject to this convention each given formula will hold for each value that the indices occurring in it can take, unless otherwise is indicated.

The properties of the numbers a_{ij} may be summarized by

(2.1.3) $$\vec{a}_i \stackrel{d}{=} (a_{i1},\ldots,a_{ik})' \in \{(a_1,\ldots,a_k)' \mid a_j \in \{0,\ldots,n-1\},\ \sum_{j=1}^{k} a_j = n\}.$$

The number N_i of different words for E_i is

$$(2.1.4) \qquad N_i = \frac{n!}{a_{i1}!\, a_{i2}! \ldots a_{ik}!} .$$

Let

$$(2.1.5) \qquad \mathcal{R}_i \stackrel{d}{=} \{1,\ldots,N_i\}$$

and let

$$(2.1.6) \qquad \Omega_i \stackrel{d}{=} \{\pi_{i1},\ldots,\pi_{iN_i}\}$$

denote the set of words pertaining to E_i, in lexicographical order. Then

$$(2.1.7) \qquad \Omega \stackrel{d}{=} \Omega_1 \times \ldots \times \Omega_m$$

is the set of all possible outcomes of the composite experiment

$$(2.1.8) \qquad E \stackrel{d}{=} (E_1,\ldots,E_m).$$

To complete the mathematical model we shall use, we look at the class of all probability distributions on Ω, with $E_1,\ldots,E_m$ independent. Let $\mathcal{P}_i$ be the class of probability distributions on Ω_i

$$(2.1.9) \qquad \mathcal{P}_i \stackrel{d}{=} \{(p_1,\ldots,p_{N_i})' \,|\, p_r \geq 0,\ r \in \mathcal{R}_i,\ \sum_{r=1}^{N_i} p_r = 1\}$$

and let $\underline{\omega}_i$ be random on Ω_i with

$$(2.1.10) \qquad P(\underline{\omega}_i = \pi_{ir}) = p_{ir},\ r \in \mathcal{R}_i,\ \vec{p}_i \stackrel{d}{=} (p_{i1},\ldots,p_{iN_i})' \in \mathcal{P}_i.$$

Then, according to the independence of $E_1,\ldots,E_m$ we have

$$(2.1.11) \qquad P((\underline{\omega}_1,\ldots,\underline{\omega}_m) = (\pi_{1r_1},\ldots,\pi_{mr_m})) = \prod_{i=1}^{m} p_{ir_i}$$

with $r_i \in \mathcal{R}_i$ and $\vec{p}_i \in \mathcal{P}_i$, and this is the class of probability distributions we consider. It will be indicated by

$$(2.1.12) \qquad \mathcal{P} = \mathcal{P}_1 \times \ldots \times \mathcal{P}_m.$$

In order to formulate the hypotheses about $\mathcal{P}$ which we want to consider, we introduce parameters Δ_{ir} as follows

$$(2.1.13)\qquad \Delta_{ir} \overset{d}{=} p_{ir} - \frac{1}{N_i} \qquad\qquad (r \in \mathcal{R}_i),$$

with, obviously

$$(2.1.14)\qquad -\frac{1}{N_i} \le \Delta_{ir} \le 1 - \frac{1}{N_i}\ ; \qquad \sum_{r=1}^{N_i} \Delta_{ir} = 0.$$

Let

$$(2.1.15)\qquad \mathcal{D}_i \overset{d}{=} \{(\Delta_1,\dots,\Delta_{N_i})' \mid -\frac{1}{N_i} \le \Delta_r \le 1 - \frac{1}{N_i},\ r \in \mathcal{R}_i,\ \sum_{r=1}^{N_i} \Delta_r = 0\},$$

and

$$(2.1.16)\qquad \vec{\Delta}_i \overset{d}{=} (\Delta_{i1},\dots,\Delta_{iN_i})' \in \mathcal{D}_i,$$

then every element of

$$(2.1.17)\qquad \mathcal{D} \overset{d}{=} \mathcal{D}_1 \times \dots \times \mathcal{D}_m$$

specifies a distribution from $\mathcal{P}$ and v.v. .

The hypothesis to be tested is

$$(2.1.18)\qquad H_0:\quad \forall_i \quad \vec{\Delta}_i = \vec{0}\ .$$

The widest class of alternative hypotheses is of course

$$(2.1.19)\qquad H_1:\quad \exists_i \quad \vec{\Delta}_i \ne \vec{0}\ ,$$

but this class is too amorphous for our purposes. In order to formulate a useful subclass of H_1, we first introduce elementary random variables on Ω_i, which are used for building up test-statistics.

Let, for all i, j and ν, $t_{ij}^{(\nu)}: \Omega_i \to \{0,1\}$ be defined as

$$(2.1.20)\qquad t_{ij}^{(\nu)}(\pi) \overset{d}{=} \begin{cases} 1 \text{ if in } \pi\ C_j \text{ occurs in the } \nu\text{'th place;} \\ 0 \text{ otherwise.} \end{cases}$$

The following relations are then easily proved

$$(2.1.21)\quad \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir}) = N_i \frac{a_{ij}}{n},$$

$$(2.1.22)\quad \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, t_{ij}^{(\mu)}(\pi_{ir}) = N_i \frac{a_{ij}(a_{ij}-1)}{n(n-1)} \qquad (\nu\neq\mu),$$

$$(2.1.23)\quad \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, t_{il}^{(\mu)}(\pi_{ir}) = N_i \frac{a_{ij}a_{il}}{n(n-1)} \qquad (\nu\neq\mu, j\neq l).$$

Now let

$$(2.1.24)\quad \delta_{ijl}^{(\nu,\mu)} \overset{d}{=} \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, t_{il}^{(\mu)}(\pi_{ir})\, \Delta_{ir}$$

and

$$(2.1.25)\quad \delta_{ij}^{(\nu,\mu)} \overset{d}{=} \delta_{ijj}^{(\nu,\mu)}; \qquad \delta_{ij}^{(\nu)} \overset{d}{=} \delta_{ijj}^{(\nu,\nu)}.$$

Let the random variables induced by $\mathcal{P}_i$ and (2.1.20) be denoted by $\underline{t}_{ij}^{(\nu)}$, then we have, under $\mathcal{P}_i$,

$$(2.1.26)\quad P(\underline{t}_{ij}^{(\nu)} = 1) = \frac{a_{ij}}{n} + \delta_{ij}^{(\nu)},$$

$$(2.1.27)\quad P(\underline{t}_{ij}^{(\nu)} = \underline{t}_{ij}^{(\mu)} = 1) = \frac{a_{ij}(a_{ij}-1)}{n(n-1)} + \delta_{ij}^{(\nu,\mu)} \qquad (\nu\neq\mu),$$

$$(2.1.28)\quad P(\underline{t}_{ij}^{(\nu)} = \underline{t}_{il}^{(\mu)} = 1) = \frac{a_{ij}a_{il}}{n(n-1)} + \delta_{ijl}^{(\nu,\mu)} \qquad (\nu\neq\mu, j\neq l).$$

Notice that

$$(2.1.29)\quad P(\underline{t}_{ij}^{(\nu)} = \underline{t}_{il}^{(\nu)} = 1) = 0, \qquad (j\neq l).$$

The proof of these relations follows from the fact that the left-hand members are equal to the expected value of the product of the r.v.'s occurring in these expressions. For instance for (2.1.28) we have, using (2.1.13), (2.1.23) and (2.1.24)

$$P(\underline{t}_{ij}^{(\nu)} = \underline{t}_{il}^{(\mu)} = 1) = E\underline{t}_{ij}^{(\nu)}\underline{t}_{il}^{(\mu)} = \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, t_{il}^{(\mu)}(\pi_{ir})\, p_{ir} =$$

$$= \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, t_{il}^{(\mu)}(\pi_{ir})\, \left(\frac{1}{N_i} + \Delta_{ir}\right) =$$

$$= \frac{a_{ij}a_{il}}{n(n-1)} + \delta_{ijl}{}^{(\nu,\mu)} .$$

The other two relations are special cases.

Note that under H_0 all δ's are equal to 0. Thus if

$$(2.1.30) \qquad H_0': \quad \forall_{ijl\nu\mu} \quad \delta_{ijl}{}^{(\nu,\mu)} = 0,$$

then

$$(2.1.31) \qquad H_0 \Rightarrow H_0'$$

but not the other way around.

We can now tentatively formulate the alternative hypotheses we wish to consider.

We shall say that place ν has a *preference* for or an *aversion* from character C_j respectively if

$$(2.1.32) \qquad \sum_{i=1}^{m} \delta_{ij}{}^{(\nu)} > 0 \qquad \text{or} \qquad < 0.$$

An aversion thus is the same as a negative preference. Now it is easily verified by means of (2.1.26) that

$$(2.1.33) \qquad \sum_{\nu=1}^{n} \delta_{ij}{}^{(\nu)} = 0 \; ; \qquad \sum_{j=1}^{k} \delta_{ij}{}^{(\nu)} = 0 \; ;$$

thus preferences cannot occur in one place only, they are automatically balanced by aversions in other places and vice versa. In fact, preferences as defined above are relative preferences of the places with respect to the preferences of other places, not preferences in an absolute sense.

These considerations lead us to consider the following class of alternative hypotheses

$$(2.1.34) \qquad H_1': \quad \exists_{j\nu} \quad \left| \sum_{i=1}^{m} \delta_{ij}{}^{(\nu)} \right| > 0.$$

The statistic

$$(2.1.35) \qquad \sum_{i=1}^{m} (\underline{t}_{ij}{}^{(\nu)} - \frac{a_{ij}}{n})$$

has according to (2.1.26) the expected value $\sum_{i=1}^{m} \delta_{ij}^{(\nu)}$ and is thus obviously a good building stone for a test-statistic.

Defining

$$(2.1.36) \qquad \underline{f}_j^{(\nu)} \overset{d}{=} \sum_{i=1}^{m} \underline{t}_{ij}^{(\nu)} \ ; \qquad a_{+j} \overset{d}{=} \sum_{i=1}^{m} a_{ij} \ ,$$

we have

$$(2.1.37) \qquad E(\underline{f}_j^{(\nu)} | H_0) = E(\underline{f}_j^{(\nu)} | H_0') = \frac{a_{+j}}{n}$$

and an intuitively attractive test-statistic is

$$(2.1.38) \qquad \sum_{\nu=1}^{n} \sum_{j=1}^{k} (\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})^2 / \frac{a_{+j}}{n} \ .$$

This statistic has the form of the traditional chi-squared statistic: it will assume large values under H_1' and large terms will indicate the preferences and aversions which cause the sum to be large.

It would be too much, however, to expect this statistic to have a chi-squared distribution and it will be shown later that it has a more complicated one (under H_0 as well as under H_1'), which can nevertheless be approximated by means of a modified chi-squared distribution.

The choice of a quadratic form in the $\underline{f}_j^{(\nu)}$ will be shown in chapter 4 to be indicated by the simultaneous asymptotic normality of the $\underline{f}_j^{(\nu)}$ and other theoretical considerations. Several degrees of generalisation of (2.1.38) can then be considered. The most promising one is

$$(2.1.39) \qquad \sum_{\nu=1}^{n} \sum_{j=1}^{k} g_j (\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})^2 ,$$

where the g_j are weighing coefficients for the categories, which will generally (as in (2.1.38)) depend on the a_{+j}/n, but which may also express the experimenter's emphasis on certain characters as compared to others. One might also choose the weights dependent, not only on the characters, but also on the places: $g_j^{(\nu)}$ instead of g_j. We do not, however, elaborate this case in this thesis. In every trial every place occurs exactly once, but the frequencies of the characters may be different from trial to trial. In the applications which led to the development of our tests the places were equivalent: changing their order should have no influence on the experimental situation. Therefore, although the general theory developed later also covers this case, we do not, at this moment, aim at generalisations where different weights are attached to the places.

A further generalisation is to allow cross-terms in the test-statistic

(2.1.40) $$\sum_{\nu=1}^{n} \sum_{j=1}^{k} \sum_{l=1}^{k} g_{jl}(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})(\underline{f}_l^{(\nu)} - \frac{a_{+l}}{n}) .$$

The behaviour of such a test-statistic is far more complicated than that of (2.1.39) and as the result of theoretical considerations (mostly of an asymptotic character) the form (2.1.39) will emerge as the most useful one. The choice of weighing coefficients will be considered in chapters 4 and 5. Some special cases are treated in chapter 6.

The most general quadratic form is. of course

(2.1.41) $$\sum_{\nu=1}^{n} \sum_{\mu=1}^{n} \sum_{j=1}^{k} \sum_{l=1}^{k} g_{jl}^{(\nu,\mu)}(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})(\underline{f}_l^{(\mu)} - \frac{a_{+l}}{n}) .$$

Although test-statistics of this generality are difficult to interpret and therefore of little practical use, the theory which will be developed in later chapters will completely cover this general case. For practical purposes specialisation to the form (2.1.39) is recommended and special attention is paid to this test-statistic and to (2.1.36).

As will appear later, tests based on (2.1.39) will, under acceptable conditions for the g_j, be consistent for $m \to \infty$ if

(2.1.42) $$\exists_{j\nu} \quad |\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \delta_{ij}^{(\nu)}| \to \infty \qquad \text{as} \qquad m \to \infty.$$

This holds e.g. for (2.1.38).

2.2. NOTATION AND SIMPLE RESULTS

Notational conventions.

If X_i denotes any quantity (scalar, vector, r.v., matrix etc.) indexed by the variable i which ranges (for instance) over the index set $\{1,\ldots,m\}$, we shall frequently use the derived quantities X_+, X_* and $X_.$, defined by

(2.2.1) $$X_+ = \sum_{i=1}^{m} X_i ,$$

(2.2.2) $$X_* = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} X_i = X_+/\sqrt{m} ,$$

$$(2.2.3)\qquad X_{\bullet} = \frac{1}{m}\sum_{i=1}^{m} X_i = X_+/m \; .$$

Note that X_+, X_* and $X_{\bullet}$ all depend on m, though this is not apparent from this notation. Sometimes a fixed number m_0 of experiments is considered. To distinguish the quantities X_+ etc., which are defined for general m, from the equivalent quantities for m_0, we shall write

$$(2.2.4)\qquad X_{\oplus} = \sum_{i=1}^{m_0} X_i \; ,$$

and so on.

Furthermore, I_n denotes the identity matrix of order n, $0_{n,k}$ is the $n \times k$ matrix consisting of zero's ($0_n \stackrel{d}{=} 0_{n,n}$), and $1_{n,k}$ the $n \times k$ matrix of one's ($1_n \stackrel{d}{=} 1_{n,n}$). We use the symbol $\otimes$ to denote the Kronecker product of matrices (RAO (1973)).

We consider the r.v.'s defined by (2.1.20) as

$$(2.2.5)\qquad \underline{t}_{ij}^{(\nu)} = \begin{cases} 1 & \text{if in the word obtained at the i'th trial, the character } C_j \text{ occurs in the } \nu\text{'th place;} \\ 0 & \text{otherwise.} \end{cases}$$

Let, for all i, j, ν,

$$(2.2.6)\qquad \underline{\tilde{t}}_{ij}^{(\nu)} \stackrel{d}{=} \underline{t}_{ij}^{(\nu)} - \frac{a_{ij}}{n} \; ,$$

and let

$$(2.2.7)\qquad \vec{\underline{t}}_i \stackrel{d}{=} (\underline{\tilde{t}}_{i1}^{(1)},\ldots,\underline{\tilde{t}}_{ik}^{(1)};\underline{\tilde{t}}_{i1}^{(2)},\ldots,\underline{\tilde{t}}_{ik}^{(2)};\ldots;\underline{\tilde{t}}_{i1}^{(n)},\ldots,\underline{\tilde{t}}_{ik}^{(n)})' .$$

Furthermore, we shall consider

$$(2.2.8)\qquad \vec{\underline{t}}_* \stackrel{d}{=} \frac{1}{\sqrt{m}}\sum_{i=1}^{m} \vec{\underline{t}}_i$$

We shall use the following real, symmetric, $k \times k$ matrix of weighing factors

$$(2.2.9)\qquad G \stackrel{d}{=} \begin{pmatrix} g_{11} & g_{12} & \cdot & \cdot & \cdot & g_{1k} \\ g_{21} & g_{22} & \cdot & \cdot & \cdot & g_{2k} \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ g_{k1} & g_{k2} & \cdot & \cdot & \cdot & g_{kk} \end{pmatrix} , \qquad (g_{jl}=g_{lj}).$$

Let

(2.2.10) $\quad Q \overset{d}{\equiv} I_n \otimes G.$

The test-statistic, defined in (2.1.40), may then, apart from a factor m^{-1}, be written as

$$(2.2.11) \quad \underline{v} \equiv \underline{v}(G) \equiv \vec{\underline{t}}_*' Q \vec{\underline{t}}_* \equiv \frac{1}{m} \sum_{\nu=1}^{n} \sum_{j=1}^{k} \sum_{l=1}^{k} g_{jl}(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})(\underline{f}_l^{(\nu)} - \frac{a_{+l}}{n}).$$

Though it is not essential w.r.t. our problem, we shall wish, for practical reasons, that

$$(2.2.12) \quad \vec{\underline{t}}_*' Q \vec{\underline{t}}_* \geq 0,$$

with probability one. This means that we have to choose G such that Q is non-negative definite (n.n.d). (A $k \times k$ matrix Q is n.n.d iff $\vec{x}' Q \vec{x} \geq 0$ for each $\vec{x} \in \mathbb{R}^k$).

In most of the theory it is irrelevant whether Q has the structure as in (2.2.10) or is an arbitrary n.n.d, real symmetric matrix. So from now on we shall suppose that Q is an arbitrary real, symmetric, n.n.d matrix.

We define the test function $\underline{\varphi}_{m,Q}$:

$$(2.2.13) \quad \underline{\varphi}_{m,Q} = \begin{cases} 1 \text{ if } \vec{\underline{t}}_*' Q \vec{\underline{t}}_* \geq k_{1-\alpha}(m,Q); \\ 0 \text{ otherwise,} \end{cases}$$

where $k_{1-\alpha}(m,Q)$ is determined as the smallest value such that

$$(2.2.14) \quad E(\underline{\varphi}_{m,Q} | H_0) \leq \alpha.$$

Sometimes we shall randomise the test for theoretical purposes, i.e.

$$(2.2.15) \quad \underline{\varphi}_{m,Q} = \begin{cases} 1 \\ \gamma(m,Q) \text{ if } \vec{\underline{t}}_*' Q \vec{\underline{t}}_* \begin{matrix} > \\ = \\ < \end{matrix} k_{1-\alpha}'(m,Q) \\ 0 \end{cases},$$

where $k_{1-\alpha}'(m,Q)$ is the highest possible outcome of $\vec{\underline{t}}_*' Q \vec{\underline{t}}_*$ which is smaller than $k_{1-\alpha}(m,Q)$, and $\gamma(m,Q)$ is determined such that in (2.2.14) the equality sign holds. It will be clear from the context whether we use $\underline{\varphi}_{m,Q}$ defined by (2.2.15) or by (2.2.13). The decision rule for the resulting level-α test is derived from $\underline{\varphi}_{m,Q}$ in the usual way.

We define the vector of expectations of $\underline{t}_{ij}^{(\nu)}$ in the same way as (2.2.7)

$$\vec{\delta}_i \overset{d}{=} (\delta_{i1}^{(1)},\dots,\delta_{ik}^{(1)};\delta_{i1}^{(2)},\dots,\delta_{ik}^{(2)};\dots;\delta_{i1}^{(n)},\dots,\delta_{ik}^{(n)})' \,. \tag{2.2.16}$$

It follows directly from (2.1.26), that

$$E\vec{\underline{t}}_* = \vec{\delta}_* \,, \tag{2.2.17}$$

which reduces, under H_0, to

$$E(\vec{\underline{t}}_*|H_0) = \vec{0} \,. \tag{2.2.18}$$

It is useful to define

$$\vec{\underline{u}}_i \overset{d}{=} \vec{\underline{t}}_i - \vec{\delta}_i \,, \qquad (\vec{\underline{u}}_* \equiv \vec{\underline{t}}_* - \vec{\delta}_*). \tag{2.2.19}$$

The matrix of variances and covariances of the components of a vector of r.v.'s $\vec{\underline{x}}$ will be called the dispersion matrix of $\vec{\underline{x}}$, and will be denoted by

$$D(\vec{\underline{x}}). \tag{2.2.20}$$

In particular we define for each i

$$\Sigma_{1i} \overset{d}{=} D(\vec{\underline{t}}_i) \,, \tag{2.2.21}$$

$$\Sigma_{0i} \overset{d}{=} D(\vec{\underline{t}}_i|H_0) \,. \tag{2.2.22}$$

The entries of Σ_{1i} and Σ_{0i} may be found from the following moments, which may be derived from (2.1.26),...,(2.1.28), together with the obvious relation (cf.(2.1.29))

$$P(\underline{t}_{ij}^{(\nu)} = \underline{t}_{il}^{(\nu)} = 1) = 0, \qquad (j\neq l). \tag{2.2.23}$$

The moments are

$$\sigma^2(\underline{t}_{ij}^{(\nu)}) = \frac{a_{ij}}{n} - \frac{a_{ij}^2}{n^2} - 2\,\delta_{ij}^{(\nu)}\frac{a_{ij}}{n} + \delta_{ij}^{(\nu)} - (\delta_{ij}^{(\nu)})^2 \,, \tag{2.2.24}$$

(2.2.25) $\quad \mathrm{cov}(\underline{t}_{ij}{}^{(\nu)}, \underline{t}_{ij}{}^{(\mu)}) = \qquad (\nu \neq \mu)$

$$= -\frac{a_{ij}(n-a_{ij})}{n^2(n-1)} + \delta_{ij}{}^{(\nu,\mu)} - \frac{a_{ij}}{n}(\delta_{ij}{}^{(\nu)} + \delta_{ij}{}^{(\mu)}) - \delta_{ij}{}^{(\nu)}\delta_{ij}{}^{(\mu)},$$

(2.2.26) $\quad \mathrm{cov}(\underline{t}_{ij}{}^{(\nu)}, \underline{t}_{il}{}^{(\nu)}) = \qquad (j \neq l)$

$$= -\frac{a_{ij}a_{il}}{n^2} - \frac{a_{ij}}{n}\delta_{il}{}^{(\nu)} - \frac{a_{il}}{n}\delta_{ij}{}^{(\nu)} - \delta_{ij}{}^{(\nu)}\delta_{il}{}^{(\nu)},$$

(2.2.27) $\quad \mathrm{cov}(\underline{t}_{ij}{}^{(\nu)}, \underline{t}_{il}{}^{(\mu)}) = \qquad (j \neq l, \nu \neq \mu)$

$$= \frac{a_{ij}a_{il}}{n^2(n-1)} + \delta_{ijl}{}^{(\nu,\mu)} - \frac{a_{ij}}{n}\delta_{il}{}^{(\mu)} - \frac{a_{il}}{n}\delta_{ij}{}^{(\nu)} - \delta_{ij}{}^{(\nu)}\delta_{il}{}^{(\mu)}.$$

Σ_{0i} follows from Σ_{1i} by omitting all terms containing a δ (cf. (2.1.18)). Let, for $n \geq 2$,

$$N \stackrel{d}{=} \begin{pmatrix} 1 & -\frac{1}{n-1} & \cdots & -\frac{1}{n-1} \\ -\frac{1}{n-1} & 1 & \cdots & -\frac{1}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n-1} & -\frac{1}{n-1} & \cdots & 1 \end{pmatrix}, \qquad (2.2.28)$$

then

$$N^2 = \frac{n}{n-1} N, \qquad (2.2.29)$$

and N is of rank n-1, with eigenvalues 0, and $\frac{n}{n-1}$ with multiplicity n-1. Furthermore, let

$$K_i = \frac{1}{n^2} \begin{pmatrix} a_{i1}(n-a_{i1}) & -a_{i1}a_{i2} & \cdots & -a_{i1}a_{ik} \\ -a_{i2}a_{i1} & a_{i2}(n-a_{i2}) & \cdots & -a_{i2}a_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{ik}a_{i1} & -a_{ik}a_{i2} & \cdots & a_{ik}(n-a_{ik}) \end{pmatrix}. \qquad (2.2.30)$$

We then have, as can easily be verified

$$(2.2.31)\qquad D(\vec{\underline{t}}_i|H_0) = \Sigma_{0i} = N \otimes K_i \ ,$$

$$(2.2.32)\qquad D(\vec{\underline{t}}_*|H_0) = \Sigma_{0\cdot} = N \otimes K_\cdot \ .$$

Notice that $\Sigma_{0\cdot}$ is singular, because the sums over rows and column's of $K_\cdot$ and N are zero. More generally, Σ_{0i} and Σ_{1i} are singular because, for each fixed i, the following $n+k-1$ linear relationships hold true for the $n \times k$ random variables $\underline{t}_{ij}^{(\nu)}$, both under H_0 and H_1,

$$(2.2.33)\qquad \underline{t}_{i+}^{(\nu)} \equiv 1, \qquad \underline{t}_{ij}^{(+)} \equiv a_{ij}, \qquad \nu=1,\ldots,n;\ j=1,\ldots,k.$$

REMARK 2.2.1. Not only the singularity of $\Sigma_{0\cdot}$, but also its rank will play a part in the considerations. What can be said about the rank of $\Sigma_{0\cdot}$? Let's first consider the determinant of a matrix which has the same structure as K_i,

$$(2.2.34)\qquad \begin{vmatrix} a_1(n-a_1) & -a_1a_2 & \cdots & -a_1a_j \\ -a_2a_1 & a_2(n-a_2) & \cdots & -a_2a_j \\ \vdots & \vdots & \ddots & \vdots \\ -a_ja_1 & -a_ja_2 & \cdots & a_j(n-a_j) \end{vmatrix} = a_1a_2\ldots a_j\, n^{j-1}\, (n - \sum_{l=1}^{j} a_l).$$

Using this relation, it can easily be shown that K_i is singular. Also the rank of K_i may now be found easily. Let k_i be the number of positive a_{ij}'s at the i'th experiment. Then, again using (2.2.34) it follows that

$$(2.2.35)\qquad \mathrm{rank}(K_i) = k_i - 1.$$

Moreover, because in our case we have $n - \sum_{l=1}^{j} a_l \geq 0$ for $j=1,\ldots,k$, (2.2.34) is non-negative for each j, $j=1,\ldots,k$. It follows that the matrix K_i is non-negative definite.

For $\mathrm{rank}(\Sigma_{0i})$ we find

(2.2.36) $\quad \operatorname{rank}(\Sigma_{0i}) = \operatorname{rank}(N \otimes K_i) = (\operatorname{rank} N)(\operatorname{rank} K_i) = (n-1)(k_i-1).$

Now consider the matrix K_+. We have

(2.2.37) $\quad \vec{x}' K_+ \vec{x} = \vec{x}' K_1 \vec{x} + \vec{x}' K_2 \vec{x} + \ldots + \vec{x}' K_m \vec{x}.$

If r denotes the rank of K_+, there exist k-r linearly independent vectors $\vec{x}$, such that $K_+\vec{x} = 0$, or $\vec{x}' K_+ \vec{x} = 0$. But this implies, because the matrices K_i are non-negative definite that $\vec{x}' K_i \vec{x} = 0$ for all i. This means that

(2.2.38) $\quad r = \operatorname{rank}(K_+) \geq \max_i \operatorname{rank}(K_i) = \max_i (k_i-1).$

Because in any case $\operatorname{rank}(K_+) \leq k-1$ we find

(2.2.39) $\quad \max_i (k_i-1) \leq \operatorname{rank}(K_+) \leq k-1$

and similar bounds for $\operatorname{rank}(\Sigma_{0.})$. It follows that $\operatorname{rank}(\Sigma_{0.})$ is not a fixed number, but has to be determined in each separate case.

For the expectation of $\underline{v}$ we find

(2.2.40) $\quad E\underline{v}(Q) = \operatorname{trace} Q\Sigma_{1.} + \vec{\delta}_*' Q \vec{\delta}_*$,

which reduces, under H_0, to

(2.2.41) $\quad E(\underline{v}(Q)|H_0) = \operatorname{trace} Q\Sigma_{0.}$.

And when $Q = I_n \otimes G$,

(2.2.42) $\quad E(\underline{v}(Q)|H_0) = \operatorname{trace}(I_n \otimes G)(N \otimes K_.) = \frac{n}{m} \sum_{i=1}^{m} \operatorname{trace} GK_i.$

2.3. SOME ASYMPTOTIC CONSIDERATIONS

The choice of the weighing coefficients g_{jl} in the matrix G, or, more generally, the choice of the matrix Q, will largely be determined by the nature of the asymptotic distribution of $\underline{v}(Q)$ as $m \to \infty$.

Because the vectors $\vec{\underline{t}}_1$, $\vec{\underline{t}}_2$, ... are not identically distributed, it is necessary to impose some asymptotic conditions. Therefore we only consider infinite sequences of experiments for which the following 2 assumptions hold.

ASSUMPTION 1. For all j and l the following finite limits exist,

$$(2.3.1) \qquad \lim_{m\to\infty} a_{\cdot j} = a_j > 0, \qquad \text{(say)},$$

$$(2.3.2) \qquad \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} (a_{ij} - \frac{a_{+j}}{m})(a_{il} - \frac{a_{+l}}{m}) = d_{jl}, \qquad \text{(say)}.$$

The alternatives we consider satisfy

ASSUMPTION 2. For all j,l,ν and μ the following limits exist,

$$(2.3.3) \qquad \lim_{m\to\infty} \delta_{\cdot jl}^{(\nu,\mu)} = \zeta_{jl}^{(\nu,\mu)}, \qquad \text{(say)},$$

$$(2.3.4) \qquad \lim_{m\to\infty} \delta_{*j}^{(\nu)} = \delta_j^{(\nu)}, \qquad \text{(say)}.$$

In (2.3.4), but not in (2.3.3), we accept $|\delta_j^{(\nu)}| = \infty$. As a special case of (2.3.3) we have

$$(2.3.5) \qquad \zeta_j^{(\nu)} \overset{d}{=} \zeta_{jj}^{(\nu,\nu)} = \lim_{m\to\infty} \delta_{\cdot j}^{(\nu)}.$$

The vectors $\vec{\delta}$ and $\vec{\zeta}$ with components $\delta_j^{(\nu)}$ and $\zeta_j^{(\nu)}$ are constructed as in (2.2.16).

At first sight, these conditions may seem to be very strong. It is indeed very easy to construct examples of infinite sequences of experiments that do not satisfy these assumptions. However, we have to bear in mind that for statistical purposes, the asymptotic distributions are only necessary to provide a good approximation for the finite situation. Furthermore, there exist situations for which the conditions are trivially fulfilled. For instance in the case that $\vec{a}_1 = \vec{a}_2 = \vec{a}_3 = \ldots$ and $\vec{\Delta}_1 = \vec{\Delta}_2 = \vec{\Delta}_3 = \ldots$. Or, more generally, if

$$(\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_m) = (\vec{a}_{m+1},\vec{a}_{m+2},\ldots,\vec{a}_{2m}) = (\vec{a}_{2m+1},\vec{a}_{2m+2},\ldots,\vec{a}_{3m}) = \ldots$$

and

$$(\vec{\Delta}_1,\vec{\Delta}_2,\ldots,\vec{\Delta}_m) = (\vec{\Delta}_{m+1},\vec{\Delta}_{m+2},\ldots,\vec{\Delta}_{2m}) = (\vec{\Delta}_{2m+1},\vec{\Delta}_{2m+2},\ldots,\vec{\Delta}_{3m}) = \ldots$$

So if we have m_0 observations, we could think of an infinite sequence of experiments, where the whole block of m_0 experiments is repeated infinitely many times. Then we would have, for instance,

$$(2.3.6) \qquad \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} a_{ij} = \frac{1}{m_0} \sum_{i=1}^{m_0} a_{ij} = a_{\circledast j} \, ,$$

etcetera. In this way, all the limiting values are equal to the values in the finite case. It may be expected that the approximations derived from the asymptotic distributions are then fairly good.

We leave it to the reader to verify that under assumption 1,

$$(2.3.7) \qquad \lim_{m\to\infty} D(\vec{\underline{t}}_* | H_0) = \lim_{m\to\infty} \Sigma_{0\bullet} = \Sigma_0 \, , \qquad \text{(say)},$$

exists. Under assumptions 1 & 2

$$(2.3.8) \qquad \lim_{m\to\infty} D(\vec{\underline{t}}_*) = \lim_{m\to\infty} \Sigma_{1\bullet} = \Sigma_1 \, , \qquad \text{(say)},$$

exists as well as

$$(2.3.9) \qquad \lim_{m\to\infty} E(\vec{\underline{t}}_*) = \lim_{m\to\infty} \vec{\delta}_* = \vec{\delta} \, ,$$

where the components of $\vec{\delta}$ may be $+\infty$ or $-\infty$. Also

$$(2.3.10) \qquad \lim_{m\to\infty} K_\bullet = K \, , \qquad \text{(say)},$$

exists, and we have

$$(2.3.11) \qquad \Sigma_0 = N \otimes K.$$

The alternatives satisfying assumption 2 determine a subset A of $\mathcal{D}$, the set of all possible alternatives, with (cf.(2.1.17)),

$$(2.3.12) \qquad \mathcal{D} = \mathop{\times}_{i=1}^{m} \mathcal{D}_i \, ,$$

for infinite sequences of experiments. So we have

$$(2.3.13) \qquad A = \{d \in \mathcal{D} | d \text{ satisfies assumption 2}\}.$$

It is convenient to split up A still further. Define

$$(2.3.14)\quad A_1 \overset{d}{=} \{a \in A \mid \vec{\delta}'\vec{\delta} = \infty\},$$

$$(2.3.15)\quad A_2 \overset{d}{=} \{a \in A \mid 0 < \vec{\delta}'\vec{\delta} < \infty\},$$

$$(2.3.16)\quad A_3 \overset{d}{=} \{a \in A \mid \vec{\delta}'\vec{\delta} = 0\}.$$

Note that for alternatives in $A_2 \cup A_3$ we have

$$(2.3.17)\quad \vec{\zeta} = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \vec{\delta}_i = \vec{0}\ .$$

When an expectation is taken with respect to a particular alternative $a \in A$, we shall sometimes write E_a, and we shall write E_0 for the expectation under H_0.

We shall establish the following results:

i. When Q is non-singular, $\underline{v}(Q)$ is consistent against each alternative in A_1. When Q is singular, $\underline{v}(Q)$ may, or may not, be consistent against each alternative in A_1. In a number of cases (depending on the structure of Q and the particular alternative) it can be shown that the asymptotic distribution of $\underline{v}(Q)$ is the (standard) normal distribution (after a proper transformation).

ii. For alternatives in A_2, $\underline{v}(Q)$ has asymptotically a non-central χ^2-distribution, or the distribution of a linear combination of independent non-central χ^2-variables. The test based on $\underline{v}(Q)$ is not consistent, but its asymptotic power may still considerably exceed the level of significance.

iii. For alternatives in A_3, $\underline{v}(Q)$ has asymptotically a central χ^2-distribution or the distribution of a linear combination of central χ^2-variables. The test is not consistent and the asymptotic power remains close to the level of significance.

2.4. CONTIGUOUS ALTERNATIVES

In order to find an asymptotic expression for the power function, we need to consider contiguous alternatives, that is, alternatives that converge to H_0 at a certain rate.

Let $a \in A_1$ be a fixed alternative, and let this alternative be determined by the sequence of vectors $\vec{\Delta}_1, \vec{\Delta}_2, \ldots$. Define for each θ, with $0 \le \theta \le 1$, a_θ as the alternative that is determined by the sequence of vectors

(2.4.1) $\theta\vec{\Delta}_1, \theta\vec{\Delta}_2, \theta\vec{\Delta}_3, \ldots$.

Clearly

(2.4.2) $a_\theta \in A_1$ when $0 < \theta \le 1$.

Let

(2.4.3) $\Theta \overset{d}{=} [0,1], \qquad \Theta' \overset{d}{=} (0,1]$.

Next, consider a sequence $\{\theta_m\}_{m=1}^{\infty}$ of values in Θ' such that

(2.4.4) $m\cdot\theta_m^2 \to \eta \ge 0, \qquad$ as $m \to \infty$.

We shall call the sequence of alternatives

(2.4.5) $\{a_{\theta_m}\}_{m=1}^{\infty}$

a contiguous alternative to H_0.

Quantities associated with a contiguous alternative will carry a superscript c. In this way we have the following straightforward results

$$\lim_{m\to\infty} \delta^{c\,(\nu,\mu)}_{\cdot j1} = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, t_{i1}^{(\mu)}(\pi_{ir})\, \theta_m \Delta_{ir} =$$

$$= \lim_{m\to\infty} \theta_m\, \delta_{\cdot j1}^{(\nu,\mu)} = 0.$$

Furthermore,

$$\lim_{m\to\infty} \delta^{c\,(\nu)}_{*j} = \lim_{m\to\infty} \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir})\, \theta_m \Delta_{ir} =$$

$$= \lim_{m\to\infty} \theta_m \sqrt{m}\, \delta_{\cdot j}^{(\nu)} = \sqrt{\eta}\, \zeta_j^{(\nu)}.$$

So

$$\lim_{m\to\infty} \vec{\delta}^{c}_{*} = \sqrt{\eta}\, \vec{\zeta}\ .$$

When we take an expectation or determine a dispersion matrix under a contiguous alternative we shall write E_c and D_c respectively. We have

$$E_c \underline{\vec{t}}_* \to \sqrt{\eta}\, \vec{\zeta}\ , \tag{2.4.6}$$

$$D_c(\underline{\vec{t}}_*) \to \Sigma_0\ . \tag{2.4.7}$$

2.5. AN IMPORTANT PRACTICAL CASE

In this section we describe a class of experimental situations which is important for practical purposes and which can be reduced, by imposing a condition, to the situation of section 2.1.

We consider, again, a sequence

$$E' \stackrel{d}{=} (E_1', E_2', \ldots, E_m') \tag{2.5.1}$$

of m independent experiments. The result of each experiment is a word of length n consisting of characters from the fixed set $\{C_1, \ldots, C_k\}$, each of these characters being available for each place. The number of possible words thus is k^n and the set Ω_i' of these words is the same for all i. The set of possible outcomes for E' then is

$$\Omega' \stackrel{d}{=} \Omega_1' \times \ldots \times \Omega_m'\ . \tag{2.5.2}$$

For each i, $n \times k$ random variables $\underline{x}_{ij}^{(\nu)}$ are defined by means of the functions

$$x_{ij}^{(\nu)}(\omega') \stackrel{d}{=} \begin{cases} 1 \text{ if in } \omega' \ C_j \text{ occurs in the } \nu\text{'th place;} \\ 0 \text{ otherwise,} \end{cases} \qquad (\omega' \in \Omega_i) \tag{2.5.3}$$

and probabilities

$$p_{ij}^{(\nu)} \stackrel{d}{=} P(\underline{x}_{ij}^{(\nu)} = 1). \tag{2.5.4}$$

Thus in the random vector $(\underline{x}_{i1}^{(\nu)}, \ldots, \underline{x}_{ik}^{(\nu)})'$ one of the components assumes the value 1 (indicating the ν'th character of the word) and the others the value 0.

