Experimental design and statistical evaluation of limiting dilution assays

Citation for published version (APA):

Strijbosch, L. W. G. (1989). *Experimental design and statistical evaluation of limiting dilution assays*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19890421ls>

Document status and date:

Published: 01/01/1989

DOI:

[10.26481/dis.19890421ls](https://doi.org/10.26481/dis.19890421ls)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

**EXPERIMENTAL DESIGN AND STATISTICAL EVALUATION
OF LIMITING DILUTION ASSAYS**

PROEFSCHRIFT

ter verkrijging van de graad van
doctor aan de Rijksuniversiteit Limburg te Maastricht,

op gezag van

de Rector Magnificus, Prof.Dr. F.I.M. Bonke,
volgens het besluit van het College van Dekanen,

in het openbaar te verdedigen op
vrijdag, 21 april 1989 om 16.00 uur

door

LEONARDUS WILHELMUS GERARDUS STRIJBOSCH

geboren te Helmond

Promotor : Prof.Dr. W. Albers (UT)
Co-promotor : Dr. R.J.M.M. Does (RL)

Beoordelingscommissie :

Prof.Dr. G. Kootstra, voorzitter (RL)
Prof.Dr. B.B. van der Genugten (KUB)
Prof.Dr. F.F.H. Rutten (RL)
Prof. R. van Strik (EUR)
Prof.Dr.Ir. O.J. Vrieze (RL)

*Financial support by the Dutch Kidney Foundation ("Nier Stichting Nederland")
for the publication of this thesis is gratefully acknowledged.*

*aan Liesbeth, Willem, Manus en Eefje
aan mijn moeder
in dankbare herinnering aan mijn vader*

CONTENTS

page

Chapter 1 INTRODUCTION

1.1 The theory of limiting dilution assays	9
1.2 Main issues of limiting dilution assays examined in this dissertation	12

Chapter 2 EXPERIMENTAL DESIGN

2.1 Introduction	15
2.2 The basis of the experimental design	16
2.3 A new procedure for devising an experimental design	
2.3.1 Introduction	17
2.3.2 Graphical explanation and derivation of the design formulae	19
2.3.3 Proof of the correctness of the design formulae	21
2.4 Practical applications	22

Chapter 3 DESIGN REFINEMENTS

3.1 Introduction	29
3.2 The Fisher information	30
3.3 The choice of the sufficiently informative range	31
3.4 The choice of the numbers of replicates	32
3.5 A program for the design of dilution assays	33

Chapter 4 DESIGN GENERALIZATIONS

4.1 Introduction	39
4.2 The one-parameter biological model $y = \beta x$	40
4.3 The two-parameter biological model $y = \alpha + \beta x$	
4.3.1 Introduction	41

4.3.2	The case $\alpha = \alpha_0$ and $\beta \in [\beta_1, \beta_2]$	42
4.3.3	The case $\alpha \in [\alpha_1, \alpha_2]$ and $\beta = \beta_0$	43
4.3.4	The case $\alpha \in [\alpha_1, \alpha_2]$ and $\beta \in [\beta_1, \beta_2]$	44
4.3.5	Applications of the proposed designs	49

Chapter 5 CONVENTIONAL ESTIMATION METHODS

5.1	Introduction	51
5.2	Notation	54
5.3	The maximum likelihood (ML) method	
5.3.1	Introduction	54
5.3.2	The ML estimation procedure	55
5.3.3	Reducing the bias of the ML estimator from infinity to the desired order	56
5.4	The minimum chi-square (MC) method	
5.4.1	Introduction	57
5.4.2	The MC estimation procedure	58
5.5	Confidence intervals	59

Chapter 6 BIAS-REDUCING ESTIMATION METHODS

6.1	Introduction	61
6.2	The jackknife (three versions)	
6.2.1	Introduction	62
6.2.2	The iid case	63
6.2.3	Two non-iid cases	63
6.2.4	Reducing the bias with jackknifing in the non-iid cases	65
6.3	The bootstrap (four versions)	
6.3.1	Introduction	66
6.3.2	The iid case	67
6.3.3	Three non-iid cases	68
6.4	Other bias-reducing methods	69

Chapter 7 A MONTE CARLO STUDY FOR THE COMPARISON OF THE ESTIMATORS

7.1	Introduction	71
7.2	The simulation program	73
7.3	Results	
7.3.1	Introduction	74
7.3.2	Comparison of the ML estimator with the jackknife versions	75
7.3.3	Comparison of the ML estimator with the bootstrap versions	78
7.3.4	Comparison of the ML estimator with other bias-reducing estimators	79
7.3.5	Final comparison	81
APPENDIX	A PASCAL program for the simulation of limiting dilution experiments and the comparison of estimation methods	93
REFERENCES		105
SUMMARY		113
SAMENVATTING (Summary in Dutch)		115
DANKBETUIGING		119
CURRICULUM VITAE		120

1 INTRODUCTION

1.1 The theory of limiting dilution assays

Limiting and Serial Dilution Assays (LDA's, SDA's) are applicable in many areas of bioresearch, such as public hygiene, cell biology, immunology, bacteriology and virology (see Taswell (1987) for a review). In general, the aim of such an analysis is to estimate either the frequency of a particular cell type among other but similar cells (LDA), or the density of particular cells or microorganisms per unit volume of solution, such as bacteria and viruses capable of (un)limited proliferation or having infective properties for certain cells (SDA). LDA's, SDA's and other related assays known as fermentation tube tests, coliform density tests, etc., can all be treated by the same statistical theory, though the underlying biological processes and laboratorium techniques differ. In both LDA and SDA the unknown parameter is commonly estimated by using the "single-hit Poisson model" with binary (positive or negative) data produced by samples taken from different dilutions. The assumptions underlying this model are well-known (see Finney (1978, p. 425), Taswell (1981), Miller (1982), Lefkovits and Waldmann (1984)) and will be described briefly using the terminology of LDA first. Fig. 1.1 illustrates the setup of a limiting dilution experiment. A test preparation contains numerous cells of which an unknown proportion ϕ has a certain property, for example, immunocompetency. From this test preparation, m different dilutions are prepared. Then, from dilution i , n_i replicate cultures are taken such that the numbers of cells in the replicate cultures of dilution i are independent Poisson distributed variables with the same mean number x_i of cells tested. A fraction ϕ of those cells has the intended property. The use of the Poisson distribution for the number of cells in a sample is justified by the assumptions that the total number of cells in the test preparation is sufficiently large, the suspension has been stirred well before taking any sample or dilution and the sample volume of the cells is extremely small when compared to the total volume of the sample. A further assumption is that a positive response is obtained for a replicate culture

if and only if at least one cell of the specific type is present ("single-hit").

Fig. 1.1. The setup of a limiting dilution experiment. The number of dilutions is denoted by m ; x_i is the (mean) number of cells tested in a replicate culture of dilution i ; φx_i is the (mean) number of cells of the specific type contained in that culture; φ is the parameter to be estimated.

For SDA, φ is the density of a specific cell or microorganism per unit volume and the x_i 's, $i=1, \dots, m$, are dilution factors. For both LDA and SDA, (φx_i) is the expected number of relevant cells for a culture in group i while $\exp(-\varphi x_i)$, representing the zero term of the Poisson distribution that quantifies the number of relevant cells in the cultures, is the probability of a negative response. Hereinafter the terminology of LDA will prevail. LDA's are especially useful in the field of cellular immunology. The assay is often used for the estimation of T lymphocytes with specific properties such as cytotoxic properties or cytokine producing properties. It has been found to be a useful technique for studying (among others) the kinetics of specific T cell responses during an ongoing immune response. The culture techniques used must be adapted to the high demands of the assay system since accurate frequency analysis demands cell culture conditions that permit the growth of 'every' cell with the defined properties. In most experimental studies it is, of course, not possible to establish whether the culture conditions really allow the growth of every cell. Usually a constant fraction of such cells responds under the conditions used.

Experimentally, the simplest approach is to distinguish between those cultures which respond and those which fail to respond, corresponding to cultures which contain either at least one relevant cell or no relevant

cells. Thus the dose-response model $P = \exp(-\phi x)$ (single-hit Poisson model) is used to describe the relation between the mean number of cells tested per replicate culture and the number of negatively responding cultures per group. Culture groups for which $\phi x \gg 1$ will most surely respond positively (value of P is near 0), culture groups for which $\phi x \ll 1$ are likely to respond negatively (value of P is near 1). It is clear that a useful determination of the unknown parameter in dilution analysis requires a setup whereby $\phi x \sim 1$ for at least some culture groups (value of P is intermediate). At least some culture groups should be informative for values of ϕ in a certain range. This means that advance information about ϕ must be included in the design. No simple and widely accepted methods providing design characteristics on the basis of the researcher's criteria and a priori information about the unknown parameter can be found in existing literature.

Many different statistical methods are reported and used for the analysis of data from LDA's and SDA's (see Loyer (1981), Taswell (1981)). Some of these (e.g. the least squares method, which fits a straight line $y=a+bx$ where y is minus the natural logarithm of the fraction of negatively responding cultures in a group and x is the corresponding (mean) number of cells tested in those cultures), are statistically far from ideal but easily executed, while others (like the maximum likelihood and minimum chi-square methods) are statistically more appropriate but much less easy to apply. However, with the present widespread availability of (micro-)computers and software, ease of computation is no longer a valid enough argument for preferring the simpler statistical methods. Since McGrady (1915) described an estimation method that determines an estimator $\hat{\phi}$ ("the most probable number") maximizing the probability of obtaining the specific arrangement of positive and negative replicate cultures observed, the maximum likelihood method has been used by many authors. However, after the simulation study of Taswell (1981), who compared the maximum likelihood method and the minimum chi-square method and found that the last method should be preferred, a controversy arose with respect to the use of those methods in dilution analysis (see Fazekas de St. Groth (1982)).

LDA and SDA belong to the general area of qualitative response (QR) models. Most available textbooks on QR models are written for disciplines such as biology, psychology, sociology or engineering. For example, see Cox (1970), Finney (1971), Plackett (1974), Haberman (1974, 1978, and 1979),

Bishop, Fienberg and Holland (1975), and Daganzo (1979). Basic facts about applying QR models in general and, in particular to economics, can be found in a survey by Amemiya (1981).

1.2 Main issues of limiting dilution assays examined in this dissertation

Statisticians can contribute to the execution of a LDA or a SDA in at least three ways. In the first place they can provide validity tests that detect deviations from the single-hit Poisson model for parameter-estimation assays with a known culture system where this model is assumed to apply most of the time. Taswell (1984) discusses principles of bioassay validity tests. However, validity tests will not be discussed in this dissertation. Secondly, they can help the researcher construct an experimental design which will take advantage of existing a priori information. This hopefully precludes experimentation producing useless data and enables the adjustment of the precision of an estimator. And thirdly, they can advise on the statistical techniques to be used. The latter two topics form the subject of this study.

In many applications of dilution analysis the assays are very expensive and time-consuming, and, in some circumstances, also unrepeatable. In these cases it is of vital interest to carefully choose an experimental design and an estimation method minimizing bias and/or mean square error. Recent research has been performed on design problems (Loyer (1981), Taswell (1987), Strijbosch, Buurman, Does et al. (1987), Strijbosch, Does and Buurman (1988) and Strijbosch, Does and Albers (1988)). And, Monte Carlo studies have been made on the choice of the statistical procedure to be used (Salama, Koch and Tolley (1978), Loyer (1981), Taswell (1981), Strijbosch, Buurman, Does et al. (1987), Does, Strijbosch and Albers (1988) and Strijbosch and Does (1988)). Note that there is a difficulty in comparing the results of these Monte Carlo experiments because of the absence of generally accepted design methods; most authors used different experimental designs when generating simulation results. The controversy between Taswell (1981) and Fazekas de St. Groth (1982) emerged partly as a consequence of this kind of difficulty. It is obvious that the statistical properties of the possible estimators are dependent on the designs used. If these properties do not hold with other possible designs, it could easily occur

that one author finds that estimator 1 is better than estimator 2 while another author finds the contrary.

The results on experimental design described by Strijbosch, Buurman, Does et al. (1987) are presented in Chapter 2. Refinements upon this design method and a program to perform these techniques are described by Strijbosch, Does and Buurman (1988) and presented in Chapter 3. The methods of Chapter 2 and 3 pertain to the experimental design for LDA's and SDA's. As is described by Strijbosch, Does and Albers (1988), the basic design method of Section 2.3 can be generalized for some other related and widely used biological models, such as the logistic regression model which is most common in bioassay. These generalizations are presented in Chapter 4. Chapter 2 to 4 provide formulae for the number of culture or dose groups and for the (mean) number of cells, dilution factors or doses to be used. These formulae incorporate a lower and an upper bound for the unknown parameter(s), as well as some meaningful design parameters chosen by the researcher. The second main subject to be dealt with will be commonly used estimation methods (Chapter 5) and some bias-reducing methods (Chapter 6) when applied to dilution analyses as described by Strijbosch, Buurman, Does et al. (1987), Does, Strijbosch and Albers (1988) and Strijbosch and Does (1988). In these studies Monte Carlo experiments have been performed in order to compare maximum likelihood and minimum chi-square methods on the one hand and bias-reducing methods on the other hand. These simulation results are presented and discussed in Chapter 7. Finally, the relevant part of the PASCAL program used for the simulations is given in the Appendix.

2 EXPERIMENTAL DESIGN

2.1 Introduction

An important issue in dilution analysis is the experimental design, i.e., the choice of the number of dilutions, the number of replicates in each dilution and the (mean) number of cells/organisms to be used per replicate. Assay designs are generally set up starting from a particular a priori estimate of the unknown parameter. However, depending on the type of experiment, a priori knowledge of the unknown parameter varies significantly. Taking into account the often limited availability of cells/organisms to be tested and the very labourious and costly experimental procedures, it is obviously mandatory that an experimental design make optimal use of available means. In spite of the fact that dilution assays have been used since the beginning of this century, the development of design methods has been somewhat limited. Some of the recent suggestions for the experimental design of a dilution assay will be discussed. The design suggestions made by Finney (1978, p. 435) for dilution series can be used in practice although ready-to-use formulae are not given (see Subsection 2.3.1). Loyer (1981, p. 53) proposes a complicated procedure for the determination of an experimental design which is not suitable for general use. A design determined according to his method has the property that the probability of obtaining all fertile samples (i.e. only positive responses) is not larger than a user-supplied risk for all values of the unknown parameter which are smaller than a user-supplied maximal value. There are several objections to this procedure which will be discussed in Section 2.4. Next, without giving any design formulae, Taswell (1987) suggests that "the goal of optimization methods should be to minimize the error of the estimates subject to the constraint of a chosen level of the noninformative assay probability" which is defined as the probability of obtaining either all responses negative or all responses positive.

In practice, both symmetrical and asymmetrical designs have been used for LDA's (see e.g. Maryanski, Cerottini and Brunner (1980), Maryanski, MacDonald and Cerottini (1980), Vose and Bonnard (1983)) without explicit

knowledge of the properties of the designs. In this Chapter a useful and simple method for setting up a design, on the basis of the researcher's criteria and a priori information regarding the unknown parameter φ , will be presented. Section 2.2 describes which basic considerations suggested and derived in Section 2.3 underlie the design method for dilution series. Finally, Section 2.4 discusses the practical application of the proposed method. An example of a design mentioned in the literature is compared with the design which would have been obtained by the proposed design method.

2.2 The basis of the experimental design

This dissertation will describe a frequency determination of so-called single-hit phenomena. Culture conditions must be chosen carefully to fulfil all the demands of single-hit responses, i.e., the relevant cells must be the only limiting cells in the system, and all relevant cells must grow out to be colonies the properties of which must be detectable in a yes/no way. Such cells can be B cells producing antibodies, or T cells with cytotoxic or lymphokine-producing properties. The numbers of negative replicates, the total numbers of replicates per dilution and the (mean) numbers of cells per culture are used to determine the frequency of the limiting cell in the population. The statistical procedure used for the estimation of the unknown frequency is based on the assumption that the numbers of cells per culture have a Poisson distribution. It can simply be deduced that in the case of a frequency φ of a certain cell, the probability P of a culture with x cells responding negatively is $\exp(-\varphi x)$. This dose-response relationship $P = \exp(-\varphi x)$ is the starting point in the development of the experimental design presented in Section 2.3.

The need for such a development lies in the fact that in practice it turns out to be difficult to design experiments of limited size (e.g. 7 dilutions and 20 replicates) in which the design parameters are chosen to incorporate information available in advance about the unknown parameter and such that the design is optimal in some sense. Experiments for the determination of the unknown frequency φ of a certain well-defined cell in a population of cells by limiting dilution techniques can only be performed if a reasonable a priori estimate of that frequency is available. The design of the experiment is thus dependent on such point estimate $\tilde{\varphi}$ or (more general)

an interval estimate $[\varphi_1, \varphi_2]$ representing the researcher's boundaries for possible values of φ . Another aspect which should play a substantial role in the construction of a design is the fact that each result of the experiment (depending on the true and unknown value of φ) must contain a sufficient number of groups producing fractions of negative responders not too close to zero or one.

2.3 A new procedure for devising an experimental design

2.3.1 Introduction

The design problem is split into two parts. In the first part the number of dilutions and the corresponding (mean) numbers of cells tested are determined, whereas in the second part the number of replicate cultures are chosen. The design method proposed in Subsection 2.3.2 concerns only part one. Part two is dealt with in Chapter 3. The approach of Chapter 3 is more heuristic than that of Chapter 2.

Single-dose designs, i.e. designs containing one set of experiments with a fixed mean number of cells per culture, are discussed by many authors, including Fisher (1922), Finney (1978, p. 435) and Fazekas de St. Groth (1982). Fisher (1922) calculated that the optimal dose $x=1.59/\varphi$ where φ is the unknown frequency, which corresponds to a fraction of 0.203 negative responders. For this value the Fisher information is maximized (see Section 3.2). This suggests that selecting $x=1.59/\tilde{\varphi}$, where $\tilde{\varphi}$ represents the researcher's best a priori guess for φ , would provide the most efficient single-dose design. According to Fazekas de St. Groth (1982) the fraction of negative cultures must meet certain criteria. He stated that the experiment is efficient when the negative fraction is between 0.10 and 0.37; this range comprises levels of information (in the sense of the Fisher information) above 90% of the maximum. This corresponds to a suggestion of Finney (1978, p. 428) which is interpreted as follows : if sufficient a priori information is available and if more than one dilution is used, select the values $x_1, \dots, x_m$ such that $0.9 < \tilde{\varphi}x_i < 2.5$, $i=1, \dots, m$, corresponding to negative fractions ranging from 0.08 to 0.41 when $\varphi=\tilde{\varphi}$. This design suggestion could easily produce a setup which fails in providing useful data for the estimation of φ . If the a priori value $\tilde{\varphi}$ deviates too much from φ , all the

replicates might be either negative or positive. In such cases the experiment obviously does not yield any relevant information; nevertheless, one-sided confidence intervals for the unknown parameter can be determined (see Section 5.5). Table 2.1 provides the numerical information which is used in the discussion above. For example the second row has been obtained by solving $I(\varphi|m=1)/I_{\text{opt}}(\varphi|m=1) = (\varphi x)^2 (\exp(1.594) - 1) / \{(\exp(\varphi x) - 1)(1.594)^2\} = 0.95$ for φx , where $I(\varphi|m=1)$ denotes the Fisher information for the single-dose design and $I_{\text{opt}}(\varphi|m=1)$ denotes the maximal value of $I(\varphi|m=1)$.

Efficiency relative to optimal φx	Number of relevant cells in a culture (φx)	Probability of negative response ($\exp(-\varphi x)$)	Number of relevant cells in a culture (φx)	Probability of negative response ($\exp(-\varphi x)$)
1.00	1.594	0.203	1.594	0.203
0.95	1.165	0.312	2.101	0.122
0.90	1.004	0.367	2.346	0.096
0.85	0.885	0.413	2.552	0.078
0.80	0.789	0.455	2.741	0.065
0.75	0.706	0.494	2.922	0.054
0.70	0.632	0.532	3.100	0.045
0.65	0.566	0.568	3.278	0.038
0.60	0.506	0.603	3.459	0.031
0.55	0.450	0.638	3.646	0.026
0.50	0.398	0.672	3.841	0.021
0.40	0.302	0.739	4.266	0.014
0.30	0.217	0.805	4.773	0.008
0.20	0.139	0.870	5.433	0.004
0.10	0.067	0.935	6.474	0.002

Table 2.1. Efficiency relative to optimal φx (=1.594) and corresponding values for the (mean) number φx in a replicate culture and the probability of obtaining a negative response $\exp(-\varphi x)$ when using a single-dose design.