Now, moreover, the experimental situation considered implies that the characters of each of the words are chosen independently: denoting the random element of Ω_i' by $\underline{\omega}_i'$ this means that the probability of the word $(C_{j_1},\dots,C_{j_n})$ is equal to

$$(2.5.5) \qquad P(\underline{\omega}_i' = (C_{j_1},\dots,C_{j_n})) = \prod_{\nu=1}^{n} p_{ij_\nu}{}^{(\nu)} = \prod_{\nu=1}^{n}\prod_{j=1}^{k} \{p_{ij}{}^{(\nu)}\}^{x_{ij}{}^{(\nu)}(\omega_i')}$$

with

$$(2.5.6) \qquad \omega_i' = (C_{j_1},\dots,C_{j_n}).$$

For the whole sequence E' we get, with $\omega' = (\omega_1',\dots,\omega_m') \in \Omega'$

$$(2.5.7) \qquad P(\underline{\omega}' = \omega') = \prod_{i=1}^{m}\prod_{\nu=1}^{n}\prod_{j=1}^{k} \{p_{ij}{}^{(\nu)}\}^{x_{ij}{}^{(\nu)}(\omega_i')}.$$

The number of parameters in this model is so large that reduction is imperative. This can be achieved by imposing a condition of the following character. Let

$$(2.5.8) \qquad \underline{a}_{ij} \overset{d}{=} \sum_{\nu=1}^{n} \underline{x}_{ij}{}^{(\nu)}$$

then the condition is

$$(2.5.9) \qquad A \overset{d}{=} \forall_{ij} \quad \underline{a}_{ij} = a_{ij},$$

where, in applications, the a_{ij} are the values assumed by $\underline{a}_{ij}$. (See remark 2.5.1.). Applying A, the set Ω' is reduced to its subset Ω given by (2.1.6) and (2.1.7) and we have for all i and $\pi_{ir} \in \Omega_i$

$$(2.5.10) \qquad P(\underline{\omega}_i' = \pi_{ir}|A) = \frac{\prod_{j=1}^{k}\prod_{\nu=1}^{n} \{p_{ij}{}^{(\nu)}\}^{x_{ij}{}^{(\nu)}(\pi_{ir})}}{\sum_{s=1}^{N_i}\prod_{j=1}^{k}\prod_{\nu=1}^{n} \{p_{ij}{}^{(\nu)}\}^{x_{ij}{}^{(\nu)}(\pi_{is})}}$$

If we call this p_{ir}, as in (2.1.10), the conditional situation is identical to the situation in section 2.1., with

$$(2.5.11) \qquad \underline{t}_{ij}{}^{(\nu)} \equiv (\underline{x}_{ij}{}^{(\nu)}|A),$$

and the parameters Δ_{ir} and $\delta_{ijl}{}^{(\nu,\mu)}$ are functions of the $p_{ij}{}^{(\nu)}$ and of the a_{ij} from (2.5.9). We can now apply the methods of section 2.1.; in particular we can test the hypothesis

$$(2.5.12) \quad H_0^*: \quad \forall_{ij} \quad p_{ij}^{(1)} = \ldots = p_{ij}^{(n)}$$

against the alternative

$$(2.5.13) \quad H_1^*: \quad \exists_{j\nu} \quad \left| \sum_{i=1}^{m} (p_{ij}^{(\nu)} - \frac{1}{n} \sum_{\nu=1}^{n} p_{ij}^{(\nu)}) \right| > 0 \ ,$$

using one of the test-statistics from section 2.1. This is possible because evidently $H_0^* \Rightarrow H_0$: under A and H_0^* all words in Ω_i have the same probability. It is less evident that $H_1^{*\prime}$ corresponds to H_1' of (2.1.34), because of the complicated character of the $\delta_{ij}^{(\nu)}$. They depend not only, in a rather complicated way, on the $p_{ij}^{(\nu)}$, but also on the a_{ij} and are, therefore, in the unconditional situation random variables. As a matter of fact, these random variables have not even been defined yet. To remedy this omission, we start from (2.1.26), which can now be written as

$$(2.5.14) \quad \delta_{ij}^{(\nu)} = P(\underline{t}_{ij}^{(\nu)} = 1) - \frac{a_{ij}}{n} = P(\underline{x}_{ij}^{(\nu)} = 1|A) - \frac{a_{ij}}{n} =$$
$$= E(\underline{x}_{ij}^{(\nu)}|a_{ij}) - \frac{a_{ij}}{n} \ .$$

So if we define

$$(2.5.15) \quad \underline{\delta}_{ij}^{(\nu)} \stackrel{d}{=} E(\underline{x}_{ij}^{(\nu)}|\underline{a}_{ij}) - \frac{\underline{a}_{ij}}{n} \ ,$$

we have, as we should

$$(2.5.16) \quad \underline{\delta}_{ij}^{(\nu)}|A \equiv \delta_{ij}^{(\nu)} \ ,$$

while moreover

$$(2.5.17) \quad E\underline{\delta}_{ij}^{(\nu)} = EE(\underline{x}_{ij}^{(\nu)}|\underline{a}_{ij}) - \frac{E\underline{a}_{ij}}{n} = E\underline{x}_{ij}^{(\nu)} - \frac{E\underline{a}_{ij}}{n} =$$
$$= p_{ij}^{(\nu)} - \frac{1}{n} \sum_{\nu=1}^{n} p_{ij}^{(\nu)} \ .$$

Now formula (2.1.34) for H_1' is later justified by the consistency of the test for (2.1.42). The analogon of (2.1.42) is now

$$(2.5.18)\qquad \exists_{j\nu}\quad \left|\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\left(p_{ij}^{(\nu)}-\frac{1}{n}\sum_{\nu=1}^{n}p_{ij}^{(\nu)}\right)\right|\to\infty\quad\text{for } m\to\infty$$

and if this condition holds we have according to (2.5.17)

$$(2.5.19)\qquad \exists_{j\nu}\quad \left|\frac{1}{\sqrt{m}}\sum_{i=1}^{m}E\underline{\delta}_{ij}^{(\nu)}\right|\to\infty\ .$$

Since, however, the $\underline{\delta}_{ij}^{(\nu)}$ only assume values between -1 and 1 and since, for any fixed j and ν, $\underline{\delta}_{1j}^{(\nu)}$, $\underline{\delta}_{2j}^{(\nu)}$, ... are independent this means that (see theorem 4.5.6.)

$$(2.5.20)\qquad \exists_{j\nu}\quad \left|\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\underline{\delta}_{ij}^{(\nu)}\right|\overset{P}{\to}\infty\quad\text{for } m\to\infty.$$

Thus, the consistency of the test based on test-statistic (2.1.39) for the unconditional situation when (2.5.18) is satisfied, follows from the consistency in the conditional situation based on (2.1.42).

For future use we now give analogous definitions for $\underline{\delta}_{ij}^{(\nu,\mu)}$ and $\underline{\delta}_{ijl}^{(\nu,\mu)}$, based on (2.1.27) and (2.1.28),

$$(2.5.21)\qquad \underline{\delta}_{ij}^{(\nu,\mu)}\overset{d}{=}E(\underline{x}_{ij}^{(\nu)}\underline{x}_{ij}^{(\mu)}|\underline{a}_{ij})-\frac{\underline{a}_{ij}(\underline{a}_{ij}-1)}{n(n-1)}\qquad(\nu\neq\mu),$$

$$(2.5.22)\qquad \underline{\delta}_{ijl}^{(\nu,\mu)}\overset{d}{=}E(\underline{x}_{ij}^{(\nu)}\underline{x}_{il}^{(\mu)}|\underline{a}_{ij},\underline{a}_{il})-\frac{\underline{a}_{ij}\underline{a}_{il}}{n(n-1)}\qquad(j\neq l,\nu\neq\mu).$$

We have

$$(2.5.23)\qquad E\underline{\delta}_{ij}^{(\nu,\mu)}=\qquad(\nu\neq\mu),$$

$$=p_{ij}^{(\nu)}p_{ij}^{(\mu)}-\frac{1}{n(n-1)}\left\{\left(\sum_{\mu=1}^{n}p_{ij}^{(\mu)}\right)^2-\sum_{\mu=1}^{n}(p_{ij}^{(\mu)})^2\right\},$$

$$(2.5.24)\qquad E\underline{\delta}_{ijl}^{(\nu,\mu)}=\qquad(j\neq l,\nu\neq\mu),$$

$$=p_{ij}^{(\nu)}p_{il}^{(\mu)}-\frac{1}{n(n-1)}\left\{\left(\sum_{\nu=1}^{n}p_{ij}^{(\nu)}\right)\left(\sum_{\mu=1}^{n}p_{il}^{(\mu)}\right)-\sum_{\nu=1}^{n}p_{ij}^{(\nu)}p_{il}^{(\nu)}\right\}.$$

with

$$(2.5.25)\qquad E\underline{\delta}_{ij}^{(\nu,\mu)}=E\underline{\delta}_{ijj}^{(\nu,\mu)}.$$

Sometimes the unconditional analogon of (2.1.39), namely

$$\frac{1}{m}\sum_{j=1}^{k} g_j \sum_{\nu=1}^{n}\{\sum_{i=1}^{m}(\underline{x}_{ij}^{(\nu)} - \frac{\underline{a}_{ij}}{n})\}^2, \tag{2.5.26}$$

which we shall call $\underline{w}(G)$, or more generally $\underline{w}(Q)$ (cf.(2.2)), will be considered. (The factor $\frac{1}{m}$, as in (2.2.11), serves asymptotic purposes; for a test-statistic such a factor is irrelevant). The distribution of $\underline{w}(Q)$ depends on the nuisance parameters $p_{ij}^{(\nu)}$ and is therefore unknown, also under H_0^*. This makes (2.5.26) unfit to be used as an unconditional test-statistic. The asymptotic distribution of $\underline{w}(Q)$ is, under certain conditions, nevertheless the same as that of $\underline{v}(Q)$.

<u>REMARK 2.5.1</u>. In section 2.1. the conditions in (2.1.3) imply that $a_{ij} = n$ never occurs. In the unconditional experimental situation the probability of such an occurrence is equal to

$$\sum_{j=1}^{k}\prod_{\nu=1}^{n} p_{ij}^{(\nu)} \tag{2.5.27}$$

and is thus positive.

It is clear that such an experiment, where all characters of a word are the same, cannot contribute to finding differences between the $p_{ij}^{(\nu)}$ and that the experiment is then useless and had better be left out of consideration.

What is the effect of the deletion of such observations? To obtain m 'useful' observations a random number $\underline{\ell}_m$ of observations will have to be taken, i.e. a sequence

$$E' = (E_1', E_2', \ldots, E_{\underline{\ell}_m}') \tag{2.5.28}$$

of experiments has to be performed. Let

$$\underline{\ell}_i \overset{d}{=} \text{number of the i'th 'useful' experiment,} \tag{2.5.29}$$

where a 'useful' experiment is an experiment which does not result in an outcome where all characters are equal.

Deletion of "useless" experiments yields the sequence

$$E'' = (E_{\underline{\ell}_1}', E_{\underline{\ell}_2}', \ldots, E_{\underline{\ell}_m}') \tag{2.5.30}$$

of m independent experiments. Let

$$(2.5.31)\qquad I \overset{d}{=} (\underline{\iota}_1 = \iota_1 \wedge \underline{\iota}_2 = \iota_2 \wedge \ldots \wedge \underline{\iota}_m = \iota_m).$$

We shall now impose the condition I, where in applications, as in (2.5.9), the ι_i are the values assumed by the $\underline{\iota}_i$. Thus, given I, we consider the sequence

$$(2.5.32)\qquad E''(I) = (E'_{\iota_1}, E'_{\iota_2}, \ldots, E'_{\iota_m})$$

of m independent experiments.

The i'th experiment of $E''(I)$, $E''_i(I)$, has as set of possible outcomes

$$(2.5.33)\qquad \Omega''_i \overset{d}{=} \Omega'_{\iota_i} \setminus \{(c_1,\ldots,c_1), (c_2,\ldots,c_2),\ldots, (c_k,\ldots,c_k)\},$$

i.e. Ω''_i is the same set for each i. The set of possible outcomes for $E''(I)$ is then of course

$$(2.5.34)\qquad \Omega'' = \Omega''_1 \times \Omega''_2 \times \ldots \times \Omega''_m ,$$

while the conditional probabilities are of course proportional to the unconditional ones (cf.(2.5.5)),

$$(2.5.35)\qquad P(\underline{\omega}''_i = (c_{j_1},\ldots,c_{j_n})) = P(\underline{\omega}'_{\iota_i} = (c_{j_1},\ldots,c_{j_n})\,|\,I) = \\ = \frac{\prod_{\nu=1}^{n} p_{\iota_i j_\nu}^{(\nu)}}{1 - \sum_{j=1}^{k} \prod_{\nu=1}^{n} p_{\iota_i j}^{(\nu)}} .$$

Because, after conditioning on I, the probabilities of the 'deleted' experiments have become irrelevant, we may as well renumber the sequence of experiments, i.e. we shall replace ι_i by i throughout, in particular

$$(2.5.36)\qquad p_{\iota_i j}^{(\nu)} \to p_{ij}^{(\nu)} .$$

If we then write

(2.5.37) $$p_i \overset{d}{=} 1 - \sum_{j=1}^{k} \prod_{\nu=1}^{n} p_{ij}^{(\nu)},$$

and define $x_{ij}^{(\nu)}: \Omega_i'' \to \{0,1\}$ as in (2.5.3), we can write (2.5.35) as

(2.5.38) $$P(\underline{\omega}_i'' = \omega_i) = \frac{1}{p_i} \sum_{\nu=1}^{n} \sum_{j=1}^{k} \{p_{ij}^{(\nu)}\}^{x_{ij}^{(\nu)}(\omega_i)}.$$

A further conditioning on A reduces the set Ω'' to its subset Ω and we have for all i and $\pi_{ir} \in \Omega_i$

(2.5.39) $$P(\underline{\omega}_i'' = \pi_{ir} | A) = \frac{\frac{1}{p_i} \prod_{\nu=1}^{n} \prod_{j=1}^{k} \{p_{ij}^{(\nu)}\}^{x_{ij}^{(\nu)}(\pi_{ir})}}{\sum_{s=1}^{N_i} \frac{1}{p_i} \prod_{\nu=1}^{n} \prod_{j=1}^{k} \{p_{ij}^{(\nu)}\}^{x_{ij}^{(\nu)}(\pi_{is})}} = \frac{\prod_{\nu=1}^{n} \prod_{j=1}^{k} \{p_{ij}^{(\nu)}\}^{x_{ij}^{(\nu)}(\pi_{ir})}}{\sum_{s=1}^{N_i} \prod_{\nu=1}^{n} \prod_{j=1}^{k} \{p_{ij}^{(\nu)}\}^{x_{ij}^{(\nu)}(\pi_{is})}}.$$

This is quite the same as (2.5.10), so from here on we may proceed as from (2.5.10) on, the only difference being that the $p_{ij}^{(\nu)}$ are now not the original ones, because some experiments have been deleted. This has no influence, however, on the consistency of the test, because when (2.5.18) holds, it also holds for a sub-sequence. Furthermore it would be nice if for each finite m,

(2.5.40) $$P(\underline{\ell}_m < \infty) = 1.$$

The reader is referred to existing probability theory on this problem.

We conclude that we may safely delete observations for which $\underline{a}_{ij} = n$.

<u>REMARK 2.5.2</u>. Similarly, a category which does not occur in any of the experiments should (and can) be left out of consideration.

CHAPTER 3

SURVEY OF THEOREMS USED

3.1. DEFINITIONS AND THEOREMS ABOUT MATRICES

Consider a $q \times s$ matrix A of any rank r. A *generalised inverse* (or a *g-inverse*) of A is a $s \times q$ matrix, denoted by A^-, that satisfies

(3.1.1) $\quad A A^- A = A.$

If A^- furthermore satisfies

(3.1.2) $\quad A^- A A^- = A^-$

then A^- is called a *reflexive generalised inverse* of A. The notion of g-inverse is discussed extensively in RAO (1973).

We shall apply the notion of g-inverse in particular to real symmetric matrices. Let A be a real symmetric matrix of order q. By

(3.1.3) $\quad A = P \Lambda P'$

we denote the *canonical reduction* of A. That is, Λ is the diagonal matrix of eigenvalues of A and P is the matrix of the corresponding eigenvectors of A. Because A is real and symmetric, the eigenvalues are all real, and we shall always suppose that they occur in decreasing order ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q$). Furthermore, we shall always take the eigenvectors orthonormal, i.e.

(3.1.4) $\quad P'P = PP' = I_q.$

Now let A moreover be non-negative definite (n.n.d) and let it have rank r. (Dispersion matrices are non-negative definite). Then A has precisely r positive eigenvalues and zero as eigenvalue with multiplicity q-r. Let Λ_+ be the $r \times r$ diagonal matrix of the first r (positive) eigenvalues of

A, and P_+ the $q \times r$ matrix of corresponding eigenvectors. Partition the matrices Λ and P as follows

$$(3.1.5) \qquad \Lambda = \begin{pmatrix} \Lambda_+ & 0 \\ 0 & 0 \end{pmatrix},$$

$$(3.1.6) \qquad P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix},$$

where P_{11} has order r, so

$$(3.1.7) \qquad P_+ = \begin{pmatrix} P_{11} \\ P_{21} \end{pmatrix}.$$

It can easily be verified, using these partitionings, that

$$(3.1.8) \qquad A = P\Lambda P' = P_+\Lambda_+P'_+$$

We shall call $P_+\Lambda_+P'_+$ the *positive canonical reduction* of A.

REMARK 3.1.1. Although we speak about "the" canonical reduction of A, the reduction is not entirely unique. Any eigenvector $\vec{p}$ from P may, for instance, be replaced by $-\vec{p}$, without affecting the identity $A = P\Lambda P'$. Or when A has two equal eigenvalues, the corresponding eigenvectors may be exchanged, so that P is changed but not $P\Lambda P'$. The reduction is unique when all the eigenvalues of A are distinct and positive, and when we make the diagonal elements of P positive (SRIVASTAVA & KHATRI (1979), p.19).

Because the implications of this non-uniqueness are minor, we shall maintain the terminology. (The question arises again only in theorem 3.2.1. of the next section).

Because in the positive canonical reduction of A, the column vectors in P_+ are still orthonormal, we have

$$(3.1.9) \qquad P'_+P_+ = I_r .$$

A natural way to obtain a g-inverse of A is then

$$(3.1.10) \qquad A^- = P_+\Lambda_+^{-1}P'_+ ,$$

because it apparently satisfies (3.1.1). A^- is even reflexive because it also satisfies (3.1.2). We shall call g-inverses defined as in (3.1.10) *natural generalised inverses* or *ng-inverses*.

Any $q \times s$ matrix B such that

$$(3.1.11) \qquad A = BB'$$

will be called a *square-root* of A. The set of square roots of A is non-empty as follows from the following lemma.

<u>LEMMA 3.1.1</u>. *Any real, symmetric, n.n.d matrix* A *has at least one square-root.*

<u>PROOF</u>. Let $A = P_+ \Lambda_+ P_+'$ be the positive canonical reduction of A. Let

$$(3.1.12) \qquad L \overset{d}{=} \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_r} \end{pmatrix} .$$

and

$$(3.1.13) \qquad B \overset{d}{=} P_+ L \ .$$

Then $BB' = P_+ LLP_+' = P_+ \Lambda_+ P_+' = A$, so B is the required square root. □

The square root of A defined by (3.1.13) will be called a *natural square-root* of A; it is a $q \times r$ matrix; r is the smallest value which s can assume in a square-root B. Square-roots are uniquely determined, up to an orthonormal transformation:

<u>LEMMA 3.1.2</u>. *Let* A *be a real, symmetric n.n.d matrix of order* q. *Let* $B: q \times s_1$ *and* $C: q \times s_2$ *be two square-roots of* A, *with* $s_2 \geq s_1$. *Then there exists an orthonormal matrix* $U: s_2 \times s_2$, *such that*

$$(3.1.14) \qquad C = (B\ 0)U.$$

<u>PROOF</u>. SRIVASTAVA & KHATRI (1979), p.20. □

LEMMA 3.1.3. *Let* A *be a real, symmetric n.n.d matrix of order* q. *Let* Q *be a* $q \times q$ *matrix. Let* B *be a square-root of* A. *Then the non-zero eigenvalues of* QA *and* B'QB *are the same.*

PROOF. Let λ be a non-zero eigenvalue of QA with eigenvector $\vec{p}$. Then $QA\vec{p} = \lambda\vec{p} \Rightarrow QBB'\vec{p} = \lambda\vec{p} \Rightarrow B'QB(B'\vec{p}) = \lambda(B'\vec{p})$. It follows that λ is also an eigenvalue of B'QB, with eigenvector $B'\vec{p}$.

Now, let μ be a non-zero eigenvalue of B'QB, with eigenvector $\vec{q}$. Then $B'QB\vec{q} = \mu\vec{q} \Rightarrow QBB'(QB\vec{q}) = \mu(QB\vec{q}) \Rightarrow QA(QB\vec{q}) = \mu(QB\vec{q})$. So μ is also an eigenvalue of QA, with eigenvector $QB\vec{q}$. □

COROLLARY 3.1.1. *Let* Q *and* A *be real, symmetric, n.n.d matrices of order* q. *Let* B *and* C *be arbitrary square-roots of* A. *Then*

i. *The non-zero eigenvalues of* B'QB, C'QC *and* QA *are the same.*

ii. rank(B'QB) = rank(C'QC) = r, *where* r = *number of non-zero eigenvalues of* QA.

PROOF. *i.* Follows directly from lemma 3.1.3.

ii. The matrix B'QB is a real, symmetric, n.n.d matrix. (It is the dispersion matrix of $B'\underline{\vec{x}}$, when $\underline{\vec{x}}$ has a q-variate distribution with dispersion matrix Q). The eigenvalues of B'QB are then non-negative. The rank of B'QB is moreover equal to the number of non-zero eigenvalues. The rest of the statement follows now from *i.* □

REMARK 3.1.1. It follows from corollary 3.1.1. that, when the eigenvalues of QA have to be calculated, we can always take the natural square-root of A, B, and calculate the eigenvalues of B'QB. When A is not of full rank, B'QB is of smaller order than QA. It may then be easier to compute the eigenvalues of B'QB instead of those of QA. Moreover, it may be an advantage that B'QB is symmetric while QA is not.

The following $s \times q$ matrix is often convenient

$$(3.1.15) \qquad D_{s,q} \overset{d}{=} (I_s \; 0_{q-s}) \qquad q > s.$$

Note that, for instance

$$(3.1.16) \qquad D_{r,q} \Lambda D'_{r,q} = \Lambda_+.$$

We shall also use the following $(n+k)\times(n\cdot k)$ matrix of rank $n+k-1$*),

$$(3.1.17)\qquad F = \begin{pmatrix} 11\ldots1 & 00\ldots0 & \ldots & 00\ldots0 \\ 00\ldots0 & 11\ldots1 & \ldots & 00\ldots0 \\ \vdots & \vdots & \ddots & \vdots \\ 00\ldots0 & 00\ldots0 & \ldots & 11\ldots1 \\ I_k & I_k & \ldots & I_k \end{pmatrix}.$$

The linear space spanned by the column's of a matrix X will be denoted by $M(X)$ and the linear space of solutions of the equation $X\vec{z} = \vec{0}$, the null-space of X, will be denoted by $N(X)$. Note that

$$(3.1.18)\qquad \vec{\delta}_i \in N(F), \qquad \vec{\delta}_* \in N(F),$$

where $\vec{\delta}_i$ and $\vec{\delta}_*$ were defined in section 2.2.

LEMMA 3.1.4. *Let* A *be a* $k\times k$ *matrix and* B *an* $n\times n$ *matrix. Let the* ρ*'th eigenvalue of* A *be* λ_ρ *and a corresponding eigenvector* $(p_{1\rho},\ldots,p_{k\rho})'$*; let the* τ*'th eigenvalue of* B *be* μ_τ *and a corresponding eigenvector* $\vec{q}_\tau$*. Then the set of eigenvalues of the matrix* $A\otimes B$ *is equal to*

$$\{x \mid x = \lambda_\rho\mu_\tau,\ \rho = 1,\ldots,k,\ \tau = 1,\ldots,n\}.$$

An eigenvector corresponding to the eigenvalue $\lambda_\rho\mu_\tau$ *is*

$$(p_{1\rho}\vec{q}_\tau',\ p_{2\rho}\vec{q}_\tau',\ \ldots\ ,\ p_{k\rho}\vec{q}_\tau').$$

PROOF. ANDERSON (1958), p.348. □

Note that in particular it follows from lemma 3.1.4. that the eigenvalues of $A\otimes B$ and $B\otimes A$ are the same.

LEMMA 3.1.5. *Let* A *be a real, symmetric, n.n.d matrix of order* q *and rank* r, *and let* $\lambda_1,\ldots,\lambda_r$ *be the positive eigenvalues of* $\frac{n}{n-1}A$. *With* N *as defined by (2.2.28), the eigenvalues of* $A\otimes N$ *then are* $\lambda_1,\ldots,\lambda_r$*, all > 0, each with multiplicity* $n-1$*, and 0 with multiplicity* $q+(n-1)(q-r)$.

---

*) The symbols n and k have, in later applications, the same meaning as in chapter 2; the symbol q will usually be $n\cdot k$.

PROOF. The eigenvalues of N are 0 and $\frac{n}{n-1}$ with multiplicity n-1. The result now follows from lemma 3.1.4. □

LEMMA 3.1.6. *Let* A *be a real, symmetric, n.n.d matrix of order* q *with eigenvalues* $\lambda_1,\ldots,\lambda_q$. *Then*

$$\lambda_q \le \frac{\vec{x}'A\vec{x}}{\vec{x}'\vec{x}} \le \lambda_1 \tag{3.1.19}$$

for each $q\times 1$ *vector* $\vec{x}'$.

PROOF. SRIVASTAVA & KHATRI (1979), p.21. □

LEMMA 3.1.7. *Let* A *be a real, symmetric, n.n.d matrix of order* q. *If* $(\vec{x}_m)_{m=1}^{\infty}$ *is a sequence of vectors of* q *components such that* $\vec{x}_m'\vec{x}_m \to \infty$ *as* $m\to\infty$, *then*

$$\lim_{m\to\infty} \frac{\vec{x}_m'\vec{x}_m}{(\vec{x}_m' A \vec{x}_m)^{\frac{1}{2}}} = \infty . \tag{3.1.20}$$

PROOF. Let $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_q \ge 0$ be the eigenvalues of A. If $\lambda_1 = \lambda_2 = \ldots = \lambda_q = 0$ then $\vec{x}_m' A \vec{x}_m = 0$ for each m and (3.1.20) follows immediately. So suppose that $\lambda_1 > 0$. From lemma 3.1.6. it follows that, for each m,

$$\vec{x}_m' A \vec{x}_m \le \lambda_1 \vec{x}_m'\vec{x}_m, \qquad \text{or,}$$

$$(\lambda_1 \vec{x}_m'\vec{x}_m)^{-\frac{1}{2}} \le (\vec{x}_m' A \vec{x}_m)^{-\frac{1}{2}}, \qquad \text{or,}$$

$$\frac{(\vec{x}_m'\vec{x}_m)^{\frac{1}{2}}}{\sqrt{\lambda_1}} \le \frac{\vec{x}_m'\vec{x}_m}{(\vec{x}_m' A \vec{x}_m)^{\frac{1}{2}}} .$$

Because the lefthand-side of this inequality diverges to ∞, the righthand-side necessarily also diverges to ∞. □

3.2. DISTRIBUTION OF QUADRATIC FORMS IN NORMAL VARIATES

Various theorems are known about the distribution of quadratic forms in normal variates. However, most of the theorems concern necessary and sufficient conditions for such a quadratic form to be distributed as a (non-) central χ^2-distribution. The following theorem gives the distribution of a n.n.d quadratic form in normal variates for the general case.

This theorem is known for non-singular dispersion matrix Σ. We give a simple proof that includes the case of a singular Σ.

THEOREM 3.2.1. *Let* $\underline{\vec{x}} \sim N_q(\vec{\mu},\Sigma)$. *Let* Q *be a real, symmetric, n.n.d* $q\times q$ *matrix. Then there exist numbers* $r \in \mathbb{N}$, $c \in \mathbb{R}$ *and vectors* $\vec{\lambda} \in \mathbb{R}^r$, $\vec{\omega} \in \mathbb{R}^r$ *such that*

$$(3.2.1) \qquad \underline{\vec{x}}'Q\underline{\vec{x}} \equiv c + \sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \omega_\tau)^2$$

with $\underline{\vec{u}} \sim N_r(\vec{0},I_r)$, *i.e.* $\underline{u}_1,\ldots,\underline{u}_r$ *are independent and each has a standard normal distribution.*

Let B *be an arbitrary square-root of* Σ. *Explicit values of* r, c, $\vec{\lambda}$ *and* $\vec{\omega}$ *may then be calculated from*

$$(3.2.2) \qquad r = \text{rank}(B'QB),$$

$$(3.2.3) \qquad \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r \text{ are the positive eigenvalues of } B'QB,$$

If we furthermore denote the positive canonical reduction of $B'QB$ *by* $P_+\Lambda_+P_+'$, *we have*

$$(3.2.4) \qquad \vec{\omega} = \Lambda_+^{-1}P_+'B'Q\vec{\mu},$$

$$(3.2.5) \qquad c = \vec{\mu}'(Q - QBP_+\Lambda_+^{-1}P_+'B'Q)\vec{\mu}.$$

The values of r, $\lambda_1,\ldots,\lambda_r$ *and* c *are independent of the particular choice of square-root of* Σ.

PROOF. We first prove the existence of the four quantities r, c, $\vec{\lambda}$ and $\vec{\omega}$. When B is a $q\times s$ square-root of Σ ($BB' = \Sigma$), the stochastic vector $\underline{\vec{x}}$ may be represented as $\underline{\vec{x}} \equiv \vec{\mu} + B\underline{\vec{y}}$, with $\underline{\vec{y}} \sim N_s(\vec{0},I_s)$. Then

$$\underline{\vec{x}}'Q\underline{\vec{x}} \equiv (\vec{\mu}+B\underline{\vec{y}})'Q(\vec{\mu}+B\underline{\vec{y}}) \equiv \underline{\vec{y}}'B'QB\underline{\vec{y}} + 2\vec{\mu}'QB\underline{\vec{y}} + \vec{\mu}'Q\vec{\mu}.$$

The matrix $B'QB$ is a real, symmetric, n.n.d matrix, being the dispersion matrix of $B'\underline{\vec{s}}$, when $\underline{\vec{s}} \sim N_q(\vec{0},Q)$. Taking $r = \text{rank}(B'QB)$, it follows that $B'QB$ has exactly r positive eigenvalues and zero as eigenvalue of multiplicity $s-r$. Let the eigenvalues of $B'QB$ be $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_s = 0$. Let $P\Lambda P' = B'QB$ be the canonical reduction of $B'QB$ and let

$P_+\Lambda_+P_+'$ be the positive canonical reduction of B'QB.

We now define $\vec{\underline{v}} \stackrel{d}{=} P'\vec{\underline{y}}$, or equivalently $\vec{\underline{y}} \equiv P\vec{\underline{v}}$, because $PP' = I_s$. Then $E\vec{\underline{v}} = \vec{0}$, and

$$D(\vec{\underline{v}}) = D(P'\vec{\underline{y}}) = P'D(\vec{\underline{y}})P = P'I_sP = P'P = I_s.$$

So $\vec{\underline{v}} \sim N_s(\vec{0},I_s)$. Take $\vec{\underline{u}} \stackrel{d}{=} D_{r,s}\vec{\underline{v}}$, so $\vec{\underline{u}}$ is the vector of the first r coordinates of $\vec{\underline{v}}$. Of course, $\vec{\underline{u}} \sim N_r(\vec{0},I_r)$. It follows that

$$\vec{\underline{y}}'B'QB\vec{\underline{y}} \equiv \vec{\underline{y}}'P\Lambda P'\vec{\underline{y}} \equiv \vec{\underline{v}}'\Lambda\vec{\underline{v}} \equiv \sum_{\tau=1}^{s} \lambda_\tau \underline{v}_\tau^2 \equiv \sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 ,$$

because $\lambda_{r+1} = \ldots = \lambda_s = 0$.

For the second term, $2\vec{\mu}'QB\vec{\underline{y}}$, we may write

$$\vec{\mu}'QB\vec{\underline{y}} \equiv \vec{\mu}'QBP\vec{\underline{v}} \equiv \sum_{\tau=1}^{s} d_\tau \underline{v}_\tau ,$$

with $\vec{d}' = \vec{\mu}'QBP$.

Let $\vec{p}_\tau$ be the eigenvector from P corresponding to the eigenvalue λ_τ. The eigenvalues $\lambda_{r+1},\ldots,\lambda_s$ are equal to zero. Therefore $B'QB\vec{p}_\tau = \vec{0}$, for $\tau = r+1,\ldots,s$. Because Q is a real, symmetric, n.n.d matrix, we may write $Q = TT'$ for some square-root T of Q. We then have

$$B'QB\vec{p}_\tau = \vec{0} \Rightarrow \vec{p}_\tau'B'QB\vec{p}_\tau = 0 \Rightarrow \vec{p}_\tau'B'TT'B\vec{p}_\tau = 0 \Rightarrow \|T'B\vec{p}_\tau\| = 0 \Rightarrow$$

$$T'B\vec{p}_\tau = \vec{0} \Rightarrow TT'B\vec{p}_\tau = \vec{0} \Rightarrow QB\vec{p}_\tau = \vec{0}.$$

Therefore, the last s-r column's of the matrix QBP contain only zero's, from which it follows that $d_{r+1} = \ldots = d_s = 0$. So we have

$$\vec{\mu}'QB\vec{\underline{y}} \equiv \sum_{\tau=1}^{s} d_\tau \underline{v}_\tau \equiv \sum_{\tau=1}^{r} d_\tau \underline{u}_\tau.$$

Let $\omega_\tau \stackrel{d}{=} d_\tau/\lambda_\tau$ for $\tau = 1,\ldots,r$, or

$$\vec{\omega} = \Lambda_+^{-1}D_{r,s}\vec{d} = \Lambda_+^{-1}D_{r,s}P'B'Q\vec{\mu} = \Lambda_+^{-1}P_+'B'Q\vec{\mu}.$$

Then it follows that

$$\vec{\mu}'QB\vec{\underline{y}} \equiv \sum_{\tau=1}^{r} \lambda_\tau\omega_\tau\underline{u}_\tau .$$

The first two terms together now give

$$\vec{\underline{y}}'B'QB\vec{\underline{y}} + 2\vec{\mu}'QB\vec{\underline{y}} \equiv \sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 + 2\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau .$$

If we add $\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2$ to both sides, the righthand-side may be written as $\sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \omega_\tau)^2$. But

$$\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2 = \vec{\omega}'\Lambda_+\vec{\omega} = \vec{\mu}'QBP_+\Lambda_+^{-1}\Lambda_+\Lambda_+^{-1}P_+'B'Q\vec{\mu} = \vec{\mu}'QBP_+\Lambda_+^{-1}P_+'B'Q\vec{\mu} .$$

So we have finally

$$\vec{\underline{x}}'Q\vec{\underline{x}} \equiv \sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \omega_\tau)^2 + \vec{\mu}'QBP_+\Lambda_+^{-1}P_+'B'Q\vec{\mu} .$$

This establishes the existence of the quantities c, r, $\vec{\lambda}$ and $\vec{\omega}$, and the formulae (3.2.2)-(3.2.5).

We shall now show that r, c and $\vec{\lambda}$ are independent of the particular choice of B. Let C be another square-root of Σ, and suppose that C is used instead of B. From corollary 3.1.1. it follows that rank(C'QC) = rank(B'QB) = = r, and furthermore that the non-zero eigenvalues of C'QC and B'QB are the same. Furthermore, because $\vec{\underline{x}}'Q\vec{\underline{x}} \geq 0$ and $\sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \omega_\tau)^2 \geq 0$ (with probability one), c is the 'minimum value' that $\vec{\underline{x}}'Q\vec{\underline{x}}$ can assume ($\vec{\underline{u}}$ has a positive density in $-\vec{\omega}$), which is of course independent of the choice of square root of Σ. This completes the proof. □

<u>COROLLARY 3.2.1</u>. *Let* $\vec{\underline{x}} \sim N_q(\vec{0},\Sigma)$. *Let* Q *be a real, symmetric, n.n.d* $q \times q$ *matrix. Then*

(3.2.6) $$\vec{\underline{x}}'Q\vec{\underline{x}} \equiv \sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2$$

with $\vec{\underline{u}} \sim N_r(\vec{0},I_r)$, $\lambda_1,\ldots,\lambda_r$ *the positive eigenvalues of* $Q\Sigma$.

<u>PROOF</u>. This follows from theorem 3.2.1. and lemma 3.1.3. □

<u>REMARK 3.2.1</u>. For the matrix B of theorem 3.2.1., we may take the natural square-root of Σ as defined in section 3.1.

That the vector $\vec{\omega}$ is not necessarily uniquely determined, and does not have to be, may be understood from the (obvious) fact that for r = 3 and $\lambda_1 = \lambda_2$ the r.v.'s
$\underline{w}_1 \overset{d}{=} \lambda_1(\underline{u}_1 + \omega_1)^2 + \lambda_2(\underline{u}_2 + \omega_2)^2 + \lambda_3(\underline{u}_3 + \omega_3)^2$, and

$\underset{\sim}{w}_2 \overset{d}{=} \lambda_1(\underset{\sim}{u}_1 - \omega_2)^2 + \lambda_2(\underset{\sim}{u}_2 + \omega_1)^2 + \lambda_3(\underset{\sim}{u}_3 - \omega_3)^2$

have the same distribution. The non-uniqueness follows from the fact that the canonical reduction is not uniquely determined. (See remark 3.1.1.).

All the same, it is possible to get the same $\vec{\omega}$, even when using two different square-roots of Σ. Suppose that C has been used instead of B, and that C is a $q \times t$ matrix with $t \geq q$. C gives the same r, $\vec{\lambda}$ and c as follows from theorem 3.2.1. By lemma 3.1.2. there exists an orthonormal matrix U such that $C = (B|0)U$, or in greater detail,

$$C = \left(\begin{array}{c|c} B & 0 \end{array}\right)\left(\begin{array}{c|c} U_{11} & U_{12} \\ \hline U_{21} & U_{22} \end{array}\right)\begin{array}{l} s \\ t-s \end{array}$$

$$\begin{array}{cccc} s & t-s & s & t-s \end{array}$$

Define $V' = (U_{11}|U_{12})$. Then $C = BV'$ and $V'V = I_s$. Also $C'QC = VB'QBV' = VP_+\Lambda_+P_+'V'$, where $P_+\Lambda_+P_+'$ is the positive canonical reduction of $B'QB$. It may easily be verified that $VP_+\Lambda_+P_+'V'$ is "a" positive canonical reduction of $C'QC$. If we would have used this reduction for the calculation of $\vec{\omega}$, then:

$$\vec{\omega}_C = \Lambda_+^{-1}P_+'V'C'Q\vec{\mu} = \Lambda_+^{-1}P_+'V'VB'Q\vec{\mu} = \Lambda_+^{-1}P_+'I_sB'Q\vec{\mu} = \Lambda_+^{-1}P_+'B'Q\vec{\mu} = \vec{\omega}_B \ .$$

So in this case the $\vec{\omega}$'s would be the same.