When a useful a priori guess is not available, Finney (1978, p. 437) states that the experimenter must decide between which lower (φ_L) and upper (φ_U) bound φ (almost) certainly lies. Then he must plan an assay with values for the mean numbers of cells per culture ranging from $2/\varphi_L$ through $1/(2\varphi_U)$. Furthermore, Finney (1978) suggests that the dilution factor ("2 and 4 are preferable to 10") should be as small as is practicable. These design proposals, however, lack an indication of the number of dilutions which should be used. Since a geometrical series is (rightly) recommended for $x_1, \dots, x_m$, m should obviously be dependent on φ_U/φ_L . In Subsection 2.3.2 a

simple and efficient design method for the one-parameter model is developed which is based on the researcher's criteria. Compared with the suggestions of Finney (1978) mentioned above, it can be seen as a more elaborated concept which produces a setup with known and suitable properties. According to this design method, certain criteria have to be chosen by the researcher: a lower bound φ_1 and an upper bound φ_2 for the unknown parameter φ ; a lower bound P_1 and an upper bound P_2 for a fraction of negative responders to be considered sufficiently informative; and d , a number of dilutions which should lead to sufficiently informative fractions within the fixed bounds P_1 and P_2 .

2.3.2 Graphical explanation and derivation of the design formulae

The researcher must specify an interval estimate $[\varphi_1, \varphi_2]$ for φ based on existing a priori information. There must be strong reasons supporting the notion that the true value φ is between these bounds. Only fractions of negative responders between certain values P_1 and P_2 ($P_1 < P_2$) are considered sufficiently informative by the researcher. The aim is to have, per experiment (on the average), d dilutions that yield sufficiently informative fractions. Then, using the formulae (2.1)-(2.4), the number of dilutions m , the mean number of cells x_1 in the replicate cultures of the first dilution, and finally the dilution factor c can be determined :

$$(2.1) \quad m = d + \text{ENTIER}(\log(\varphi_2/\varphi_1)/\log(c)),$$

$$(2.2) \quad x_1 = \{(\log P_1)(\log P_2)c^{1-m}/(\varphi_1\varphi_2)\}^{1/2},$$

$$(2.3) \quad c = \{(\log P_1)/(\log P_2)\}^{1/d},$$

$$(2.4) \quad x_i = x_1 c^{i-1}, i=2, \dots, m.$$

('ENTIER' means 'the integer part of' and log is the natural logarithm function).

These design formulae can be explained in a heuristic way assisted by the graphical illustration in Fig. 2.1 (see Strijbosch, Does and Buurman (1987)).

Fig. 2.1. The construction of a design method based on a lower boundary φ_1 and an upper boundary φ_2 for the unknown parameter φ , and on a sufficiently informative range $[P_1, P_2]$ for the proportion of negative responders in individual culture groups. It is assumed here that, for each value of $\varphi \in [\varphi_1, \varphi_2]$, three culture groups produce sufficiently informative data.

In this Figure the dose-response relationship $P = \exp(-\varphi x)$ is presented as a straight line through the origin with slope φ . Suppose that $d=3$, and x_1, x_2, x_3 are given values satisfying the condition $-\log(P_2) < \varphi_2 x_1 < \varphi_2 x_2 < \varphi_2 x_3 \leq -\log(P_1)$. The value of x_4 can then be found by letting φ decrease from φ_2 to φ_0 which is determined by the point $(x_1, -\log(P_2))$. For $\varphi \in (\varphi_0, \varphi_2]$, it follows that $-\log(P_2) < \varphi x_1 < \varphi x_2 < \varphi x_3 \leq -\log(P_1)$, while $\varphi x_1 \leq -\log(P_2)$ for $\varphi \leq \varphi_0$. Thus a new value x_4 has to be added to the series in order to satisfy the design requirements mentioned at the end of Subsection 2.3.1. It follows that $x_4 = -\log(P_1)/\varphi_0$. Since $\varphi_0 = -\log(P_2)/x_1$, it is clear that x_4 is neither dependent on x_2 nor x_3 . Rather x_4 is a function of x_1 : $x_4 = x_1 \log(P_1)/\log(P_2)$. In general it follows that $x_{i+d} = x_i \log(P_1)/\log(P_2)$, so that the choice of a geometrical series $x_i = x_1 c^{i-1}$ with $c = [\log(P_1)/\log(P_2)]^{1/d}$ is obvious. Note that in Fig. 2.1 $x_1, \dots, x_7$ satisfy

$x_i = x_{i-d} c^d$, $i=4, \dots, 7$, but do not satisfy (2.4). The geometrical series (2.4) is chosen for convenience.

The remaining unknown design characteristics x_1 and m can be found by demanding two conditions. The first one is a condition of symmetry, i.e. on a logarithmic scale the interval $[x_1, x_m]$ must lie symmetrically within the interval $[-\log(P_2)/\varphi_2, -\log(P_1)/\varphi_1]$. The lower bound of the last interval is the least possible value for x_1 , and, analogously, the upper bound is the largest value for x_m . Thus : $x_1 x_m = \log(P_1) \log(P_2) / (\varphi_1 \varphi_2)$. As $x_1 x_m = x_1^2 c^{m-1}$, (2.2) is obtained. Finally, (2.1) is explained by observing that typically $x_1 \approx -\log(P_2)/\varphi_2$, while $x_{m-d+1} \approx -\log(P_2)/\varphi_1$, and hence $c^{m-d} = x_{m-d+1}/x_1 \approx \varphi_2/\varphi_1$. This satisfies the second condition being that m be sufficiently large.

In the next Subsection (2.3.3) this procedure is proven to be mathematically correct.

2.3.3 Proof of the correctness of the design formulae

The following proposition will be proved.

For each $\varphi \in [\varphi_1, \varphi_2]$, exactly d values of $x_1, \dots, x_m$ as defined by (2.1) - (2.4) satisfy the condition $P_1 \leq \exp(-\varphi x_i) < P_2$.

Let $x_0 = x_1/c$ and $x_{m+1} = x_m c$. It suffices to prove that for each $\varphi \in [\varphi_1, \varphi_2]$ exactly one integer i_φ exists which satisfies the following conditions :

$$(2.5) \quad 1 \leq i_\varphi \leq m-d+1,$$

$$(2.6) \quad x_{i_\varphi} > -\log(P_2)/\varphi,$$

$$(2.7) \quad x_{i_\varphi+d-1} \leq -\log(P_1)/\varphi,$$

$$(2.8) \quad x_{i_\varphi-1} \leq -\log(P_2)/\varphi,$$

$$(2.9) \quad x_{i_\varphi+d} > -\log(P_1)/\varphi.$$

From (2.1) and (2.3) it can be shown that (2.6) and (2.7) as well as (2.8) and (2.9) are equivalent to

$$(2.10) \quad 1 < x_{i_{\varphi}} \varphi / (-\log(P_2)) \leq c.$$

From (2.1) it follows that

$$(2.11) \quad m = d + \log(\varphi_2/\varphi_1)/\log(c) - \epsilon, \text{ for some } \epsilon, 0 \leq \epsilon < 1.$$

Thus, from (2.2)

$$(2.12) \quad x_1 = -\log(P_2) c^{(1+\epsilon)/2} / \varphi_2.$$

From (2.3), (2.10) and (2.12) it follows

$$(2.13) \quad (1-\epsilon)/2 + \log(\varphi_2/\varphi)/\log(c) < i_{\varphi} \leq (3-\epsilon)/2 + \log(\varphi_2/\varphi)/\log(c).$$

Hence exactly one integer i_{φ} satisfies (2.13). Furthermore, it is obvious from (2.11) that

$$(2.14) \quad 0 \leq \log(\varphi_2/\varphi)/\log(c) \leq m-d+\epsilon.$$

From (2.13) and (2.14) can be concluded that (2.5) is valid.

2.4 Practical applications

An application of (2.1)-(2.4)

Suppose that an experiment has to be set up to estimate the frequency of a certain well-defined cell in a population of cells. The frequency of that cell is a priori estimated to be smaller than $1/2000$ ($=\varphi_1$) and larger than $1/200$ ($=\varphi_2$), and for each value of the unknown parameter φ within these bounds 3 ($=d$) groups of replicate cultures are required yielding fractions of negative cultures larger than 0.15 ($=P_1$) and smaller than 0.70 ($=P_2$). Applying the design formulae (2.1)-(2.4) leads to the following experimental design : $m = 7$ groups of replicate cultures with respectively 98, 171, 298,

520, 908, 1585 and 2767 ($x_1, \dots, x_7$) cells tested per culture. In order to clarify the procedure Table 2.2 and Fig. 2.2 are presented. Table 2.2 gives the probabilities of a negative response for all combinations of the seven x_i 's used and five different values of the unknown frequency φ . As a result of the procedure, exactly three probabilities in the range 0.15 through 0.70 are obtained for each frequency φ between φ_1 and φ_2 .

φ	x_1	x_2	x_3	x_4	x_5	x_6	x_7
	98	171	298	520	908	1585	2767
1/200	(0.61)	(0.43)	(0.23)	0.07	0.01	0.00	0.00
1/300	0.72	(0.57)	(0.37)	(0.18)	0.05	0.01	0.00
1/500	0.82	0.71	(0.55)	(0.35)	(0.16)	0.04	0.00
1/1000	0.91	0.84	0.74	(0.59)	(0.40)	(0.20)	0.06
1/2000	0.95	0.92	0.86	0.77	(0.64)	(0.45)	(0.25)

Table 2.2. Some function values of $f_i(\varphi) = \exp(-\varphi x_i)$, $i=1, \dots, m$, representing the probabilities of a negative response. The following design parameters are used : $\varphi_1=1/2000$, $\varphi_2=1/200$, $d=3$, $P_1=0.15$ and $P_2=0.70$, resulting in $m=7$. The values between P_1 and P_2 appear between parentheses.