The non-uniqueness of $\vec{\omega}$ is otherwise not very important, because whatever the choice of square-root of Σ and the choice of canonical reduction is, the resulting distribution is of course always the same.

<u>REMARK 3.2.2.</u> When $\underset{\sim}{\vec{x}}$ has a positive density in $\vec{0}$, the minimum value that $\underset{\sim}{\vec{x}}'Q\underset{\sim}{\vec{x}}$ can assume is 0, from which it follows that the constant c, defined by (3.2.5) is equal to 0. Formally this can be verified as follows. When $\underset{\sim}{\vec{x}}$ can assume the value $\vec{0}$, then, in the representation $\underset{\sim}{\vec{x}} \equiv \vec{\mu} + B\underset{\sim}{\vec{y}}$, the vector $\underset{\sim}{\vec{y}}$ can assume a value $\vec{y}$ such that $\vec{\mu} + B\vec{y} = \vec{0}$, or $\vec{\mu} = -B\vec{y}$. It follows that

$$c = \vec{\mu}'(Q - QBP_+\Lambda_+^{-1}P_+'B'Q)\vec{\mu} = +\vec{y}'B'(Q - QBP_+\Lambda_+^{-1}P_+'B'Q)B\vec{y} =$$

$$= \vec{y}'B'QB\vec{y} - \vec{y}'B'QBP_+\Lambda_+^{-1}P_+'\ B'QB\vec{y} =$$

$$= \vec{y}'B'QB\vec{y} - \vec{y}'\underbrace{P_+\Lambda_+P_+'P_+\Lambda_+^{-1}}_{I}P_+'B'QB\vec{y} = \vec{y}'B'QB\vec{y} - \vec{y}'B'QB\vec{y} = 0.$$

REMARK 3.2.3. Although $\vec{\omega}$ does depend, in a way, on the choice of square-root of Σ, certain functions of $\vec{\omega}$ are independent of it. Consider the case $c = 0$. (We have $c = 0$ in our situation). We have

$$E\underline{\vec{x}}'Q\underline{\vec{x}} = \text{trace } Q\Sigma + \vec{\mu}'Q\vec{\mu} = \sum_{\tau=1}^{r} \lambda_\tau + \vec{\mu}'Q\vec{\mu}$$

and

$$E \sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \omega_\tau)^2 = \sum_{\tau=1}^{r} \lambda_\tau + \sum_{\tau=1}^{r} \lambda_\tau\omega_\tau^2.$$

It follows that

(3.2.7) $$\sum_{\tau=1}^{r} \lambda_\tau\omega_\tau^2 = \vec{\mu}'Q\vec{\mu}$$

Alternatively, this may be derived from (3.2.5).

Another interesting identity, which we shall use is

(3.2.8) $$\sum_{\tau=1}^{r} \lambda_\tau^2\omega_\tau^2 = \vec{\omega}'\Lambda_+\Lambda_+\vec{\omega} = \vec{\mu}'QBP_+\Lambda_+^{-1}\Lambda_+\Lambda_+\Lambda_+^{-1}P_+'B'Q\vec{\mu} = \vec{\mu}'QBP_+P_+'B'Q\vec{\mu} =$$

$$= \vec{\mu}'QBB'Q\vec{\mu} = \vec{\mu}'Q\Sigma Q\vec{\mu} .$$

The following theorem is closely related to the preceding one. It gives criteria for (3.2.1) to have a (non-) central χ^2-distribution, without having to calculate eigenvalues and -vectors explicitly. Though it may be found (without proof) in various places, we quote it here for completeness. The proof can be found in RAYNER & LIVINGSTONE (1965). The theorem is also given by RAO (1973). We quote JOHNSON & KOTZ (1970).

THEOREM 3.2.2. *Let* $\underline{\vec{x}} \sim N_q(\vec{\mu},\Sigma)$. *Let* Q *be a real, symmetric* $q \times q$ *matrix and assume that*

(3.2.9) $\text{trace } Q\Sigma = r \neq 0.$

The quadratic form

(3.2.10) $\underline{\vec{x}}' Q \underline{\vec{x}}$

has a (non-) central χ^2*-distribution iff*

(3.2.11) $\Sigma Q \Sigma Q \Sigma = \Sigma Q \Sigma,$

(3.2.12) $\vec{\mu}' Q \Sigma Q \vec{\mu} = \vec{\mu}' Q \vec{\mu},$

(3.2.13) $\Sigma Q \Sigma Q \vec{\mu} = \Sigma Q \vec{\mu}.$

The number of degrees of freedom is then

(3.2.14) $\text{trace } Q\Sigma = r \in \mathbb{N}$

and the non-centrality parameter is

(3.2.15) $\frac{1}{2} \vec{\mu}' Q \vec{\mu}\,.$

PROOF. RAYNER & LIVINGSTONE (1965). □

3.3. THE DISTRIBUTIONS OF $\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2$ AND $\sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \omega_\tau)^2$.

The distributions of $\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2$ and $\sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \omega_\tau)^2$, with $\underline{\vec{u}} \sim N(\vec{0}, I_r)$, $\lambda_1, \ldots, \lambda_r \in \mathbb{R}_+$, $\omega_1, \ldots, \omega_r \in \mathbb{R}$ have been studied extensively. The reader is referred to JOHNSON & KOTZ (1970), and KOTZ, JOHNSON & BOYD (1967a,b).

We list here some of the facts that we used. Let

(3.3.1) $\underline{Q} \overset{d}{=} Q(\underline{\vec{u}}, \vec{\lambda}, \vec{\omega}) \overset{d}{=} \sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \omega_\tau)^2,$

(3.3.2) $\underline{Q}_0 \overset{d}{=} Q(\underline{\vec{u}}, \vec{\lambda}, \vec{0}).$

Moments.

(3.3.3) $E\underline{Q} = \sum_{\tau=1}^{r} \lambda_\tau + \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2 \,,$

$$(3.3.4) \qquad \sigma^2(\underline{Q}) = 2\sum_{\tau=1}^{r}\lambda_\tau^2 + 4\sum_{\tau=1}^{r}\lambda_\tau^2\omega_\tau^2 ,$$

$$(3.3.5) \qquad \mu_3(\underline{Q}) = 8\sum_{\tau=1}^{r}\lambda_\tau^3 + 24\sum_{\tau=1}^{r}\lambda_\tau^3\omega_\tau^2 .$$

The moments of $\underline{Q}_0$ follow by omitting the terms containing ω_τ's. If $\lambda_1 = \lambda_2 = \dots = \lambda_r = 1$, the moments reduce to those of the (non-) central χ^2-distribution.

Asymptotic expansions.

The distribution function of $\underline{Q} \equiv \sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \omega_\tau)^2$,

$$(3.3.6) \qquad F(z;\vec{\lambda};\vec{\omega}) \overset{d}{=} P(\underline{Q} \le z)$$

may be represented in an infinite series of central χ^2-distributions

$$(3.3.7) \qquad F(z;\vec{\lambda};\vec{\omega}) = \sum_{k=0}^{\infty} a_k P(\underline{\chi}^2[r+2k] \le \tfrac{z}{\beta}),$$

with coefficients a_k, recursively defined by

$$(3.3.8) \qquad a_0 \overset{d}{=} \exp\{-\tfrac{1}{2}\sum_{\tau=1}^{r}\omega_\tau^2\}\cdot \prod_{\tau=1}^{r}(\tfrac{\beta}{\lambda_\tau})^{\frac{1}{2}},$$

$$(3.3.9) \qquad a_k \overset{d}{=} \frac{1}{k}\sum_{\ell=0}^{k-1} b_{k-\ell}a_\ell, \qquad k = 1,2,\dots ,$$

$$(3.3.10) \qquad \gamma_\tau \overset{d}{=} 1 - \beta/\lambda_\tau, \qquad \tau = 1,2,\dots,r,$$

$$(3.3.11) \qquad b_k \overset{d}{=} \tfrac{1}{2}k\sum_{\tau=1}^{r}\omega_\tau^2\gamma_\tau^{k-1} + \tfrac{1}{2}\sum_{\tau=1}^{r}(1-k\omega_\tau^2)\gamma_\tau^k, \qquad k = 1,2,\dots ,$$

$\underline{\chi}^2[\nu]$ a r.v. with a χ^2-distribution with ν degrees of freedom, and $\beta > 0$ a suitably chosen constant. The choice of β leaves some room to influence the rate of convergence of the series. A good choice is

$$(3.3.12) \qquad \beta = \frac{2\lambda_1\lambda_r}{\lambda_1+\lambda_r},$$

so that $|\gamma_\tau| < 1$ for $\tau = 1,\dots,r$. In our calculations in chapter 9 we always take β as defined by (3.3.12).

The series is uniformly convergent for any bounded interval of $z > 0$.

A bound for the error ($E_N(z)$) resulting from the truncation of (3.3.7) after the N'th term is given by

$$(3.3.13) \qquad E_N(z) \le a_0 \left\{ \frac{\Gamma(\frac{r}{2}+N+1)}{\Gamma(\frac{r}{2})(N+1)!} \right\} \left\{ \frac{\mu^{N+1}}{(1-\mu)^{r/2+N}} \right\} P(\underline{\chi}^2[r+2N+2] \le \frac{(1-\mu)z}{\beta}),$$

for $0 < \mu < 1$, and

$$(3.3.14) \qquad E_N(z) \le a_0 \frac{z}{\beta} f_{[r]}(\frac{z}{\beta}) \exp\{\frac{z\mu}{2\beta}\} (\frac{z\mu}{2\beta})^{N+1} \frac{1}{(N+1)!}$$

for $\mu > 1$, $f_{[\nu]}$ is the density of the χ^2-distribution with ν degrees of freedom, while

$$(3.3.15) \qquad \mu \overset{d}{=} \frac{1}{2} \sum_{\tau=1}^{r} \frac{\omega_\tau^2 \beta}{\lambda_\tau} + \max \left|1 - \frac{\beta}{\lambda_\tau}\right|.$$

This expansion can fruitfully be used for computer calculation, when a subroutine program for the χ^2-distribution is available. (See also the examples in chapter 9).

Approximations.

When no computer is available, the distribution of $\underline{Q}_0$ may be approximated, using the same method as proposed in chapter 1, by an adapted χ^2-distribution, i.e. the distribution of $b\underline{\chi}^2[\nu]$, where b and ν are chosen to make the first two moments agree with those of $\underline{Q}_0$, i.e.

$$(3.3.15) \qquad b = \left(\sum_{\tau=1}^{r} \lambda_\tau^2 \right) / \left(\sum_{\tau=1}^{r} \lambda_\tau \right),$$

$$(3.3.16) \qquad \nu = \left(\sum_{\tau=1}^{r} \lambda_\tau \right)^2 / \left(\sum_{\tau=1}^{r} \lambda_\tau^2 \right).$$

An improvement is possible if we use $a + b\underline{\chi}^2[\nu]$ instead of $b\underline{\chi}^2[\nu]$. We have in that case,

$$(3.3.17) \qquad a = \sum_{\tau=1}^{r} \lambda_\tau - \left(\sum_{\tau=1}^{r} \lambda_\tau^2 \right)^2 / \left(\sum_{\tau=1}^{r} \lambda_\tau^3 \right) ; \quad b = \left(\sum_{\tau=1}^{r} \lambda_\tau^3 \right) / \left(\sum_{\tau=1}^{r} \lambda_\tau^2 \right),$$

$$(3.3.18) \qquad \nu = \left(\sum_{\tau=1}^{r} \lambda_\tau^2 \right)^3 / \left(\sum_{\tau=1}^{r} \lambda_\tau^3 \right)^2.$$

However, for positive a, this approximation assigns the value 0 to $P(\underline{Q}_0 \le z)$, for all $0 \le z \le a$, so this approximation does not work well for small z.

The most simple approximation to the distribution of $\underline{Q}$ is the distribution of $\underline{\chi}^2[\nu,\delta^2]$ with

$$(3.3.19)\qquad \nu = \sum_{\tau=1}^{r} \lambda_\tau \ ; \qquad \delta^2 = \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2 \ ,$$

so that only the first moments of $\underline{Q}$ and $\underline{\chi}^2[\nu,\delta^2]$ agree. When we take

$$(3.3.20)\qquad \nu = -\tfrac{1}{2}\sigma^2(\underline{Q}) + 2E(\underline{Q}),$$

$$(3.3.21)\qquad \delta^2 = \tfrac{1}{2}\sigma^2(\underline{Q}) - E(\underline{Q}),$$

the first two moments of $\underline{Q}$ and $\underline{\chi}^2[\nu,\delta^2]$ agree.

And when we take $c\underline{\chi}^2[\nu,\delta^2]$ to approximate $\underline{Q}$, with

$$(3.3.22)\qquad c = \frac{2\sigma^2 + \sqrt{4\sigma^4 - 2\mu_3 E}}{4E} \ ,$$

$$(3.3.23)\qquad \delta^2 = \frac{\sigma^2 - 2cE}{2c^2} \ ,$$

$$(3.3.24)\qquad \nu = \frac{E}{c} - \delta^2,$$

then the first three moments of $c\underline{\chi}^2[\nu,\delta^2]$ agree with those of $\underline{Q}$. ($E = E(\underline{Q})$, $\sigma^2 = \sigma^2(\underline{Q})$, $\mu_3 = \mu_3(\underline{Q})$).

The last approximation, however, is only possible when

$$(3.3.25)\qquad 4\sigma^4(\underline{Q}) - 2\mu_3(\underline{Q})E(\underline{Q}) > 0 \ .$$

All these approximations necessitate the use of a table of the non-central χ^2-distribution. The reader is referred to JOHNSON & KOTZ (1970), p.137, for a survey of existing tables and approximations of the non-central χ^2-distribution.

We shall now prove a lemma concerning a slightly more general situation, but which we shall apply to the distribution of $\underline{Q}$.

<u>LEMMA 3.3.1</u>. *Let* $\underline{u}$ *be a r.v. with an absolutely continuous distribution. Suppose that the density* f *of* $\underline{u}$ *is symmetric with respect to 0 and that* f *is strictly decreasing, continuously differentiable and positive on* $[0,\infty)$. *Let* $\underline{u}_1,\ldots,\underline{u}_r$ *be independent and identically distributed as* $\underline{u}$. *Let*

$\lambda_1,\ldots,\lambda_r \in \mathbb{R}_+,\ \omega_1,\ldots,\omega_r \in \mathbb{R},\ \exists_{\tau\in\{1,\ldots,r\}}: \omega_\tau \neq 0.$

The function

$$H_r(t,z) \stackrel{d}{=} P(\sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + t\omega_\tau)^2 \le z), \qquad t \ge 0,$$

is then strictly decreasing in t *on* $[0,\infty)$ *for each* $z \in \mathbb{R}_+$.

PROOF. Without loss of generality we may suppose that $\omega_\tau \ge 0$ for $\tau = 1,\ldots,r$, where at least one of the inequalities is strict. Let, for $\tau = 1,\ldots,r$,

$$\underline{y}_\tau \stackrel{d}{=} \lambda_\tau(\underline{u}_\tau + t\omega_\tau)^2$$

with density $g_\tau(t,y)$ and distribution function $G_\tau(t,y)$, and

$$\underline{z}_r \stackrel{d}{=} \sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + t\omega_\tau)^2$$

with density $h_r(t,z)$ and distribution function $H_r(t,z)$.

We shall prove the lemma by induction. First consider the case $r = 1$. We have

$$H_1(t,z) = G_1(t,z) = P(\lambda_1(\underline{u}_1 + t\omega_1)^2 \le z) =$$

$$= P(-\sqrt{\frac{z}{\lambda_1}} \le \underline{u}_1 + t\omega_1 \le \sqrt{\frac{z}{\lambda_1}}) = \int_{-\sqrt{\frac{z}{\lambda_1}} - t\omega_1}^{\sqrt{\frac{z}{\lambda_1}} - t\omega_1} f(u)du .$$

So

$$\frac{\partial H_1(t,z)}{\partial t} = -\omega_1 f(\sqrt{\frac{z}{\lambda_1}} - t\omega_1) + \omega_1 f(-\sqrt{\frac{z}{\lambda_1}} - t\omega_1) < 0, \qquad \text{for } z > 0,$$

because f is symmetric w.r.t. 0 and strictly decreasing on $[0,\infty)$ and $\omega_1 \neq 0$. It follows that $H_1(t,z)$ is strictly decreasing in t for each fixed $z > 0$.

Next, suppose that $H_r(t,z)$ is strictly decreasing in t for each $z > 0$. Then we have for r+1,

$$H_{r+1}(t,z) = P(\underline{z}_{r+1} \le z) = P(\underline{z}_r + \underline{y}_{r+1} \le z) = P(\underline{z}_r + \underline{y}_1 \le z) =$$

$$= \int_0^z g_1(t,x)H_r(t,z-x)dx.$$

Now we have

$$\frac{\partial H_{r+1}(t,z)}{\partial t} = \int_0^z H_r(t,z-x)\ \frac{\partial}{\partial t}\, g_1(t,x)dx + \int_0^z g_1(t,x)\ \frac{\partial}{\partial t}\, H_r(t,z-x)dx.$$

Now $g_1(t,x) > 0$ for $x \in [0,z]$ and $\frac{\partial}{\partial t} H_r(t,z-x) < 0$ for fixed z-x by the induction hypothesis. It follows that the second term is negative.

For the first term we can write, using partial integration,

$$\int_0^z H_r(t,z-x)\ \frac{\partial}{\partial t}\, g_1(t,x)dx =$$

$$= \left[\ H_r(t,z-x)\ \frac{\partial}{\partial t}\, G_1(t,x)\ \right]_0^z + \int_0^z h_r(t,z-x)\ \frac{\partial}{\partial t}\, G_1(t,x)dx =$$

$$= \qquad 0 \qquad + \int_0^z \underbrace{h_r(t,z-x)}_{>0}\ \underbrace{\frac{\partial}{\partial t}\, G_1(t,x)dx}_{<0} < 0$$

So $H_{r+1}(t,z)$ is also strictly decreasing in t for each $z > 0$. It is left to the reader to verify that all operations used were permissable. □

3.4. MULTIVARIATE CENTRAL LIMIT THEOREM

Because of the fact that in our problem the vectors $\vec{\underline{t}}_i$ do not have identical distributions, we need a multivariate C.L.T. for unequal components. The most general form of C.L.T. that we need is a theorem for triangular array's. Because we have not been able to find a suitable reference in the literature, we give this theorem here with its proof. The proof is based on theorem's 27.2 and 29.4 of BILLINGSLEY (1979). The former is a Lindeberg-type C.L.T. for triangular array's of random variables, while the latter provides a standard way in which limit theorems for random vectors can be derived from corresponding theorems about random variables. (The Cramer-Wold device).

<u>THEOREM 3.4.1</u>. *Let* $(\underset{\sim}{\vec{x}}_{m,1},\ldots,\underset{\sim}{\vec{x}}_{m,m})$ *be a triangular array of* q*-dimensional random vectors such that for each* m, *the vectors* $\underset{\sim}{\vec{x}}_{m,1},\ldots,\underset{\sim}{\vec{x}}_{m,m}$ *are independent, and,*

$$E\underset{\sim}{\vec{x}}_{m,i} = \vec{0} \qquad i = 1,\ldots,m \,, \tag{3.4.1}$$

$$D(\underset{\sim}{\vec{x}}_{m,i}) = \Sigma_{m,i} \qquad i = 1,\ldots,m \,. \tag{3.4.2}$$

Suppose that

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \Sigma_{m,i} = \Sigma \neq 0_q \,, \tag{3.4.3}$$

and that for every $\varepsilon > 0$,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \int_{\|\vec{x}\| > \varepsilon\sqrt{m}} \|\vec{x}\|^2 \, dF_{m,i}(\vec{x}) = 0 \tag{3.4.4}$$

where $F_{m,i}$ *is the distribution function of* $\underset{\sim}{\vec{x}}_{m,i}$.

Then

$$\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \underset{\sim}{\vec{x}}_{m,i} \overset{\sim}{\to} N(\vec{0},\Sigma), \qquad \text{as } m\to\infty. \tag{3.4.5}$$

<u>PROOF</u>. Let $\vec{\lambda} = (\lambda_1,\ldots,\lambda_q)'$ be an arbitrary q-dimensional vector. Define the following random variables

$$\underset{\sim}{y}_{m,i} \overset{d}{=} \vec{\lambda}'\underset{\sim}{\vec{x}}_{m,i} \,.$$

Then $(\underset{\sim}{y}_{m,1},\ldots,\underset{\sim}{y}_{m,m})$, $m = 1,2,\ldots$ is a triangular array of one-dimensional random variables with the following properties for each m and $i = 1,\ldots,m$:

i. the variables $\underset{\sim}{y}_{m,1},\ldots,\underset{\sim}{y}_{m,m}$ are independent;

ii. $E\underset{\sim}{y}_{m,i} = E\vec{\lambda}'\underset{\sim}{\vec{x}}_{m,i} = 0$;

iii. $\sigma^2(\underset{\sim}{y}_{m,i}) = \sigma^2(\vec{\lambda}'\underset{\sim}{\vec{x}}_{m,i}) = \vec{\lambda}'\Sigma_{m,i}\vec{\lambda}$.

Define

$$s_m^2 \overset{d}{=} \sum_{i=1}^{m} \sigma^2(\underset{\sim}{y}_{m,i}) = \vec{\lambda}'(\sum_{i=1}^{m} \Sigma_{m,i})\vec{\lambda} \,.$$

Note that (3.4.3) gives

$$\lim_{m\to\infty} \frac{s_m^2}{m} = \vec{\lambda}'(\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \Sigma_{m,i})\vec{\lambda} = \vec{\lambda}'\Sigma\vec{\lambda}.$$

We shall proceed to show that the Lindeberg-condition of theorem 27.2 of BILLINGSLEY (1979) is satisfied for the just defined variables $\underline{y}_{m,i}$. Let $G_{m,i}$ be the distribution function of $\underline{y}_{m,i}$. Then we have, by the Cauchy-Schwartz inequality, for every $\varepsilon > 0$,

$$\frac{1}{s_m^2} \sum_{i=1}^{m} \int_{|y| > \varepsilon s_m} y^2 dG_{m,i}(y) =$$

$$= \frac{1}{s_m^2} \sum_{i=1}^{m} \int_{|\vec{\lambda}'\vec{x}| > \varepsilon s_m} (\vec{\lambda}'\vec{x})^2 dF_{m,i}(\vec{x}) \leq$$

$$\leq \frac{1}{s_m^2} \vec{\lambda}'\vec{\lambda} \sum_{i=1}^{m} \int_{|\vec{\lambda}'\vec{x}| > \varepsilon s_m} (\vec{x}'\vec{x}) dF_{m,i}(\vec{x}) =$$

$$= \frac{\|\vec{\lambda}\|^2}{s_m^2} \sum_{i=1}^{m} \int_{|\vec{\lambda}'\vec{x}| > \varepsilon s_m} \|\vec{x}\|^2 dF_{m,i}(\vec{x}).$$

For every $\varepsilon > 0$, there exists a $\varepsilon' > 0$, such that for sufficiently large m ($m \geq m'(\varepsilon)$),

$$\{\vec{x} \in \mathbb{R}^q | \; |\vec{\lambda}'\vec{x}| > \varepsilon s_m\} \subset \{\vec{x} \in \mathbb{R}^q | \; \|\vec{x}\|^2 > (\varepsilon'\sqrt{m})^2\} .$$

So we have, for $m \geq m'(\varepsilon)$,

$$\frac{\|\vec{\lambda}\|^2}{s_m^2/m} \frac{1}{m} \sum_{i=1}^{m} \int_{|\vec{\lambda}'\vec{x}| > \varepsilon s_m} \|\vec{x}\|^2 dF_{m,i}(\vec{x}) \leq$$

$$\leq \frac{\|\vec{\lambda}\|^2}{s_m^2/m} \frac{1}{m} \sum_{i=1}^{m} \int_{\|\vec{x}\| > \varepsilon'\sqrt{m}} \|\vec{x}\|^2 dF_{m,i}(\vec{x}).$$

Now, $\vec{\lambda}$ is fixed, s_m^2/m converges to a constant which may be taken to be unequal to zero (treat the case $\vec{\lambda}'\Sigma\vec{\lambda} = 0$ separately). With (3.4.4) it then follows that, for every $\varepsilon > 0$,

$$\lim_{m\to\infty} \frac{1}{s_m^2} \sum_{i=1}^{m} \int_{|y|>\varepsilon s_m} y^2 dG_{m,i}(y) = 0 .$$

Applying the above mentioned theorem 27.2, it follows that

$$\frac{1}{s_m^2} \sum_{i=1}^{m} \underline{y}_{m,i} \overset{\sim}{\to} N(0,1)$$

as $m\to\infty$, for every $\vec{\lambda}\in \mathbb{R}^q$. If we take $\underline{\vec{x}} \sim N(\vec{0},\Sigma)$, this result may be written as

$$\frac{1}{s_m^2} \vec{\lambda}'(\sum_{i=1}^{m} \underline{\vec{x}}_{m,i}) \overset{L}{\to} \frac{1}{(\vec{\lambda}'\Sigma\vec{\lambda})^{\frac{1}{2}}} \vec{\lambda}'\underline{\vec{x}}$$

as $m\to\infty$, for every $\vec{\lambda}\in \mathbb{R}^q$.

Because $s_m/\sqrt{m} \to (\vec{\lambda}'\Sigma\vec{\lambda})^{\frac{1}{2}}$, as $m\to\infty$, we have also

$$\vec{\lambda}'(\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \underline{\vec{x}}_{m,i}) \overset{L}{\to} \vec{\lambda}'\underline{\vec{x}} ,$$

as $m\to\infty$, for every $\vec{\lambda}\in \mathbb{R}^q$.

The proof is now completed with theorem 29.4 of BILLINGSLEY (1979), which states that $\underline{\vec{x}}_m \overset{L}{\to} \underline{\vec{x}}$ iff $\vec{\lambda}'\underline{\vec{x}}_m \overset{L}{\to} \vec{\lambda}'\underline{\vec{x}}$ for each $\vec{\lambda}\in \mathbb{R}^q$. □

CHAPTER 4

CONSISTENCY, ASYMPTOTIC DISTRIBUTIONS AND POWER

4.1. CONSISTENCY

A sequence of level-α tests $\{\underline{\varphi}_{m,Q}\}$ is consistent against a fixed alternative $\mathcal{a} \in A$ (cf. section 2.3.), iff for $m \to \infty$

(4.1.1) $\quad E_{\mathcal{a}}\underline{\varphi}_{m,Q} \to 1.$

It is desirable that (4.1.1) holds for each $\alpha \in (0,1)$, so we shall call $\{\underline{\varphi}_{m,Q}\}$ consistent only if this is the case.

The class of alternatives against which $\{\underline{\varphi}_{m,Q}\}$, based on $\underline{v}(Q) \equiv \vec{\underline{t}}_*'Q\vec{\underline{t}}_*$, is consistent depends on the choice of Q. In this section we shall determine this class.

In section 4.3. we shall prove that, under H_0 and assumption 1, the distribution of $\vec{\underline{t}}_*'Q\vec{\underline{t}}_*$ converges to a fixed distribution. It follows that the sequence of critical values $\{k_{1-\alpha}(m,Q)\}$ is at least bounded. Therefore (4.1.1) holds for every $\alpha \in (0,1)$ iff

(4.1.2) $\quad P_{\mathcal{a}}(\vec{\underline{t}}_*'Q\vec{\underline{t}}_* \geq M) \to 1 \qquad$ for each $M \in \mathbb{R}$.

<u>THEOREM 4.1.1</u>. *A necessary and sufficient condition for* $\{\underline{\varphi}_{m,Q}\}$ *to be consistent against a fixed alternative* $\mathcal{a} \in A$, *is that*

(4.1.3) $\quad \lim_{m\to\infty} \vec{\delta}_*'Q\vec{\delta}_* = \infty$

for this alternative. The class of alternatives for which $\{\underline{\varphi}_{m,Q}\}$ *is consistent forms a subclass of* A_1.

<u>PROOF</u>. The test $\{\underline{\varphi}_{m,Q}\}$ is consistent against $\mathcal{a} \in A$ iff (4.1.2) holds. From the Cantelli inequality (RAO (1973)) it follows that for a sequence of random variables $\{\underline{y}_m\}$,

(4.1.4) $\quad P(\underline{y}_m \geq M) \to 1 \quad$ for each $M \in \mathbb{R}$

iff

(4.1.5) $\quad E\underline{y}_m \to \infty \quad$ and $\quad E\underline{y}_m/\sigma(\underline{y}_m) \to \infty$.

First suppose that *not* $\lim_{m\to\infty} \vec{\delta}_*'Q\vec{\delta}_* = \infty$. This means that $\vec{\delta}_*'Q\vec{\delta}_*$ has a finite limit point, $d \geq 0$, say. If we take $\underline{y}_m \overset{d}{=} \vec{\underline{t}}_*'Q\vec{\underline{t}}_*$, with $E_a\underline{y}_m =$ trace $Q\Sigma_{1.} + \vec{\delta}_*'Q\vec{\delta}_*$, there exists a subsequence $\{\underline{y}_{m_k}\}_{k=1}^{\infty}$, such that

(4.1.6) $\quad E\underline{y}_{m_k} \to \text{trace } Q\Sigma_1 + d < \infty.$

So (4.1.5) is not satisfied for this subsequence, and therefore $\lim_{m\to\infty} P(\underline{y}_m \geq M)$ does not exist. It follows that $\{\underline{\varphi}_{m,Q}\}$ is not consistent.

Next, suppose that $\lim_{m\to\infty} \vec{\delta}_*'Q\vec{\delta}_* = \infty$. With $\vec{\underline{u}}_* \equiv \vec{\underline{t}}_* - \vec{\delta}_*$, we have

(4.1.7) $\quad \vec{\underline{t}}_*'Q\vec{\underline{t}}_* \equiv \vec{\underline{u}}_*'Q\vec{\underline{u}}_* + 2\vec{\delta}_*'Q\vec{\underline{u}}_* + \vec{\delta}_*'Q\vec{\delta}_*.$

Take this time $\underline{y}_m \overset{d}{=} 2\vec{\delta}_*'Q\vec{\underline{u}}_* + \vec{\delta}_*'Q\vec{\delta}_*$, then $E\underline{y}_m = \vec{\delta}_*'Q\vec{\delta}_*$, $\sigma^2(\underline{y}_m) = 4\vec{\delta}_*'Q\Sigma_{1.}Q\vec{\delta}_*$, and it follows from lemma 3.1.7. that (4.1.5) is satisfied. The consistency now follows from

(4.1.8) $\quad P_a(\vec{\underline{t}}_*'Q\vec{\underline{t}}_* \geq M) \geq P_a(\underline{y}_m \geq M)$

which is true because Q is n.n.d, and therefore $\vec{\underline{u}}_*'Q\vec{\underline{u}}_* \geq 0$ with probability one.

Furthermore, $\vec{\delta}_*'Q\vec{\delta}_* \to \infty$ implies $\vec{\delta}_*'\vec{\delta}_* \to \infty$ by lemma 3.1.6. This completes the proof. □

<u>THEOREM 4.1.2</u>. *A sufficient condition for* $\{\underline{\varphi}_{m,Q}\}$ *to be consistent against each alternative in* A_1, *is that* Q *is non-singular.*

<u>PROOF</u>. When Q is non-singular, all its eigenvalues are positive. It follows from lemma 3.1.3. that then $\vec{\delta}_*'Q\vec{\delta}_* \to \infty \iff \vec{\delta}_*'\vec{\delta}_* \to \infty$. The proof is completed with theorem 4.1.1. □

The question naturally arises now if there exist a Q and an $a \in A_1$ for which $\vec{\delta}_*'Q\vec{\delta}_* \not\to \infty$. The above theorem shows that Q has to be singular if this is to be true. A trivial example is furnished if we take $Q \overset{d}{=} I_n \otimes 1_k$, with rank n, because $\vec{\delta}_*' Q \vec{\delta}_* = 0$ ($\underline{\vec{t}}_*' Q \underline{\vec{t}}_* \equiv 0$) for each m and each $a \in A_1$ for this choice of Q.

A more important question is the following. Does there exist, given a singular Q, an alternative $a \in A_1$ for which $\vec{\delta}_*' Q \vec{\delta}_* \not\to \infty$, i.e. an alternative for which $\{\underline{\varphi}_{m,Q}\}$ is not consistent. The answer is: not always; there are singular matrices Q such that $\{\underline{\varphi}_{m,Q}\}$ is consistent against all alternatives in A_1.

Recall that the vectors $\vec{\delta}_i$ and $\vec{\delta}_*$ are elements of the null-space $N(F)$ of the matrix F defined in section 3.1. Conversely, under certain circumstances, any element of $N(F)$ may, apart from a constant factor, occur as a vector $\vec{\delta}_i$, as follows from the following lemma.

<u>LEMMA 4.1.1</u>. *Let* $n \geq k \geq 4$. *Let* $\vec{\eta}$ *be an arbitrary element from* $N(F)$. *Then for any experiment* E_i, *with* $a_{i1} > 0$, $a_{i2} > 0, \ldots, a_{ik} > 0$, *there exist constants* $\Delta_{i1}, \ldots, \Delta_{iN_i}$, *and a constant* c, *such that*

$$\vec{\delta} \overset{d}{=} \vec{\eta}/c = E\underline{\vec{t}}_i. \tag{4.1.9}$$

<u>PROOF</u>. For the experiment E_i, with $a_{i1} > 0, \ldots, a_{ik} > 0$, we have if $n \geq k \geq 4$,

$$N_i = \frac{n!}{a_{i1}!\, a_{i2}! \ldots a_{ik}!} \geq \frac{n!}{(n-(k-1))!} = n(n-1)\ldots(n-k+2) \geq nk+1.$$

Introduce the variables $x_1, \ldots, x_{N_i}$ and solve the following $nk+1$ equations (writing $\vec{\eta} = (\eta_1^{(1)}, \ldots, \eta_k^{(1)}, \eta_1^{(2)}, \ldots, \eta_k^{(2)}, \ldots, \eta_1^{(n)}, \ldots, \eta_k^{(n)})'$).

$$\eta_j^{(\nu)} = \sum_{r=1}^{N_i} t_{ij}^{(\nu)}(\pi_{ir}) x_r, \qquad \sum_{r=1}^{N_i} x_r = 0.$$

Because the number of variables (N_i) exceeds the number of equations ($nk+1$), there is always a solution.

Next, choose a constant c such that the solutions x_r satisfy

$$-\frac{1}{N_i} \leq \frac{x_r}{c} \leq 1 - \frac{1}{N_i}.$$

Then we may take $\Delta_{ir} = x_r/c$, from which follows that $E\underline{\vec{t}}_i = \vec{\eta}/c = \vec{\delta}$. □

We then have the following theorem.

THEOREM 4.1.3. *Let* $n \geq k \geq 4$. *Let* Q *be a singular (real, symmetric and n.n.d)* $nk \times nk$ *matrix, with* $\text{rank } Q < (n-1)(k-1)$. *Then there exists an alternative* $a \in A_1$ *such that* $\lim_{m\to\infty} \vec{\delta}_*'\vec{\delta}_* = \infty$, *but* $\lim_{m\to\infty} \vec{\delta}_*'Q\vec{\delta}_* = 0$.

PROOF. Because $\text{rank } Q < (n-1)(k-1)$, $N(Q) \cap N(F) \neq \{\vec{0}\}$. So take $\vec{\eta} \in N(Q) \cap N(F)$, $\vec{\eta} \neq \vec{0}$. By lemma 4.1.1., there now exist constants c and $\Delta_{i1},\ldots,\Delta_{iN_i}$ and an associated experiment E_i such that $E\underline{\vec{t}}_i = \vec{\eta}/c \overset{d}{=} \vec{\delta}_i$.
Also $\vec{\delta}_i \in N(Q) \cap N(F)$. It follows that

$$\vec{\delta}_*'\vec{\delta}_* = \left(\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \vec{\delta}_i\right)'\left(\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \vec{\delta}_i\right) = \frac{m}{c^2}\, \vec{\eta}'\vec{\eta} \to \infty,$$

but

$$\vec{\delta}_*'Q\vec{\delta}_* = \frac{m}{c^2}\, \vec{\eta}'Q\vec{\eta} = 0$$

for each m, because $\vec{\eta} \in N(Q)$. □

This theorem shows that, if the test is not directed against a very specialised alternative, test-statistics with a matrix Q with rank smaller than $(n-1)(k-1)$ should be avoided, because there are then always alternatives that cannot be detected.

When rank Q is larger than or equal to $(n-1)(k-1)$, two different situations can occur. We show, by examples, that it is possible that there still is an alternative for which $\{\underline{\varphi}_{m,Q}\}$ is not consistent (example 4.1.1) and it is also possible that, even though Q is singular, $\{\underline{\varphi}_{m,Q}\}$ is consistent against all alternatives in A_1 (example 4.1.2).

EXAMPLE 4.1.1. Take $n = k = 4$. Choose a sequence E_1, E_2, ... with $a_{i1} > 0$, $a_{i2} > 0,\ldots,$ $a_{ik} > 0$. Define a c and $\Delta_{i1},\ldots,\Delta_{iN_i}$ such that

$$\vec{\delta}_i \overset{d}{=} E\underline{\vec{t}}_i = c(1,-1,0,0;-1,1,0,0;0,0,0,0;0,0,0,0)'$$

This is possible by lemma 4.1.1. This defines an alternative a to H_0 with

$$\vec{\delta}_*'\vec{\delta}_* = mc^2 \cdot 4 \to \infty, \qquad \text{as } m \to \infty,$$

so $a \in A_1$.

Take

$$Q \stackrel{d}{=} \left(\begin{array}{cccccc|c} 1 & 1 & 0 & 0 & 1 & 1 & \\ 1 & 1 & 0 & 0 & 1 & 1 & \\ 0 & 0 & 1 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 1 & 0 & 0 & 0_{6,10} \\ 1 & 1 & 0 & 0 & 1 & 1 & \\ 1 & 1 & 0 & 0 & 1 & 1 & \\ \hline & & & 0_{10,6} & & & I_{10} \end{array}\right)$$

Q is a real, symmetric, n.n.d matrix of order 16, with rank 13, which is larger than $(n-1)(k-1) = 3 \cdot 3 = 9$. Obviously $\vec{\delta}_*' Q \vec{\delta}_* = 0$ for each m. It follows that for this particular alternative $\{\underline{\varphi}_{m,Q}\}$ is not consistent.

<u>EXAMPLE 4.1.2</u>. Take $n \geq 2$ and $k \geq 2$ arbitrary. Let

$$Q \stackrel{d}{=} \left(\begin{array}{c|c} I_{nk-k} & 0_{nk-k,k} \\ \hline 0_{k,nk-k} & 0_{k,k} \end{array}\right)$$

Obviously, Q is a real, symmetric, n.n.d matrix of order nk, with rank $nk-k > (n-1)(k-1)$. Consider an arbitrary alternative $a \in A_1$. Now

$$(4.1.10) \qquad \vec{\delta}_*' Q \vec{\delta}_* = \sum_{\nu=1}^{n-1} \sum_{j=1}^{k} \{\delta_{*j}^{(\nu)}\}^2 .$$

Because $a \in A_1$, we have $\vec{\delta}_*' \vec{\delta}_* \to \infty$, so at least one component of $\vec{\delta}_*$ must (in absolute value) tend to infinity. Due to the linear relationships $\delta_{*+}^{(\nu)} = 0$ for all ν, and $\delta_{*j}^{(+)} = 0$ for all j, this means that

i. at least four components must in absolute value tend to infinity;
ii. those four components cannot all be among $\delta_{*1}^{(n)}, \delta_{*2}^{(n)}, \ldots, \delta_{*k}^{(n)}$.