In Fig. 2.2 the functions $\exp(-\varphi x_i)$, $i=1, \dots, m$, for $\varphi \in [\varphi_1, \varphi_2]$, as well as the horizontal lines at P_1 and P_2 , are presented. Fig. 2.2 illustrates the property of the design, i.e. a vertical line for each $\varphi \in [\varphi_1, \varphi_2]$ each intersecting exactly d functions at points between P_1 and P_2 .

A comparison with Loyer's design

From the design procedure proposed by Loyer (1981) the first dilution (x_1) of a geometrical series of dilutions is determined as a function of five parameters : the number of replicates (n), the number of culture groups (m), the dilution factor (c), an upper bound for the unknown parameter (φ_2) and the maximum acceptable probability of obtaining only positive responses (ρ). The formula used is :

$$(2.15) \quad \prod_{i=1}^m (1 - \exp(-\varphi_2 x_1 c^{i-1}))^n = \rho.$$

Fig. 2.2. Graph of the functions $f_i(\varphi)=\exp(-\varphi x_i)$, $i=1,\dots,m$, representing the probabilities of a negative response, and $g_i(\varphi)=P_i$, $i=1,2$. The following design parameters are used : $\varphi_1=1/2000$, $\varphi_2=1/200$, $d=3$, $P_1=0.15$ and $P_2=0.70$, yielding $m=7$.

The value of x_1 is determined from iterative computation. Apart from being complicated, there are several objections to this procedure. In the first place, the proposed design method demands that the number of dilutions as well as the number of replicates per dilution be chosen in advance. Secondly, the values of the (mean) number of cells tested per replicate are dependent on the number of replicates while this is generally not desirable. Furthermore, only an upper bound for the unknown parameter is included in the procedure. It is not consistent to use a design method which prohibits the occurrence of only positive responses but does not prohibit the event of only negative responses. The maximum probability of this event is given by $\prod_{i=1}^m \exp(-n\varphi_1 x_i) = \exp(-n\varphi_1 x_1 \sum_{i=1}^m c^{i-1})$, where φ_1 is the lower bound for the unknown parameter. Table 2.3 illustrates the properties of Loyer's design procedure through a comparison with parameter values for the procedure (2.1)-(2.4) producing approximately equal values for c , m , and x_1 . Note that

in some cases other choices for P_1 and P_2 would be desirable (see also Section 3.3). In Table 2.3 the lower bound φ_1 is calculated according to $\varphi_1 = \varphi_2 c^{d-m} + \epsilon$, ($\epsilon > 0$). When ϵ is sufficiently small, P_1 and P_2 can be determined from $(\log P_1)^2 \approx x_1^2 c^{2d-1} \varphi_2^2$ and $(\log P_2)^2 \approx x_1^2 c^{-1} \varphi_2^2$, respectively. Thus P_2 is dependent on d only through ϵ . In the Table $\epsilon = 0.01$ is chosen.

LOYER'S-----DESIGN-----PROCEDURE						PARAMETER VALUES FOR (2.1)- (2.4) PRODUCING EQUAL c, m, x_1			
INPUT			OUTPUT						
n	m	ϵ	c	φ_2	x_1	d	φ_1	P_1	P_2
6	3	0.01	2	8	0.109	1	1.01	0.417	0.646
						2	2.01	0.175	0.647
						3	4.01	0.031	0.647
12	3	0.01	2	8	0.166	1	1.01	0.262	0.512
						2	2.01	0.069	0.513
						3	4.01	0.005	0.514
24	3	0.01	2	8	0.234	1	1.01	0.153	0.391
						2	2.01	0.024	0.392
						3	4.01	0.0005	0.392

Table 2.3. Comparison of some designs obtained from Loyer's design procedure with possible combinations of parameter values for the procedure (2.1)-(2.4) producing approximately equal c, m and x_1 .

Estimating the density of rope spores in potato flour

The design parameters in the next example are chosen such that a design results which is described in the literature. Finney (1978), Section 20.6, presents a design, referring to Fisher and Yates (1963), for the estimation of the density of rope spores (*B. mesentericus*) in potato flour. The experiment used 5 replicates, each of 1 c.c., for each of 10 dilutions obtained in a geometrical fashion with $x_1 = 1/128$ g. per 100 c.c. = 0.0078125 g. per 100 c.c., and $x_i = x_1 2^{i-1}$. The corresponding numbers of negative responders were 5,5,3,3,2,1,0,0,0,0, respectively, leading to the estimate $\hat{\varphi} = 7.66$ rope spores per 1/100 gram. Approximately the same (serial) dilution series could be obtained using (2.1)-(2.4) with, for example, $\varphi_1 = 0.40$, $\varphi_2 = 52.9$, $d = 3$, $P_1 = 0.10$ and $P_2 = 0.75$, leading to $c = 2.0003$, $m = 10$ and

$x_1=0.007814$. This exercise makes clear under which assumptions the described design was valid.

The relation between m and P_1, P_2, d

The design procedure forces one to choose values for P_1 and P_2 . According to Fazekas de St. Groth (1982), it is considered better not to choose P_1 and P_2 symmetrically around 0.203 which is the optimal value for when $P_1=P_2$ and $\varphi_1=\varphi_2$ and there are large numbers of replicates. Suggestions for the determination of P_1 and P_2 in the general case can be found in Section 3.3. However, it seems sensible to choose P_1 always smaller than 0.203 and P_2 always larger. Furthermore, making P_1 larger and/or P_2 smaller will usually increase m (but never decrease m). The same is true for making φ_1 smaller and/or φ_2 larger. Increasing d will always increase m . From practical experience in limiting dilution assays it appears that $d=3$, $P_1=0.15$, $P_2=0.70$, are reasonable choices, whereas the a priori boundaries on φ can often be chosen so that $\varphi_2/\varphi_1=10$. The influence of d and the choices for P_1 and P_2 on the number of culture groups is illustrated in Table 2.4.

P_1	P_2	d	m	P_1	P_2	d	m
0.05	0.80	1	1	0.10	0.75	1	2
0.05	0.80	2	3	0.10	0.75	2	4
0.05	0.80	3	5	0.10	0.75	3	6
0.05	0.80	4	7	0.10	0.75	4	8
P_1	P_2	d	m	P_1	P_2	d	m
0.15	0.70	1	2	0.17	0.50	1	3
0.15	0.70	2	4	0.17	0.50	2	6
0.15	0.70	3	7	0.17	0.50	3	10
0.15	0.70	4	9	0.17	0.50	4	13

Table 2.4. The dependence of the number m of culture groups on the values of the design parameters P_1 , P_2 and d . The quotient φ_2/φ_1 is taken as 10.

(2.1)-(2.4) applied in practice

The design method presented here has been applied (see for example Doveren, Buurman, van der Linden et al. (1986)) and has been found useful, especially when low numbers of cells were available for LDA's such as in experiments with cells obtained from fine needle biopsies of rejecting kidney allografts. A typical experiment from the study of Doveren, Buurman, van der Linden et al. (1986) is presented in Table 2.5.

Culture group	Cells tested per replicate	Number of replicates	Number of negative cultures	Fraction of negative cultures
1	176	24	24	1.00
2	289	24	24	1.00
3	473	24	23	0.96
4	775	24	18	0.75
5	1271	24	21	0.87
6	2084	24	15	0.63
7	3416	24	11	0.46
8	5600	24	6	0.25

Table 2.5. Experimental data from a LDA. Responder cells from a kidney allograft were obtained by fine-needle aspiration biopsy 8 days after canine kidney transplantation. The responder cells from the recipient were cultured with stimulator cells from the donor for 10 days and assayed for specific cytotoxicity against PHA blasts of the donor dog. The design parameters chosen were $\phi_1=1/4000$, $\phi_2=1/300$, $P_1=0.15$, $P_2=0.65$, and $d=3$. The procedure to calculate the number of culture groups and numbers of cells tested per replicate produced $m=8$, $x_g=5584$ and $c=1.6391$. For convenience x_g and c have been replaced by 5600 and 100/61 respectively. Replicate cultures were scored as negative for CTL-P if the ^{51}Cr release value was smaller than three times the standard deviation above mean background level.

3 DESIGN REFINEMENTS

3.1 Introduction

In this Chapter the design method for dilution series will be completed with some suggestions for tuning the design parameters P_1 and P_2 when the number of dilutions has already been determined. Guidelines and formulae for choosing these parameters are given in Section 3.3. A method for determining the (equal) number of replicates $n=n_i$, $i=1, \dots, m$, on the basis of a user-supplied coefficient of variation is given in Section 3.4. Section 3.2 describes the role the Fisher information plays in design considerations. The application of the proposed design methods is straightforward, although the computer program described in Section 3.5 will expediate the setup of a design in the laboratory.

There are several uncertainties in choosing the design parameters. Of course, it is difficult to fix values for φ_1 and φ_2 . Perhaps a more natural approach is to base a design on a 'bell-shaped' a priori distribution rather than on a uniform distribution as was conceived in (2.1)-(2.4). Furthermore, it is unclear how P_1 and P_2 should be chosen. In general, values like $P_1=0.15$ and $P_2=0.70$ will suffice (see Fazekas de St. Groth (1982)), but it would be more ideal, for example, to choose P_1 and P_2 as close as possible to some value P_0 which is the best choice for a single-dose design with a certain number of replicates (see Subsection 2.3.1 and Section 3.2). Formula (2.1) admits two alternative refinements that could be important in constructing a design for a specific experiment. Both refinements make use of the fact that either φ_2/φ_1 might be larger or c might be smaller without altering m because in (2.1) ENTIER is used. With equal m and c , every φ_1^* and φ_2^* satisfying $c^{m-d} \leq \varphi_2^*/\varphi_1^* < c^{m-d+1}$ could be chosen instead of φ_1 and φ_2 . This results in the same x_1 only when $\varphi_1^*\varphi_2^* = \varphi_1\varphi_2$ which is assumed from now on. Analogously, with equal m and φ_2/φ_1 , P_1^* and P_2^* could be chosen instead of P_1 and P_2 whenever $(\varphi_2/\varphi_1)^{1/(m-d+1)} < c^* = [\log(P_1^*)/\log(P_2^*)]^{1/d} \leq (\varphi_2/\varphi_1)^{1/(m-d)}$. These are important tools. It follows that, with an equal number of dilutions, the properties of a design are valid for a larger

interval of φ , and, alternatively, it is possible to choose other, maybe better, values for P_1 and P_2 .

3.2 The Fisher information

An important criterium in design considerations is the Fisher information defined by $I(\varphi) = \text{var}_{\varphi} (\partial \log L(\varphi) / \partial \varphi)$ where $\log L(\varphi)$ denotes the log likelihood function (see Bickel and Doksum (1977, p. 127), and Subsection 5.3.2). It follows that for an assay with mean numbers of cells $x_1, \dots, x_m$, numbers of replicates $n_1, \dots, n_m$ and φ as the true value of the parameter, the Fisher information is given by

$$(3.1) \quad I(\varphi) = \sum_{i=1}^m n_i x_i^2 / (e^{\varphi x_i} - 1).$$

The asymptotic variance of the maximum likelihood estimator $\hat{\varphi}_{ML}$ (see Subsection 5.3.2) equals the reciprocal of $I(\varphi)$ (cf. formula (3.2)). Since φ is unknown, $I(\varphi)$ is unknown too. However, an approximation of $I(\varphi)$ is obtained as a by-product of the determination of the maximum likelihood estimate (see Subsection 5.3.2). For experiments with one single dose, maximizing $I(\varphi)$ (or minimizing $CV(\varphi)$ to be defined in (3.2)) produces $\varphi x = 1.59$ as noted earlier. This suggests a choice for x when a good point estimate of φ is available. It should be noted here that small sample research conducted by Halvorson and Ziegler (1933) shows that for small numbers of replicates the optimal number of cells should be determined according to $1.2 < \varphi x < 1.5$ rather than to $\varphi x = 1.59$. This suggests that the optimal single-dose fraction P_0 of negative responders should be dependent on the number of replicates and is generally larger than 0.203. However, this subject is not investigated any further in this study. If the a priori information is an interval estimate, it is clear that the assay must contain several different dilutions as determined, for example, by formulae (2.1)-(2.4). With these formulae, the design parameters P_1 and P_2 can be used in order to influence the function $CV(\varphi)$ on $[\varphi_1, \varphi_2]$ in a desirable way.

3.3 The choice of the sufficiently informative range

Better values for P_1 and P_2 should be obtained when the maximum of the asymptotic value of the coefficient of variation on $[\varphi_1, \varphi_2]$ is smaller and hence the estimator varies less. The coefficient of variation is defined by the quotient of the standard error of the ML estimate and φ . Accordingly, let $CV(\varphi)$ be defined by

$$(3.2) \quad CV(\varphi) = 1/(\varphi\sqrt{I(\varphi)}).$$

Both objectives, viz., a smaller $CV(\varphi)$ and less varying estimator $\hat{\varphi}_{ML}$, can be achieved by defining

$$(3.3) \quad P_1^* = P_0^{sq} \text{ and } P_2^* = P_0^{s/q},$$

where $P_0 = 0.203$ (or, depending on the number of replicates, a larger value), and $q = (c_+^*)^{d/2}$ with $c_+^* = (\varphi_2/\varphi_1)^{1/(m-d+1)} + \epsilon$ ($\epsilon > 0$, but small), and s some suitable transformation of q . The addition of a small number ϵ is mandatory, since with $c = (\varphi_2/\varphi_1)^{1/(m-d+1)}$ one extra culture group becomes necessary. However, ϵ should be sufficiently small to attain maximal gain for the alternatives P_1^* and P_2^* . Usually, $\epsilon = 0.01$ can be chosen. Trying several obvious transformations as proposed by Tukey (1977, p. 91), $s = 1/\sqrt{q}$ seems to be suitable for the purpose of equalizing $CV(\varphi)$ as much as possible provided that $n_i = n$, $i = 1, \dots, m$. This result has been obtained from empirical observation rather than algebraic analysis. Table 3.1 illustrates this observational process with a few numerical results. In this Table the transformation which is used in (3.3) is indicated by p where $s = q^p$. The number ϵ is substituted by the value 0.01. The resulting values for P_1^* and P_2^* are given. Moreover, based on the design obtained when substituting P_i^* by P_i^* , $i = 1, 2$, (3.2) is evaluated in φ_1 , $(\varphi_1 + \varphi_2)/2$ and φ_2 . As Table 3.1 shows, the maximum value of $CV(\varphi)$ on $[\varphi_1, \varphi_2]$ is minimal when $p = -0.50$.

d=2	P_1	P_2	c	CV(0.001)	CV(0.0055)	CV(0.01)
	0.150	0.700	2.306	0.274	0.223	0.239
p	P_1^*	P_2^*	c_+^*	CV(0.001)	CV(0.0055)	CV(0.01)
-0.75	0.145	0.662	2.164	0.269	0.219	0.238
-0.50	0.096	0.606	2.164	0.253	0.224	0.248
-0.25	0.058	0.545	2.164	0.240	0.230	0.261
d=3	P_1	P_2	c	CV(0.001)	CV(0.0055)	CV(0.01)
	0.150	0.700	1.746	0.203	0.179	0.188
p	P_1^*	P_2^*	c_+^*	CV(0.001)	CV(0.0055)	CV(0.01)
-0.75	0.150	0.626	1.595	0.199	0.170	0.184
-0.50	0.104	0.572	1.595	0.189	0.173	0.191
-0.25	0.067	0.515	1.595	0.181	0.177	0.200

Table 3.1. Motivation of the choice for the transformation function s in formula (3.3). The parameters of the designs used are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $n=12$, and $d=2$ or 3 , resulting in $m=4$ or 7 , respectively.

3.4 The choice of the numbers of replicates

The second part of the design problem is the determination of the numbers of replicates for each dilution. Experiments in Fazekas de St. Groth (1982) and Strijbosch, Buurman, Does et al. (1987) indicate that, in practice, values of the order of 20 suffice. Nevertheless, more precise considerations can be given for the determination of the number of replicates in general. Assuming that $n_i=n$, $i=1,\dots,m$, a value of n can be derived from the requirement that the coefficient of variation $CV(\varphi)$ is not larger than a predetermined value. Thus the maximum value of $CV(\varphi)$ on $[\varphi_1, \varphi_2]$ is needed for the determination of n . From graphical analysis using various designs according to (2.1)-(2.4), it seems that as a natural consequence of the design method, for $\varphi_1 \leq \varphi \leq \varphi_2$, the quantity $CV(\varphi)$ attains its maximum at φ_1 or φ_2 provided that $n_i=n$, $i=1,\dots,m$ (exceptions to this situation only occur for very extreme designs like the one obtained for $\varphi_1=0.001$, $\varphi_2=0.500$, $d=2$, $P_1=0.01$ and $P_2=0.99$ (resulting in $m=4$, $c=21.406$ and

$x_1=0.097$)). Furthermore, $CV(\varphi_1)$ and $CV(\varphi_2)$ are approximately equal when using the P_1^* and P_2^* alternatives according to (3.3). A simple method to determine the (equal) number of replicates is therefore given by formula

$$(3.4) \quad n = 1 + \text{ENTIER} \left\{ \left[\text{CoV}^2 \sum_{i=1}^m \varphi_0^2 x_i^2 / (e^{\varphi_0 x_i} - 1) \right]^{-1} \right\},$$

where CoV denotes the user-supplied value of the coefficient of variation, and $\varphi_0 = \varphi_1$ when $CV(\varphi_1) \geq CV(\varphi_2)$, and $\varphi_0 = \varphi_2$ otherwise.

3.5 A program for the design of dilution assays

The setup of a design according to the suggestions in Chapter 2 and this Chapter can be simplified by appropriate software. In Strijbosch, Does and Buurman (1988) two programs (DESIGN and EVALUATE) are described that can be used on a common microcomputer with operating system MS-DOS or PC-DOS. These programs are both applicable to LDA and SDA without restriction. The program DESIGN is useful in the laboratory prior to the final determination of the setup for a particular experiment and will be described in more detail. The program gives the researcher the opportunity to set up different designs, to improve a design according to several suggestions, to make a graph of a design (see Fig. 2.2) and to evaluate the results of artificial data obtained from a random experiment corresponding with the chosen design, until he is convinced of having a suitable design for a particular experiment. The main options of this program are described next. DESIGN is completely menu-driven and has three different menu's. In the Main Menu the user can choose from the options D (go to the Design Menu), I (make an illustration), S (go to the Simulation Menu), H (ask for help) and Q (Quit the program).

In the Design Menu, the five design parameters $\varphi_1, \varphi_2, P_1, P_2$, and d , which are initialized as in Fig. 3.1, may be altered. When the user is satisfied with the entered values, he can get the resulting design determined through (2.1)-(2.4) by pressing the R-key (Recalculate). When

$$(3.5) \quad m > d + \log(\varphi_2/\varphi_1)/\log(c) - 1,$$

DESIGN MENU	
Options	
M : Main menu.	
L : change Lower bound phi1.	
U : change Upper bound phi2.	
d : change d.	
1 : change P1.	
2 : change P2.	
H : Help.	
F : replace $\phi_1 \setminus 2$ by $\phi_1^* \setminus 2$.	
P : replace $P_1 \setminus 2$ by $P^* \setminus 2$.	
Choose :	

CURRENT VALUES
phi1 = 0.000500
phi2 = 0.005000
d = 3
P1 = 0.150
P2 = 0.700
Resulting Design
m = 7
x[1] = 97.811856
c = 1.745584

Alternatives
phi*1= 0.000393
phi*2= 0.006365
P*1 = 0.105
P*2 = 0.568

Fig. 3.1. The Design Menu in the program DESIGN. Current values of the design parameters can be adjusted by choosing one or more options.

values ϕ_1^* and ϕ_2^* exist which satisfy $\phi_1^* < \phi_1 < \phi_2 < \phi_2^*$ and $\phi_1^* \phi_2^* = \phi_1 \phi_2$ and result (with equal d, P_1 and P_2) in the same design (i.e. equal m, x_1 , and c). In that case the ultimate ϕ^* -values are displayed in a special box with the heading "Alternatives". Another way to exploit the situation in (3.5) is a reduction of the P-range which can improve the design. In the event of (3.5), (this will generally be the case), in addition to the ϕ^* -values, the P^* -values as can be obtained from (3.3) are displayed as alternatives in the "Alternatives"-box. When (3.5) holds with an equality sign (for example as a result of choosing the ultimate ϕ^* -values by pressing the F-key when (3.5) is true)), it is still possible that the optimal P^* -values as determined by (3.3) are not equal to the present P-values. In that case only these P^* -values are displayed as alternatives. In Fig. 3.1 a typical screen of the Design Menu is displayed.