It follows that at least two terms of the sum (4.1.10) must tend to infinity, and so $\vec{\delta}_*' Q \vec{\delta}_* \to \infty$.

Because $a \in A_1$ was arbitrary, we have

$$\vec{\delta}_*' \vec{\delta}_* \to \infty \Rightarrow \vec{\delta}_*' Q \vec{\delta}_* \to \infty \qquad \text{for each } a \in A_1 .$$

From theorem 4.1.1. it then follows that $\{\underline{\varphi}_{m,Q}\}$ is consistent against each $a \in A_1$.

4.2. ASYMPTOTIC MULTIVARIATE NORMALITY

In this section we shall establish the asymptotic normality of $\underline{\vec{t}}_*$, under the assumptions 1 & 2 of section 2.3. Under these conditions, the dispersion matrix of $\underline{\vec{t}}_*$ converges to a limit as $m \to \infty$.

Recall that

$$\Sigma_1 \overset{d}{=} \lim_{m\to\infty} \Sigma_{1\bullet} = \lim_{m\to\infty} D(\underline{\vec{t}}_*), \tag{4.2.1}$$

$$\Sigma_0 \overset{d}{=} \lim_{m\to\infty} \Sigma_{0\bullet} = \lim_{m\to\infty} D(\underline{\vec{t}}_* | H_0), \tag{4.2.2}$$

$$\vec{\delta} \overset{d}{=} \lim_{m\to\infty} \vec{\delta}_* = \lim_{m\to\infty} E\underline{\vec{t}}_* , \tag{4.2.3}$$

$$\vec{\zeta} \overset{d}{=} \lim_{m\to\infty} \vec{\delta}_\bullet . \tag{4.2.4}$$

<u>THEOREM 4.2.1</u>. *Under the assumptions 1 & 2 of section 2.3., we have for* $m \to \infty$,

$$\underline{\vec{u}}_* \equiv \underline{\vec{t}}_* - \vec{\delta}_* \overset{\sim}{\to} N(\vec{0},\Sigma_1). \tag{4.2.5}$$

<u>PROOF</u>. $\underline{\vec{u}}_1, \underline{\vec{u}}_2, \ldots$ is a sequence of independent $n \times k$ dimensional random vectors, with $E\underline{\vec{u}}_i = \vec{0}$ and dispersion matrix $D(\underline{\vec{u}}_i) = D(\underline{\vec{t}}_i)$. Under the assumptions 1 & 2 we have

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} D(\underline{\vec{u}}_i) = \lim_{m\to\infty} D(\underline{\vec{u}}_*) = \lim_{m\to\infty} D(\underline{\vec{t}}_*) = \Sigma_1.$$

Because there exists a constant c, such that $\|\underline{\vec{u}}_i\|^2 \le c$ with probability one, for $i = 1,2,\ldots$, the 'Lindeberg' condition of theorem 3.4.1. is trivially fulfilled. (Adaptation of theorem 3.4.1. to ordinary sequences of random vectors is straightforward). Therefore, all the conditions of theorem 3.4.1. are fulfilled and the result follows. □

<u>COROLLARY 4.2.1</u>. *Under the assumptions 1 & 2 of section 2.3. and* H_0, *we have for* $m \to \infty$,

$$\underline{\vec{t}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_0). \tag{4.2.6}$$

<u>PROOF</u>. Under H_0, $\vec{\delta} = \vec{0}$ and $D(\underline{\vec{t}}_*) \to \Sigma_0$. □

<u>COROLLARY 4.2.2</u>. *Under the assumptions 1 & 2, we have for alternatives in* $A_2 \cup A_3$,

$$(4.2.7) \qquad \underset{-}{\vec{t}}_* \overset{\sim}{\to} N(\vec{\delta},\Sigma_1).$$

<u>PROOF</u>. Because $\underset{-}{\vec{t}}_* \equiv \underset{-}{\vec{u}}_* + \vec{\delta}_*$, $\underset{-}{\vec{u}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_1)$ by theorem 4.2.1., $\vec{\delta}_* \to \vec{\delta}$ by assumption 2, the result follows from a Cramer-type theorem. □

For contiguous alternatives we have the following theorem.

<u>THEOREM 4.2.2</u>. *Under contiguous alternatives, as defined in section 2.4., we have*

$$(4.2.8) \qquad \underset{-}{\vec{t}}_* \overset{\sim}{\to} N(\sqrt{\eta}\,\vec{\zeta},\Sigma_0).$$

<u>PROOF</u>. The proof is analogous to the proof of theorem 4.2.1., this time using the triangular-array method explicitly. □

4.3. ASYMPTOTIC DISTRIBUTION OF THE TEST-STATISTIC

In this section we derive the asymptotic distribution of $\underset{-}{v}(Q) \equiv \underset{-}{\vec{t}}_*'Q\underset{-}{\vec{t}}_*$.

<u>THEOREM 4.3.1</u>. *Under* H_0, *under alternatives from* $A_2 \cup A_3$ *and under contiguous alternatives, we have*

$$(4.3.1) \qquad \underset{-}{v}(Q) \overset{L}{\to} \underset{-}{\vec{x}}'Q\underset{-}{\vec{x}}$$

where

$$(4.3.2) \qquad \underset{-}{\vec{x}} \sim N(\vec{0},\Sigma_0), \text{ under } H_0;$$

$$(4.3.3) \qquad \underset{-}{\vec{x}} \sim N(\vec{\delta},\Sigma_1), \text{ under alternatives from } A_2 \cup A_3;$$

$$(4.3.4) \qquad \underset{-}{\vec{x}} \sim N(\sqrt{\eta}\,\vec{\zeta},\Sigma_0), \text{ under contiguous alternatives.}$$

<u>PROOF</u>. In all three cases we have $\underset{-}{\vec{t}}_* \overset{L}{\to} \underset{-}{\vec{x}}$, where $\underset{-}{\vec{x}}$ has one of the distributions (4.3.2) - (4.3.4). Because (.)'Q(.) is a continuous function, the result follows. □

Note that the distributions of $\underline{\vec{x}}'Q\underline{\vec{x}}$ follow from theorem 3.2.1.

For alternatives in A_1, the situation is somewhat different. Although $\{\underline{\varphi}_{m,Q}\}$ is consistent against $a \in A_1$ when $\vec{\delta}_*'Q\vec{\delta}_* \to \infty$, the asymptotic distribution (a.d.) does not exist in all cases. We shall give some examples where the a.d. does exist.

THEOREM 4.3.2. *For alternatives in* A_1, $\underline{v}(Q) \equiv \underline{\vec{t}}_*'Q\underline{\vec{t}}_*$ *has the following a.d.*

$$\text{(4.3.5)} \qquad \frac{\underline{\vec{t}}_*'Q\underline{\vec{t}}_* - \vec{\delta}_*'Q\vec{\delta}_*}{\sqrt{m}} \overset{\sim}{\to} N(0,\sigma^2),$$

where

$$\text{(4.3.6)} \qquad \sigma^2 = 4\vec{\zeta}'Q\Sigma_1 Q\vec{\zeta}.$$

PROOF. From (4.1.7) it follows that

$$\frac{\underline{\vec{t}}_*'Q\underline{\vec{t}}_* - \vec{\delta}_*'Q\vec{\delta}_*}{\sqrt{m}} \equiv \frac{\underline{\vec{u}}_*'Q\underline{\vec{u}}_*}{\sqrt{m}} + 2\vec{\delta}_\bullet'Q\underline{\vec{u}}_*.$$

From theorem 4.2.1. it follows that $\underline{\vec{u}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_1)$. As in theorem 4.3.1., $\underline{\vec{u}}_*'Q\underline{\vec{u}}_*$ then converges to a fixed distribution. Therefore $m^{-\frac{1}{2}}(\underline{\vec{u}}_*'Q\underline{\vec{u}}_*) \overset{P}{\to} 0$. Furthermore, from $\vec{\delta}_\bullet \to \vec{\zeta}$ it follows that $2\vec{\delta}_\bullet'Q\underline{\vec{u}}_* \overset{L}{\to} 2\vec{\zeta}'Q\underline{\vec{x}}$, where $\underline{\vec{x}} \sim N(\vec{0},\Sigma_1)$. The result follows from $E(2\vec{\zeta}'Q\underline{\vec{x}}) = 0$ and $\sigma^2(2\vec{\zeta}'Q\underline{\vec{x}}) = 4\vec{\zeta}'Q\Sigma_1 Q\vec{\zeta}$. □

Note that in many cases σ^2, defined by (4.3.6), is equal to zero. The a.d. of $\underline{v}(Q)$ is then degenerate. Other transformations may still yield a proper a.d., as is illustrated by the following theorems. Let

$$\text{(4.3.7)} \qquad \sigma_m^2 \overset{d}{=} \sigma^2(2\vec{\delta}_*'Q\underline{\vec{u}}_*) = 4\vec{\delta}_*'Q\Sigma_{1\bullet}Q\vec{\delta}_*.$$

THEOREM 4.3.3. *For alternatives in* A_1, *and matrices* Q *such that* $\sigma_m^2 \to 0$, $\underline{v}(Q) \equiv \underline{\vec{t}}_*'Q\underline{\vec{t}}_*$ *has the following a.d.*

$$\text{(4.3.8)} \qquad \underline{\vec{t}}_*'Q\underline{\vec{t}}_* - \vec{\delta}_*'Q\vec{\delta}_* \overset{L}{\to} \underline{\vec{x}}'Q\underline{\vec{x}},$$

where $\underline{\vec{x}} \sim N(\vec{0},\Sigma_1)$.

PROOF. Because $\sigma_m^2 \to 0$ and $\underline{\vec{u}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_1)$, the result follows from the identity (4.1.7). □

Note that $\sigma_m^2 \to 0$ implies that $4\vec{\zeta}'Q\Sigma_1 Q\vec{\zeta} = 0$, so the situation of theorem 4.3.3. is not covered by theorem 4.3.2.

THEOREM 4.3.4. *For alternatives in* A_1, *and matrices* Q *such that* $\sigma_m^2 \to \infty$ *and* $\lim_{m\to\infty} \vec{\delta}_*/\sigma_m$ *exists,* $\underline{v}(Q) \equiv \vec{\underline{t}}_*'Q\vec{\underline{t}}_*$ *has the following a.d.*

$$\sigma_m^{-1}(\vec{\underline{t}}_*'Q\vec{\underline{t}}_* - \vec{\delta}_*'Q\vec{\delta}_*) \overset{\sim}{\to} N(0,1). \tag{4.3.9}$$

PROOF. From (4.1.7) it follows that

$$\sigma_m^{-1}(\vec{\underline{t}}_*'Q\vec{\underline{t}}_* - \vec{\delta}_*'Q\vec{\delta}_*) \equiv \sigma_m^{-1}\vec{\underline{u}}_*'Q\vec{\underline{u}}_* + 2\sigma_m^{-1}\vec{\delta}_*'Q\vec{\underline{u}}_*.$$

Because $\vec{\underline{u}}_*'Q\vec{\underline{u}}_*$ converges to a fixed distribution, $\sigma_m^{-1}\vec{\underline{u}}_*'Q\vec{\underline{u}}_* \overset{P}{\to} 0$. Furthermore $E(2\sigma_m^{-1}\vec{\delta}_*'Q\vec{\underline{u}}_*) = 0$ and $\sigma^2(2\sigma_m^{-1}\vec{\delta}_*'Q\vec{\underline{u}}_*) = 1$. The random variable $2\sigma_m^{-1}\vec{\delta}_*'Q\vec{\underline{u}}_*$ converges because $\lim_{m\to\infty} \sigma_m^{-1}\vec{\delta}_*$ exists and $\vec{\underline{u}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_1)$. □

THEOREM 4.3.5. *For those alternatives in* A_1, *for which* $\lim_{m\to\infty} \vec{\delta}_*/\|\vec{\delta}_*\|$ *exists,* $\underline{v}(Q) \equiv \vec{\underline{t}}_*'Q\vec{\underline{t}}_*$ *has the following a.d.*

$$\|\vec{\delta}_*\|^{-1}(\vec{\underline{t}}_*'Q\vec{\underline{t}}_* - \vec{\delta}_*'Q\vec{\delta}_*) \overset{\sim}{\to} N(0,\tau^2) \tag{4.3.10}$$

where

$$\tau^2 = \lim_{m\to\infty} \|\vec{\delta}_*\|^{-2}\sigma_m^2 . \tag{4.3.11}$$

PROOF. The proof is analogous to the proof of theorem 4.3.4. □

REMARK 4.3.1. Until now we have assumed Q to be a fixed, real, symmetric, n.n.d matrix. However Q may depend on m (we shall write Q_m in that case) without affecting the results, provided that

$$\lim_{m\to\infty} Q_m = Q \tag{4.3.12}$$

exists, and Q is a real, symmetric, n.n.d matrix.

4.4. POWER OF THE TEST

After having examined the consistency of $\{\underline{\varphi}_{m,Q}\}$ and having determined the asymptotic distributions of $\underline{v}(Q)$, we shall now investigate the power of the test. We shall consider the exact power (for finite m), the asymptotic power and the asymptotic power under contiguous alternatives. The latter may be considered as an approximation to the exact power for finite m; we shall give a lower-bound for it, which can be calculated easily using only the table of the standard normal distribution.

Furthermore we shall investigate to what extent we can use the power to select a useful Q-matrix.

Recall that the critical value of the test was defined in (2.2.12). The exact power of $\underline{\varphi}_{m,Q}$ for a specified alternative $a^0 \in A$ is then, by definition

$$(4.4.1) \qquad P_{a^0}(\underline{\vec{t}}_*'Q\underline{\vec{t}}_* \geq k_{1-\alpha}(m,Q)).$$

Because it is almost impossible (except for small m) to determine the exact distribution of $\underline{\vec{t}}_*'Q\underline{\vec{t}}_*$, neither the critical value $k_{1-\alpha}(m,Q)$, nor the exact power of the test can be determined. In practice therefore, an approximate value is chosen for $k_{1-\alpha}(m,Q)$. This might be $k_{1-\alpha}(Q)$, defined by

$$(4.4.2) \qquad P(\underline{\vec{x}}'Q\underline{\vec{x}} \geq k_{1-\alpha}(Q) \,|\, \underline{\vec{x}} \sim N(\vec{0},\Sigma_0)) = \alpha,$$

i.e. the critical value of the asymptotic distribution of $\underline{\vec{t}}_*'Q\underline{\vec{t}}_*$. It might be $\chi^2_{1-\alpha}[\nu]$, with the degrees of freedom, ν, properly chosen. Or it might be a critical value resulting from some other practical approximation to the distribution of $\underline{\vec{t}}_*'Q\underline{\vec{t}}_*$ under H_0. The possible ways of determining a practical critical value are described in chapter 9. Let k be any critical value determined in one of those ways. The exact 'practical' power of the test, against $a^0 \in A$,

$$(4.4.3) \qquad P_{a^0}(\underline{\vec{t}}_*'Q\underline{\vec{t}}_* \geq k)$$

still cannot be determined, except for very small m. However, it may be estimated by simulating the distribution of $\underline{\vec{t}}_*'Q\underline{\vec{t}}_*$ under $a^0 \in A$ with the aid of a computer. Some examples of this simulation can be found in chapter 9.

The asymptotic power,

$$(4.4.4) \qquad \lim_{m\to\infty} P_{\mathcal{a}}(\vec{\underline{t}}_*'Q\vec{\underline{t}}_* \geq k)$$

is equal to 1 for consistent tests, and is apparently not a good approximation to the exact power for finite m.

Therefore, to obtain an approximation to (4.4.3), we shall proceed as follows. Suppose that a fixed number $m_0 > 0$, of experiments $(E_1,\ldots,E_{m_0})$ have been performed, with

$$(4.4.5) \qquad \vec{a}_1,\ \vec{a}_2,\ldots,\ \vec{a}_{m_0},$$

under the alternative $\mathcal{a}^0 \in A$, which is determined by

$$(4.4.6) \qquad \mathcal{a}^0 = (d_1,\ d_2,\ d_3,\ \ldots\ ), \qquad d_i \in \mathcal{D}_i.$$

Because only the first m_0 coordinates of $\mathcal{a}^0$ are of interest for the calculation of (4.4.3), we shall compare the experiment with the following infinite sequence of experiments

$$(4.4.7) \qquad E_1,\ E_2,\ldots,\ E_{m_0},\ E_1,\ E_2,\ldots,\ E_{m_0},\ldots$$

with associated $\vec{a}$ vectors

$$(4.4.8) \qquad \vec{a}_1,\ \vec{a}_2,\ldots,\ \vec{a}_{m_0},\ \vec{a}_1,\ \vec{a}_2,\ldots,\ \vec{a}_{m_0},\ldots$$

and under the alternative from A_1,

$$(4.4.9) \qquad \mathcal{a} = (d_1,\ d_2,\ldots,\ d_{m_0},\ d_1,\ d_2,\ldots,\ d_{m_0},\ldots\ ).$$

This infinite sequence of experiments clearly satisfies assumptions 1 & 2.

To stress the dependence of $\vec{\delta}_i$ and $\vec{\delta}$ on the particular chosen alternative, we shall write

$$(4.4.10) \qquad \vec{\delta}_i(\mathcal{a})$$

if we mean the vector $\vec{\delta}_i$ computed from the i'th component of $\mathcal{a}$, and

(4.4.11) $\quad \vec{\delta}(\alpha),\ \Sigma_0(\alpha),\ \Sigma_{0\bullet}(\alpha)$, etc.

likewise.

Now consider the particular contiguous alternative, derived from $\alpha \in A_1$,

(4.4.12) $\quad \{\alpha_{\theta_m}\}_{m=1}^{\infty}, \qquad m\theta_m^2 \to m_0.$

It follows from theorem 4.3.1. that

(4.4.13) $\quad \lim_{m\to\infty} P_c(\underline{\vec{t}}_*'Q\underline{\vec{t}}_* \geq k) = P(\underline{\vec{x}}'Q\underline{\vec{x}} \geq k \mid \underline{\vec{x}} \sim N(\sqrt{m_0}\ \vec{\zeta}(\alpha), \Sigma_0(\alpha))),$

where P_c means that the probability is calculated under the contiguous alternative (4.4.12).

We shall take (4.4.13) as an approximation to (4.4.3). The, intuitive, motivation is the analogy with usual statistical practice. The motivation is supported by the following facts. Notice that

(4.4.14) $\quad \vec{\zeta}(\alpha) = \lim_{m\to\infty} \vec{\delta}_{\bullet}(\alpha) = \vec{\delta}_{\circledcirc}(\alpha) = \vec{\delta}_{\circledcirc}(\alpha^0),$

and likewise,

(4.4.15) $\quad \Sigma_0(\alpha) = \Sigma_{0\circledcirc}(\alpha^0),$

so that $\vec{\zeta}(\alpha)$ and $\Sigma_0(\alpha)$ are equal to the corresponding quantities, calculated for the first m_0 experiments. Furthermore, the exact expectation of the test-statistic for the m_0 experiments is equal to

(4.4.16) $\quad E\underline{\vec{t}}_{\circledast}'Q\underline{\vec{t}}_{\circledast} = \text{trace } Q\Sigma_{1\circledcirc}(\alpha^0) + \vec{\delta}_{\circledast}'(\alpha^0)Q\vec{\delta}_{\circledast}(\alpha^0),$

while

(4.4.17) $$\lim_{m\to\infty} E\underline{\vec{t}}_*'Q\underline{\vec{t}}_* = \text{trace } Q\Sigma_0 + \lim_{m\to\infty} \theta_m^2\ \vec{\delta}_*'(\alpha)Q\vec{\delta}_*(\alpha) =$$

$$= \text{trace } Q\Sigma_0 + m_0\vec{\zeta}'(\alpha)Q\vec{\zeta}(\alpha) =$$

$$= \text{trace } Q\Sigma_0 + \vec{\delta}_{\circledast}'(\alpha^0)Q\vec{\delta}_{\circledast}(\alpha^0),$$

and

$$(4.4.18)\qquad E\underline{\vec{x}}'Q\underline{\vec{x}} = \text{trace } Q\Sigma_0 + \vec{\delta}_{\oplus}'(a^0)Q\vec{\delta}_{\oplus}(a^0).$$

It follows that at least the non-centrality parts of the exact expectation and of the expectation of the approximation are equal. This motivates the choice of this particular contiguous alternative (4.4.12).

From theorem 3.2.1. it follows that (4.4.13) is equal to

$$(4.4.19)\qquad P(\sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \sqrt{m_0}\omega_\tau)^2 \geq k)$$

with $\underline{\vec{u}} \sim N(\vec{0}, I_r)$, $\Sigma_0 = BB'$, $r = \text{rank } B'QB$, $\lambda_1,\ldots,\lambda_r$ the (positive) eigenvalues of $B'QB$ and $\vec{\omega} = \Lambda_+^{-1}P_+'B'Q\vec{\zeta}(a)$. The reader is referred to section 3.2. for details. The asymptotic expansion of KOTZ, JOHNSON & BOYD (1967b), as described in section 3.3., may be used for the actual calculation of (4.4.13) (or (4.4.19)), or any approximation to this distribution. Some examples can be found in chapter 9.

For all these calculations a table of the non-central χ^2-distribution or the aid of a computer is necessary. (The approximations involve the non-central χ^2-distribution).

However, there exists a lower-bound for (4.4.13) that can be calculated easily. It follows from the following theorem.

<u>THEOREM 4.4.1</u>. *Let* $\underline{\vec{u}} \sim N(\vec{0}, I_r)$, $\lambda_1,\ldots,\lambda_r \in \mathbb{R}_+$, $\omega_1,\ldots,\omega_r \in \mathbb{R}_+ \cup \{0\}$, *then*

$$(4.4.20)\qquad P(\sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \omega_\tau)^2 \leq z) \leq P(\frac{-\nu\sqrt{z} - \nu^2}{w} \leq \underline{u} \leq \frac{\nu\sqrt{z} - \nu^2}{w}), \qquad \forall\ z \geq 0.$$

where

$$\nu \stackrel{d}{=} \{\sum_{\tau=1}^{r} \lambda_\tau\omega_\tau^2\}^{\frac{1}{2}}, \qquad w \stackrel{d}{=} \{\sum_{\tau=1}^{r} \lambda_\tau^2\omega_\tau^2\}^{\frac{1}{2}}$$

and $\underline{u} \sim N(0,1)$.

<u>PROOF</u>.

$$P(\sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \omega_\tau)^2 \leq z) = P(\sum_{\tau=1}^{r} \lambda_\tau\underline{u}_\tau^2 + 2\sum_{\tau=1}^{r} \lambda_\tau\omega_\tau\underline{u}_\tau + \sum_{\tau=1}^{r} \lambda_\tau\omega_\tau^2 \leq z).$$

By the Cauchy-Schwartz inequality we have

$$(\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau)^2 \equiv (\sum_{\tau=1}^{r} \sqrt{\lambda_\tau}\underline{u}_\tau \sqrt{\lambda_\tau}\omega_\tau)^2 \le (\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2)(\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2).$$

Writing $\nu = \{\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2\}^{\frac{1}{2}}$ and $\mathcal{W} = \{\sum_{\tau=1}^{r} \lambda_\tau^2 \omega_\tau^2\}^{\frac{1}{2}}$, we have

$$\begin{aligned}
&P(\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 + 2\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau + \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2 \le z) = \\
&= P(\nu^2 \sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 + 2\nu^2 \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau + \nu^4 \le z\nu^2) \le \\
&\le P((\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau)^2 + 2\nu^2 \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau + \nu^4 \le z\nu^2) = \\
&= P(\{\sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau + \nu^2\}^2 \le z\nu^2) = \\
&= P(-\nu\sqrt{z} \le \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau + \nu^2 \le \nu\sqrt{z}) = \\
&= P(-\nu\sqrt{z} - \nu^2 \le \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau \underline{u}_\tau \le \nu\sqrt{z} - \nu^2) = \\
&= P(\frac{-\nu\sqrt{z} - \nu^2}{\mathcal{W}} \le \underline{u} \le \frac{\nu\sqrt{z} - \nu^2}{\mathcal{W}}),
\end{aligned}$$

where $\underline{u} \sim N(0,1)$. □

When we apply theorem 4.4.1. to (4.4.19) we get

$$\begin{aligned}
(4.4.21)\quad &P(\sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \sqrt{m_0}\omega_\tau)^2 \ge k) \ge \\
&\ge 1 - P(\frac{-\sqrt{m_0}\nu\sqrt{k} - m_0\nu^2}{\mathcal{W}\sqrt{m_0}} \le \underline{u} \le \frac{\sqrt{m_0}\nu\sqrt{k} - m_0\nu^2}{\mathcal{W}\sqrt{m_0}}) = \\
&= 1 - P(\frac{-\nu\sqrt{k} - \sqrt{m_0}\nu^2}{\mathcal{W}} \le \underline{u} \le \frac{\nu\sqrt{k} - \sqrt{m_0}\nu^2}{\mathcal{W}}) .
\end{aligned}$$

From remark 3.2.3. and (4.4.14), (4.4.15) it follows moreover that

$$(4.4.22)\quad \nu^2 = \vec{\zeta}'(a)Q\vec{\zeta}(a) = \vec{\delta}_{\circledast}'(a^0)Q\vec{\delta}_{\circledast}(a^0) ,$$

$$(4.4.23)\quad \mathcal{W}^2 = \vec{\zeta}'(a)Q\Sigma_0(a)Q\vec{\zeta}(a) = \vec{\delta}_{\circledast}'(a^0)Q\Sigma_{0\circledast}(a^0)Q\vec{\delta}_{\circledast}(a^0).$$

Therefore, it is not even necessary to calculate the λ_τ's and ω_τ's to determine this lower bound (the righthand side of (4.4.21)) of the power of the test for finite m_0.

Notice that this lower bound may also be used to make a quick estimate of the number of observations necessary to achieve a given power against a given alternative $a^0 \in A$.

The last question that remains to be answered is whether we can use the approximation to the powerfunction to select an 'optimal' Q-matrix. There are two approaches to this problem. Suppose that we have a fixed alternative $a^0 \in A$ and a fixed m_0, so that $\vec{\zeta}(a)$ and $\Sigma_0(a)$ are fixed and known. We would then want to select the matrix Q that gives the highest power against this alternative. Alternatively, suppose that the number of observations is not fixed in advance, and that we want to achieve a given power against a^0. We then would select a matrix Q that would need the least observations to do this. However, the alternative a as defined in (4.4.9) would then depend on m_0, and so also $\vec{\zeta}(a)$ and $\Sigma_0(a)$. The situation is then rather complicated. Both approaches may not be equivalent.

We do not pursue this problem here any further, because we come back to it in the next chapter.

However, one step towards simplification of matters can already be made. In order to keep the critical value k (exact, or resulting from some approximation) in the neighbourhood of the critical values of the χ^2-distribution, we shall choose Q, without loss of generality, in such a way that for the a.d.,

$$(4.4.24) \qquad \sum_{\tau=1}^{r} \lambda_\tau = r = \text{rank } B'QB$$

where $BB' = \Sigma_0$.

4.5. ASYMPTOTIC DISTRIBUTIONS IN THE UNCONDITIONAL CASE

Although the unconditional version of our test-statistic, previously defined (in (2.5.26)) as

$$(4.5.1) \qquad \underline{w}(G) \stackrel{d}{=} \frac{1}{m} \sum_{j=1}^{k} g_j \sum_{\nu=1}^{n} \{ \sum_{i=1}^{m} (\underline{x}_{ij}^{(\nu)} - \frac{\underline{a}_{ij}}{n}) \}^2$$

is unfit to be used as a test-statistic, it is nevertheless interesting to investigate its asymptotic distribution. To this end we introduce similar notation as in section 2.2. (See also section 2.5.). Let

$$(4.5.2) \qquad \dot{\underline{x}}_{ij}^{(\nu)} \stackrel{d}{=} \underline{x}_{ij}^{(\nu)} - \frac{\underline{a}_{ij}}{n} ,$$

$$(4.5.3)\qquad \vec{\underline{x}}_i \overset{d}{=} (\dot{\underline{x}}_{i1}^{(1)},\dots,\dot{\underline{x}}_{ik}^{(1)};\dot{\underline{x}}_{i1}^{(2)},\dots,\dot{\underline{x}}_{ik}^{(2)};\dots;\dot{\underline{x}}_{i1}^{(n)},\dots,\dot{\underline{x}}_{ik}^{(n)})'.$$

With $Q = I_n \otimes G$, or a more general Q, $\underline{w}(G)$ may be written as

$$(4.5.4)\qquad \underline{w}(Q) \equiv \vec{\underline{x}}_*' Q \vec{\underline{x}}_* .$$

Let

$$(4.5.5)\qquad \varepsilon_{ij}^{(\nu)} \overset{d}{=} E\dot{\underline{x}}_{ij}^{(\nu)} = p_{ij}^{(\nu)} - \frac{1}{n}\sum_{\mu=1}^{n} p_{ij}^{(\mu)} = p_{ij}^{(\nu)} - p_{ij}^{(\cdot)} .$$

Define $\vec{\varepsilon}_i$ as $\vec{\delta}_i$ in (2.2.16). Then

$$(4.5.6)\qquad E\vec{\underline{x}}_* = \vec{\varepsilon}_*$$

which reduces, under H_0^*, to $E(\vec{\underline{x}}_*|H_0^*) = \vec{0}$. Let

$$(4.5.7)\qquad \Sigma_{1i} \overset{d}{=} D(\vec{\underline{x}}_i) ,$$

$$(4.5.8)\qquad \Sigma_{0i} \overset{d}{=} D(\vec{\underline{x}}_i|H_0^*)$$

The entries of Σ_{1i} and Σ_{0i} may be found from

$$(4.5.9)\qquad \sigma^2(\dot{\underline{x}}_{ij}^{(\nu)}) = \frac{n-2}{n}(p_{ij}^{(\nu)} - \{p_{ij}^{(\nu)}\}^2) + \frac{1}{n^2}\sum_{\mu=1}^{n}(p_{ij}^{(\mu)} - \{p_{ij}^{(\mu)}\}^2),$$

$$(4.5.10)\qquad \sigma^2(\dot{\underline{x}}_{ij}^{(\nu)}|H_0^*) = \frac{n-1}{n}(p_{ij} - \{p_{ij}\}^2),$$

$$(4.5.11)\qquad \operatorname{cov}(\dot{\underline{x}}_{ij}^{(\nu)},\dot{\underline{x}}_{il}^{(\nu)}) = \frac{n-2}{n}(-p_{ij}^{(\nu)}p_{il}^{(\nu)}) + \frac{1}{n^2}\sum_{\mu=1}^{n}(-p_{ij}^{(\mu)}p_{il}^{(\mu)}),$$

$$(4.5.12)\qquad \operatorname{cov}(\dot{\underline{x}}_{ij}^{(\nu)},\dot{\underline{x}}_{il}^{(\nu)}|H_0^*) = \frac{n-1}{n}(-p_{ij}p_{il}), \qquad (j\neq l),$$

and the independence of the variables $\underline{x}_{ij}^{(\nu)}$ (with respect to the indices i and ν).

Furthermore, let

$$(4.5.13)\qquad L_i \overset{d}{=} \frac{n-1}{n}\begin{pmatrix} p_{i1}(1-p_{i1}) & -p_{i1}p_{i2} & \cdots & -p_{i1}p_{ik} \\ -p_{i2}p_{i1} & p_{i2}(1-p_{i2}) & \cdots & -p_{i2}p_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{ik}p_{i1} & -p_{ik}p_{i2} & \cdots & p_{ik}(1-p_{ik}) \end{pmatrix},$$

then

$$(4.5.14)\qquad \Sigma_{0i} = D(\vec{\underline{x}}_i|H_0^*) = N \otimes L_i,$$

$$(4.5.15)\qquad \Sigma_{0\cdot} = D(\vec{\underline{x}}_*|H_0^*) = N \otimes L_\cdot .$$

We shall obtain the a.d. of $\underline{w}(Q)$, under H_0^*, under the following assumption.

ASSUMPTION 3.

$$(4.5.16)\qquad \lim_{m\to\infty} p_{\cdot j} = p_j \,, \qquad \text{(say)},$$

$$(4.5.17)\qquad \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m}(p_{ij}-p_{\cdot j})(p_{il}-p_{\cdot l}) = e_{jl}\,, \qquad \text{(say)}.$$

Furthermore, we shall consider alternatives that satisfy

ASSUMPTION 4.

$$(4.5.18)\qquad \lim_{m\to\infty} p_{\cdot j}^{(\nu)} = p_j^{(\nu)}, \qquad \text{(say)},$$

$$(4.5.19)\qquad \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m}(p_{ij}^{(\nu)}-p_{\cdot j}^{(\nu)})(p_{il}^{(\nu)}-p_{\cdot l}^{(\nu)}) = e_{jl}^{(\nu)}, \qquad \text{(say)},$$

$$(4.5.20)\qquad \lim_{m\to\infty} p_{*j}^{(\nu)} - p_{*j}^{(\cdot)} = \varepsilon_j^{(\nu)}, \qquad \text{(say)},$$

where $|\varepsilon_j^{(\nu)}|$ may be ∞.

Let the vector $\vec{\varepsilon}$, with components $\varepsilon_j^{(\nu)}$ be constructed as in (2.2.16).

REMARK 4.5.1. The remarks that we have made on the plausability of the assumptions 1 & 2 in section 2.3. apply here also. Notice that assumption 4 & H_0^* imply assumption 3.

From (4.5.18) it follows that $\lim_{m\to\infty} p_{\cdot j}^{(\nu)} - p_{\cdot j}^{(\cdot)} = p_j^{(\nu)} - p_j^{(\cdot)}$. For those j and ν for which $p_j^{(\nu)} - p_j^{(\cdot)} \neq 0$, the $\varepsilon_j^{(\nu)}$ of (4.5.19) satisfy $|\varepsilon_j^{(\nu)}| = \infty$ and hence $\vec{\varepsilon}'\vec{\varepsilon} = \infty$. Only when $p_j^{(\nu)} - p_j^{(\cdot)} = 0$ for each j and ν we can have $|\varepsilon_j^{(\nu)}| < \infty$ for each j and ν, or $\vec{\varepsilon}'\vec{\varepsilon} < \infty$. However in that case the alternative clearly converges to H_0^*, and is not very interesting to us. So we shall consider mostly alternatives for which $\vec{\varepsilon}'\vec{\varepsilon} = \infty$, being the equivalent of the class of alternatives for which $\vec{\delta}'\vec{\delta} = \infty$ in the conditional case. For the sake of completeness, some results are also given for the case that $\vec{\varepsilon}'\vec{\varepsilon} < \infty$.

Under the assumption 3, we have

$$\lim_{m\to\infty} D(\vec{\underline{x}}_*|H_0^*) = \lim_{m\to\infty} \Sigma_{0\cdot} = \Sigma_0 , \qquad \text{(say)}, \tag{4.5.21}$$

while under assumption 4, the following limits exist

$$\lim_{m\to\infty} D(\vec{\underline{x}}_*) = \lim_{m\to\infty} \Sigma_{1\cdot} = \Sigma_1 , \qquad \text{(say)}, \tag{4.5.22}$$

$$\lim_{m\to\infty} E\vec{\underline{x}}_* = \lim_{m\to\infty} \vec{\varepsilon}_* = \vec{\varepsilon}. \tag{4.5.23}$$

It is again useful to define

$$\vec{\underline{z}}_i \stackrel{d}{=} \vec{\underline{x}}_i - \vec{\varepsilon}_i . \tag{4.5.24}$$

<u>THEOREM 4.5.1</u>. *Under the assumption 4, we have for* $m\to\infty$,

$$\vec{\underline{z}}_* \stackrel{\sim}{\to} N(\vec{0},\Sigma_1). \tag{4.5.25}$$

<u>PROOF</u>. The proof is analogous to the proof of theorem 4.2.1. □

<u>COROLLARY 4.5.1</u>. *Under assumption 4 & H_0^*, we have for* $m\to\infty$,

$$\vec{\underline{x}}_* \stackrel{\sim}{\to} N(\vec{0},\Sigma_0). \tag{4.5.26}$$

<u>PROOF</u>. As in corollary 4.2.1. □

<u>COROLLARY 4.5.2</u>. *Under the assumption 4, we have for* $\vec{\varepsilon}'\vec{\varepsilon} < \infty$,

$$\vec{\underline{x}}_* \stackrel{\sim}{\to} N(\vec{\varepsilon},\Sigma_1). \tag{4.5.27}$$

PROOF. As in corollary 4.2.2. □

THEOREM 4.5.2. *Under assumption 4 we have, under* H_0^* *and under alternatives such that* $\vec{\varepsilon}'\vec{\varepsilon} < \infty$,

(4.5.28) $\underline{w}(Q) \xrightarrow{L} \vec{\underline{x}}'Q\vec{\underline{x}}$

where

(4.5.29) $\vec{\underline{x}} \sim N(\vec{0},\Sigma_0)$, *under* H_0^*,

(4.5.30) $\vec{\underline{x}} \sim N(\vec{\varepsilon},\Sigma_1)$, *under alternatives such that* $\vec{\varepsilon}'\vec{\varepsilon} < \infty$.

PROOF. As in theorem 4.3.1. □

When we consider a sequence of experiments E_1', E_2', ... with probabilities $p_{ij}^{(\nu)}$ that satisfy assumption 4, it is a natural question to investigate whether from assumption 4 it follows that the limits of the quantities of assumption 1 & 2 exist (or at least almost surely, because these quantities are now random variables). While assumption 4 is sufficient to ensure the convergence of the distributions in the unconditional case, it is not sufficient to ensure convergence in the conditional case. We then need an additional assumption, which, however, is as plausible as the others, because it also involves the convergence of the arithmetic mean of a sequence. See also the remarks in section 2.3.

ASSUMPTION 5. For all j, l, ν and μ the following limits exist

(4.5.31) $\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} (p_{ij}^{(\nu)} - p_{\cdot j}^{(\nu)})(p_{il}^{(\mu)} - p_{\cdot l}^{(\mu)}) = e_{jl}^{(\nu,\mu)}$, (say).

Assumption 1 is implied by the assumptions 4 & 5 in the sense of the following theorem.

THEOREM 4.5.3. *Consider a sequence* E_1', E_2',... *satisfying assumptions 4 & 5. Then*

(4.5.32) $\underline{a}_{\cdot j} \xrightarrow{a.s.} p_j^{(+)}$,

(4.5.33) $\frac{1}{m} \sum_{i=1}^{m} (\underline{a}_{ij} - \underline{a}_{\cdot j})^2 \xrightarrow{a.s.} \sum_{\nu=1}^{n} p_j^{(\nu)}(1 - p_j^{(\nu)}) + \sum\sum_{\nu\neq\mu} e_{jj}^{(\nu,\mu)}$, (4.5.34) $$\frac{1}{m}\sum_{i=1}^{m}(\underline{a}_{ij}-\underline{a}_{\bullet j})(\underline{a}_{il}-\underline{a}_{\bullet l}) \xrightarrow{\text{a.s.}} -\sum_{\nu=1}^{n} p_j^{(\nu)} p_l^{(\nu)} + \sum_{\nu\neq\mu}\sum e_{jl}^{(\nu,\mu)} \qquad (j\neq l).$$

PROOF. The results follow with the strong law of the large numbers (RAO (1973), p.114) and the fact that $\sigma^2(\underline{a}_{ij})$, $\sigma^2(\underline{a}_{ij}^2)$ and $\sigma^2(\underline{a}_{ij}\underline{a}_{il})$ are all bounded, uniformly in i. □

For the first part of assumption 2 we have an equivalent theorem.

THEOREM 4.5.4. *Let* E_1', E_2', ... *be a sequence satisfying assumptions 4 & 5. Then*

(4.5.35) $$\underline{\delta}_{\bullet jl}^{(\nu,\mu)} \xrightarrow{\text{a.s.}} e_{jl}^{(\nu,\mu)} + p_j^{(\nu)} p_l^{(\mu)} - \sum_{\nu\neq\mu}\sum (e_{jl}^{(\nu,\mu)} + p_j^{(\nu)} p_l^{(\mu)}) \qquad (j\neq l, \nu\neq\mu);$$

(4.5.36) $$\underline{\delta}_{\bullet j}^{(\nu,\mu)} \xrightarrow{\text{a.s.}} e_{jj}^{(\nu,\mu)} + p_j^{(\nu)} p_j^{(\mu)} - \sum_{\nu\neq\mu}\sum (e_{jj}^{(\nu,\mu)} + p_j^{(\nu)} p_j^{(\mu)}) \qquad (\nu\neq\mu);$$

(4.5.37) $$\underline{\delta}_{\bullet j}^{(\nu)} \xrightarrow{\text{a.s.}} p_j^{(\nu)} - \frac{1}{n}\sum_{\nu=1}^{n} p_j^{(\nu)}.$$

PROOF. As in theorem 4.5.3. □

In the second part of assumption 2, (2.3.4), it is assumed that $\lim_{m\to\infty} \delta_{*j}^{(\nu)} = \delta_j^{(\nu)}$, where $|\delta_j^{(\nu)}|$ may be ∞. We have

THEOREM 4.5.5. *Let* E_1', E_2', ... *be a sequence satisfying assumptions 4 & 5. For each pair* (j,ν) *such that*

(4.5.38) $$p_j^{(\nu)} - p_j^{(\bullet)} \neq 0,$$

we have

(4.5.39) $$|\underline{\delta}_{*j}^{(\nu)}| \xrightarrow{\text{a.s.}} \infty.$$

PROOF. The result follows directly from (4.5.37) and (4.5.38). □

It is, however, possible that $p_j^{(\nu)} = p_j^{(\cdot)}$ for some j and ν. The a.s. convergence of $\underset{\sim}{\delta}_{*j}^{(\nu)}$ is then not guaranteed. But, for the consistency of the test in the conditional situation it is sufficient that $\vec{\delta}'\vec{\delta} = \infty$. We shall now show that (2.5.18) is sufficient to ensure the consistency of the test.

THEOREM 4.5.6. *A sufficient condition such that*

$$(4.5.40) \qquad \vec{\underset{\sim}{\delta}}_*'\vec{\underset{\sim}{\delta}}_* \xrightarrow{\text{a.s.}} \infty$$

is that

$$(4.5.41) \qquad \exists_{j\nu} \quad \left|\frac{1}{\sqrt{m}}\sum_{i=1}^{m}(p_{ij}^{(\nu)} - p_{ij}^{(\cdot)})\right| \to \infty \qquad \textit{for } m \to \infty$$

PROOF. Consider a j and ν for which (4.5.41) holds. Take $\underset{\sim}{x}_i \overset{d}{=} \underset{\sim}{\delta}_{ij}^{(\nu)}$, $\mu_i \overset{d}{=} p_{ij}^{(\nu)} - p_{ij}^{(\cdot)}$ and $\tilde{\underset{\sim}{x}}_i \overset{d}{=} \underset{\sim}{x}_i - \mu_i$. Then $E\underset{\sim}{x}_i = \mu_i$ and $E\tilde{\underset{\sim}{x}}_i = 0$. Furthermore, the variables $\tilde{\underset{\sim}{x}}_1$, $\tilde{\underset{\sim}{x}}_2$, ... are independent. Because $\underset{\sim}{\delta}_{ij}^{(\nu)}$ can only assume values between -1 and 1 for each i, $\sigma^2(\underset{\sim}{\delta}_{ij}^{(\nu)}) = \sigma^2(\tilde{\underset{\sim}{x}}_i)$ is bounded, uniformly in i. We have

$$\underset{\sim}{\delta}_{*j}^{(\nu)} = \underset{\sim}{x}_* = \frac{1}{\sqrt{m}}(\tilde{\underset{\sim}{x}}_1+\ldots+\tilde{\underset{\sim}{x}}_m) + \frac{1}{\sqrt{m}}(\mu_1+\ldots+\mu_m).$$

From assumption 4 (4.5.18) it follows that $\lim_{m\to\infty} m^{-1}(\mu_1+\ldots+\mu_m)$ exists and therefore

$$\frac{1}{\sqrt{m}}(\mu_1+\ldots+\mu_m) = O(\sqrt{m}) \qquad \text{for } m \to \infty$$

From the strong law of the large numbers (RAO (1973), p.114) it follows that

$$\frac{1}{m}(\tilde{\underset{\sim}{x}}_1+\ldots+\tilde{\underset{\sim}{x}}_m) \xrightarrow{\text{a.s.}} 0 \qquad \text{as } m \to \infty$$

and therefore

$$\frac{1}{\sqrt{m}}(\tilde{\underset{\sim}{x}}_1+\ldots+\tilde{\underset{\sim}{x}}_m) = o(\sqrt{m}) \quad \text{a.s.} \quad \text{as } m \to \infty$$

It follows that $m^{-\frac{1}{2}}(\tilde{\underset{\sim}{x}}_1+\ldots+\tilde{\underset{\sim}{x}}_m)$ is negligible with respect to $m^{-\frac{1}{2}}(\mu_1+\ldots+\mu_m)$ (a.s.). So (4.5.41) gives

$$|\underline{\delta}_{*j}^{(\nu)}| \xrightarrow{a.s.} \infty$$

and therefore

$$\vec{\underline{\delta}}_*' \vec{\underline{\delta}}_* \xrightarrow{a.s.} \infty.$$

□

Finally, some remarks must be made concerning the power of the test in the unconditional case. Recall that the proposed test is carried out conditionally. For the interesting alternatives in the unconditional case, (4.5.41) holds. Theorem 4.5.6. then gives that the (conditional) test is consistent (a.s.). This means also that the unconditional asymptotic power is equal to 1. In section 4.4. we have described for the conditional case what to do to approximate the power for finite m. In the unconditional case, the situation is even more complicated. Not only are the $\underline{a}_{ij}$ now random variables, but also the critical values are now random. However, a possibly crude approximation can be obtained by computing $E\underline{\delta}_{ij}^{(\nu)}$ and $E\underline{a}_{ij}$ from the $p_{ij}^{(\nu)}$ of the alternative considered and substituting these values into the formulas of section 4.4. for the conditional power.

CHAPTER 5

ASYMPTOTIC RELATIVE EFFICIENCIES

5.1. PITMAN EFFICIENCIES

To make a comparison possible between the different tests that we get for different choices of the matrix Q, we shall investigate the asymptotic relative Pitman efficiency (ARPE) and the asymptotic relative Bahadur efficiency (ARBE), for two different consistent tests based on $\underline{v}(Q_1)$ and $\underline{v}(Q_2)$.

We shall start in this section to give a definition of ARPE as this is given by ROTHE (1979). Then we shall give some theorems of Rothe and in the next section we shall apply his theory to our situation.

We quote from ROTHE (1979).

Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability distributions on a space (Ω, F), where Θ is an interval (finite or infinite) on the real line containing zero. Furthermore, $\{\underline{\varphi}_m\}$ is a sequence of level-α tests $(\alpha > 0)$ for H_0: $\theta = 0$ against H_1: $\theta \in \Theta \setminus \{0\}$. Take $\Theta' = \Theta \setminus \{0\}$. We shall assume that for every $\theta \neq 0$,

(5.1.1) $\quad E_\theta(\underline{\varphi}_m) \geq \alpha,$

(5.1.2) $\quad \lim_{m\to\infty} E_\theta(\underline{\varphi}_m) = 1.$

Usually, $\underline{\varphi}_m$ is a test based on m observations. Now the question arises how many observations are necessary to achieve a given power $\beta \in (\alpha, 1)$. So for $0 < \alpha < \beta < 1$, we define a function $N: \Theta' \to \mathbb{N}$, which is called a Pitman efficiency function for β (β - PEF), if

(5.1.3) $\quad E_\theta(\underline{\varphi}_{N(\theta)}) \geq \beta,$

(5.1.4) $\quad E_\theta(\underline{\varphi}_{N(\theta)-1}) < \beta,$

where $\underline{\varphi}_0 \equiv \alpha$.

Further, let

(5.1.5) $\qquad \underline{N_\beta}(\theta) = \inf\{n \in \mathbb{N} \,|\, E_\theta(\underline{\varphi}_n) \geq \beta\}$,

(5.1.6) $\qquad \overline{N_\beta}(\theta) = \inf\{n \in \mathbb{N} \,|\, E_\theta(\underline{\varphi}_m) \geq \beta \text{ for all } m \geq n\}$.

Now let Π denote the collection of all sequences $\{\theta_m\}$, with $\theta_m \in \Theta'$, $\theta_m \to 0$. Let $\{\underline{\varphi}_m^{(i)}\}$, $i = 1,2$ be two sequences of level-α tests with β-PEF $\underline{N_\beta}^{(i)}$, $\overline{N_\beta}^{(i)}$, respectively. Then

$$(5.1.7) \qquad e_{12}^- \overset{d}{=} \inf_\Pi \liminf_{m \to \infty} \frac{\underline{N_\beta}^{(2)}(\theta_m)}{\overline{N_\beta}^{(1)}(\theta_m)}$$

resp.

$$(5.1.8) \qquad e_{12}^+ \overset{d}{=} \sup_\Pi \limsup_{m \to \infty} \frac{\overline{N_\beta}^{(2)}(\theta_m)}{\underline{N_\beta}^{(1)}(\theta_m)}$$

are the lower (resp. upper) ARPE. If $e_{12}^- = e_{12}^+ = e_{12}$ (say) then e_{12} is the ARPE of $\{\underline{\varphi}_m^{(1)}\}$ with respect to $\{\underline{\varphi}_m^{(2)}\}$.

Now if the following three conditions are satisfied, a general theorem about e_{12} is applicable.

<u>CONDITION A</u>. There is a strictly increasing and bijective function $H: [0,\infty) \to [\alpha,1)$ such that for sequences $\{\theta_m\}$ in Θ satisfying $m\theta_m^2 \to \eta \geq 0$, as $m\to\infty$, we have

$$\lim_{m\to\infty} E_{\theta_m}(\underline{\varphi}_m) = H(\eta).$$

<u>CONDITION B</u>. For every $m \in \mathbb{N}$, the function Ψ_m: $\theta \to E_\theta(\underline{\varphi}_m)$ is continuous at $\theta = 0$.

<u>CONDITION C</u>. For every sequence $\{\theta_m\} \in \Pi$ such that $m\theta_m^2 \to \infty$, we have

$$E_{\theta_m}(\underline{\varphi}_m) \to 1.$$

THEOREM 5.1.1. *Let* $\{\underline{\varphi}_m^{(i)}\}$, $i = 1,2$ *be level-*α *test-sequences satisfying conditions A (with functions* H_i*, respectively), B and C. Then the ARPE of* $\{\underline{\varphi}_m^{(1)}\}$ *with respect to* $\{\underline{\varphi}_m^{(2)}\}$ *exists and is equal to*

$$(5.1.9) \qquad e_{12}(\beta) = \frac{H_2^{-1}(\beta)}{H_1^{-1}(\beta)}, \qquad \beta \in (\alpha,1).$$

PROOF. ROTHE (1979). □

Note that the function $H(\eta)$ of condition A is precisely the asymptotic power function of $\{\underline{\varphi}_{m,Q}\}$ under contiguous alternatives.

5.2. DETERMINATION OF "ARPE" IN OUR CASE

We shall apply the theory of section 5.1. in our situation using the contiguous alternatives as defined in section 2.4. So $a \in A_1$ is a fixed alternative, $\{\theta_m\}_{m=1}^{\infty}$ is a sequence in Θ' such that $m\theta_m^2 \to \eta \geq 0$, and $\{a_{\theta_m}\}_{m=1}^{\infty}$ is the associated contiguous alternative.

We consider matrices Q such that (4.1.3) is satisfied for $a \in A_1$, i.e. $\underline{v}(Q)$ is consistent against a. We use $\underline{\varphi}_{m,Q}$ as defined by (2.2.15).

In view of (4.1.3), (5.1.1) will hold, at least from a certain index m_1 on. When $\{\underline{\varphi}_{m,Q}\}$ is consistent against $a \in A_1$, it is also consistent against a_θ, for each θ, $0 < \theta \leq 1$, as follows easily from theorem 4.1.1. Therefore (5.1.2) is also satisfied. We shall now proceed to verify conditions A, B and C of the preceding section.

Condition A.

Let $\{\theta_m\}$ be a sequence in Θ such that $m\theta_m^2 \to \eta$. It follows from theorem 4.1.2 that $\underline{\vec{t}}_* \overset{\sim}{\to} N(\sqrt{\eta}\,\vec{\zeta},\Sigma_0)$ and from theorem 4.3.1. that $\underline{\vec{t}}_*'Q\underline{\vec{t}}_* \overset{L}{\to} \underline{\vec{x}}'Q\underline{\vec{x}}$, where $\underline{\vec{x}} \sim N(\sqrt{\eta}\,\vec{\zeta},\Sigma_0)$, under the contiguous alternative $\{a_{\theta_m}\}$. Then

$$\lim_{m\to\infty} E_{\theta_m}(\underline{\varphi}_{m,Q}) = P(\underline{\vec{x}}'Q\underline{\vec{x}} \geq k_{1-\alpha}(Q)),$$

where $k_{1-\alpha}(Q)$ is defined in (4.4.2).

The question is now whether

$$(5.2.1) \qquad H(\eta) \overset{d}{=} P(\underline{\vec{x}}'Q\underline{\vec{x}} \geq k_{1-\alpha}(Q))$$

is a strictly increasing function of η. We have

$$H(\eta) = P(\underline{\vec{x}}'Q\underline{\vec{x}} \geq k_{1-\alpha}(Q)) = P(\sum_{\tau=1}^{r} \lambda_\tau(\underline{u}_\tau + \sqrt{\eta}\,\omega_\tau)^2 \geq k_{1-\alpha}(Q)),$$

for some $r \in \mathbb{N}$, $\lambda_1 \geq \ldots \geq \lambda_r > 0$ and $\omega_1,\ldots,\omega_r \in \mathbb{R}$, $\underline{\vec{u}} \sim N(\vec{0}, I_r)$. Clearly $H(0) = \alpha$, $\lim_{\eta\to\infty} H(\eta) = 1$.

The fact that $H(\eta)$ is strictly increasing follows from lemma 3.3.1. So condition A is fulfilled, and $H(\eta)$ has a unique inverse $H^{-1}(\beta)$ for each $\beta \in (\alpha,1)$.

Condition B.

This follows from the fact that the exact distribution of $\underline{\vec{t}}_*$ depends on θ in a continuous way, when θ is in a sufficiently small neighbourhood of 0.

Condition C.

Condition C follows from the fact that for sequences $\{\theta_m\}$ such that $m\theta_m^2 \to \infty$, $E_{\theta_m} \underline{y}_m \to \infty$ and $E_{\theta_m} \underline{y}_m / \sigma_{\theta_m}(\underline{y}_m) \to \infty$, with $\underline{y}_m \overset{d}{=} 2\vec{\delta}_*'(a_{\theta_m})Q\underline{\vec{u}}_* + \vec{\delta}_*'(a_{\theta_m})Q\vec{\delta}_*(a_{\theta_m})$. The rest of the arguments are similar to those of theorem 4.1.1.

It follows that the conditions of theorem 5.1.1. are satisfied. The ARPE of $\{\underline{\varphi}_{m,Q_1}\}$ with respect to $\{\underline{\varphi}_{m,Q_2}\}$ is thus given by

$$(5.2.2) \qquad e_{12}(\beta) = \frac{H_2^{-1}(\beta)}{H_1^{-1}(\beta)}, \qquad \beta \in (\alpha,1).$$

With

$$(5.2.3) \qquad H_i(\eta) \overset{d}{=} P(\underline{\vec{x}}'Q_i\underline{\vec{x}} \geq k_{1-\alpha}(Q_i)), \qquad i = 1,2,$$

and $\underline{\vec{x}} \sim N(\sqrt{\eta}\,\vec{\zeta}, \Sigma_0)$.

No explicit formula can be given for $H_i^{-1}(\beta)$, though the inverse may be determined numerically. See chapter 9. In general, $e_{12}(\beta)$ will be dependent on α, β, Q_1, Q_2 and the particular alternative $a \in A_1$.

Furthermore, it turns out that $H(\eta)$ is nothing else than the asymptotic power of the test under contiguous alternatives, i.e. $H(m_0)$ is equal to (4.4.19), which we used as an approximation to the power of the test for m_0 trials.

When two matrices Q_1 and Q_2 are compared, the criterion (5.2.2) selects as best test the one that first reaches β, as $\sqrt{\eta}\,\vec{\zeta}$ tends away from $\vec{0}$. When $Q_1 \neq Q_2$ would imply $H_1(\eta) \leq H_2(\eta)$ for all η, or $H_1(\eta) \geq H_2(\eta)$ for all η, then a "best" Q could be selected and it would be independent of β. But this would be the same Q that would have been selected when we would have taken the Q that gives, for a fixed m_0, the highest asymptotic power against $\vec{\zeta}$. So in that case, the concept of Pitman-efficiency does not add anything that we do not already know. However, $Q_1 \neq Q_2$ does not always imply that $H_1 \leq H_2$ or $H_1 \geq H_2$ and so the selection of a "best" Q does depend on β.

All this means that the Pitman-efficiency is not very helpful in selecting a 'good' Q.

When, however, we approximate $H(\eta)$ by a function $H^*(\eta)$, using the approximation (3.3.19), i.e.

$$(5.2.4) \qquad H^*(\eta) \stackrel{d}{=} P(\underline{\chi}^2[\nu,\delta^2] \geq \chi^2_{1-\alpha}[\nu]),$$

with

$$(5.2.5) \qquad \nu = \sum_{\tau=1}^{r} \lambda_\tau, \qquad \delta^2 = \eta \sum_{\tau=1}^{r} \lambda_\tau \omega_\tau^2,$$

we have (ROTHE (1979))

$$(5.2.6) \qquad H^{*-1}(\beta) = \frac{c^2(\nu,\alpha,\beta)}{\delta^2},$$

where $c^2(\nu,\alpha,\beta)$ is the (uniquely determined) non-centrality parameter such that the β-fractile of $\underline{\chi}^2[\nu,c^2(\nu,\alpha,\beta)]$ and the α-fractile of $\underline{\chi}^2[\nu]$ coincide.

When we now compare Q_1 and Q_2, with

$$(5.2.7) \qquad \text{rank } B'Q_1B = \text{rank } B'Q_2B = r,$$

we may choose Q_1 and Q_2 such that (4.4.24) holds, i.e.

$$(5.2.8) \qquad \nu_1 = \nu_2 = \sum_{\tau=1}^{r} \lambda_\tau^{(1)} = \sum_{\tau=1}^{r} \lambda_\tau^{(2)} = r.$$

Then it follows from (3.2.7) that

$$(5.2.9)\qquad \delta_1^2 = \eta \sum_{\tau=1}^{r} \lambda_\tau^{(1)} (\omega_\tau^{(1)})^2 = \eta\, \vec{\zeta}' Q_1 \vec{\zeta}\ ,$$

$$(5.2.10)\qquad \delta_2^2 = \eta \sum_{\tau=1}^{r} \lambda_\tau^{(2)} (\omega_\tau^{(2)})^2 = \eta\, \vec{\zeta}' Q_2 \vec{\zeta}\ .$$

Then we may approximate the ARPE as follows

$$(5.2.11)\qquad e_{12}(\beta) = \frac{H_2^{-1}(\beta)}{H_1^{-1}(\beta)} \approx \frac{H_2^{*-1}(\beta)}{H_1^{*-1}(\beta)} = \frac{c^2(r,\alpha,\beta)}{\delta_2^2} \times \frac{\delta_1^2}{c^2(r,\alpha,\beta)} = \frac{\delta_1^2}{\delta_2^2} = \frac{\vec{\zeta}' Q_1 \vec{\zeta}}{\vec{\zeta}' Q_2 \vec{\zeta}}\ ,$$

which is independent of r, α and β. Moreover, this apporoximate value of $e_{12}(\beta)$ corresponds to the usual ARPE in the case of χ^2-distributions. Note that this approximate value (e_{12}^* (say)) may be calculated directly, without having to calculate eigenvalues etc.

When we would use e_{12}^* as a criterion to select a Q-matrix, we would choose the one that maximises the "non-centrality parameter", $\vec{\zeta}' Q \vec{\zeta}$, in accordance with usual practice.

5.3. BAHADUR EFFICIENCIES

The reader is referred to BAHADUR (1960) for his concept of comparisons of asymptotic slopes for the comparison of the different tests that we get for different choices of Q. We give here only short definitions of the "standard sequence" and the "Bahadur slope" of his article.

In order to compute the "Bahadur slope" for a sequence of test-statistics $\{\underline{T}_m\}_{m=1}^{\infty}$, this sequence has to be a "standard sequence", i.e. it has to satisfy the following three conditions.

i. There exists a continuous probability distribution function F such that, under H_0,

$$(5.3.1)\qquad \lim_{m\to\infty} P(\underline{T}_m < x \mid H_0) = F(x).$$

ii. There exists a constant a, $0 < a < \infty$, such that

$$(5.3.2)\qquad \log(1 - F(x)) = -\frac{ax^2}{2}(1 + o(1)), \qquad \text{as } x \to \infty.$$

iii. There exists a function b on H_1, with $0 \le b < \infty$, such that for each $\theta \in H_1$

$$(5.3.3) \qquad \lim_{m\to\infty} P_\theta\left(\left|\frac{\underline{T}_m}{\sqrt{m}} - b(\theta)\right| > \varepsilon\right) = 0 \qquad \text{for every } \varepsilon > 0.$$

The asymptotic Bahadur slope is now defined to be

$$(5.3.4) \qquad G(\theta) = a\{b(\theta)\}^2.$$

The asymptotic relative Bahadur efficiency (ARBE) for two standard sequences $\{\underline{T}_m^{(1)}\}_{m=1}^{\infty}$ and $\{\underline{T}_m^{(2)}\}_{m=1}^{\infty}$ is then defined as

$$(5.3.5) \qquad E_{12}(\theta) = \frac{G_1(\theta)}{G_2(\theta)} = \frac{a_1\{b_1(\theta)\}^2}{a_2\{b_2(\theta)\}^2}\,.$$

5.4. DETERMINATION OF "ARBE" IN OUR CASE

We apply the theory of ARBE for tests against alternatives in A_1, i.e. instead of $\theta \in H_1$, we shall write $a \in A_1$.

THEOREM 5.4.1. *When* Q *is chosen such that* $\vec{\delta}_*'Q\vec{\delta}_* \to \infty$ *for each* $a \in A_1$, *then* $\{(\underline{v}(Q))^{\frac{1}{2}}\}_{m=1}^{\infty} = \{(\vec{\underline{t}}_*'Q\vec{\underline{t}}_*)^{\frac{1}{2}}\}_{m=1}^{\infty}$ *is a standard sequence for testing* H_0.

PROOF. We verify the three conditions for a standard sequence.

i. By theorem 4.3.1., we have, under H_0, $\underline{v}(Q) \xrightarrow{L} \vec{\underline{x}}'Q\vec{\underline{x}}$, with $\vec{\underline{x}} \sim N(\vec{0},\Sigma_0)$, and so $(\underline{v}(Q))^{\frac{1}{2}} \xrightarrow{L} (\vec{\underline{x}}'Q\vec{\underline{x}})^{\frac{1}{2}}$, and $(\vec{\underline{x}}'Q\vec{\underline{x}})^{\frac{1}{2}}$ has a continuous distribution. This proves *i*.

ii. Using theorem 3.2.1., we have

$$1 - F(x) = 1 - P((\vec{\underline{x}}'Q\vec{\underline{x}})^{\frac{1}{2}} < x) = 1 - P(\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 < x^2),$$

with $\vec{\underline{u}} \sim N(\vec{0},I_r)$ and $\lambda_1 \ge \ldots \ge \lambda_r > 0$.
Note that

$$P(\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 > x^2) \le P(\lambda_1 \sum_{\tau=1}^{r} \underline{u}_\tau^2 > x^2) = P((\underline{\chi}^2[r])^{\frac{1}{2}} > \frac{x}{\sqrt{\lambda_1}})\ ;$$

$$P(\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 > x^2) \ge P(\lambda_1 \underline{u}_1^2 > x^2) = P((\underline{\chi}^2[1])^{\frac{1}{2}} > \frac{x}{\sqrt{\lambda_1}}).$$

So

$$\log P((\underline{\chi}^2[1])^{\frac{1}{2}} > \frac{x}{\sqrt{\lambda_1}}) \le \log(1-F(x)) \le \log P((\underline{\chi}^2[r])^{\frac{1}{2}} > \frac{x}{\sqrt{\lambda_1}}).$$

BAHADUR (1960) showed for χ-distributions that (5.3.2) is satisfied with $a = 1$, regardless of the number of degrees of freedom. It follows that

$$-\frac{x^2}{2\lambda_1}(1+o(1)) \le \log(1-F(x)) \le -\frac{x^2}{2\lambda_1}(1+o(1)),$$

and so

$$\log(1-F(x)) = -\frac{x^2}{2\lambda_1}(1+o(1)).$$

Hence (5.3.2) is satisfied in this case, with $a = \frac{1}{\lambda_1}$. This proves *ii*.

iii.

$$\frac{1}{m}\,\vec{\underline{t}}_*'Q\vec{\underline{t}}_* \equiv \frac{1}{m}\,\vec{\underline{u}}_*'Q\vec{\underline{u}}_* + \frac{2}{m}\,\vec{\delta}_*'Q\vec{\underline{u}}_* + \frac{1}{m}\,\vec{\delta}_*'Q\vec{\delta}_*.$$

By theorem 4.2.1. and 4.3.1., $\vec{\underline{u}}_*'Q\vec{\underline{u}}_* \overset{L}{\to} \vec{\underline{x}}'Q\vec{\underline{x}}$, so

$$\frac{1}{m}\,\vec{\underline{u}}_*'Q\vec{\underline{u}}_* \overset{P}{\to} 0.$$

The expectation of $\frac{2}{m}\,\vec{\delta}_*'Q\vec{\underline{u}}_*$ is zero, and its variance is equal to

$$\frac{4}{m^2}\,\vec{\delta}_*'Q\Sigma_{1.}Q\vec{\delta}_* = \frac{4}{m}\,\vec{\delta}_.'Q\Sigma_{1.}Q\vec{\delta}_. \;.$$

By assumption 2, $\vec{\delta}_.$ and $\Sigma_{1.}$ converge to a finite limit. Therefore

$$\lim_{m\to\infty} \frac{4}{m}\,\vec{\delta}_.'Q\Sigma_{1.}Q\vec{\delta}_. = 0.$$

It follows that

$$\frac{2}{m}\,\vec{\delta}_*'Q\vec{\underline{u}}_* \overset{P}{\to} 0.$$

Furthermore,

$$\lim_{m\to\infty} \frac{1}{m} \vec{\delta}_*'Q\vec{\delta}_* = \vec{\delta}_.'Q\vec{\delta}_. = \vec{\zeta}'Q\vec{\zeta} .$$

It follows that,

$$\frac{1}{m} \vec{\underline{t}}_*'Q\vec{\underline{t}}_* \overset{P}{\to} \vec{\zeta}'Q\vec{\zeta} ,$$

and with Slutzky's theorem,

$$(\frac{1}{m} \vec{\underline{t}}_*'Q\vec{\underline{t}}_*)^{\frac{1}{2}} \overset{P}{\to} (\vec{\zeta}'Q\vec{\zeta})^{\frac{1}{2}}.$$

This proves *iii*.

It follows from *i*., *ii*. and *iii*. that $(\vec{\underline{t}}_*'Q\vec{\underline{t}}_*)^{\frac{1}{2}}$ is a standard sequence. □

We find that the ARBE of $\underline{v}(Q_1)$ w.r.t. $\underline{v}(Q_2)$ is equal to

$$\text{(5.4.1)} \qquad E_{12}(a) = \frac{\frac{1}{\lambda_1^{(1)}} \vec{\zeta}'Q_1\vec{\zeta}}{\frac{1}{\lambda_1^{(2)}} \vec{\zeta}'Q_2\vec{\zeta}} .$$

This is almost equal to e_{12}^*. In the case that $\lambda_1^{(1)} = \lambda_1^{(2)}$ we even have $E_{12} = e_{12}^*$. This supports the use of e_{12}^* as a measure of relative efficiency. Notice that it is not surprising that the largest eigenvalue of $Q_1\Sigma_0$ and $Q_2\Sigma_0$ occur in E_{12}, because of their influence on the distribution of $\underline{v}(Q_l)$. When rank $B'Q_lB = r$, and Q_1 and Q_2 are chosen such that (4.4.24) holds, then $\lambda_1^{(1)}$ and $\lambda_1^{(2)}$ will not differ very much, so in that case $e_{12}^* \approx E_{12}$. We conclude that the measures of relative efficiency ARPE and ARBE are, in our situation, not essentially different.

CHAPTER 6

SPECIAL CASES & PRACTICE

6.1. MATRICES Q SUCH THAT THE A.D. OF THE TEST-STATISTIC IS CHI-SQUARED.

When for some reason a test-statistic is desired of which the a.d. under H_0 is chi-squared, it can be constructed as follows. Suppose that $\underline{\vec{x}} \sim N(\vec{0},\Sigma)$. Let Σ^- be any g-inverse of Σ. Then it follows from theorem 3.2.2. that the quadratic form $\underline{\vec{x}}'\Sigma^-\underline{\vec{x}}$ has a (central) χ^2-distribution, because $\Sigma\Sigma^-\Sigma\Sigma^-\Sigma = \Sigma\Sigma^-\Sigma$, by the properties of g-inverses. The number of degrees of freedom is then trace$(\Sigma^-\Sigma)$, which is equal to rank Σ, because $\Sigma^-\Sigma$ is idempotent (RAO (1973)).

By corollary 4.2.1., we have, under assumption 1 & H_0,

$$\underline{\vec{t}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_0), \qquad \Sigma_0 = N \otimes K.$$

So if we choose $Q = \Sigma_0^-$, it follows that $\underline{\vec{t}}_*'Q\underline{\vec{t}}_*$ has an asymptotic χ^2-distribution. Because of the special structure of Σ_0, we may choose Q of the form (2.2.10), $Q = I_n \otimes G$. Take $G = \frac{n-1}{n} K^-$, then $Q = I_n \otimes G$ is indeed a g-inverse of Σ_0, as is easily verified. Because the order of K is much smaller than the order of Σ_0, finding a g-inverse of K is more practicable. Of course, the natural g-inverse of K may be taken.

The number of degrees of freedom is

$$\text{trace}(\Sigma_0^-\Sigma_0) = \text{rank } \Sigma_0 \le (n-1)(k-1).$$

In practice we do not have Σ_0 at our disposal, and therefore we would choose $G = G(m) = \frac{n-1}{n} K_\bullet^-$ converging to $\frac{n-1}{n} K^-$ as $m \to \infty$. (See also remark 4.3.1).

Furthermore, if we choose $Q = \Sigma_0^-$, the approximation to the power of the test (4.4.19) becomes a non-central χ^2-distribution with trace$(\Sigma_0^-\Sigma_0)$ as degrees of freedom and non-centrality parameter

$$\sum_{\tau=1}^{n} \omega_\tau^2 = \vec{\omega}'\vec{\omega} = \vec{\zeta}'(a)\Sigma_0^- BP_+\Lambda_+^{-1}\Lambda_+^{-1}P_+'B'\Sigma_0^-\vec{\zeta}(a) = \vec{\zeta}'(a)\Sigma_0^-\Sigma_0\Sigma_0^-\vec{\zeta}(a).$$

(The positive eigenvalues of $Q\Sigma_0 = \Sigma_0^-\Sigma_0$ are all equal to 1).

It can be shown that there exists a g-inverse of Σ_0 such that the test-statistic based on it not only has the same distribution as the statistic MADANSKY (1963) proposed as a generalisation of Cochran's Q - test (see section 6.2.), but also gives the same numerical value for each $\vec{\omega} \in \Omega$. (The $\vec{\omega}$ and Ω as they were defined in section 2.1.). It follows that Madansky's test is a special case of the class of tests we investigate. The same follows automatically for Cochran's Q-test, which is in its turn a special case of Madansky's test. We return to Cochran's test in section 6.2.

Tests of this type, including the one of Madansky, have to be used with some care, because, depending on the kind of g-inverse which is chosen, Σ_0^- may not be of full rank. In view of the results of section 4.1., the test might not be consistent against each alternative in A_1.

In the cases $k = 2$ and $k = 3$, however, matrices Q can be found, such that Q is non-singular, diagonal, and such that the a.d. of $\underline{v}(Q)$ is chi-squared.

6.2. THE CASE k = 2

In the case that $k = 2$, the eigenvalues of the asymptotic distribution of $\vec{\underline{t}}_*'Q\vec{\underline{t}}_*$, under H_0, may be found by an elementary calculation, at least when we take Q of the form $Q = I_n \otimes G$. It is even more simple when we take G diagonal, i.e.

(6.2.1) $$G = \begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix}.$$

Recall that

(6.2.2) $$\Sigma_0 = N \otimes K,$$

with, because $k = 2$

(6.2.3) $$K = \begin{pmatrix} a & -a \\ -a & a \end{pmatrix},$$

(6.2.4) $$a = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \frac{a_{i1}a_{i2}}{n^2}.$$

To find the a.d. of $\vec{\underline{t}}_*'Q\vec{\underline{t}}_*$, we have to calculate the eigenvalues of

(6.2.5) $$Q\Sigma_0 = (I_n \otimes G)(N \otimes K) = N \otimes GK.$$

By lemma 3.1.5., the eigenvalues of $N \otimes GK$ can be found from the eigenvalues of $\frac{n}{n-1}$ GK. Let

(6.2.6) $$A \overset{d}{=} \frac{n}{n-1}\, a.$$

Then we have

$$\left|\frac{n}{n-1}\, GK - \lambda I_2\right| = \begin{vmatrix} g_1A - \lambda & -g_1A \\ -g_2A & g_2A - \lambda \end{vmatrix} = \begin{vmatrix} -\lambda & -g_1A \\ -\lambda & g_2A - \lambda \end{vmatrix} =$$

$$= -\lambda \begin{vmatrix} 1 & -g_1A \\ 1 & g_2A - \lambda \end{vmatrix} = -\lambda(g_2A-\lambda+g_1A).$$

It follows that the eigenvalues of $\frac{n}{n-1}$ GK are equal to

(6.2.7) $$\lambda = 0 \;\vee\; \lambda = (g_1+g_2)A = \frac{n}{n-1}(g_1+g_2)a.$$

The eigenvalues of $Q\Sigma_0$ are therefore $\lambda = (g_1+g_2)A$ with multiplicity n-1 and 0 with multiplicity n+1.

To obtain an a.d. which is chi-squared, we only have to choose g_1 and g_2 such that

(6.2.8) $$g_1 + g_2 = \frac{n-1}{n}\,\frac{1}{a}\,.$$

If we furthermore take

(6.2.9) $$g_1 > 0 \wedge g_2 > 0,$$

G, and hence Q, has full rank. The resulting test is then consistent against all alternatives in A_1.

In practical cases we take G = G(m) with

(6.2.10) $$g_1 + g_2 = \frac{n-1}{n}\,\frac{1}{a_m}$$

where

$$(6.2.11)\qquad a_m \overset{d}{=} \frac{1}{m}\sum_{i=1}^{m} \frac{a_{i1}a_{i2}}{n^2}\,.$$

From (6.2.8) or (6.2.10) it would appear that we still have a choice for g_1 and g_2. However for $k=2$, in this case

$$\begin{aligned}\underline{v}(Q) &\equiv \frac{1}{m}\sum_{\nu=1}^{n} g_1(\underline{f}_1^{(\nu)} - \frac{a_{+1}}{n})^2 + \frac{1}{m}\sum_{\nu=1}^{n} g_2(\underline{f}_2^{(\nu)} - \frac{a_{+2}}{n})^2 \equiv \\ &\equiv \frac{1}{m}\sum_{\nu=1}^{n} g_1(\underline{f}_1^{(\nu)} - \frac{a_{+1}}{n})^2 + \frac{1}{m}\sum_{\nu=1}^{n} g_2(m-\underline{f}_1^{(\nu)} - \frac{nm-a_{+1}}{n})^2 \equiv \\ &\equiv \frac{1}{m}(g_1+g_2)\sum_{\nu=1}^{n}(\underline{f}_1^{(\nu)} - \frac{a_{+1}}{n})^2.\end{aligned}$$

So with (6.2.10) this gives

$$(6.2.12)\qquad \underline{v}(Q) \equiv \frac{n-1}{n}\,\frac{\sum_{\nu=1}^{n}(\underline{f}_1^{(\nu)} - \frac{a_{+1}}{n})^2}{\frac{1}{n^2}\sum_{i=1}^{m} a_{i1}(n-a_{i1})} \equiv \frac{n(n-1)\sum_{\nu=1}^{n}(\underline{f}_1^{(\nu)} - \frac{a_{+1}}{n})^2}{n\sum_{i=1}^{m} a_{i1} - \sum_{i=1}^{m} a_{i1}^2}\,.$$

Hence in this case we obtain Cochran's Q-statistic (COCHRAN (1950)). The asymptotic distribution, under H_0, is then $\chi^2[n-1]$.

For the diagonal matrix G of (6.2.1) with g_1 and g_2 given by (6.2.10) we shall write G_2 and the corresponding statistic as $\underline{v}(G_2)$.