Choosing the I-option in the Main Menu will cause the display of a screen in which the functions $f_i(\phi) = \exp(-\phi x_i)$, $i=1, \dots, m$, representing the

probabilities of a negative response and $g_i(\varphi) = P_i$, $i=1,2$, are drawn for $\varphi \in [\varphi_1, \varphi_2]$. This graph is mainly illustrative. It shows various properties of the current design (see Fig. 2.2). Exactly d values of $f_i(\varphi)$, $i=1, \dots, m$, are between P_1 and P_2 for each φ which is the main property of the design. Fig. 2.2 shows, for example, that a wider φ -range could be chosen which maintains the same design.

In the Simulation Menu the main options are T (change True φ), n (input n_i , $i=1, \dots, m$), C (determine $n=n_i$ on the basis of a user-supplied coefficient of variation CoV according to (3.4)) and S (Start simulation). The true value of φ is initialized as $(\varphi_1 + \varphi_2)/2$ and can be adjusted by pressing the T-key. The numbers of replicate cultures are initialized as $n_i=12$, $i=1, \dots, m$, and can be changed by pressing the n-key. The user-supplied coefficient of variation is altered by pressing the C-key. The present values of n_i , $i=1, \dots, m$, and CoV always correspond to each other according to (3.4). In other words, after CoV is changed, $n_i=n$, $i=1, \dots, m$ holds, where n is obtained from (3.4); while after altering n_i , $i=1, \dots, m$, CoV will be the maximum of $CV(\varphi_1)$ and $CV(\varphi_2)$ which are obtained from (3.1) and (3.2). After pressing S the results of the simulation (r_i , $i=1, \dots, m$, denoting the numbers of negatively responding cultures) are presented, as well as the values of three different estimators ($\hat{\varphi}_{ML}$, $\hat{\varphi}_{MC}$, and $\hat{\varphi}_{Je}$, see Chapters 5 and 6) with the corresponding standard errors and 95% confidence bounds. In addition χ^2 goodness-of-fit test statistics are calculated. If this statistic is significant at the 5% level, it is starred, indicating that the assumed single-hit Poisson model with the estimated value does not fit the data. When $r_i=n_i$, $i=1, \dots, m$, or $r_i=0$, $i=1, \dots, m$, one-sided confidence intervals are presented (see Section 5.5). The results of one simulation with the design of Fig. 3.1 and with the (true) φ and $n=n_i$, $i=1, \dots, m$, as stated in Fig. 3.2 are presented in Fig. 3.3.

The described program is used to determine the number of dilutions, the number of cells tested per culture and the number of replicate cultures. The demanded accuracy of the assay and the a priori information about the value of the frequency to be estimated determine the parameters mentioned above. The degree of difficulty in determining the frequency interval in which the unknown frequency will range will depend on the type of experiment.

SIMULATION MENU

Options

M : Main menu.

T : change True phi.

n : input $n[i]$, $i=1, \dots, m$.

C : base $n=n[i]$ on Coefficient of variation (CoV).

S : Start simulation.

H : Help.

Choose :

CURRENT VALUES

phi1 = 0.000500

phi = 0.004000

phi2 = 0.005000

CoV = 0.203404

$n[1] = 12$

$n[7] = 12$

Fig. 3.2. The Simulation Menu in the program DESIGN. Current values of φ , n_i , $i=1, \dots, m$, and CoV can be changed by choosing the appropriate option. Changing CoV will change n_i , $i=1, \dots, m$, and vice versa.

i	x[i]	n[i]	r[i]	r[i]/n[i]
1	97.81186	12	7	0.5833
2	170.73880	12	6	0.5000
3	298.03891	12	7	0.5833
4	520.25193	12	0	0.0000
5	908.14339	12	0	0.0000
6	1585.24051	12	0	0.0000
7	2767.17034	12	0	0.0000

CoV = 0.203404
True phi = 0.004000

Method	Estimate	St.Error	95% Confidence Interval		Goodn-of-fit	
			Lower	Upper	Gof-X2	DF
Max.Lik	0.004080	0.000764	0.002583	0.005578	7.108	6
Jackkni	0.003987	0.000704	0.002607	0.005368	6.957	6
Min.Chi	0.003645	0.000658	0.002355	0.004935	6.702	6

Press a key

Fig. 3.3. The results of one simulation with design parameters, true φ and n_i , $i=1, \dots, m$, as stated in the Figures 3.1 and 3.2.

In specific situations where either only a higher or a lower frequency may be expected, it can be more economical to set up two designs : one for the higher and one for the lower potential range. Especially when the lower bound of the higher range is much larger than the upper bound of the lower range, the choice of two designs may lead to a more economical use of resources.

The program EVALUATE uses the same statistical procedures as DESIGN. It reads files with experimental data on disc consisting of one or more data sets $\{(x_i, n_i, r_i)_{i=1, \dots, m}\}$ and calculates the three estimates with corresponding standard errors, 95% confidence bounds and goodness-of-fit statistics for each data set in the file. The results are written to a disc file.

4 DESIGN GENERALIZATIONS

4.1 Introduction

The quality of the experimental design is crucial for the accuracy of the results of the statistical analysis. This Chapter deals with the experimental design for quantal two-parameter biomedical models which describe the relation between a binary ('quantal') response and one independent variable (dose) in such a way that the dose-response relationship can be expressed as a linear regression model $y = \alpha + \beta x$ where x is a function of the dose z and y is a function of the probability $P(x)$ of a *positive* response. A well-known example of such a model is the logit regression model resulting from the logistic function $P(x) = 1/(1 + \exp(-(\alpha + \beta x)))$ where $x = \log z$ and y is the logit of $P(x)$ (see Ashton (1972) and Finney (1978, p. 359)). This model could be used to study the toxicity of a certain drug, or substance, e.g., the relationship between the dose of an insecticide and the mortality of a target insect species.

The one-parameter biomedical models ($\alpha = 0$) such as Limiting or Serial Dilution Assays [$P(x) = 1 - \exp(-\beta x)$ where $x = z$ and $y = -\log(1 - P(x))$] form a special case. Design formulae for this case have been proposed in Chapter 2 (Subsection 2.3.2). In Section 4.2 this method will be presented in a slightly different manner in order to fit within the context of this Chapter. An experimental design for the two-parameter model is more complicated, since the a priori knowledge on two different parameters has to be incorporated. Suggestions for designs are given by Brown (1966), Hsi (1969), Tsutakawa (1972, 1980) and Finney (1978, Chapter 19), among others. These methods do not provide formulae for design characteristics based on a priori estimates of both parameters. Here the design method for the one-parameter model as proposed in Subsection 2.3.2 is generalized for the two-parameter case. Three special cases are dealt with (Subsections 4.3.2 - 4.3.4). A design corresponding to an example taken from the literature is compared with the design which would have been obtained when using the generalized design formulae in this case (Subsection 4.3.5).

4.2 The one-parameter biological model $y = \beta x$

In Subsection 2.3.2 the design formulae have been expressed in the parameters of the single-hit Poisson model using φ as the unknown parameter, and P_1 and P_2 as bounds for a fraction of *negative* responders to be considered sufficiently informative. As in the context of this Chapter the linearized models will be used, the design formulae of Subsection 2.3.2 have to be reformulated. Furthermore, in the context of the two-parameter biological models, the sufficiently informative range can best be expressed as fractions of *positive* responders (cf. Section 4.1).

The following criteria have to be chosen by the experimenter :

- (i) a lower (β_1) and an upper (β_2) boundary for the unknown frequency β ($0 < \beta_1 \leq \beta \leq \beta_2$),
- (ii) a lower (P_1) and an upper (P_2) boundary for the fractions of positive replicates considered sufficiently informative ($0 < P_1 < P_2 < 1$),
- (iii) an integer number (d) of doses which will produce sufficiently informative data as specified under (ii), ($d \geq 1$).

Let $y_1 = -\log(1-P_1)$, and $y_2 = -\log(1-P_2)$, so that $0 < y_1 < y_2$. Formulae (2.1)-(2.4) are now transformed to

$$(4.1) \quad m = d + \text{ENTIER}(\log(\beta_2/\beta_1)/\log(c)),$$

$$(4.2) \quad x_1 = \{y_1 y_2 c^{1-m}/(\beta_1 \beta_2)\}^{1/2},$$

$$(4.3) \quad c = \{y_2/y_1\}^{1/d},$$

$$(4.4) \quad x_i = x_1 c^{i-1}, \quad i = 2, \dots, m.$$

As is shown in Subsection 2.3.3, the resulting assay design has the property that for each β in the interval $[\beta_1, \beta_2]$ exactly d values of $x_1, \dots, x_m$ satisfy the following condition :

$$(4.5) \quad y_1 < \beta x_i \leq y_2, \quad i=1, \dots, m.$$

4.3 The two-parameter biological model $y = \alpha + \beta x$

4.3.1 Introduction

After a linearizing transformation many dose-response relationships can be expressed as a linear regression model with two unknown parameters : $y = \alpha + \beta x$ where $x = \log z$, $z = \text{dose}$, and $y = f(P(x))$, $P(x) = \text{probability of a positive response}$ (see Ashton (1972) and Finney(1978)). For example the logistic model

$$(4.6) \quad P(x) = (1 + \exp(-(\alpha + \beta x)))^{-1}$$

can be written as

$$(4.7) \quad y = \log(P(x)/(1-P(x))) = \alpha + \beta x.$$

As an example of the application of the logistic model, consider the study of the relationship between the dose of an insecticide and the mortality of a target insect species. A dose x (actually the logarithm of the true dose) of an insecticide is given to each of n insects in a group. In order to formulate the model it is useful to assume that each insect possesses its own tolerance against a particular insecticide and dies when the dose level exceeds the tolerance. Suppose that the tolerance T_j of the j -th insect is an independent drawing from a logistic distribution identical for all insects. Then $P(T_j < x) = (1 + \exp(-(\alpha + \beta x)))^{-1}$.

In analogy to the design method for the one-parameter case (see Section 4.2), design methods for the two-parameter case can be developed for the situations that a priori the following is available : either a point estimate for α and an interval estimate for β (Subsection 4.3.2), or a point estimate for β and an interval estimate for α (Subsection 4.3.3), or interval estimates for both α and β (Subsection 4.3.4).

4.3.2 The case $\alpha = \alpha_0$ and $\beta \in [\beta_1, \beta_2]$

Suppose that α_0 is a point estimate for α and that $[\beta_1, \beta_2]$ is an interval estimate for β . Just as in Section 4.2, assume that $0 < \beta_1 \leq \beta_2$. With respect to α_0 , the more interesting, and at the same time more complicated, case where $\alpha_0 < y_1 < y_2$ will be investigated. The interval $[y_1, y_2]$ is based on the interval $[P_1, P_2]$. The P_1 and P_2 mark a sufficiently informative range for fractions of positive responders. The design formulae (4.1)-(4.4) can then be applied directly :

$$(4.8) \quad m = d + \text{ENTIER}(\log(\beta_2/\beta_1)/\log(c)),$$

$$(4.9) \quad x_1 = \left[(y_1 - \alpha_0)(y_2 - \alpha_0)c^{1-m}/(\beta_1\beta_2) \right]^{1/2}.$$

$$(4.10) \quad c = \left[\frac{y_2 - \alpha_0}{y_1 - \alpha_0} \right]^{1/d},$$

$$(4.11) \quad x_i = x_1 c^{i-1}, \quad i=2, \dots, m.$$

The property of the resulting design is that for each $\beta \in [\beta_1, \beta_2]$ exactly d values of $x_1, \dots, x_m$ satisfy the condition $y_1 < \alpha_0 + \beta x_i \leq y_2$. Note that for the doses determined by this design we may write :

$$(4.12) \quad z_i = z_{i-1}^c,$$

which means that the original doses should be chosen equidistantly on a loglog-scale. This is remarkable, since in most literature (see Brown (1966), Tsutakawa (1972,1980), Finney (1978), Hamilton (1979,1980)) doses, equidistant on the log-scale, are recommended or used in Monte Carlo experiments. Brown (1966) proposed a procedure for using available information to design a quantal assay of the potency of a test preparation relative to a standard preparation. The probabilities of response to the two preparations for varying log-doses were considered well-approximated by cumulative normal distributions with common standard deviation σ and means denoted by μ_T and μ_S , respectively. For the potency (ρ) of the test

preparation relative to the standard preparation the relation $\log(\rho) = \mu_S - \mu_T$ holds. Brown (1966) assumed the prior information available to be represented as lower and upper bounds for the parameters σ , μ_S and $\log(\rho)$. His design proposal was based on this prior information and included suggestions for the extreme log-doses of both standard and test preparation and for the spacing of the log-doses which he assumed to be equal. Finney (1978) advocated symmetric designs in general since they are usually easier to effectuate and analyze statistically. Tsutakawa (1980) presented experimental designs for estimating an extreme percentage point of a logistic distribution when the location and scale parameters are unknown. His method has been based on a prior distribution of the parameters and a predicted value of posterior variance. The numerical results in his paper are restricted primarily to designs with equally spaced levels. As (4.8)-(4.12) show, doses equidistant on the log-scale are inappropriate if a good a priori estimate for α is available and if one requires a design which has the above-mentioned property, namely, for each $\beta \in [\beta_1, \beta_2]$ exactly d values of $x_1, \dots, x_m$ satisfy the condition $y_1 < \alpha_0 + \beta x_i \leq y_2$.

Finally, in the remaining case $y_1 \leq \alpha_0 \leq y_2$, it is easy to see that $m=d$, while the x_i 's, $i=1, \dots, d$, merely need to satisfy : $y_1 < \alpha_0 + \beta_2 x_i \leq y_2$, $i=1, \dots, d$.

4.3.3 The case $\alpha \in [\alpha_1, \alpha_2]$ and $\beta = \beta_0$

Suppose that $[\alpha_1, \alpha_2]$ is an interval estimate for α and that β_0 is a point estimate for β . As before, suppose that $[y_1, y_2]$ marks a sufficiently informative range. Then the following design formulae can be deduced from (4.1)-(4.4) :

$$(4.13) \quad m = d + \text{ENTIER}((\alpha_2 - \alpha_1)/e),$$

$$(4.14) \quad x_1 = (y_1 + y_2 - \alpha_1 - \alpha_2 + e(1 - m))/(2\beta_0),$$

$$(4.15) \quad e = (y_2 - y_1)/d,$$

$$(4.16) \quad x_i = x_1 + (i-1)e/\beta_0, \quad i=2, \dots, m.$$

This follows from the transformations $y^* = e^y$, $\beta^* = e^\alpha$, $x^* = e^{\beta_0 x}$ and applying (4.1)-(4.4) to the model $y^* = \beta^* x^*$. The property of the resulting design is that for each $\alpha \in [\alpha_1, \alpha_2]$ exactly d values of $x_1, \dots, x_m$ satisfy the condition $y_1 < \alpha + \beta_0 x_1 \leq y_2$. Note that doses are equidistant on the log-scale.

4.3.4 The case $\alpha \in [\alpha_1, \alpha_2]$ and $\beta \in [\beta_1, \beta_2]$

Fig. 4.1 illustrates the construction of the design formulae in this Subsection. Suppose that $[\alpha_1, \alpha_2]$ and $[\beta_1, \beta_2]$ are interval estimates for α and β respectively such that $\alpha_1 < \alpha_2 < y_1$, $0 < \beta_1 < \beta_2$ and $\beta_1/\beta_2 \leq (y_2 - \alpha_2)/(y_2 - \alpha_1)$.

Fig. 4.1. Illustration of the conditions for the construction of the design method in Subsection 4.3.4.

This configuration corresponds to the situation which is of primary interest and will be investigated analogously to Subsection 4.3.2. The last condition means that the slope (β^0) of the line through the points $(0, \alpha_2)$ and $((y_2 - \alpha_1)/\beta_2, y_2)$ is greater than or equal to β_1 . Then, combining the results of the two preceding Subsections, a method can be devised. Firstly, the approximate number of doses is determined. Applying (4.8)-(4.11) with

$\alpha_0 = \alpha_1$ yields m_1 and c_1 . In order to cover all remaining combinations of α and β , it is sufficient to extend the design applying (4.8)-(4.11) once more on $y = \alpha_2 + \beta x$ with $\beta \in [\beta_1^0, \beta_2]$ where $\beta_1^0 = \beta_2(y_2 - \alpha_2)/(y_2 - \alpha_1)$. This leads to values for m_2 and c_2 . Since the m_1 and m_2 doses overlap an interval corresponding with d doses, $m = m_1 + m_2 - d$ doses are needed for a design with the property that at least d values satisfy $y_1 < \alpha + \beta x_i \leq y_2$ (for any $\alpha \in [\alpha_1, \alpha_2]$ and $\beta \in [\beta_1, \beta_2]$). It is not possible to construct a design with the property that exactly d values satisfy this condition. Consider, for example, a configuration which satisfies $(y_1 - \alpha_2)/\beta_1 = (y_1 - \alpha_1)/\beta_2$. Then, assuming that we have exactly d values which satisfy the condition for (α_1, β) , $\beta \in [\beta_1, \beta_2]$, it is, as is illustrated in Fig. 4.2, easy to see that the parameters can be chosen such that (many) more than d values satisfy the condition for (α_2, β_1) .

It follows that

$$(4.17) \quad m = d + \text{ENTIER}(\log(\beta_2/\beta_1)/\log(c_1)) + \\ \text{ENTIER}(\log((y_2 - \alpha_1)/(y_2 - \alpha_2))/\log(c_2)),$$

where

$$(4.18) \quad c_1 = \{(y_2 - \alpha_1)/(y_1 - \alpha_1)\}^{1/d},$$

and

$$(4.19) \quad c_2 = \{(y_2 - \alpha_2)/(y_1 - \alpha_2)\}^{1/d}.$$

Obvious choices for x_1 and x_m are

$$(4.20) \quad x_1 = (y_2 - \alpha_2)c_2^{1-d}/\beta_2,$$

and

$$(4.21) \quad x_m = (y_1 - \alpha_1)c_1^{d-1}/\beta_1.$$

Now, it follows from the results of the Subsections 4.3.2 and 4.3.3, that the choice of x_i , $i=2, \dots, m-1$, should depend on the ratio of the lengths of the intervals $[\alpha_1, \alpha_2]$ and $[\beta_1, \beta_2]$. When the first is large and the second is small, equidistant values for x_i , $i=1, \dots, m$, are advisable, whereas in the

Fig. 4.2. A configuration which shows that it is impossible to construct a design with the property that for each $(\alpha, \beta) \in [\alpha_1, \alpha_2] * [\beta_1, \beta_2]$ exactly d values of $x_1, \dots, x_m$ satisfy the condition $y_1 < \alpha + \beta x_i \leq y_2$.

Fig. 4.3. The motivation of the choice for the weights of translation and rotation in the construction of the design formulae (4.22)-(4.24).

opposite case a geometrical series is preferable. Combining these ideas, the following design formulae are proposed :

$$(4.22) \quad x_i = \left\{ T \left[x_1 + (x_m - x_1) (i-1) / (m-1) \right] + R x_1 (x_m / x_1)^{(i-1) / (m-1)} \right\} / (T+R),$$

$$i=2, \dots, m-1,$$

where

$$(4.23) \quad T = (\alpha_2 - \alpha_1) / (y_1 + y_2 - \alpha_1 - \alpha_2),$$

and

$$(4.24) \quad R = (\beta_2 - \beta_1) / (\beta_1 + \beta_2).$$

The choice of the weights T (for translation) and R (for rotation) is illustrated in Fig. 4.3 and can be motivated as follows.