6.3. THE CASE $k=3$

In the case that $k=3$, we may proceed in the same way as for $k=2$. Let this time,

$$(6.3.1)\qquad G = \begin{pmatrix} g_1 & 0 & 0 \\ 0 & g_2 & 0 \\ 0 & 0 & g_3 \end{pmatrix}$$

and

$$(6.3.2)\qquad K = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}$$

with

$$(6.3.3) \qquad a = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \frac{a_{i1}(n-a_{i1})}{n^2}, \qquad \text{etc.,}$$

and

$$(6.3.4) \qquad A \overset{d}{=} \frac{n}{n-1} a, \qquad \text{etc.}$$

Note that

$$(6.3.5) \qquad A + D + E = D + B + F = E + F + C = 0.$$

It follows that

$$(6.3.6) \qquad AB - D^2 = BC - F^2 = DF - BE = AC - E^2 = -DC + EF = -AF + DE =$$

$$= -DC + FE = C_K \qquad \text{(say)}.$$

C_K may also be written as

$$(6.3.7) \qquad C_K = -\frac{1}{2}(AF + BE + CD).$$

Imposing the condition that the sum of the eigenvalues of $Q\Sigma_0$ must be equal to $(n-1)(k-1) = (n-1)\cdot 2$, we find for the eigenvalues of $\frac{n}{n-1}GK$,

$$(6.3.8) \qquad \lambda = 0 \ \vee \ \lambda = 1 \pm \frac{1}{2}\sqrt{4-4(k_1k_2-k_3)},$$

with

$$(6.3.9) \qquad k_1 \overset{d}{=} g_2B - g_1D; \quad k_2 \overset{d}{=} g_3C - g_1E; \quad k_3 \overset{d}{=} (g_3F-g_1D)(g_2F-g_1E).$$

Both non-zero eigenvalues are now equal to 1 if

$$(6.3.10) \qquad k_1k_2 - k_3 = 1.$$

It may be verified that this is the case only when

$$(6.3.11) \qquad g_1 = -F/C_K; \quad g_2 = -E/C_K; \quad g_3 = -D/C_K.$$

It follows that G, and hence Q, is non-singular, and the resulting test is therefore consistent against each alternative in A_1.

In practical cases we shall determine G = G(m) from $K_\bullet$ instead of K in the same way, and we shall call it $G_3 = G_3(m)$ and we shall write the corresponding statistic as $\underline{v}(G_3)$.

6.4. RECOMMENDATIONS

In section 4.1. it has been shown that the test based on the test-statistic $\underline{v}(Q) \equiv \vec{\underline{t}}_*' Q \vec{\underline{t}}_*$ is consistent against each alternative in A_1, the class of alternatives that we wish to detect, when Q is non-singular. When Q is singular, the test may, or may not, be consistent against each alternative in A_1. Because our aim was to design an overall test which is consistent against each alternative in A_1, we recommend the most simple form of test-statistic, i.e. with Q of the form $Q = I_n \otimes G$, with G diagonal with non-zero diagonal elements, so that Q is non-singular. The interpretation of the observations is easier when only quadratic terms occur in the test-statistic. Moreover, when H_0 is rejected it is possible to see from the term(s) which caused the rejection, where the preferences or aversions occurred. The drawback on the use of a diagonal G is, that for $k > 3$ it is not possible to define a G such that $\underline{v}(G)$ has a χ^2-distribution under all circumstances. But this disadvantage may be overcome by the application of a modified χ^2-approximation to the distribution of $\underline{v}(G)$.

Therefore, if there is no special interest in interaction between preferences, we recommend the use of a diagonal G. If the user attaches special weight to some categories he can adjust the weights accordingly. If there is no outside reason to weigh one category differently from others, the most "natural" weights, dependent on the number of occurrences of the categories, seem to be

$$(6.4.1) \qquad g_j = \{\frac{1}{m} \sum_{i=1}^{m} \frac{a_{ij}}{n}\}^{-1} = \{\frac{1}{m} \frac{a_{+j}}{n}\}^{-1}.$$

We shall call the diagonal matrix with these weights: G_g. The test-statistic then has the following form

$$(6.4.2) \qquad \underline{v}(G_g) \equiv \sum_{j=1}^{k} \sum_{\nu=1}^{n} \frac{(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})^2}{\frac{a_{+j}}{n}},$$

i.e. the form of the usual "goodness-of-fit" statistic.

This choice of g_j has the advantage that - as is apparent from the numerical results of chapter 9 - the approximation by means of an adapted χ^2-distribution seems to be somewhat better in this case than with other weights.

The a.d. of $\underline{v}(G_g)$, under H_0, may be determined using the methods of chapter 4, and when rank $Q\Sigma_0 = r$, with $Q = I_n \otimes G_g$, we can use a correction factor $\mathcal{C} = \mathcal{C}(m)$ to make sure that $\mathcal{C}\underline{v}(G_g)$ has

$$(6.4.3) \qquad \sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2 , \qquad \text{with} \qquad \sum_{\tau=1}^{r} \lambda_\tau = r ,$$

as asymptotic distribution. Then (4.4.24) is also satisfied.

In a special case (see section 6.5.) we have $\lambda_1 = \ldots = \lambda_r = 1$, so in that case the a.d. is $\chi^2[r]$.

In general, when $\sum_{\tau=1}^{r} \lambda_\tau = r$, the λ_τ's will not be very far away from 1, and the distribution of (6.4.3) will then closely resemble a $\chi^2[r]$ distribution. (The asymptotic expansion (3.3.7) for the distribution of $\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2$ seems to work best when the λ_τ's are not too far away from 1).

However, the exact moments of $\underline{v}(G_g)$ and of its a.d. may not be very close to each other. But, because the shape of the distribution of $\underline{v}(G_g)$ will resemble a χ^2-distribution, we can use the above mentioned approximation with a modified χ^2-distribution. The first two moments of $\underline{v}(G_g)$ are then equal to the first two moments of its approximation. The reader is referred to chapter 1 for a description of this approximation.

Because the expectation and variance have also been determined for $\underline{v}(G_1)$ and $\underline{v}(G_2)$, and in general for $\underline{v}(G)$ with diagonal G, see chapter 7, this method may also be applied to these variables.

To conclude, we recapitulate the reasons for the choice of $\underline{v}(G_g)$ as recommended test-statistic.

i. The test based on $\underline{v}(G_g)$ is consistent against each alternative in A_1;

ii. the test-statistic has a well-known, simple form, is easy to calculate and lends itself well for interpretation;

iii. in a special case, the a.d. is χ^2 and in general the a.d. will resemble a χ^2-distribution;

iv. the exact expectation and variance, under H_0, are known and a useful approximation exists, from which critical values may be determined.

6.5. ONE MORE SPECIAL CASE

In the special case that

(6.5.1) $$a_{1j} = a_{2j} = \ldots = a_{mj} \neq 0 \quad \text{for each } j,$$

the a.d. of $\frac{n-1}{n}\,\underline{v}(G_g)$, under H_0, is $\chi^2[(n-1)(k-1)]$.

This can be shown as follows. Notice that

(6.5.2) $$a_j = \lim_{m\to\infty} a_{\cdot j} = a_{ij} \quad \text{for each } i \text{ and } j.$$

Let

(6.5.3) $$H \stackrel{d}{=} \frac{1}{n}\begin{pmatrix} a_1 & a_2 & \cdot & \cdot & \cdot & a_k \\ a_1 & a_2 & \cdot & \cdot & \cdot & a_k \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ a_1 & a_2 & \cdot & \cdot & \cdot & a_k \end{pmatrix}.$$

Then from $a_+ = n$ it follows that H is idempotent, i.e.

(6.5.4) $$H^2 = H.$$

The diagonal elements of G_g, defined in (6.4.1), reduce under (6.5.1) to

(6.5.5) $$g_j = \frac{n}{a_j}$$

and K (defined in (2.3.10)) reduces to

(6.5.6) $$K = \frac{1}{n^2}\begin{pmatrix} a_1(n-a_1) & -a_1a_2 & \cdot & \cdot & \cdot & -a_1a_k \\ -a_2a_1 & a_2(n-a_2) & \cdot & \cdot & \cdot & -a_2a_k \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ -a_ka_1 & -a_ka_2 & \cdot & \cdot & \cdot & a_k(n-a_k) \end{pmatrix}.$$

The a.d. of $\frac{n-1}{n}\,\underline{v}(G_g)\,|H_0$ is determined by the eigenvalues of

$$Q\Sigma_0 = \frac{n-1}{n}(I_n \otimes G_g)(N \otimes K) = N \otimes \frac{n-1}{n}\,G_gK.$$

The eigenvalues may be found from the eigenvalues of $\frac{n}{n-1}(\frac{n-1}{n} G_g K) = G_g K$. We have

$$G_g K = \frac{n}{n^2} \begin{pmatrix} \frac{1}{a_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{a_k} \end{pmatrix} \begin{pmatrix} a_1(n-a_1) & -a_1 a_2 & \cdots & -a_1 a_k \\ -a_2 a_1 & a_2(n-a_2) & \cdots & -a_2 a_k \\ \vdots & \vdots & & \vdots \\ -a_k a_1 & -a_k a_2 & \cdots & a_k(n-a_k) \end{pmatrix} =$$

$$= \frac{1}{n} \begin{pmatrix} n-a_1 & -a_2 & \cdots & -a_k \\ -a_1 & n-a_2 & \cdots & -a_k \\ \vdots & \vdots & & \vdots \\ -a_1 & -a_2 & \cdots & n-a_k \end{pmatrix} = I_k - H.$$

Now from (6.5.4) it follows that

$$(6.5.7) \qquad (I_k - H)(I_k - H) = I_k - H - H + H^2 = I_k - H,$$

so that $I_k - H$ is also idempotent. It follows that the eigenvalues of $G_g K = I_k - H$ are either 0 or 1 (RAO (1973)). Therefore the eigenvalues of $Q\Sigma_0$ are also either 0 or 1.

The a.d. of $\frac{n-1}{n} \underline{v}(G_g)$ is then chi-squared with

$$(6.5.8) \qquad \text{trace } (Q\Sigma_0) = n \text{ trace}(\frac{n-1}{n} G_g K) = (n-1) \sum_{j=1}^{k} (1 - \frac{a_j}{n}) = (n-1)(k-1)$$

as number of degrees of freedom.

CHAPTER 7

EXPECTATION AND VARIANCE

7.1. NOTATION

In the exact expectation and variance of $\underline{v}(Q)$, which we shall derive for some special cases, the following quantities occur.

(7.1.1) $$E_j \overset{d}{=} n^{-1} \sum_{i=1}^{m} a_{ij},$$

(7.1.2) $$S_j \overset{d}{=} n^{-2} \sum_{i=1}^{m} a_{ij}(n-a_{ij}),$$

(7.1.3) $$T_j \overset{d}{=} n^{-4} \sum_{i=1}^{m} a_{ij}^2(n-a_{ij})^2,$$

(7.1.4) $$S_{jl} \overset{d}{=} n^{-2} \sum_{i=1}^{m} a_{ij}a_{il},$$

(7.1.5) $$T_{jl} \overset{d}{=} n^{-4} \sum_{i=1}^{m} a_{ij}^2 a_{il}^2.$$

7.2. EXPECTATION

The expectation of $\underline{v}(Q)$, which we already mentioned in (2.2.40),

(7.2.1) $$E\underline{v}(Q) = \mathrm{trace}(Q\Sigma_{1\cdot}) + \vec{\delta}_*'Q\vec{\delta}_*,$$

may be found as an application of the general formula for the expectation of a quadratic form (RAO (1973)).

Under H_0, it reduces to

(7.2.2) $$E(\underline{v}(Q)|H_0) = \mathrm{trace}(Q\Sigma_{0\cdot}),$$

and when $Q = I_n \otimes G$,

(7.2.3) $$E(\underline{v}(G)|H_0) = \frac{n}{m}\sum_{i=1}^{m} \mathrm{trace}(GK_i).$$

When G is moreover diagonal, we have

(7.2.4) $$E(\underline{v}(G)|H_0) = \frac{n}{m}\sum_{j=1}^{k} g_j S_j.$$

For $\underline{v}(G_2)$ we obtain

(7.2.5) $$E(\underline{v}(G_2)|H_0) = n-1,$$

which is thus also the expectation of Cochran's Q-statistic. For $\underline{v}(G_3)$ we have, using (6.3.7),

(7.2.6) $$E(\underline{v}(G_3)|H_0) = \frac{n}{m}g_1S_1 + \frac{n}{m}g_2S_2 + \frac{n}{m}g_3S_3 = -\frac{n}{m}\frac{1}{C_K}\{FS_1+ES_2+DS_3\} =$$

$$= -n\cdot\frac{1}{C_K}\frac{n-1}{n}\{FA+EB+DC\} = -(n-1)\cdot\frac{1}{C_K}\cdot -2C_K = 2(n-1).$$

Notice that for $\underline{v}(G_2)$ and $\underline{v}(G_3)$ it is not necessary to apply a correction-factor to make the test-statistic satisfy (4.4.24).

Finally we have for $\underline{v}(G_g)$

(7.2.7) $$E(\underline{v}(G_g)|H_0) = n\sum_{j=1}^{k}\frac{S_j}{E_j}.$$

In the special case that

(7.2.8) $$a_{1j} = a_{2j} = \dots = a_{mj} \qquad \text{for each } j,$$

formula (7.2.7) reduces to

(7.2.9) $$E(\underline{v}(G_g)|H_0) = n(k-1).$$

We have

LEMMA 7.2.1.

(7.2.10) $$E(\underline{v}(G_g)|H_0) \le n(k-1)$$

with equality iff (7.2.8) *holds.*

PROOF. The proof is left to the reader. □

7.3. VARIANCE

The variance of $\underline{v}(Q)$ is hard to determine in general and therefore we limit ourselves to a special case. Let $Q = I_n \otimes G$, with G diagonal. We shall determine $\sigma^2(\underline{v}(G)|H_0)$. The determination of $\sigma^2(\underline{v}(G))$ under alternatives is completely analogous, but would take too much space to reproduce here. We shall therefore suppose, throughout this section, that H_0 holds. Let

(7.3.1) $$\underline{s}_{ij}^{(\nu)} \overset{d}{=} \underline{\tilde{t}}_{ij}^{(\nu)} \equiv \underline{t}_{ij}^{(\nu)} - \frac{a_{ij}}{n},$$

(7.3.2) $$\underline{e}_j^{(\nu)} \overset{d}{=} \underline{\tilde{f}}_j^{(\nu)} \equiv \sum_{i=1}^{m} \underline{s}_{ij}^{(\nu)},$$

(7.3.3) $$\pi_{ij} \overset{d}{=} \frac{a_{ij}}{n}.$$

LEMMA 7.3.1. *We have for all* i, i_1, i_2, j *and* l, *except where otherwise indicated.*

(7.3.4) $$E\{\underline{s}_{ij}^{(1)}\}^2 = \sigma^2(\underline{t}_{ij}^{(1)}) = \pi_{ij} - \pi_{ij}^2,$$

(7.3.5) $$E\,\underline{s}_{ij}^{(1)}\underline{s}_{il}^{(1)} = \mathrm{cov}(\underline{t}_{ij}^{(1)},\underline{t}_{il}^{(1)}) = -\pi_{ij}\pi_{il}, \qquad j\neq l,$$

(7.3.6) $$E\,\underline{s}_{ij}^{(1)}\underline{s}_{ij}^{(2)} = \mathrm{cov}(\underline{t}_{ij}^{(1)},\underline{t}_{ij}^{(2)}) = -\frac{1}{n-1}\{\pi_{ij}-\pi_{ij}^2\},$$

(7.3.7) $$E\,\underline{s}_{ij}^{(1)}\underline{s}_{il}^{(2)} = \mathrm{cov}(\underline{t}_{ij}^{(1)},\underline{t}_{il}^{(2)}) = \frac{1}{n-1}\pi_{ij}\pi_{il}, \qquad j\neq l,$$

(7.3.8) $$\mathrm{cov}(\{\underline{s}_{ij}^{(1)}\}^2,\{\underline{s}_{il}^{(2)}\}^2) = -\frac{1}{n-1}\mathrm{cov}(\{\underline{s}_{ij}^{(1)}\}^2,\{\underline{s}_{il}^{(1)}\}^2),$$

(7.3.9) $$\mathrm{cov}(\{\underline{s}_{ij}^{(1)}\}^2,\{\underline{s}_{i_1 l}^{(2)}\}\{\underline{s}_{i_2 l}^{(2)}\}) = -\frac{1}{n-1}\mathrm{cov}(\{\underline{s}_{ij}^{(1)}\}^2,\{\underline{s}_{i_1 l}^{(1)}\}\{\underline{s}_{i_2 l}^{(1)}\}),$$

(7.3.10) $$\mathrm{cov}(\{\underline{s}_{i_1 j}^{(1)}\}\{\underline{s}_{i_2 j}^{(1)}\},\{\underline{s}_{il}^{(2)}\}^2) = -\frac{1}{n-1}\mathrm{cov}(\{\underline{s}_{i_1 j}^{(1)}\}\{\underline{s}_{i_2 j}^{(1)}\},\{\underline{s}_{il}^{(1)}\}^2).$$

PROOF. (7.3.4)-(7.3.7) follow from (2.2.24)-(2.2.27). We next prove (7.3.10). Write

(7.3.11) $\underline{x}^{(\nu)} \stackrel{d}{=} \{\underline{s}_{i_1j}^{(\nu)}\}\{\underline{s}_{i_2j}^{(\nu)}\}.$

Then, for each i and l,

$$\mathrm{cov}(\underline{x}^{(1)},\{\underline{s}_{il}^{(2)}\}^2) = \mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(2)}(1-2\pi_{il})) =$$

$$= (1-2\pi_{il})\mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(2)}) .$$

Observe that

$$\sum_{\nu=1}^{n} \mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(\nu)}) = \mathrm{cov}(\underline{x}^{(1)},\sum_{\nu=1}^{n} \underline{t}_{il}^{(\nu)}) = \mathrm{cov}(\underline{x}^{(1)},a_{ij}) = 0.$$

Thus, using the fact that, due to symmetry, the joint distributions of the pairs $(\underline{x}^{(1)},\underline{t}_{il}^{(2)}),\ldots,(\underline{x}^{(1)},\underline{t}_{il}^{(n)})$ are the same,

$$\mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(1)}) + (n-1)\mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(2)}) = 0.$$

Therefore

$$\mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(2)}) = -\frac{1}{n-1}\mathrm{cov}(\underline{x}^{(1)},\underline{t}_{il}^{(1)}).$$

And thus also

$$\mathrm{cov}(\underline{x}^{(1)},\{\underline{s}_{il}^{(2)}\}^2) = -\frac{1}{n-1}\mathrm{cov}(\underline{x}^{(1)},\{\underline{s}_{il}^{(1)}\}^2).$$

This proves (7.3.10). Simultaneous interchanging of j and l and (1) and (2) in (7.3.10) gives (7.3.9). (7.3.8) follows if we take

$$\underline{x}_j^{(\nu)} \stackrel{d}{=} \{\underline{s}_{ij}^{(\nu)}\}^2$$

instead of (7.3.11). □

We use the following notation

$$(7.3.12)\qquad \sum{}^{*} = \sum_{\substack{i_1=1\\ i_1\neq i_2}}^{m}\ \sum_{i_2=1}^{m}$$

LEMMA 7.3.2.

$$(7.3.13)\qquad n\,\mathrm{cov}(\textstyle\sum^{*}\{\underline{s}_{i_1j}^{(1)}\}\{\underline{s}_{i_2j}^{(1)}\},\sum^{*}\{\underline{s}_{i_1l}^{(1)}\}\{\underline{s}_{i_2l}^{(1)}\}) +$$

$$+\, n(n-1)\,\mathrm{cov}(\textstyle\sum^{*}\{\underline{s}_{i_1j}^{(1)}\}\{\underline{s}_{i_2j}^{(1)}\},\sum^{*}\{\underline{s}_{i_1l}^{(2)}\}\{\underline{s}_{i_2l}^{(2)}\}) =$$

$$= \begin{cases} \dfrac{2n^2}{n-1}\sum^{*} \pi_{i_1j}\pi_{i_1l}\pi_{i_2j}\pi_{i_2l} & \textit{if } j\neq l, \\[2ex] \dfrac{2n^2}{n-1}\sum^{*} \pi_{i_1j}(1-\pi_{i_1j})\pi_{i_2j}(1-\pi_{i_2j}) & \textit{if } j=l. \end{cases}$$

PROOF. Using independence when $i_1 \neq i_2$, $E\underline{s}_{ij}^{(\nu)} = 0$, and (7.3.5) we have for $j \neq l$,

$$\mathrm{cov}(\textstyle\sum^{*}\{\underline{s}_{i_1j}^{(1)}\}\{\underline{s}_{i_2j}^{(1)}\},\sum^{*}\{\underline{s}_{i_1l}^{(1)}\}\{\underline{s}_{i_2l}^{(1)}\}) =$$

$$= 2\textstyle\sum^{*} E\ \underline{s}_{i_1j}^{(1)}\ \underline{s}_{i_2j}^{(1)}\ \underline{s}_{i_1l}^{(1)}\ \underline{s}_{i_2l}^{(1)} =$$

$$= 2\textstyle\sum^{*} E\ \underline{s}_{i_1j}^{(1)}\ \underline{s}_{i_1l}^{(1)}\ E\ \underline{s}_{i_2j}^{(1)}\ \underline{s}_{i_2l}^{(1)} =$$

$$= 2\textstyle\sum^{*} (-\pi_{i_1j}\pi_{i_1l})(-\pi_{i_2j}\pi_{i_2l}) = 2\sum^{*} \pi_{i_1j}\pi_{i_1l}\pi_{i_2j}\pi_{i_2l}\ .$$

Also, now using (7.3.7),

$$\mathrm{cov}(\textstyle\sum^{*}\{\underline{s}_{i_1j}^{(1)}\}\{\underline{s}_{i_2j}^{(1)}\},\sum^{*}\{\underline{s}_{i_1l}^{(2)}\}\{\underline{s}_{i_2l}^{(2)}\}) =$$

$$= \frac{2}{(n-1)^2}\textstyle\sum^{*} \pi_{i_1j}\pi_{i_1l}\pi_{i_2j}\pi_{i_2l}\ .$$

These two results together give (7.3.13) if $j \neq l$. The case $j = l$ is proved analogously. □

THEOREM 7.3.1. *For diagonal* G, *we have*

(7.3.14) $$\sigma^2(\underline{v}(G)|H_0) = \frac{1}{m^2}\frac{2n^2}{n-1}\{\sum_{j=1}^{k} g_j^2(S_j^2-T_j) + \sum_{j\neq l}\sum g_j g_l (S_{jl}^2-T_{jl})\}.$$

PROOF. We have

$$\underline{v}(G) \equiv \frac{1}{m}\sum_{j=1}^{k}\sum_{\nu=1}^{n} g_j(\underline{f}_j^{(\nu)} - \frac{a_{+j}}{n})^2 \equiv \frac{1}{m}\sum_{j=1}^{k}\sum_{\nu=1}^{n} g_j\{\underline{e}_j^{(\nu)}\}^2.$$

Using permutability over the index ν we have, under H_0,

$$\sigma^2(\underline{v}(G)) = \frac{1}{m^2}\sum_{j=1}^{k}\sum_{l=1}^{k} g_j g_l \operatorname{cov}(\sum_{\nu=1}^{n}\{\underline{e}_j^{(\nu)}\}^2, \sum_{\mu=1}^{n}\{\underline{e}_l^{(\mu)}\}^2) =$$

$$= \frac{1}{m^2}\sum_{j=1}^{k}\sum_{l=1}^{k} g_j g_l\{\sum_{\nu=1}^{n}\sum_{\mu=1}^{n}\operatorname{cov}(\{\underline{e}_j^{(\nu)}\}^2,\{\underline{e}_l^{(\mu)}\}^2)\} =$$

$$= \frac{1}{m^2}\sum_{j=1}^{k}\sum_{l=1}^{k} g_j g_l\{n \operatorname{cov}(\{\underline{e}_j^{(1)}\}^2,\{\underline{e}_l^{(1)}\}^2) +$$

$$+ n(n-1)\operatorname{cov}(\{\underline{e}_j^{(1)}\}^2,\{\underline{e}_l^{(2)}\}^2)\}.$$

Next, observe that, using (7.3.8), (7.3.9) and (7.3.10),

$$n \operatorname{cov}(\{\underline{e}_j^{(1)}\}^2, \{\underline{e}_l^{(1)}\}^2) + n(n-1)\operatorname{cov}(\{\underline{e}_j^{(1)}\}^2,\{\underline{e}_l^{(2)}\}^2) =$$

$$= n \operatorname{cov}(\sum_{i=1}^{m}\{\underline{s}_{ij}^{(1)}\}^2 + \sum{}^{*}\underline{s}_{i_1 j}^{(1)}\underline{s}_{i_2 j}^{(1)}, \sum_{i=1}^{m}\{\underline{s}_{il}^{(1)}\}^2 + \sum{}^{*}\underline{s}_{i_1 l}^{(1)}\underline{s}_{i_2 l}^{(1)}) +$$

$$+ n(n-1)\operatorname{cov}(\sum_{i=1}^{m}\{\underline{s}_{ij}^{(1)}\}^2 + \sum{}^{*}\underline{s}_{i_1 j}^{(1)}\underline{s}_{i_2 j}^{(1)}, \sum_{i=1}^{m}\{\underline{s}_{il}^{(2)}\}^2 + \sum{}^{*}\underline{s}_{i_1 l}^{(2)}\underline{s}_{i_2 l}^{(2)}) =$$

$$= n \operatorname{cov}(\sum{}^{*}\underline{s}_{i_1 j}^{(1)}\underline{s}_{i_2 j}^{(1)}, \sum{}^{*}\underline{s}_{i_1 l}^{(1)}\underline{s}_{i_2 l}^{(1)}) +$$

$$+ n(n-1)\operatorname{cov}(\sum{}^{*}\underline{s}_{i_1 j}^{(1)}\underline{s}_{i_2 j}^{(1)}, \sum{}^{*}\underline{s}_{i_1 l}^{(2)}\underline{s}_{i_2 l}^{(2)}) .$$

This is just the expression of lemma 7.3.2.; the rest of the proof is simple calculation. □

For $\underline{v}(G_2)$ we obtain

(7.3.15) $$\sigma^2(\underline{v}(G_2)|H_0) = 2(n-1)(\frac{S_1^2 - T_1}{S_1^2}) ,$$

the variance of Cochran's Q-statistic (under H_0).

Substitution of (6.3.11) in (7.3.14) does not lead to a simpler form for $\sigma^2(\underline{v}(G_3)|H_0)$.

For $\underline{v}(G_g)$ we obtain

(7.3.16) $$\sigma^2(\underline{v}(G_g)|H_0) = \frac{2n^2}{n-1}\{\sum_{j=1}^{k}\frac{S_j^2-T_j}{E_j} + \sum_{j\neq l}\sum\frac{S_{jl}^2-T_{jl}}{E_jE_l}\},$$

which reduces in the special case (7.2.8) to

(7.3.17) $$\sigma^2(\underline{v}(G_g)|H_0) = \frac{2n^2}{n-1}(k-1)\frac{m-1}{m}\ .$$

CHAPTER 8

MOTIVATION OF THE CHOICE OF QUADRATIC FORMS

To derive tests for the simple hypothesis H_0: $\forall_i \; \vec{\Delta}_i = \vec{0}$, against the composite hypothesis H_1: $\exists_i \; \vec{\Delta}_i \neq \vec{0}$, with certain optimality properties, there are basically only a few methods. We shall show why two standard methods fail in our situation, and why we have therefore chosen for a third method based on asymptotic distributions.

8.1. THE NEYMAN & PEARSON FUNDAMENTAL LEMMA - METHOD

Consider the problem of testing the simple hypothesis

(8.1.1) $$H_0: \quad \forall_i \quad \vec{\Delta}_i = \vec{0},$$

against the simple alternative

(8.1.2) $$H_1: \quad \vec{\Delta}_i = \vec{\Delta}_i^{*} ,$$

where $\vec{\Delta}_1^{*},\ldots,\vec{\Delta}_m^{*}$ are fixed elements from $\mathcal{D}_1,\ldots,\mathcal{D}_m$. We have

(8.1.3) $$P(\underline{\omega}_i = \pi_{ir} | H_0) = \frac{1}{N_i} , \qquad r \in \mathcal{R}_i,$$

(8.1.4) $$P(\underline{\omega}_1 = \pi_{1r_1} \wedge \ldots \wedge \underline{\omega}_m = \pi_{mr_m} | H_0) = \prod_{i=1}^{m} \frac{1}{N_i} , \qquad r_i \in \mathcal{R}_i,$$

(8.1.5) $$P(\underline{\omega}_i = \pi_{ir} | H_1) = \frac{1}{N_i} + \Delta_{ir}^{*}, \qquad r \in \mathcal{R}_i,$$

(8.1.6) $$P(\underline{\omega}_1 = \pi_{1r_1} \wedge \ldots \wedge \underline{\omega}_m = \pi_{mr_m} | H_1) = \prod_{i=1}^{m} (\frac{1}{N_i} + \Delta_{ir_i}^{*}), \qquad r_i \in \mathcal{R}_i.$$

According to Neyman & Pearson's fundamental lemma, the most powerful test rejects H_0 for large values of the quotient of (8.1.6) and (8.1.4), i.e.

$$(8.1.7)\qquad \frac{\prod_{i=1}^{m}\left(\frac{1}{N_i}+\Delta^*_{ir_i}\right)}{\prod_{i=1}^{m}\frac{1}{N_i}} = \prod_{i=1}^{m}(1+N_i\Delta^*_{ir_i}),$$

leading to the test-statistic

$$(8.1.8)\qquad \underline{T}_1 \equiv \prod_{i=1}^{m}(1+N_i\Delta_i^*(\underline{\omega}_i)),$$

where $\Delta_i^*(.)$ is a function $\Omega_i \to \mathbb{R}$, with

$$(8.1.9)\qquad \Delta_i^*(\pi_{ir}) = \Delta^*_{ir}, \qquad r\in R_i.$$

For given $\vec{\Delta}_i^*$, this may lead to a useful test, with, after taking the logarithm,

$$(8.1.10)\qquad E\log \underline{T}_1 = \sum_{i=1}^{m}\sum_{r=1}^{N_i}\left(\frac{1}{N_i}+\Delta^*_{ir}\right)\log(1+N_i\Delta^*_{ir}),$$

$$(8.1.11)\qquad E(\log \underline{T}_1|H_0) = 0.$$

By the Central Limit Theorem, $\log \underline{T}_1$ is asymptotically normal. Critical values could be determined by enumerating the exact distribution, or can be based on the asymptotic distribution, after having computed the variance of $\log \underline{T}_1$.

If, however, we are interested in the behaviour of this test also for other (or even all) alternatives from A_1, then it is not at all clear how $E\log \underline{T}_1$ behaves under these alternatives. Moreover, we see no way to adapt it to work against other alternatives in A_1 too. Therefore, we do not pursue this method any further.

8.2. THE LIKELIHOOD-RATIO METHOD

Consider the problem of testing the simple hypothesis

$$(8.2.1)\qquad H_0:\quad \forall_i \quad \vec{\Delta}_i = \vec{0}$$

against

$$(8.2.2) \qquad H_1: \quad \exists_i \quad \vec{\Delta}_i \neq \vec{0}.$$

The likelihood-ratio test rejects H_0 for large values of

$$(8.2.3) \qquad \Lambda = \frac{\sup_{\vec{\Delta}_i} P(\underline{\omega}_1 = \pi_{1r_1} \wedge \ldots \wedge \underline{\omega}_m = \pi_{mr_m} | H_1)}{P(\underline{\omega}_1 = \pi_{1r_1} \wedge \ldots \wedge \underline{\omega}_m = \pi_{mr_m} | H_0)} =$$

$$= \frac{\sup_{\vec{\Delta}_i} \prod_{i=1}^{m} (\frac{1}{N_i} + \Delta_{ir_i})}{\prod_{i=1}^{m} \frac{1}{N_i}} = \frac{1}{\prod_{i=1}^{m} \frac{1}{N_i}} = \prod_{i=1}^{m} N_i .$$

The likelihood-ratio is in this case apparently a constant, and is therefore also unfit to produce a useful test-statistic, to test H_0 against this wide class of alternatives.

Now suppose that we only wish to consider alternatives from H_1 for which words beginning with the character C_1 have, for each i, a higher probability than the other words. This restricts the possibilities considerably and so

$$\sup_{\vec{\Delta}_i} \prod_{i=1}^{m} (\frac{1}{N_i} + \Delta_{ir_i})$$

will take a lower value if the $\vec{\Delta}_i$ may only range over these restricted alternatives. Define

$$(8.2.4) \qquad N_i' \stackrel{d}{=} \frac{(n-1)!}{(a_{i1}-1)!a_{i2}! \ldots a_{ik}!} ,$$

i.e. N_i' is the number of words with C_1 in the first position in the set of outcomes Ω_i of the i'th trial. Let

$$(8.2.5) \qquad R_i' \stackrel{d}{=} \{1,\ldots,N_i'\}.$$

Then restricted alternatives may be formulated as follows

$$(8.2.6) \qquad H_1^{**}: \forall_i \quad \Delta_{ir} > 0 \quad \text{for } r \in R_i' \quad \text{and} \quad \Delta_{ir} \leq 0 \quad \text{for } r \notin R_i' .$$

Notice that $H_1^{**} \Rightarrow H_1'$ but not conversely, so that a possible test derived in this way only works against this much smaller class of alternatives. We have

$$\text{(8.2.7)} \qquad \sup_{\vec{\Delta}_i \in H_1^{**}} \left(\frac{1}{N_i} + \Delta_{ir_i}\right) = \begin{cases} 1 \text{ if } r_i \in \mathcal{R}_i', \\ \frac{1}{N_i} \text{ if } r_i \notin \mathcal{R}_i'. \end{cases}$$

This may be written as

$$\text{(8.2.8)} \qquad \sup_{\vec{\Delta}_i \in H_1^{**}} \left(\frac{1}{N_i} + \Delta_{ir_i}\right) = \left(\frac{1}{N_i}\right)^{1-t_{i1}^{(1)}(\pi_{ir_i})}.$$

So the likelihood-ratio becomes

$$\text{(8.2.9)} \qquad \Lambda = \frac{\prod_{i=1}^{m} \left(\frac{1}{N_i}\right)^{1-t_{i1}^{(1)}(\pi_{ir_i})}}{\prod_{i=1}^{m} \left(\frac{1}{N_i}\right)}$$

and the likelihood-ratio test rejects H_0 for large values of

$$\text{(8.2.10)} \qquad \prod_{i=1}^{m} \left(\frac{1}{N_i}\right)^{1-\underline{t}_{i1}^{(1)}}$$

or, equivalently, for large values of

$$\text{(8.2.11)} \qquad \underline{T}_1^{(1)} \stackrel{d}{=} \sum_{i=1}^{m} \underline{t}_{i1}^{(1)} \log N_i.$$

Analogously we may define

$$\text{(8.2.12)} \qquad \underline{T}_j^{(\nu)} \stackrel{d}{=} \sum_{i=1}^{m} \underline{t}_{ij}^{(\nu)} \log N_i.$$

Aiming at an overall test, as we do, we do not know in advance where possible preferences occur, so we might combine these statistics to, for instance

$$\text{(8.2.13)} \qquad \underline{T}_2 \stackrel{d}{=} \max_{j,\nu} \underline{T}_j^{(\nu)},$$

to get an overall test-statistic for H_0. This is possible, because the

$\underline{T}_j^{(\nu)}$ do not depend on any particular alternative. This would lead to a test for an outlier among the characters

Let's now study the r.v.'s $\underline{T}_j^{(\nu)}$ in greater detail. We have, using the results of chapter 2,

$$(8.2.14) \qquad E\underline{T}_j^{(\nu)} = \sum_{i=1}^{m} \frac{a_{ij}}{n} \log N_i + \sum_{i=1}^{m} \delta_{ij}^{(\nu)} \log N_i ,$$

$$(8.2.15) \qquad \sigma^2(\underline{T}_j^{(\nu)}) = \sum_{i=1}^{m} \{\frac{a_{ij}}{n} - \frac{a_{ij}^2}{n^2}\} \log^2 N_i \; + $$

$$+ \sum_{i=1}^{m} \{-2\delta_{ij}^{(\nu)}\frac{a_{ij}}{n} + \delta_{ij}^{(\nu)} - (\delta_{ij}^{(\nu)})^2\} \log^2 N_i \; .$$

The expectation and variance under H_0 are found by deleting the terms containing δ's.

The variables $\underline{T}_j^{(\nu)}$ are not independent, and their covariances may be found using (2.1.27) and (2.1.28). The (marginal) a.d.'s are normal by the C.L.T. The joint a.d. of the $\underline{T}_j^{(\nu)}$ may be found using the methods of chapter 4.

The distribution of $\underline{T}_2$, however, is difficult to obtain, the exact distribution as well as the asymptotic distribution (JOHNSON & KOTZ (1972), p.44). The development of this outlier-test would be an interesting subject for further research.

The result (8.2.14) suggests the use of a test-statistic similar to the one defined in (2.1.39)

$$(8.2.16) \qquad \underline{T}_3 \stackrel{d}{=} \frac{1}{m} \sum_{j=1}^{k} g_j \sum_{\nu=1}^{n} \{ \sum_{i=1}^{m} (\underline{t}_{ij}^{(\nu)} - \frac{a_{ij}}{n}) \log N_i \}^2 \; .$$

This statistic, which gives trials with a high number of possible words more weight than $\underline{v}(G)$ does, may be treated in the same way as the statistic $\underline{v}(G)$. Its a.d. may be determined in a similar way as that of $\underline{v}(G)$, both under H_0 and under alternatives.

By these considerations we could be led to consider a class of test-statistics which is even more general than (2.1.41), of the form

$$(8.2.17) \qquad \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} \sum_{j=1}^{k} \sum_{l=1}^{k} g_{jl}^{(\nu,\mu)} \{ \sum_{i=1}^{m} G_i (\underline{t}_{ij}^{(\nu)} - \frac{a_{ij}}{n}) \} \{ \sum_{i=1}^{m} G_i (\underline{t}_{il}^{(\mu)} - \frac{a_{il}}{n}) \} .$$

with

(8.2.18) $G_i = \log N_i$

as suggested weights.

Although the analysis of the behaviour of such an extensive class of test-statistics would lead to a new and major enterprise, the suggestion that the use of $G_i = \log N_i$ would possibly increase the power of our tests in an adapted form is worthy of future consideration.

8.3. AN APPROXIMATE LIKELIHOOD-RATIO METHOD

Let $\underline{\vec{x}}$ have a q-variate normal distribution, with mean vector $\vec{\mu}$ and dispersion matrix Σ. ($\underline{\vec{x}} \sim N_q(\vec{\mu},\Sigma)$). Suppose that $\vec{\mu}$ is unknown, but that Σ is a known, fixed, positive definite matrix. For the problem of testing H_0: $\vec{\mu} = \vec{\mu}_0$ against H_1: $\vec{\mu} \neq \vec{\mu}_0$, the likelihood-ratio is equal to

$$(8.3.1) \qquad \Lambda(\vec{x}) = \frac{\sup_{H_1} (2\pi)^{-q/2}|\Sigma|^{-\frac{1}{2}}\exp\{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})\}}{\sup_{H_0} (2\pi)^{-q/2}|\Sigma|^{-\frac{1}{2}}\exp\{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})\}}.$$

Clearly,

$$(8.3.2) \qquad \sup_{H_0} \exp\{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})\} = \exp\{-\frac{1}{2}(\vec{x}-\vec{\mu}_0)'\Sigma^{-1}(\vec{x}-\vec{\mu}_0)\}.$$

Furthermore,

$$(8.3.3) \qquad \sup_{H_1} \exp\{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})\} = 1,$$

because the infimum of $(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})$ is equal to 0, at $\vec{\mu} = \vec{x}$.

So the likelihood-ratio test rejects H_0 for large values of the statistic,

$$(8.3.4) \qquad (\underline{\vec{x}}-\vec{\mu}_0)'\Sigma^{-1}(\underline{\vec{x}}-\vec{\mu}_0),$$

i.e. a quadratic form in $\underline{\vec{x}}$, where the weighing coefficients are elements of the inverse of the covariance matrix Σ. It follows easily from theorem 3.2.1. that the a.d. of this statistic is a central χ^2-distribution with q degrees of freedom under H_0, and a non-central χ^2-distribution with q degrees of freedom and non-centrality parameter $(\vec{\mu}-\vec{\mu}_0)'\Sigma^{-1}(\vec{\mu}-\vec{\mu}_0)$ under H_1. These are well-known facts.

The situation is not essentially changed when the dispersion matrix of $\vec{\underline{x}}|H_0$, Σ_0 (say), differs from the dispersion matrix of $\vec{\underline{x}}|H_1$, Σ_1 (say). The likelihood-ratio test-statistic would become

$$(8.3.5) \qquad (\vec{\underline{x}}-\vec{\mu}_0)'\Sigma_0^{-1}(\vec{\underline{x}}-\vec{\mu}_0),$$

with a central χ^2-distribution under H_0, but, in general, *not* a non-central χ^2-distribution under H_1. The distribution of (8.3.5) under H_1 may be determined with theorem 3.2.1.

The situation does change when the dispersion matrix of $\vec{\underline{x}}$ is dependent on $\vec{\mu}$, as is shown in the following example.

<u>EXAMPLE 8.3.1</u>. Suppose that $\vec{\underline{x}} \sim N_3(\vec{\mu}, \Sigma(\vec{\mu}))$, where $\vec{\mu} \in \mathbb{R}^3$ and

$$\Sigma(\vec{\mu}) = \begin{pmatrix} \mu_1^2 & 0 & 0 \\ 0 & \mu_2^2 & 0 \\ 0 & 0 & \mu_3^2 \end{pmatrix}, \ \vec{\mu} \neq \vec{0}, \qquad \Sigma(\vec{0}) = \Sigma_0,$$

with Σ_0 positive definite.

We test H_0: $\vec{\mu} = \vec{0}$ against H_1: $\vec{\mu} \neq \vec{0}$. The question is now: does

$$\sup_{H_1} \ (2\pi)^{-3/2}|\Sigma(\vec{\mu})|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{\mu})(\vec{x}-\vec{\mu})\}$$

still attain its maximum value at $\vec{\mu} = \vec{x}$ for a given $\vec{x}$?

Let $\vec{x} = (1,2,3)'$ be an observed value of $\vec{\underline{x}}$. Then we have for $\vec{\mu} = \vec{x}$

$$\Sigma(\vec{\mu}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}; \qquad |\Sigma(\vec{\mu})| = 36;$$

$$\Sigma^{-1}(\vec{\mu}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{9} \end{pmatrix} \qquad \text{and}$$

$$(2\pi)^{-3/2}|\Sigma(\vec{\mu})|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{\mu})(\vec{x}-\vec{\mu})\} =$$

$$= (2\pi)^{-3/2}36^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}\cdot 0\} = (2\pi)^{-3/2}\cdot\tfrac{1}{6}.$$

However, when we take $\vec{\mu} = (\frac{1}{2},2,\frac{5}{2})'$, we have

$$\Sigma(\vec{\mu}) = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & \frac{25}{4} \end{pmatrix} \quad ; \qquad |\Sigma(\vec{\mu})| = \frac{25}{4} \; ;$$

$$\Sigma^{-1}(\vec{\mu}) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{4}{25} \end{pmatrix} \qquad \text{and}$$

$$(2\pi)^{-3/2}|\Sigma(\vec{\mu})|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{\mu})(\vec{x}-\vec{\mu})\} =$$

$$= (2\pi)^{-3/2}(\tfrac{25}{4})^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}\cdot (\tfrac{1}{2},0,\tfrac{1}{2}) \begin{pmatrix} 4 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{4}{25} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} \} =$$

$$= (2\pi)^{-3/2}(\tfrac{2}{5})\exp\{-\tfrac{1}{2}(1+0+\tfrac{1}{25})\} = (2\pi)^{-3/2}(\tfrac{2}{5})\exp\{-\tfrac{13}{25}\}.$$

Now $(\frac{2}{5})\exp\{-\frac{13}{25}\} \approx 0.2378 > \frac{1}{6}$. It follows that the maximum is not attained at $\vec{\mu} = \vec{x}$. Therefore, in this case, the statistic

$$(\vec{\underline{x}}-\vec{\mu}_0)'\Sigma_0^{-1}(\vec{\underline{x}}-\vec{\mu}_0)$$

is *not* the likelihood-ratio statistic.

When Σ is singular, for instance with rank $r < q$, a straightforward generalisation is possible. Let $\vec{\underline{x}} \sim N_q(\vec{\mu},\Sigma)$. The density of $\vec{\underline{x}}$ can then be represented as (RAO (1973), p.528),

$$(2\pi)^{-r/2}(\lambda_1\cdot\ldots\cdot\lambda_r)^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^-(\vec{x}-\vec{\mu})\}$$

where the density is concentrated on the hyperplane

$$N'\vec{x} = N'\vec{\mu}$$

with probability one. Σ^- is any g-inverse of Σ, $\lambda_1,\ldots,\lambda_r$ are the non-zero eigenvalues of Σ and N is a $q \times (q-r)$ matrix of rank $(q-r)$ such that $N'\Sigma = 0$.

Now suppose again that $\vec{\mu}$ is unknown, but that Σ is a known, fixed, non-negative definite matrix. We want to test H_0: $\vec{\mu} = \vec{\mu}_0$ against H_1: $\vec{\mu} \neq \vec{\mu}_0$. Because Σ is fixed, the matrix N is fixed and both $\vec{\underline{x}}$ and $\vec{\mu}$ satisfy the same (q-r) linear constraints, under H_0 and under H_1. This means that the distribution of $\vec{\underline{x}}$ is concentrated on the same hyperplane under H_0 and under H_1.

The likelihood-ratio for this testing problem is then equal to

$$(8.3.6) \qquad \Lambda(\vec{x}) = \frac{\sup_{H_0} (2\pi)^{-r/2}(\lambda_1 \cdot \ldots \cdot \lambda_r)^{-\frac{1}{2}}\exp\{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^-(\vec{x}-\vec{\mu})\}}{\sup_{H_1} (2\pi)^{-r/2}(\lambda_1 \cdot \ldots \cdot \lambda_r)^{-\frac{1}{2}}\exp\{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^-(\vec{x}-\vec{\mu})\}}$$

for each $\vec{x} \in \{\vec{x} | N'\vec{x} = N'\vec{\mu}_0\}$.

It is then again clear that the likelihood-ratio test-statistic is equal to

$$(8.3.7) \qquad (\vec{\underline{x}}-\vec{\mu}_0)'\Sigma^-(\vec{\underline{x}}-\vec{\mu}_0).$$

Again it follows from theorem 3.2.2. that this statistic has, under H_0, a central χ^2-distribution with trace$(\Sigma^-\Sigma)$ = rank Σ = r as number of degrees of freedom and, under H_1, a non-central χ^2-distribution with the same number of degrees of freedom and $(\vec{\mu}-\vec{\mu}_0)'\Sigma^-(\vec{\mu}-\vec{\mu}_0)$ as non-centrality parameter.

Furthermore, it can be seen that it is possible, also in this case, to construct an example like example 8.3.1.

<u>REMARK 8.3.1</u>. It can be proved that (8.3.7) does not depend on the choice of g-inverse Σ^-.

In our testing problem, we have $\vec{\underline{t}}_* \overset{\sim}{\to} N(\vec{0},\Sigma_0)$ under H_0, where Σ_0 is singular and $\vec{\underline{t}}_* \overset{\sim}{\to} N(\vec{\delta},\Sigma_1)$ under alternatives from $A_2 \cup A_3$. Notice that both $\vec{\delta}$ and Σ_1 are dependent on the particular alternative $a \in A_2 \cup A_3$. For alternatives in A_1, we have $\vec{\delta}_*'\vec{\delta}_* \to \infty$, but the distribution of $\vec{\underline{t}}_*$ does not converge in all cases.

Nevertheless, we could suppose that for large enough m we would have

$$(8.3.8) \qquad \vec{\underline{t}}_* \approx N(\vec{0},\Sigma_{0\cdot}) \qquad \text{under } H_0$$

and

$$(8.3.9) \qquad \vec{\underline{t}}_* \approx N(\vec{\delta}_*,\Sigma_{1\cdot}) \qquad \text{under } H_1,$$

where "≈" means "is approximately distributed as" (in (8.3.8) and (8.3.9)). When we make furthermore the crude assumption that $\Sigma_{1.} \approx \Sigma_{0.}$ for all alternatives, then it follows from the preceding theory that

$$(8.3.10) \qquad \vec{\underline{t}}_*' \Sigma_{0.}^- \vec{\underline{t}}_*$$

is an "approximate" likelihood-ratio test-statistic.

The use of (8.3.10) as test-statistic is equivalent to the method proposed, somewhat summarily, by MADANSKY (1963). He does not base the choice of his statistic on likelihood considerations and considering the fact that the assumptions necessary for this justification are hardly satisfied, this is just as well. For $\Sigma_{1.} \approx \Sigma_{0.}$ is not satisfied in our case, nor is $\Sigma_{1.}$ the same for each alternative $a \in A_1$. On the contrary, both $\vec{\delta}_*$ and $\Sigma_{1.}$ depend in a rather complex way on a and in view of example 8.3.1. we conclude that even all the above conditions would not guarantee that (8.3.10) would, asymptotically, be equivalent to the likelihood-ratio test.

Furthermore, (8.3.10) is of the type (2.1.41) in section 2.1., with cross-terms, thus complicating the interpretation of the results. It seems perfectly justified, therefore, to consider a more general class of quadratic forms, presented in this thesis as

$$(8.3.11) \qquad \vec{\underline{t}}_*' Q \vec{\underline{t}}_*$$

with Q n.n.d, and to pay special attention to the diagonal case. This is what we have made our purpose to do.

CHAPTER 9

NUMERICAL RESULTS

This research would not be complete without illustrative examples of a numerical kind. Because of the huge number of parameters in our problem, it is hardly possible to cover all the situations that can occur, and therefore the results of the numerical computations that we give must merely be seen as illustrations of the theory.

There are two kinds of numerical computations that have been made. The first kind concerns the elaboration of most of the formula's that occur in the theory, for a typical practical case, like the computation of exact moments, eigenvalues etc. The second kind concerns the numerical simulation of the exact probability distributions of the test-statistics involved.

All calculations were performed on the CDC - CYBER 73 computer of SARA ("Stichting Academisch Rekencentrum Amsterdam"). Several procedures were used from the library STATAL of statistical procedures, developed by the "Mathematisch Centrum", Amsterdam, and from the library NUMAL of numerical procedures developed by the University of Amsterdam.

We start with the definition of a typical practical case, in the conditional situation.

9.1. A TYPICAL CASE

Suppose we have the following tableau of observations, i.e. with $m = 10$, $n = 5$ and $k = 3$. The table has the same structure as table 1.2.1. For shortness, the categories chosen are indicated by their numbers instead of by their names.

Table 9.1.1. Example of an observation for $m = 10$, $n = 5$ and $k = 3$.

i \ ν	1	2	3	4	5		j=1	j=2	j=3	
1	1	1	2	3	3		2	1	2	5
2	1	1	2	3	2		2	2	1	5
3	3	3	2	2	3		0	2	3	5
4	1	3	3	3	1		2	0	3	5
5	2	2	3	2	2		0	4	1	5
6	1	2	3	2	2		1	3	1	5
7	1	3	3	2	3		1	1	3	5
8	1	2	3	1	1		3	1	1	5
9	1	3	3	2	1		2	1	2	5
10	2	1	3	2	3		1	2	2	5
							14	17	19	50
j=1	7	3	0	1	3	14				
j=2	2	3	3	6	3	17				
j=3	1	4	7	3	4	19				
	10	10	10	10	10	50				

We shall test our null-hypothesis on the basis of these observations. The very first thing to do is to select a Q matrix for the test-statistic $\underline{v}(Q)$.

We shall consider four different statistics, with Q of the form

$$Q = I_n \otimes G, \tag{9.1.1}$$

and G diagonal, i.e.

$$G = \begin{pmatrix} g_1 & 0 & 0 \\ 0 & g_2 & 0 \\ 0 & 0 & g_3 \end{pmatrix} . \tag{9.1.2}$$

The test-statistic that we recommend has weighing factors as given in (6.4.1). We obtain in this case

(9.1.3) $g_1 = \frac{50}{14} = 3.5714$; $g_2 = \frac{50}{17} = 2.9412$; $g_3 = \frac{50}{19} = 2.6316$.

In order to satisfy (4.4.24), we shall modify these weighing factors a little by multiplying each with the same constant factor (0.9610), giving

(9.1.4) $g_1 = 3.4320$; $g_2 = 2.8263$; $g_3 = 2.5288$.

This has of course no influence on the performance of the test. For ease of reference, we shall call the matrix Q defined by (9.1.1), (9.1.2) and (9.1.3), Q_1, and the associated test-statistic $\underline{v}(Q_1)$ or simply $\underline{v}_1$.

A second possible choice is to take the weighing factors equal:

(9.1.5) $g_1 = g_2 = g_3 = 2.8986$,

giving a matrix Q_2 and a statistic $\underline{v}_2$. The value 2.8986 is again the result of a modification (we could otherwise have taken $g_1 = g_2 = g_3 = 1$).

A third statistic $\underline{v}_3$ may be obtained when we have the impression (before the actual observations were made) that there is a preference for C_1 in the first position. We can then give more weight to the first character by choosing (for instance)

(9.1.6) $g_1 = 4$; $g_2 = 1$; $g_3 = 0.5$,

or, after modification,

(9.1.7) $g_1 = 6.7227$; $g_2 = 1.6807$; $g_3 = 0.8403$.

The fourth and last statistic that we consider, $\underline{v}_4$, has weighing factors given by (6.3.11),

(9.1.8) $g_1 = 3.4638$; $g_2 = 3.2072$; $g_3 = 2.1809$.

This gives also a g-inverse type, or Madansky-type statistic. The Q-matrices of $\underline{v}_3$ and $\underline{v}_4$ are called Q_3 and Q_4 resp.

Recapitulating, we shall consider the following four test-statistics.

Table 9.1.2. Weighing factors of four possible test-statistics.

	weighing factors			
statistic	g_1	g_2	g_3	type
$\underset{\sim}{v}_1 \equiv \vec{\underset{\sim}{t}}_*' Q_1 \vec{\underset{\sim}{t}}_*$	3.4320	2.8263	2.5288	"χ^2"
$\underset{\sim}{v}_2 \equiv \vec{\underset{\sim}{t}}_*' Q_2 \vec{\underset{\sim}{t}}_*$	2.8986	2.8986	2.8986	"equal weights"
$\underset{\sim}{v}_3 \equiv \vec{\underset{\sim}{t}}_*' Q_3 \vec{\underset{\sim}{t}}_*$	6.7227	1.6807	0.8403	"directed"
$\underset{\sim}{v}_4 \equiv \vec{\underset{\sim}{t}}_*' Q_4 \vec{\underset{\sim}{t}}_*$	3.4638	3.2072	2.1809	"asymptotic χ^2"

To investigate the performance of the test, we have constructed 2 alternatives, which we shall call $a_{(1)}$ and $a_{(2)}$.

Alternative $a_{(1)}$. Because we are in the conditional situation, an alternative is defined by the assignment of (unequal) probabilities to each of the possible words in each of the experiments E_i. We have

Table 9.1.3. Number of possible words per experiment.

i	$\vec{a}_i$			number of possible words N_i
1	2	1	2	30
2	2	2	1	30
3	0	2	3	10
4	2	0	3	10
5	0	4	1	5
6	1	3	1	20
7	1	1	3	20
8	3	1	1	20
9	2	1	2	30
10	1	2	2	30

(Notice that there are $30 \times 30 \times 10 \times \ldots \times 30 \times 30 = 3.24 \times 10^{12}$ possible ways of obtaining a table of observations like the one in table 9.1.1.).

We have constructed an alternative in which a preference for C_1 in the first position is reflected in the fact that

$$(9.1.9) \qquad P(\underline{t}_{i1}^{(1)} = 1 | \alpha_{(1)}) = 0.8$$

for those experiments for which $a_{i1} = 2$, and

$$(9.1.10) \qquad P(\underline{t}_{i1}^{(1)} = 1 | \alpha_{(1)}) = 0.4$$

in the cases that $a_{i1} = 1$. For the rest the probabilities are spread evenly over the words. For instance, in the first experiment, the 12 words commencing with C_1 have probability $0.8/12 = 0.0\not{6}$, while the other 18 of the 30 possible words have probability $0.2/18 = 0.0\not{1}$. In the sixth experiment, the 4 words beginning with C_1 have probability $0.4/4 = 0.1$ and the other 16 $0.6/16 = 0.0375$. In the third experiment the words have the same probability as under H_0, namely 0.1. The probabilities in the other experiments were determined likewise.

<u>Alternative $\alpha_{(2)}$</u>. This alternative is more intricate, because it has been constructed to represent three relative preferences, a preference of C_1 for the first position, a preference of C_2 for the second and a preference of C_3 for the third position.

Probabilities have been assigned in the following way. Probability 0.4 has been divided evenly over all the words of the type

$$(9.1.11) \qquad C_1 \quad C_2 \quad C_3 \quad \times \quad \times$$

where × stands for an arbitrary character, i.e. words which are completely in accordance with the presumed preferences. Probability 0.3 has been distributed over all the words of one of the following types

$$(9.1.12) \qquad \begin{array}{ccccc} C_1 & C_2 & \overline{C}_3 & \times & \times \\ C_1 & \overline{C}_2 & C_3 & \times & \times \\ \overline{C}_1 & C_2 & C_3 & \times & \times \end{array}$$

where $\overline{C}_j$ means: not the character C_j. Words of the type

$$(9.1.13)\quad \begin{matrix} c_1 & \bar{c}_2 & \bar{c}_3 & \times & \times \\ \bar{c}_1 & c_2 & \bar{c}_3 & \times & \times \\ \bar{c}_1 & \bar{c}_2 & c_3 & \times & \times \end{matrix}$$

together have probability 0.2, while words of the type

$$(9.1.14)\quad \bar{c}_1 \quad \bar{c}_2 \quad \bar{c}_3 \quad \times \quad \times$$

get together probability 0.1. We shall call the types of words given by (9.1.11) - (9.1.14), type A, B, C and D respectively. When words of a certain type do not occur, types have been taken together. The assignment of probabilities is illustrated in the following tables. For the first experiment we have

Table 9.1.4. Assignment of probabilities in the first experiment, under the alternative $a_{(2)}$.

word	type	probability
c_1 c_1 c_2 c_3 c_3	C	0.2/12 = 0.0166..
c_1 c_1 c_3 c_2 c_3	B	0.3/6 = 0.05
c_1 c_1 c_3 c_3 c_2	B	0.3/6 = 0.05
c_1 c_2 c_1 c_3 c_3	B	0.3/6 = 0.05
c_1 c_2 c_3 c_1 c_3	A	0.4/2 = 0.2
c_1 c_2 c_3 c_3 c_1	A	0.4/2 = 0.2
c_1 c_3 c_1 c_2 c_3	C	0.2/12 = 0.0166..
c_1 c_3 c_1 c_3 c_2	C	0.2/12 = 0.0166..
c_1 c_3 c_2 c_1 c_3	C	0.2/12 = 0.0166..
c_1 c_3 c_2 c_3 c_1	C	0.2/12 = 0.0166..
c_1 c_3 c_3 c_1 c_2	B	0.3/6 = 0.05
c_1 c_3 c_3 c_2 c_1	B	0.3/6 = 0.05
c_2 c_1 c_1 c_3 c_3	D	0.1/10 = 0.01
c_2 c_1 c_3 c_1 c_3	C	0.2/12 = 0.0166..
c_2 c_1 c_3 c_3 c_1	C	0.2/12 = 0.0166..
c_2 c_3 c_1 c_1 c_3	D	0.1/10 = 0.01
c_2 c_3 c_1 c_3 c_1	D	0.1/10 = 0.01
c_2 c_3 c_3 c_1 c_1	C	0.2/12 = 0.0166..

C_3	C_1	C_1	C_2	C_3	D	0.1/10 = 0.01
C_3	C_1	C_1	C_3	C_2	D	0.1/10 = 0.01
C_3	C_1	C_2	C_1	C_3	D	0.1/10 = 0.01
C_3	C_1	C_2	C_3	C_1	D	0.1/10 = 0.01
C_3	C_1	C_3	C_1	C_2	C	0.2/12 = 0.0166..
C_3	C_1	C_3	C_2	C_1	C	0.2/12 = 0.0166..
C_3	C_2	C_1	C_1	C_3	C	0.2/12 = 0.0166..
C_3	C_2	C_1	C_3	C_1	C	0.2/12 = 0.0166..
C_3	C_2	C_3	C_1	C_1	B	0.3/6 = 0.05
C_3	C_3	C_1	C_1	C_2	D	0.1/10 = 0.01
C_3	C_3	C_1	C_2	C_1	D	0.1/10 = 0.01
C_3	C_3	C_2	C_1	C_1	D	0.1/10 = 0.01

In the third experiment, the character C_1 does not occur, so words of the type A do not occur. In such cases we have given probability 0.7 to the words of type B, as is illustrated in the following table

Table 9.1.5. Assignment of probabilities in the third experiment, under the alternative $a_{(2)}$.

word					type	probability
C_2	C_2	C_3	C_3	C_3	B	0.7/3 = 0.233..
C_2	C_3	C_2	C_3	C_3	D	0.1/3 = 0.033..
C_2	C_3	C_3	C_2	C_3	C	0.2/4 = 0.05
C_2	C_3	C_3	C_3	C_2	C	0.2/4 = 0.05
C_3	C_2	C_2	C_3	C_3	C	0.2/4 = 0.05
C_3	C_2	C_3	C_2	C_3	B	0.7/3 = 0.233..
C_3	C_2	C_3	C_3	C_2	B	0.7/3 = 0.233..
C_3	C_3	C_2	C_2	C_3	D	0.1/3 = 0.033..
C_3	C_3	C_2	C_3	C_2	D	0.1/3 = 0.033..
C_3	C_3	C_3	C_2	C_2	C	0.2/4 = 0.05

The probabilities in the other experiments were determined in the same way.

9.2. ASYMPTOTIC DISTRIBUTIONS UNDER H_0 AND CRITICAL VALUES

The a.d. of $\underline{v}_i$, under H_0, is given by

$$\sum_{\tau=1}^{r} \lambda_\tau \underline{u}_\tau^2, \tag{9.2.1}$$

where $\lambda_1,\ldots,\lambda_r$ are the non-zero eigenvalues of $Q_i\Sigma_0$. For the actual calculations we work with $\Sigma_{0\bullet}$. Because $Q = I_n \otimes G$, $\Sigma_{0\bullet} = N \otimes K_\bullet$, the non-zero eigenvalues of $Q\Sigma_{0\bullet}$ are equal to the non-zero eigenvalues of $\frac{n}{n-1}\, GK_\bullet$, each of which must be taken with multiplicity (n-1). We have in our example

$$K_\bullet = \begin{pmatrix} 0.168 & -0.068 & -0.100 \\ -0.068 & 0.176 & -0.108 \\ -0.100 & -0.108 & 0.208 \end{pmatrix}. \tag{9.2.2}$$

The eigenvalues calculated for the four statistics are given in the following table.

Table 9.2.1. Eigenvalues for the a.d. of $\underline{v}_i$, $i = 1,\ldots,4$.

	eigenvalues	
statistic	$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$	$\lambda_5 = \lambda_6 = \lambda_7 = \lambda_8$
$\underline{v}_1$	1.0716	0.9284
$\underline{v}_2$	1.1328	0.8672
$\underline{v}_3$	1.5333	0.4667
$\underline{v}_4$	1.0000	1.0000

Let $\underline{Q}_0^{(1)},\ldots,\underline{Q}_0^{(4)}$ be random variables of the type (3.3.2), with eigenvalues as in table 9.2.1., i.e. their distributions are the a.d.'s of $\underline{v}_1,\ldots,\underline{v}_4$.

The distribution-functions of the a.d. of $\underline{v}_i$ can be calculated using (3.3.7). Some results are given in the following table.

Table 9.2.2. Distribution functions of $\underset{\sim}{Q}_0^{(1)},\dots,\underset{\sim}{Q}_0^{(4)}$, the a.d.'s of $\underset{\sim}{v}_1,\dots,\underset{\sim}{v}_4$, under H_0.

z	$P(\underset{\sim}{Q}_0^{(1)} \le z)$	$P(\underset{\sim}{Q}_0^{(2)} \le z)$	$P(\underset{\sim}{Q}_0^{(3)} \le z)$	$P(\underset{\sim}{Q}_0^{(4)} \le z)$
0	0.0000	0.0000	0.0000	0.0000
2	0.0191	0.0194	0.0283	0.0190
4	0.1434	0.1448	0.1770	0.1429
6	0.3535	0.3551	0.3888	0.3528
8	0.5669	0.5679	0.5836	0.5665
10	0.7350	0.7350	0.7316	0.7350
12	0.8485	0.8478	0.8333	0.8488
14	0.9179	0.9170	0.8991	0.9182
16	0.9573	0.9564	0.9402	0.9576
18	0.9785	0.9778	0.9650	0.9788
20	0.9895	0.9890	0.9798	0.9897
22	0.9950	0.9947	0.9884	0.9951
24	0.9976	0.9974	0.9934	0.9977

Notice that the last column of table 9.2.2. gives the distribution-function of the χ^2-distribution with 8 degrees of freedom.

Using an iterative zero-searching procedure, critical values of $\underset{\sim}{Q}_0^{(1)},\dots,\underset{\sim}{Q}_0^{(4)}$ were obtained (we shall call this method of obtaining critical values: "method A"). The results are given in the following table.

Table 9.2.3. Critical values of the distributions of $\underset{\sim}{Q}_0^{(1)},\dots,\underset{\sim}{Q}_0^{(4)}$. (Method A).

α	$Q^{(1)}_{0,1-\alpha}$	$Q^{(2)}_{0,1-\alpha}$	$Q^{(3)}_{0,1-\alpha}$	$Q^{(4)}_{0,1-\alpha}$
0.1000	13.3730	13.4016	14.0336	13.3616
0.0500	15.5293	15.5824	16.6753	15.5073
0.0250	17.5685	17.6505	19.2285	17.5345
0.0100	20.1432	20.2682	22.5142	20.0902
0.0050	22.0240	22.1850	24.9520	21.9550
0.0025	23.8545	24.0545	27.3545	23.7745
0.0010	26.2245	26.4745	30.4745	26.1245

Notice again that the last column contains the critical values of the $\chi^2[8]$ - distribution.

Most program-libraries of numerical methods contain procedures to calculate eigenvalues. They will, however, probably not contain a procedure to calculate the distribution of $\underset{\sim}{Q}_0^{(i)}$. The possible user of our methods thus has to write a program for these distributions and critical values himself.

To avoid this, he can use the approximation to the distribution of $\underset{\sim}{Q}_0$, which we described in section 3.3. and use a table of the χ^2-distribution. We have done this ("method B") for the approximation using two adapted moments. The correction factor (b) and the degrees of freedom (ν) are given by (3.3.15) and (3.3.16) respectively. We found

Table 9.2.4. Approximate critical values for the distributions of $\underset{\sim}{Q}_0^{(1)}, \ldots, \underset{\sim}{Q}_0^{(3)}$, obtained from an approximation with two adapted moments. (Method B).

		$\underset{\sim}{Q}_0^{(1)}$	$\underset{\sim}{Q}_0^{(2)}$	$\underset{\sim}{Q}_0^{(3)}$
	b	1.0051	1.0176	1.2844
	ν	7.9592	7.8613	6.2286
		critical values		
α		$k_{B,1-\alpha}^{(1)}$	$k_{B,1-\alpha}^{(2)}$	$k_{B,1-\alpha}^{(3)}$
0.1000		13.3754	13.4090	14.0786
0.0500		15.5283	15.5795	16.6105
0.0250		17.5627	17.6313	19.0232
0.0100		20.1278	20.2193	22.0879
0.0050		21.9996	22.1084	24.3374
0.0025		23.8262	23.9522	26.5416
0.0010		26.1855	26.3342	29.4001

The determination of critical values may also be based on the exact moments of $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_4$, which can be calculated from (7.2.2) and (7.3.14). ("Method C"). We have

Table 9.2.5. Exact moments of $\underline{v}_1,\ldots,\underline{v}_4$, under H_0.

i	$E\underline{v}_i \mid H_0$	$\sigma^2(\underline{v}_i \mid H_0)$	c	η
1	8.0000	14.0337	1.1401	9.1209
2	8.0000	14.2424	1.1234	8.9873
3	8.0000	17.8405	0.8968	7.1747
4	8.0000	13.9529	1.1467	9.1737

The last two column's contain c as defined by (1.4.4) and η, defined by (1.4.5). Using the method described in section 1.4. we find the following approximate critical values of the distribution of $\underline{v}_i$.

Table 9.2.6. Approximate critical values for the distributions of $\underline{v}_1,\ldots,\underline{v}_4$, using the exact moments. (Method C).

α	$k^{(1)}_{C,1-\alpha}$	$k^{(2)}_{C,1-\alpha}$	$k^{(3)}_{C,1-\alpha}$	$k^{(4)}_{C,1-\alpha}$
0.1000	13.0182	13.0558	13.6633	13.0035
0.0500	14.9880	15.0445	15.9684	14.9659
0.0250	16.8411	16.9164	18.1542	16.8118
0.0100	19.1685	19.2683	20.9187	19.1296
0.0050	20.8615	20.9797	22.9409	20.8155
0.0025	22.5099	22.6465	24.9177	22.4568
0.0010	24.6345	24.7952	27.4753	24.5721

All the critical values, calculated from the a.d. of $\underline{v}_i$, from an approximation to the a.d. or from the exact moments of the $\underline{v}_i$ may be used as *approximate* critical values for the performance of the test.

The *exact* critical values would have to be based on the exact distribution of $\underline{v}_i$ which is unavailable to us. See also section 9.6.

The outcomes of $\underline{v}_i$, $i=1,\ldots,4$, for the data of table 9.1.1. are as given in the following table.

Table 9.2.7. Outcomes of the four test-statistics for the data of table 9.1.1.

statistic	outcome
$\underline{v}_1$	17.24
$\underline{v}_2$	16.46
$\underline{v}_3$	22.49
$\underline{v}_4$	17.03

The outcomes are significant at the 5% level for all four tests. Actually, the data of table 9.1.1. were obtained from a simulation of the experiment under the alternative $a_{(1)}$. This explains the fact that the outcome of $\underline{v}_3$ is the highest of the four, because this statistic was designed especially to work against $a_{(1)}$.

9.3. SIMULATION RESULTS (UNDER H_0)

For each of the four statistics considered, we obtained 1000 pseudo-observations, under H_0, by generating for each experiment E_i a pseudo-random word. The words of each experiment were combined and an outcome of $\underline{v}_i$ was calculated. In this way we were able to make (pseudo-) estimates of the right-tail probabilities of the critical values of the preceding sections. These results thus also give an impression of the actual level of significance of the tests as compared to the nominal level α.

Table 9.3.1. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{A,1-\alpha}$ of method A, under H_0, obtained by simulation. (See remark 9.3.1.).

α	$k^{(1)}_{A,1-\alpha}$	e.r.t.p	$k^{(2)}_{A,1-\alpha}$	e.r.t.p	$k^{(3)}_{A,1-\alpha}$	e.r.t.p	$k^{(4)}_{A,1-\alpha}$	e.r.t.p
0.1000	13.3730	0.073	13.4016	0.107	14.0336	0.100	13.3616	0.104
0.0500	15.5293	0.038	15.5824	0.058	16.6753	0.038	15.5073	0.035
0.0250	17.5685	0.015	17.6505	0.017	19.2285	0.020	17.5345	0.015
0.0100	20.1432	0.004	20.2682	0.009	22.5142	0.005	20.0902	0.007
0.0050	22.0240	0.003	22.1850	0.002	24.9520	0.003	21.9550	0.002
0.0025	23.8545	0.000	24.0545	0.002	27.3545	0.001	23.7745	0.000
0.0010	26.2245	0.000	26.4745	0.001	30.4745	0.000	26.1245	0.000

Table 9.3.2. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k_{B,1-\alpha}^{(i)}$ of method B, under H_0, obtained by simulation.

α	$k_{B,1-\alpha}^{(1)}$	e.r.t.p	$k_{B,1-\alpha}^{(2)}$	e.r.t.p	$k_{B,1-\alpha}^{(3)}$	e.r.t.p
0.1000	13.3754	0.103	13.4090	0.084	14.0786	0.086
0.0500	15.5283	0.055	15.5795	0.043	16.6105	0.031
0.0250	17.5627	0.019	17.6313	0.017	19.0232	0.014
0.0100	20.1278	0.011	20.2193	0.006	22.0879	0.004
0.0050	21.9996	0.006	22.1084	0.003	24.3374	0.001
0.0025	23.8262	0.001	23.9522	0.001	26.5416	0.001
0.0010	26.1855	0.001	26.3342	0.001	29.4001	0.000

Table 9.3.3. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k_{C,1-\alpha}^{(i)}$ of method C, under H_0, obtained by simulation.

α	$k_{C,1-\alpha}^{(1)}$	e.r.t.p	$k_{C,1-\alpha}^{(2)}$	e.r.t.p	$k_{C,1-\alpha}^{(3)}$	e.r.t.p	$k_{C,1-\alpha}^{(4)}$	e.r.t.p
0.1000	13.0182	0.115	13.0558	0.103	13.6633	0.101	13.0035	0.095
0.0500	14.9880	0.060	15.0445	0.061	15.9684	0.046	14.9659	0.044
0.0250	16.8411	0.034	16.9164	0.030	18.1542	0.024	16.8118	0.017
0.0100	19.1685	0.012	19.2683	0.011	20.9187	0.009	19.1296	0.005
0.0050	20.8615	0.005	20.9797	0.006	22.9409	0.003	20.8155	0.003
0.0025	22.5099	0.003	22.6465	0.002	24.9177	0.000	22.4568	0.002
0.0010	24.6345	0.002	24.7952	0.000	27.4753	0.000	24.5721	0.000

REMARK 9.3.1. Due to high costs of computer time, the simulations have not been made for each α separately. Therefore, the estimates of the right-tail probabilities in tables 9.3.1., 9.3.2. and 9.3.3. are dependent columnwise.

Inspection of the tables 9.3.1., 9.3.2. and 9.3.3. shows that in almost all cases the approximate critical values are slightly too high. This means that the actual level of the test is lower than the nominal level α. We shall call such tests "timid" (such tests are usually called "conservative"), in contrast with the tests where the actual level is higher than the nominal level α, which we shall call "bold". When critical values are used which are obtained by method C for $\underline{v}_1$, the test that we recommend, "bold" tests are obtained. Of course we have to take the inaccuracy into account resulting from the fact that we have only estimates of the right-tail probabilities at our disposal. The general tendency is however clear enough.

Furthermore it seems that the estimates in table 9.3.3. (method C) are generally closer to the nominal values of α than in the other two cases. Therefore we recommend method C for the approximation of the critical values in all cases.

Another impression of the goodness of the approximation using the exact moments of $\underline{v}$ (method C) may be obtained from the following figures.

Figure 9.3.1. Pseudo-emperical distribution function of 1000 simulated observations of $\underline{v}_1$ (dashed line) and distribution function of an adapted χ^2-distribution. (The same simulation results as for table 9.3.3.).

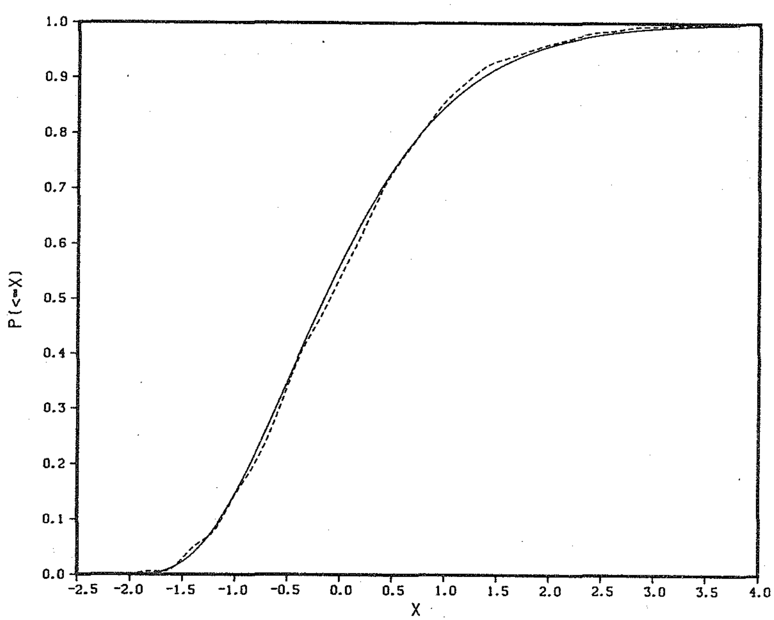


Figure 9.3.2. Pseudo-emperical distribution function of 1000 simulated observations of $\underline{v}_2$ (dashed line) and distribution function of an adapted χ^2-distribution. (The same simulation results as for table 9.3.3.).

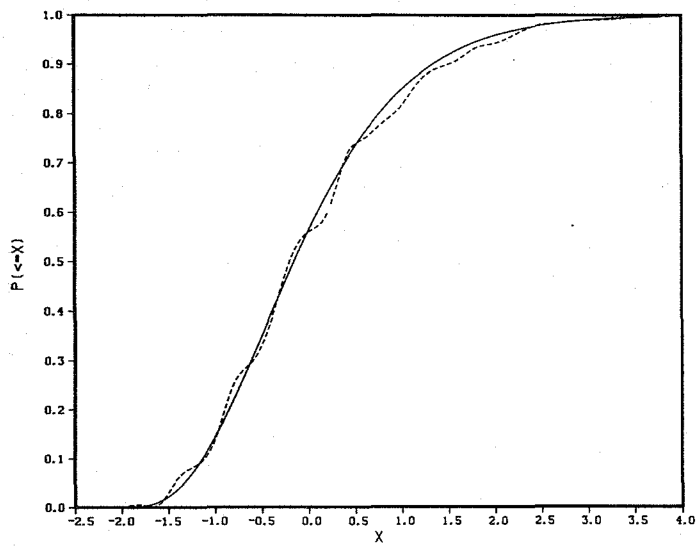


Figure 9.3.3. Pseudo-emperical distribution function of 1000 simulated observations of $\underline{v}_3$ (dashed line) and distribution function of an adapted χ^2-distribution. (The same simulation results as for table 9.3.3.).

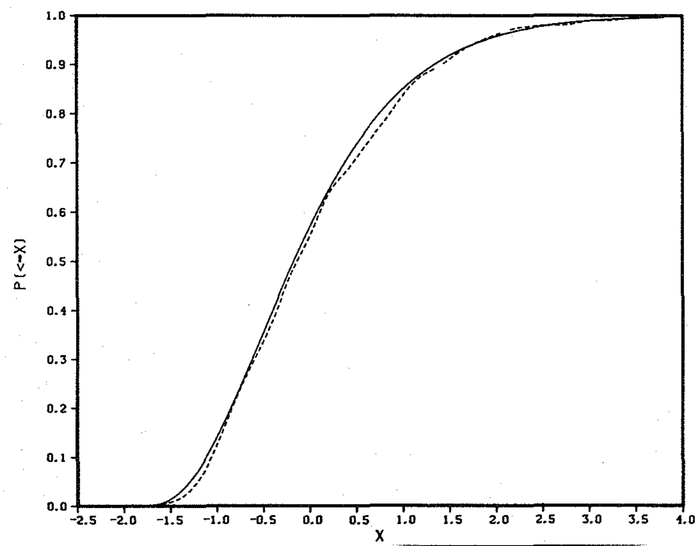
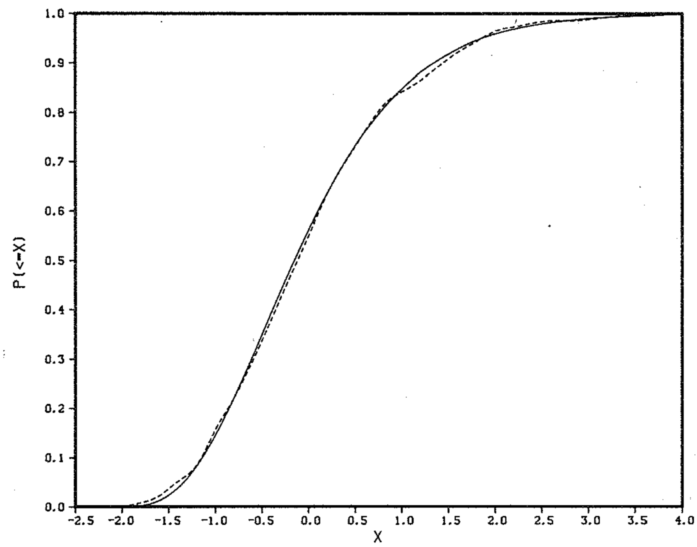


Figure 9.3.4. Pseudo-emperical distribution function of 1000 simulated observations of $\underline{v}_4$ (dashed line) and distribution function of an adapted χ^2-distribution. (The same simulation results as for table 9.3.3.).



9.4. SIMULATION RESULTS (UNDER ALTERNATIVES) & POWER

To make an estimate of the power of the four tests considered, we also generated 1000 pseudo-observations of $\underline{v}_1,\ldots,\underline{v}_4$, under each of the two alternatives $a_{(1)}$ and $a_{(2)}$ that were defined in section 9.1. Estimates of the right-tail probabilities of the (approximate) critical values of section 9.2. are given in the tables 9.4.3. - 9.4.8.

In each case the estimate is compared with the approximation to the power of the test as calculated from formula (4.4.19),

$$(9.4.1) \qquad P\left(\sum_{\tau=1}^{r} \lambda_\tau (\underline{u}_\tau + \sqrt{m_0}\,\omega_\tau)^2 \geq k\right).$$

So, in addition to the eigenvalues $\lambda_1,\ldots,\lambda_r$ which are the same as those under H_0 (table 9.2.1.), we need the vectors $\vec{\omega}$ defined by (cf. section 4.4.)

$$(9.4.2) \qquad \vec{\omega} = \Lambda_+^{-1} P_+' B' Q \vec{\zeta}(a),$$

for each of the four choices of Q and each of the two choices of a. For those readers who might wish to check the calculations, we give the eight vectors $\vec{\omega}$ in the following tables.

Table 9.4.1. Components of the vectors $\vec{\omega}$ for $Q_1,\ldots,Q_4$ and $a_{(1)}$.

τ	$\omega_\tau^{(1)}(a_{(1)})$	$\omega_\tau^{(2)}(a_{(1)})$	$\omega_\tau^{(3)}(a_{(1)})$	$\omega_\tau^{(4)}(a_{(1)})$
1	0.1868	0.1377	-0.5025	-0.4055
2	0.1932	0.2357	-0.2277	0.2431
3	0.0126	-0.0805	-0.2990	-0.2049
4	-0.3861	-0.0883	-0.0896	0.1983
5	-0.2305	0.2825	-0.0036	0.1464
6	-0.1950	0.1950	0.0017	0.0916
7	-0.2643	-0.3372	-0.0137	-0.1758
8	-0.1412	-0.2865	0.0218	0.3625

Table 9.4.2. Components of the vectors $\vec{\omega}$ for $Q_1,\ldots,Q_4$ and $a_{(2)}$.

τ	$\omega_\tau^{(1)}(a_{(2)})$	$\omega_\tau^{(2)}(a_{(2)})$	$\omega_\tau^{(3)}(a_{(2)})$	$\omega_\tau^{(4)}(a_{(2)})$
1	0.1905	-0.2085	-0.5484	-0.6021
2	0.5163	0.3100	-0.0277	0.3325
3	-0.4854	-0.6994	-0.3583	0.3345
4	-0.2673	0.0661	-0.2832	0.5011
5	0.2177	0.3071	-0.1960	-0.3165
6	-0.5087	0.3168	-0.3130	0.0530
7	-0.4392	-0.5438	0.5709	0.4182
8	-0.1266	0.0396	0.4031	-0.1050

We now have

Table 9.4.3. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{A,1-\alpha}$ of method A, under $a_{(1)}$, obtained by simulation. The values in brackets give the approximate power (a.p.) calculated from (4.4.19).

α	$k^{(1)}_{A,1-\alpha}$	e.r.t.p (a.p.)	$k^{(2)}_{A,1-\alpha}$	e.r.t.p (a.p.)	$k^{(3)}_{A,1-\alpha}$	e.r.t.p (a.p.)	$k^{(4)}_{A,1-\alpha}$	e.r.t.p (a.p.)
0.1000	13.3730	0.304 (0.355)	13.4016	0.281 (0.331)	14.0336	0.435 (0.436)	13.3616	0.298 (0.354)
0.0500	15.5293	0.184 (0.240)	15.5824	0.153 (0.217)	16.6753	0.290 (0.314)	15.5073	0.167 (0.239)
0.0250	17.5685	0.113 (0.159)	17.6505	0.071 (0.139)	19.2285	0.185 (0.221)	17.5345	0.088 (0.158)
0.0100	20.1432	0.048 (0.091)	20.2682	0.029 (0.076)	22.5142	0.077 (0.135)	20.0902	0.043 (0.090)
0.0050	22.0240	0.030 (0.058)	22.1850	0.017 (0.047)	24.9520	0.038 (0.092)	21.9550	0.024 (0.058)
0.0025	23.8545	0.015 (0.037)	24.0545	0.007 (0.029)	27.3545	0.023 (0.062)	23.7745	0.013 (0.037)
0.0010	26.2245	0.006 (0.020)	26.4745	0.000 (0.015)	30.4745	0.010 (0.036)	26.1245	0.005 (0.020)

Table 9.4.4. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{B,1-\alpha}$ of method B, under $a_{(1)}$, obtained by simulation. The values in brackets give the approximate power (a.p.) calculated from (4.4.19).

α	$k^{(1)}_{B,1-\alpha}$	e.r.t.p (a.p.)	$k^{(2)}_{B,1-\alpha}$	e.r.t.p (a.p.)	$k^{(3)}_{B,1-\alpha}$	e.r.t.p (a.p.)	$k^{(4)}_{B,1-\alpha}$	e.r.t.p (a.p.)
0.1000	13.3754	0.335 (0.355)	13.4090	0.283 (0.330)	14.0786	0.398 (0.434)	13.3616	0.322 (0.354)
0.0500	15.5283	0.192 (0.240)	15.5795	0.172 (0.217)	16.6105	0.276 (0.317)	15.5073	0.191 (0.239)
0.0250	17.5627	0.109 (0.159)	17.6313	0.083 (0.140)	19.0232	0.180 (0.228)	17.5345	0.117 (0.158)
0.0100	20.1278	0.051 (0.091)	20.2193	0.037 (0.077)	22.0879	0.092 (0.145)	20.0902	0.053 (0.090)
0.0050	21.9996	0.022 (0.059)	22.1084	0.026 (0.048)	24.3374	0.059 (0.102)	21.9550	0.026 (0.058)
0.0025	23.8262	0.013 (0.037)	23.9522	0.010 (0.030)	26.5416	0.038 (0.071)	23.7745	0.012 (0.037)
0.0010	26.1855	0.003 (0.020)	26.3342	0.004 (0.016)	29.4001	0.016 (0.044)	26.1245	0.003 (0.020)

Table 9.4.5. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{C,1-\alpha}$ of method C, under $a_{(1)}$, obtained by simulation. The values in brackets give the approximate power (a.p.) calculated from (4.4.19).

α	$k^{(1)}_{C,1-\alpha}$	e.r.t.p (a.p.)	$k^{(2)}_{C,1-\alpha}$	e.r.t.p (a.p.)	$k^{(3)}_{C,1-\alpha}$	e.r.t.p (a.p.)	$k^{(4)}_{C,1-\alpha}$	e.r.t.p (a.p.)
0.1000	13.0182	0.365 (0.377)	13.0558	0.309 (0.352)	13.6633	0.405 (0.455)	13.0035	0.339 (0.376)
0.0500	14.9880	0.216 (0.266)	15.0445	0.209 (0.242)	15.9684	0.272 (0.344)	14.9659	0.220 (0.265)
0.0250	16.8411	0.128 (0.185)	16.9164	0.134 (0.164)	18.1542	0.175 (0.257)	16.8118	0.130 (0.184)
0.0100	19.1685	0.071 (0.113)	19.2683	0.079 (0.096)	20.9187	0.095 (0.173)	19.1296	0.073 (0.112)
0.0050	20.8615	0.039 (0.077)	20.9797	0.044 (0.064)	22.9409	0.052 (0.127)	20.8155	0.041 (0.076)
0.0025	22.5099	0.020 (0.052)	22.6465	0.021 (0.041)	24.9177	0.031 (0.093)	22.4568	0.018 (0.051)
0.0010	24.6345	0.007 (0.031)	24.7952	0.011 (0.024)	27.4753	0.020 (0.060)	24.5721	0.007 (0.030)

Table 9.4.6. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{A,1-\alpha}$ of method A, under $a_{(2)}$, obtained by simulation. The values in brackets give the approximate power (a.p.) calculated from (4.4.19).

α	$k^{(1)}_{A,1-\alpha}$	e.r.t.p (a.p.)	$k^{(2)}_{A,1-\alpha}$	e.r.t.p (a.p.)	$k^{(3)}_{A,1-\alpha}$	e.r.t.p (a.p.)	$k^{(4)}_{A,1-\alpha}$	e.r.t.p (a.p.)
0.1000	13.3730	0.694 (0.761)	13.4016	0.724 (0.761)	14.0336	0.619 (0.669)	13.3616	0.701 (0.760)
0.0500	15.5293	0.590 (0.651)	15.5824	0.607 (0.651)	16.6753	0.482 (0.535)	15.5073	0.575 (0.650)
0.0250	17.5685	0.478 (0.543)	17.6505	0.478 (0.543)	19.2285	0.370 (0.415)	17.5345	0.464 (0.542)
0.0100	20.1432	0.349 (0.412)	20.2682	0.360 (0.412)	22.5142	0.239 (0.285)	20.0902	0.332 (0.412)
0.0050	22.0240	0.281 (0.328)	22.1850	0.293 (0.327)	24.9520	0.166 (0.210)	21.9550	0.261 (0.327)
0.0025	23.8545	0.208 (0.256)	24.0545	0.208 (0.255)	27.3545	0.121 (0.151)	23.7745	0.201 (0.256)
0.0010	26.2245	0.122 (0.181)	26.4745	0.150 (0.179)	30.4745	0.067 (0.096)	26.1245	0.147 (0.180)

Table 9.4.7. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{B,1-\alpha}$ of method B, under $a_{(2)}$, obtained by simulation. The values in brackets give the approximate power (a.p.) calculated from (4.4.19).

α	$k^{(1)}_{B,1-\alpha}$	e.r.t.p (a.p.)	$k^{(2)}_{B,1-\alpha}$	e.r.t.p (a.p.)	$k^{(3)}_{B,1-\alpha}$	e.r.t.p (a.p.)	$k^{(4)}_{B,1-\alpha}$	e.r.t.p (a.p.)
0.1000	13.3754	0.680 (0.761)	13.4090	0.716 (0.761)	14.0786	0.637 (0.666)	13.3616	0.681 (0.760)
0.0500	15.5283	0.588 (0.651)	15.5795	0.588 (0.651)	16.6105	0.491 (0.538)	15.5073	0.571 (0.650)
0.0250	17.5627	0.470 (0.543)	17.6313	0.469 (0.544)	19.0232	0.378 (0.424)	17.5345	0.464 (0.542)
0.0100	20.1278	0.327 (0.413)	20.2193	0.347 (0.414)	22.0879	0.265 (0.300)	20.0902	0.363 (0.412)
0.0050	21.9996	0.262 (0.329)	22.1084	0.275 (0.330)	24.3374	0.171 (0.227)	21.9550	0.271 (0.327)
0.0025	23.8262	0.187 (0.257)	23.9522	0.208 (0.258)	26.5416	0.117 (0.169)	23.7745	0.205 (0.256)
0.0010	26.1855	0.130 (0.182)	26.3342	0.121 (0.183)	29.4001	0.070 (0.113)	26.1245	0.138 (0.180)

Table 9.4.8. Estimates of the right-tail probabilities (e.r.t.p.) of the critical values $k^{(i)}_{C,1-\alpha}$ of method C, under $a_{(2)}$, obtained by simulation. The values in brackets give the approximate power (a.p.) calculated from (4.4.19).

α	$k^{(1)}_{C,1-\alpha}$	e.r.t.p (a.p.)	$k^{(2)}_{C,1-\alpha}$	e.r.t.p (a.p.)	$k^{(3)}_{C,1-\alpha}$	e.r.t.p (a.p.)	$k^{(4)}_{C,1-\alpha}$	e.r.t.p (a.p.)
0.1000	13.0182	0.735 (0.788)	13.0558	0.729 (0.779)	13.6633	0.663 (0.687)	13.0035	0.705 (0.778)
0.0500	14.9880	0.623 (0.680)	15.0445	0.629 (0.679)	15.9684	0.528 (0.570)	14.9659	0.601 (0.679)
0.0250	16.8411	0.518 (0.581)	16.9164	0.527 (0.581)	18.1542	0.417 (0.464)	16.8118	0.500 (0.580)
0.0100	19.1685	0.403 (0.460)	19.2683	0.433 (0.460)	20.9187	0.305 (0.344)	19.1296	0.398 (0.459)
0.0050	20.8615	0.319 (0.379)	20.9797	0.326 (0.379)	22.9409	0.222 (0.271)	20.8155	0.317 (0.378)
0.0025	22.5099	0.260 (0.307)	22.6465	0.251 (0.308)	24.9177	0.169 (0.211)	22.4568	0.252 (0.306)
0.0010	24.6345	0.183 (0.229)	24.7952	0.169 (0.230)	27.4753	0.114 (0.148)	24.5721	0.180 (0.228)

REMARK 9.4.1. The same remark applies as in remark 9.3.1. Notice furthermore that the last column of table 9.4.4. contains the same critical values as the last column of table 9.4.3. The estimates, however, are independent of the estimates of table 9.4.3. They are given to make a better comparison possible.

According to these results it appears that $\underline{v}_3$ has the highest power against $a_{(1)}$, which is in agreement with the fact that Q_3 was especially chosen so that $\underline{v}_3$ would react on this kind of alternative. Furthermore, $\underline{v}_1$ and $\underline{v}_4$ are equally good, though $\underline{v}_1$ seems to perform slightly better than $\underline{v}_4$.

Against $a_{(2)}$ it is $\underline{v}_2$ that appears to work best, which is again not surprising because there are preferences for three positions in $a_{(2)}$. Again $\underline{v}_1$ and $\underline{v}_4$ are competetive while now $\underline{v}_3$ seems to work worst.

It seems that the highest power is obtained when the critical values have been determined by method C. The results may be misleading, however, because we have not made a correction for the fact that the actual levels of the tests are not equal to the nominal levels. Therefore we compare "bold" and "timid" tests which possibly gives a distorted picture of the situation. The results are nevertheless supported by the results of Pitman-efficiencies (section 9.5.).

Finally we observe that the approximate power of the tests as calculated from (4.4.19) is too high throughout. Again for the critical values of method C, the results are closest to the estimates from the simulation. Because no calculations have been made for a higher number of experiments than 10, it is not clear whether the approximation improves as m gets larger. This should therefore be a subject for further (numerical) research.

9.5. PITMAN & BAHADUR EFFICIENCIES

The asymptotic relative Pitman efficiency (ARPE) of $\underline{v}(Q_i)$ with respect to $\underline{v}(Q_j)$ is equal to (cf. (5.2.2))

$$(9.5.1) \qquad e_{ij}(\beta) = \frac{H_j^{-1}(\beta)}{H_i^{-1}(\beta)}, \qquad \beta \in (\alpha,1),$$

where $H_i(\eta)$ is given by (5.2.3). Using an iterative zero-searching procedure the inverse values of $H(\eta)$ were calculated. Because $e_{ij}(\beta)$ depends on α,β

and a, we cannot give a complete survey of the results. Therefore, we shall only give the results for $\alpha = 0.05$ and $\beta = 0.25$, 0.50 and 0.75 and for $a_{(1)}$ and $a_{(2)}$. For other values of α and β the ARPE's show generally the same pattern.

Table 9.5.1. ARPE's for the 4 tests considered, with $\alpha = 0.05$ and $\beta = 0.25$ for $a_{(1)}$.

$H^{-1}(0.25)$		10.43	11.55	7.85	10.47
	test no.	1	2	3	4
10.43	1	1.000	1.107	0.753	1.004
11.55	2	0.903	1.000	0.680	0.906
7.85	3	1.329	1.471	1.000	1.334
10.47	4	0.996	1.103	0.750	1.000

Table 9.5.2. ARPE's for the 4 tests considered, with $\alpha = 0.05$ and $\beta = 0.50$ for $a_{(1)}$.

$H^{-1}(0.50)$		20.82	22.68	16.35	20.89
	test no.	1	2	3	4
20.82	1	1.000	1.089	0.785	1.003
22.68	2	0.918	1.000	0.721	0.921
16.35	3	1.273	1.387	1.000	1.278
20.89	4	0.997	1.086	0.783	1.000

Table 9.5.3. ARPE's for the four tests considered, with $\alpha = 0.05$ and $\beta = 0.75$ for $a_{(1)}$.

$H^{-1}(0.75)$		33.69	36.25	27.28	36.33
	test no.	1	2	3	4
33.69	1	1.000	1.076	0.810	1.078
36.25	2	0.929	1.000	0.753	1.002
27.28	3	1.235	1.329	1.000	1.332
36.33	4	0.927	0.998	0.751	1.000

Table 9.5.4. ARPE's for the four tests considered, with $\alpha = 0.05$ and $\beta = 0.25$ for $a_{(2)}$.

$H^{-1}(0.25)$		3.74	3.73	4.70	3.75
	test no.	1	2	3	4
3.74	1	1.000	0.997	1.257	1.003
3.73	2	1.003	1.000	1.260	1.005
4.70	3	0.796	0.794	1.000	0.798
3.75	4	0.997	0.995	1.253	1.000

Table 9.5.5. ARPE's for the four tests considered, with $\alpha = 0.05$ and $\beta = 0.50$ for $a_{(2)}$.

$H^{-1}(0.50)$		7.46	7.45	9.32	7.47
	test no.	1	2	3	4
7.46	1	1.000	0.999	1.249	1.001
7.45	2	1.001	1.000	1.251	1.003
9.32	3	0.800	0.799	1.000	0.802
7.47	4	0.999	0.997	1.248	1.000

Table 9.5.6. ARPE's for the four tests considered, with $\alpha = 0.05$ and $\beta = 0.75$ for $a_{(2)}$.

$H^{-1}(0.75)$		12.06	12.06	14.94	12.08
	test no.	1	2	3	4
12.06	1	1.000	1.000	1.239	1.002
12.06	2	1.000	1.000	1.239	1.002
14.94	3	0.807	0.807	1.000	0.809
12.08	4	0.998	0.998	1.237	1.000

These results confirm clearly the simulation results of the preceding section. I.e. $\underline{v}_3$ (the "directed" test) performs best against $a_{(1)}$. Second best is $\underline{v}_1$ (the "χ^2-type" test), though only slightly better than $\underline{v}_4$ (the "asymptotic χ^2-type" test). The "equal weights" test, $\underline{v}_2$ performs definitely worse against $a_{(1)}$.

The situation under $a_{(2)}$ is different, fully in accordance with our expectations. As in the simulation results, the "equal weights" test performs best against $a_{(2)}$. Again $\underline{v}_1$ is second best and is slightly better than $\underline{v}_4$. This time the directed test, $\underline{v}_3$ performs worst.

The approximation to $e_{ij}(\beta)$, (cf. (2.5.11)),

$$e^*_{ij} = \frac{\vec{\zeta}' Q_i \vec{\zeta}}{\vec{\zeta}' Q_j \vec{\zeta}} \tag{9.5.2}$$

is much easier to use, because no eigenvalues etc. have to be calculated. We have

Table 9.5.7. Components of the vectors $\vec{\zeta}$ for $a_{(1)}$ and $a_{(2)}$.

ν			$\vec{\zeta}^{(1)}$	$\vec{\zeta}^{(2)}$
		1	0.2600	0.2683
1	j	2	-0.1033	-0.1020
		3	-0.1567	-0.1663
		1	-0.0650	-0.1200
2	j	2	0.0258	0.2783
		3	0.0392	-0.1583
		1	-0.0650	-0.1463
3	j	2	0.0258	-0.1870
		3	0.0392	0.3333
		1	-0.0650	-0.0010
4	j	2	0.0258	0.0053
		3	0.0392	-0.0043
		1	-0.0650	-0.0010
5	j	2	0.0258	0.0053
		3	0.0392	-0.0043

Table 9.5.8. Approximate ARPE's for the four tests considered, calculated according to (5.2.11), for $a_{(1)}$.

$\vec{\zeta}'Q_l\vec{\zeta}$		0.405	0.373	0.616	0.402
	test no.	1	2	3	4
0.405	1	1.000	1.086	0.657	1.007
0.373	2	0.921	1.000	0.606	0.928
0.616	3	1.521	1.561	1.000	1.532
0.402	4	0.993	1.078	0.653	1.000

Table 9.5.9. Approximate ARPE's for the four tests considered, calculated according to (5.2.11), for $a_{(2)}$.

$\vec{\zeta}'Q_l\vec{\zeta}$		1.132	1.144	1.069	1.125
	test no.	1	2	3	4
1.132	1	1.000	0.990	1.059	1.006
1.144	2	1.011	1.000	1.070	1.017
1.069	3	0.944	0.934	1.000	0.950
1.125	4	0.994	0.983	1.052	1.000

The reader may judge for himself whether he thinks these approximations good enough for his purposes. In any case, the general tendency is the same as in the 'exact' ARPE cases.

Finally, the asymptotic relative Bahadur efficiency (ARBE) is equal to (cf. (5.4.1))

$$E_{ij}(a) = \frac{\frac{1}{\lambda_1^{(i)}}\vec{\zeta}'Q_i\vec{\zeta}}{\frac{1}{\lambda_1^{(j)}}\vec{\zeta}'Q_j\vec{\zeta}} \tag{9.5.3}$$

Using the data of tables 9.1.2., 9.2.1., 9.5.8. and 9.5.9. we find

Table 9.5.10. ARBE's for the four tests considered, calculated according to (5.4.1), for $a_{(1)}$.

$\frac{1}{\lambda_1^{(i)}}\vec{\zeta}'Q_i\vec{\zeta}$		0.378	0.329	0.402	0.402
	test no.	1	2	3	4
0.378	1	1.000	1.149	0.940	0.940
0.329	2	0.870	1.000	0.818	0.818
0.402	3	1.063	1.222	1.000	1.000
0.402	4	1.063	1.222	1.000	1.000

Table 9.5.11. ARBE's for the four tests considered, calculated according to (5.4.1), for $a_{(2)}$.

$\frac{1}{\lambda_1^{(i)}}\vec{\zeta}'Q_i\vec{\zeta}$		1.056	1.010	0.697	1.125
	test no.	1	2	3	4
1.056	1	1.000	1.046	1.515	0.939
1.010	2	0.956	1.000	1.449	0.898
0.697	3	0.660	0.690	1.000	0.620
1.125	4	1.065	1.114	1.614	1.000

It is clear that the simulation results of section 9.4. are more in accordance with the ARPE's than with the ARBE's. The ARPE therefore seems to be the better measure of asymptotic relative efficiencies in our case.

9.6. CONCLUDING REMARKS

By computer generation of all possible $\Pi_{i=1}^{m} N_i$ different combinations of words and calculation of $\underline{v}(Q)$ for each combination, it is in principle possible to obtain the exact distribution of $\underline{v}(Q)$. However, the number of possibilities becomes soon prohibitive. For instance in our example we have

$$(9.6.1) \qquad \prod_{i=1}^{m} N_i = 3.24 \times 10^{12}.$$

So only for relatively small m, it can be done in practice. The interested reader is referred to DIK (1979), which shows that, under H_0, the number of possibilities can be reduced a little by symmetry arguments, and which gives some results, under H_0, for the case that $Q = I_n \times G_g$.

He may also find there results of simulations of $\underline{w}(Q)$, i.e. in the unconditional situation. Some remarks are made there also on the effects of the deletion of "useless" observations.

We are aware that the numerical examples that we have given do not, in any way, cover all the possible situations that can occur. But the examples given support clearly the recommendations that we have given in this thesis.

REFERENCES

[1] ANDERSON, T.W. (1958), *An introduction to multivariate statistical analysis*, John Wiley & Sons, New York.

[2] BAHADUR, R.R. (1960), *Stochastic comparison of tests*, Ann. Math. Statist. 31, 276-295.

[3] BILLINGSLEY, P. (1979), *Probability & measure*, John Wiley & Sons, New York.

[4] COCHRAN, W.G. (1950), *The comparison of percentages in matched samples*, Biometrika 37, 256-266.

[5] DIK, J.J. (1979), *A test for the equality of preferences*, Report 79-09, Institute for applied mathematics, University of Amsterdam.

[6] JOHNSON, N.L. & S. KOTZ (1970), *Distributions in statistics: Continuous univariate distributions - 2*, John Wiley & Sons, New York.

[7] JOHNSON, N.L. & S. KOTZ (1972), *Distributions in statistics: Continuous multivariate distributions*, John Wiley & Sons, New York.

[8] KNEEPKENS, H. (1975), *De voornaamste kop op de voorpagina's van een vijftal landelijke nederlandse dagbladen in de eerste twee maanden van 1964 en 1965*, Masters thesis, Institute for applied mathematics, University of Amsterdam.

[9] KOTZ, S., N.L. JOHNSON & D.W. BOYD (1967a), *Series representations of distributions of quadratic forms in normal variables, I, Central case*, Ann. Math. Statist. 38, 823-837

[10] KOTZ, S., N.L. JOHNSON & D.W. BOYD (1967b), *Series representations of distributions of quadratic forms in normal variables, II, Non-central case*, Ann. Math. Statist. 38, 838-848.

[11] MADANSKY, A. (1963), *Tests of homogeneity for correlated samples*, Journal of the American Statistical Association 58, 97-119.

[12] RAO, C.R. (1973), *Linear statistical inference and its applications, Second edition*, John Wiley & Sons, New York.

[13] RAYNER, A.A. & D. LIVINGSTONE (1965), *On the distribution of quadratic forms in singular normal variables*, S. African J. Agric. Sci. 8, 357-370.

[14] ROTHE, G. (1979), *Some properties of the asymptotic relative Pitman efficiency*, Preprint of ROTHE (1981), Department of statistics, University of Dortmund.

[15] ROTHE, G. (1981), *Some properties of the asymptotic relative Pitman efficiency*, To appear in Ann. Math. Statist.

[16] SRIVASTAVA, M.S. & C.G. KHATRI (1979), *An introduction to multivariate statistics*, North-Holland, New York.

SAMENVATTING

Dit proefschrift handelt over een practische statistische situatie. Ons doel is geweest om de onderzoeksresultaten ook werkelijk toegankelijk te maken voor toepassers van de statistiek, en met dit doel in gedachten is dit werk geschreven. Wij geven bijvoorbeeld een aantal methoden voor het berekenen van (benaderingen van) kritieke waarden voor de voorgestelde toetsingsgrootheden.

De practische situatie betreft het ontdekken van verschillen in voorkeuren of afkeren tussen individuen, wanneer de waarnemingen de (herhaalde) keuzen zijn die zij hebben gedaan. Stel bijvoorbeeld dat n personen mogen kiezen uit k verschillende merken chocolade. Alle personen zouden dezelfde voorkeuren kunnen hebben voor bepaalde merken, mogelijkerwijs variërend in de loop der tijd, maar het is een eventueel verschil tussen de personen met betrekking tot hun voorkeuren dat we willen ontdekken.

In het eerste hoofdstuk wordt het practische probleem en de door ons aanbevolen oplossing uiteengezet. Alle gegevens die nodig zijn voor het uitvoeren van de methode in de praktijk worden in dit hoofdstuk gegeven.

De basis van de oplossing van het probleem vormt een vector $\vec{\underline{t}}_*$ van waarneembare stochastische variabelen, waarvan de asymptotische normaliteit onder bepaalde voorwaarden wordt vastgesteld. De klasse van kwadratische vormen in $\vec{\underline{t}}_*$

$$\mathcal{T} = \{\vec{\underline{t}}_*'Q\vec{\underline{t}}_* \mid Q \text{ niet-negatief definiet}\}$$

wordt dan beschouwd als een mogelijke klasse van practische toetsingsgrootheden. Het gebruik van kwadratische vormen wordt uitgebreid gemotiveerd, zowel intuïtief (sectie 2.1.) als theoretisch (hoofdstuk 8).

Twee problemen deden zich voor bij het bepalen van de asymptotische verdeling van $\vec{\underline{t}}_*'Q\vec{\underline{t}}_*$. Het eerste was de singulariteit van de dispersie-matrix van $\vec{\underline{t}}_*$ (ook in de limiet). Het tweede probleem was de betrekkelijke willekeurigheid van de matrix Q. Slechts voor enkele speciale keuzen van Q zou de toetsingsgrootheid (asymptotisch) een χ^2-verdeling kunnen opleveren. Beide problemen worden opgelost door stelling 3.2.1. welke de verdeling geeft (in expliciete vorm) van een kwadratische vorm in normaal verdeelde variabelen, ook in het geval dat de dispersie-matrix van de normaal verdeelde variabelen singulier is. Met behulp van deze stelling worden de limietverdelingen van $\underline{\vec{t}}_*'\underline{Q}\underline{\vec{t}}_*$ bepaald (onder de nulhypothese en onder alternatieven) (sectie 4.3.).

Een gebruikelijke methode om het probleem van een singuliere dispersie-matrix op te lossen is het definiëren van een transformatie naar een lager dimensionale ruimte waarin de dispersie-matrix van de (getransformeerde) variabelen niet singulier is. Dit geeft meestal aanleiding tot gecompliceerde toetsingsgrootheden en maakt de werking van de toets vager. Met behulp van Rao's theorie over g-inversen van matrices (RAO (1973)) wordt aangetoond dat zo'n transformatie onnodig is (hoofdstuk 6). MADANSKY (1963) gebruikte de methode van transformatie naar een lagere dimensie toen hij een generalisatie van Cochran's Q-toets voorstelde (COCHRAN (1950)). Zowel de toets van Madansky als die van Cochran kan gezien worden als een speciaal geval van de toetsen die we onderzochten (hoofdstuk 6).

De consistentie en het onderscheidingsvermogen van de toetsen uit $\mathcal{T}$ worden beschouwd in hoofdstuk 4. De asymptotische relatieve werkzaamheid van paren van toetsen uit $\mathcal{T}$ volgens Pitman en Bahadur wordt bepaald in hoofdstuk 5. Geen van beide methoden geeft een duidelijke indicatie welke Q gebruikt moet worden opdat de toets tegen meerdere alternatieven een hoog onderscheidingsvermogen heeft.

Een toetsingsgrootheid van het χ^2 type wordt daarom voorgesteld voor gebruik in de praktijk, in hoofdzaak gemotiveerd op intuïtieve en practische gronden (sectie 6.4.). De aanbevelingen worden ondersteund door de resultaten van numerieke simulatie die we geven in hoofdstuk 9. In dat hoofdstuk wordt ook aangetoond dat de toets gericht kan worden op een speciaal alternatief door een geschikte keuze van de matrix Q.

Teneinde een goede en eenvoudige benadering te verkrijgen voor de verdeling onder H_0, wordt tenslotte de verwachting en variantie (onder H_0) van de toetsingsgrootheid bepaald voor enkele speciale keuzen van Q (hoofdstuk 7).

STELLINGEN

bij het proefschrift *Tests for Preference*

van J.J. Dik.

I.

Het gebruik van wegingsfactoren die onderling eenvoudige quotiënten bezitten voor de toetsingsgrootheid $\underset{\sim}{\vec{t}}_*' Q \underset{\sim}{\vec{t}}_*$, met $Q = I_n \otimes G$, van dit proefschrift, kan een nadelige invloed hebben op de nauwkeurigheid van de voorgestelde benaderingen.

II.

Het toetsingsprobleem dat R.L. Anderson beschrijft in het hieronder genoemde artikel is, althans onder H_0, identiek aan het toetsingsprobleem in dit proefschrift voor het speciale geval dat $k = n$ en $a_{ij} = 1$ voor elke i en j. Dit proefschrift bevat daarom een alternatief bewijs voor de asymptotiek van Anderson's toetsingsgrootheid (onder H_0).

ANDERSON, R.L. (1959), *Use of contingency tables in the analysis of consumer preference studies,* Biometrics 15, 582-590.

III.

In de laatste paragraaf van zijn artikel toont SCHACH (1979) aan dat de suggestie die KANNEMANN (1976) doet om het probleem van knopen op te lossen bij de toets van ANDERSON (1959) niet leidt tot een asymptotische χ^2-verdeling. Als (vage) oplossing voor dit probleem stelt hij voor: "numerieke diagonalisatie van de som van de (voorwaardelijke) dispersiematrices, en de constructie van een overeenkomstige toetsingsgrootheid met een asymptotische χ^2-verdeling", wat dit ook moge betekenen.

Dit probleem kan echter worden opgelost door de limietverdeling te bepalen met behulp van stelling 3.2.1. van dit proefschrift, of door gebruik te maken van g-inversen zoals beschreven in dit proefschrift. De eerste oplossingsmethode geeft dan weliswaar geen χ^2-verdeling, maar evenals in dit proefschrift zouden geschikte benaderingen kunnen worden gevonden. Het bezwaar van de tweede oplossing (en waarschijnlijk ook van de oplossing die Schach bedoelt) is dat een andere toetsingsgrootheid wordt gedefinieerd. Het streven naar een χ^2-verdeling dient stellig niet altijd voorop te staan.

ANDERSON, R.L. (1959), *Use of contingency tables in the analysis of consumer preference studies*, Biometrics 15, 582-590.

KANNEMANN, K. (1976), *An incidence test for k related samples*, Biom. Zeitschrift 18, 3-11.

SCHACH, S. (1979), *An alternative to the Friedman test with certain optimality properties*, Ann. of Statist. 7, 537-550.

IV.

De methode die PATIL (1975) geeft voor het berekenen van de exacte verdeling van Cochran's Q-toets geeft geen wezenlijke verkorting van de berekeningen t.o.v. het uitschrijven van alle mogelijkheden. Bovendien is de methode slecht overdraagbaar naar het meer algemene geval dat beschreven wordt in dit proefschrift.

PATIL, K.D. (1975), *Cochran's Q-test: Exact distribution*, Journal of the American Statistical Association 70, 186-189.

V.

Met betrekking tot de Poisson-verdeling geldt de volgende ongelijkheid. Laat $\underline{v}_\lambda \sim \text{Poisson}(\lambda)$. Voor gehele a en b met $0 \leq a < b-1$, en reele μ en λ met $0 < \mu < \lambda$, zodanig dat

$$P(\underline{v}_\mu \leq a) = P(\underline{v}_\lambda \geq b)$$

geldt de volgende ongelijkheid,

$$P(\underline{v}_\lambda \leq a) > P(\underline{v}_\mu \geq b).$$

ZIJM, W.H.M. & J.J. DIK (1979), *Two inequalities for Poisson-probabilities*, Indagationes Mathematicae A 82, 87-94.

VI.

Het is merkwaardig dat voor de (reeds lang bekende) asymptotiek van de multihypergeometrische verdeling in de literatuur geen eenvoudig exact bewijs te vinden is, behalve in enkele speciale gevallen. Het wordt hoog tijd dat in deze lacune wordt voorzien.

VII.

Stelling 3.2.1. van dit proefschrift biedt de mogelijkheid om in een $k \times r$ tabel met vaste marges cellen samen te nemen anders dan koloms- of rij gewijs. Het is dan mogelijk om de toets speciaal te richten op afwijkingen van de nulhypothese die zich bijvoorbeeld in de diagonaal cellen van de tabel voordoen.

VIII.

Bij het berekenen van tweezijdige betrouwbaarheidsintervallen voor de parameters van de Poisson-verdeling en van de binomiale verdeling wordt soms voorgesteld om, indien een uitkomst 0 is verkregen, slechts een rechterbetrouwbaarheidsgrens te berekenen met α i.p.v. $\frac{1}{2}\alpha$. Deze methode is onjuist.

IX.

Een goed model voor de verdeling van het aantal woorden per zin bij Nederlandse roman-schrijvers is de gegeneraliseerde hypergeometrische verdeling, type IV, volgens de indeling van Kemp & Kemp (JOHNSON & KOTZ (1969)).

De parameters van de verdeling kunnen dienen voor karakterisering van de stijl van de schrijvers. Tevens is een interpretatie van dit model mogelijk.

JOHNSON, N.L. & S. KOTZ (1969), *Distributions in statistics: Discrete distributions*, John Wiley & Sons, New York.

X.

Buiten de reeds bekende voordelen van het onderstrepen van stochastische variabelen, zoals de duidelijke herkenbaarheid van parameters en stochasten en het zuiniger gebruik van letters, zeker in teksten waarin matrices voorkomen, heeft het onderstrepen van stochastische variabelen nog het voordeel dat bij mondelinge explicatie waarbij een schoolbord wordt gebruikt de docent niet (noodgedwongen) van de notatie van een boek hoeft af te wijken.

XI.

In de toename van de mogelijkheden om statistische gegevens per computer te laten verwerken door statistische bibliotheek-procedures zoals SPSS schuilt het gevaar dat steeds meer waarnemingen worden "verguld", d.w.z. dat de waarde van de door de computer afgedrukte resultaten niet meer is dan uiterlijke schijn.

XII.

In de wiskundige omgangstaal verdient het de voorkeur om de binomiaal coëfficient $\binom{n}{k}$ uit te spreken als "k uit n" in plaats van het (onzinnige) "n boven k", of, erger, "n over k".