The design formulae of Subsection 4.3.2 ($\alpha_1 = \alpha_2 = \alpha_0$) could be found by studying the effect of rotation of the line $y = \alpha_0 + \beta x$ around the point $(0, \alpha_0)$, on the log-doses. Accordingly, the design formulae of Subsection 4.3.3 ($\beta_1 = \beta_2 = \beta_0$) could be found by studying the effect of translation of the line $y = \alpha + \beta_0 x$ from $y = \alpha_1 + \beta_0 x$ to $y = \alpha_2 + \beta_0 x$ on the log-doses. These effects must be quantified somehow in order to determine the weights R and T. Since the effect of rotation is different for y_1 and y_2 , all effects are considered for $\tilde{y} = (y_1 + y_2) / 2$ only. Furthermore, the effects of rotation for $\alpha = \alpha_1$ and $\alpha = \alpha_2$, respectively, are accumulated, as well as are the effects of translation for $\beta = \beta_1$ and $\beta = \beta_2$, respectively. Thus, the effect of rotation can be expressed as $(x_3 - x_1) + (x_4 - x_2)$, whereas the effect of translation is $(x_4 - x_3) + (x_2 - x_1)$. Since $x_1 = (\tilde{y} - \alpha_2) / \beta_2$, $x_2 = (\tilde{y} - \alpha_1) / \beta_2$, $x_3 = (\tilde{y} - \alpha_2) / \beta_1$ and $x_4 = (\tilde{y} - \alpha_1) / \beta_1$, the quotient of R and T is :

$$(4.25) \quad \frac{R}{T} = \frac{(\beta_2 - \beta_1) (y_1 + y_2 - \alpha_1 - \alpha_2)}{(\alpha_2 - \alpha_1) (\beta_1 + \beta_2)}.$$

Hence, (4.23) and (4.24) represent possible choices for R and T.

For this design we have found that for each $(\alpha, \beta) \in [\alpha_1, \alpha_2] * [\beta_1, \beta_2]$ the number of values $x_1, \dots, x_m$ which satisfy the condition $y_1 < \alpha + \beta x_1 \leq y_2$ is approximately d. However, in the vicinity of (α_1, β_1) and (α_2, β_2) this number

Fig. 4.4a : $\alpha=-8$.

Fig. 4.4b : $\beta=5$.

Fig. 4.4c : $\beta=0.5$.

Fig. 4.4d : $\alpha=-2$.

Fig. 4.4. Graph of the functions $f_i(\alpha, \beta) = (1 - \exp(-(\alpha + \beta x_i)))^{-1}$ at each extreme of either parameter interval. The x_i 's, $i=1, \dots, m$, follow from formulae (4.17)-(4.24) with design parameters $\alpha_1=-8$, $\alpha_2=-2$, $\beta_1=0.5$, $\beta_2=5$, $P_1=0.2$, $P_2=0.8$ and $d=3$ resulting in $m=23$.

is often $d-1$, while near (α_2, β_1) it can be much larger than d . If not both conditions $\alpha_1 < \alpha_2 < y_1$ and $\beta_1/\beta_2 \leq (y_2 - \alpha_2)/(y_2 - \alpha_1)$ are satisfied, alternatives for the formulae (4.17)-(4.24) are possible applying the same concepts. In the next Subsection a design taken from the literature will be compared with the design obtained when applying one of the design methods proposed in Section 4.3.

Analogous to Fig. 2.2 for the one-parameter design, Figures 4.4a-4.4d illustrate the properties of the two-parameter design according to the formulae (4.17)-(4.24). The design parameters used are $\alpha_1 = -8$, $\alpha_2 = -2$, $\beta_1 = 0.5$, $\beta_2 = 5$, $P_1 = 0.2$, $P_2 = 0.8$ and $d = 3$ resulting in $m = 23$. The functions $f_i(\alpha, \beta) = (1 + \exp(-\alpha + \beta x_i))^{-1}$, $i = 1, \dots, m$, are drawn for $\alpha = -8$ and $\beta \in [0.5, 5]$ (Fig. 4.4a), $\alpha \in [-8, -2]$ and $\beta = 5$ (Fig. 4.4b), $\alpha \in [-8, -2]$ and $\beta = 0.5$ (Fig. 4.4c), $\alpha = -2$ and $\beta \in [0.5, 5]$ (Fig. 4.4d).

4.3.5 Applications of the proposed designs

In this Subsection it will be investigated which parameter choices in the design methods will approximately reproduce the design used in an example taken from the literature, thus throwing some light on the circumstances under which such designs are suitable. The example illustrates the design method of Subsection 4.3.4.

Berkson

As an alternative to the ML method for the estimation of the parameters α and β , Berkson (1944) proposed the criterium "Minimum Logit x^2 ". This modified definition of x^2 only applies to the logistic function and enables minimization to be accomplished in one stage without iteration. In order to illustrate his method Berkson (1944) presented in his Table I an experiment with doses 40, 60, 80, 100, 120, 140, 160, 180, 200, 250, 300. From the recorded experiment, assuming the logistic model (4.6), estimates $\hat{\alpha} = -10.33$ and $\hat{\beta} = 5.66$ are obtained. Since he uses the transformation $x = \log_{10}(z)$, $\hat{\beta} = 5.66/\log(10) = 2.46$ must be used here. Applying (4.17)-(4.24) with $\alpha_1 = -11$, $\alpha_2 = -10$, $\beta_1 = 2$, $\beta_2 = 3$, $d = 4$, $P_1 = 0.2$ and $P_2 = 0.8$, doses 22, 27, 34, 43, 55, 72, 95, 126, 172, 238, 335 result. Note that none of the design methods discussed in this Chapter could produce equidistant doses.

Alternative choices for x_1 and x_m

When it is important that for all considered parameter values α and β , the proportions of positive responders in an experiment vary from P_1 to P_2 , better choices for x_1 and x_m (instead of (4.20) and (4.21), respectively) are $x_1 = (y_1 - \alpha_2)/\beta_2$ and $x_m = (y_2 - \alpha_1)/\beta_1$ (see Fig. 4.1). This is the case, for example when the Spearman-Kärber method is applied to estimate the ED_{50} (median effective dose) which refers to the log-dose that on the average produces a response in 50% of the subjects. The corresponding statistic (see Spearman (1908) and Kärber (1931)) is given by

$$(4.26) \quad \sum_{i=1}^m (p_{i+1} - p_i)(x_i + x_{i+1})/2,$$

where p_i denotes the proportion of positively responding subjects at x_i , and thus $(p_{i+1} - p_i)$ is an estimator of the proportion of subjects whose tolerances lie between x_i and x_{i+1} . Conditions for the application of the Spearman-Kärber method to estimate ED_{50} are that $p_1=0$, $p_m=1$, the intervals between the x_i , $i=1, \dots, m$, are not too wide, and the underlying tolerance distribution is symmetric. When for some i the quantity $(p_{i+1} - p_i)$ is negative (due to sampling variation), groups should be combined in order to clear the data from such irregularities prior to the application of the formula. However, not much is known about the small sample properties of the Spearman-Kärber statistic when applied to experimental data as obtained from a design with unequally spaced log-doses with respect to other commonly used statistics such as the ML statistic.

Estimating the ED_γ

In estimating the ED_γ where γ is close to 0 or 1, it is considered important (see Wetherill (1963)) to locate the doses around the ED_γ . This can easily be achieved by choosing P_1 and P_2 such that $P_1 < \gamma < P_2$ and applying the most appropriate design method of this Chapter.

5 CONVENTIONAL ESTIMATION METHODS

5.1 Introduction

In this Chapter two conventional statistical procedures, i.e. the Maximum Likelihood (ML) method and the Minimum Chi-square (MC) method, for the evaluation of the experimental data in dilution analysis will be described. Most of the literature published on dilution series has been devoted to methods for estimating the unknown parameter. Recent references to this literature are Finney (1978), Salama, Koch and Tolley (1978), Taswell (1981), Loyer (1981), Fazekas de St. Groth (1982), Loyer and Hamilton (1984) and Taswell (1987) among others.

Finney (1947, 1951 and 1978, p. 430) proposed a ML estimation procedure that solves the ML equation with iterative approximation based on a log-log transformation (Mather (1949)). From this method the same estimated value of the unknown parameter is obtained as from the straightforward ML method presented in Section 5.3. However, different confidence bounds are found, since Finney's method produces an estimator of $\text{var}(\log(\hat{\varphi}))$, while the straightforward ML method produces an estimator of $\text{var}(\hat{\varphi})$.

Taswell (1981) compared four statistical procedures for the estimation of the unknown parameter in dilution series with Monte Carlo simulation. These methods are the Least Squares (LS) method, the Weighted Mean (WM) method, the ML method and the MC method. A formula for the LS estimate of φ can be obtained by applying the standard unweighted least squares regression procedure on $\log(p) = -\varphi x$ where $p=r/n$ and r is the number of negative responders. As noted by Taswell (1981) and Fazekas de St. Groth (1982) there are several reasons why this formerly used method should not be applied in dilution analysis. In the first place, the usual assumptions of the linear regression model are not satisfied. Another reason is that the use of $p_i = r_i/n_i$ without incorporating n_i in the statistical procedure results in a loss of information. Lefkovits and Waldmann (1984) did not see objections in using the LS method instead of the ML or MC method arguing that "in most biological experiments we are not trying to find some physical constants whose definition has to be absolutely precise". However, in full agreement

with Taswell (1987), the fact that biological experiments are not physical experiments with accurate adjustments of instruments but rather represent roughly approximate assessments of biological activities does not justify indifference with respect to the choice of the estimation method to be used. With an inaccurate statistical method as the LS method experimental results can generally only become worse. Some problems of the LS method are resolved by the WM method which is obtained by weighting the individual estimates $\hat{p}_i = -\log(p_i)/x_i$ with terms related to their respective variances and taking their mean. However, this method is also statistically inappropriate. Taswell (1981) found that the four estimators can be ranked in the following order by decreasing bias and mean square error : LS, WM, ML and MC. There is no doubt that the last two methods are preferable to the other two. However, regarding the comparison of the ML and MC method, the meaning of the conclusions of Taswell (1981) cannot be fully understood, since it is not entirely clear which experimental designs he used in the simulations (see Chapter 7). Obviously, the properties of statistical methods are dependent on the experimental design used.

Loyer (1981) compared the expected values and exact mean square errors of the ML estimator with those of some alternative estimators which will be discussed briefly. An estimator ($\hat{p}_C$) described by Thomas (1955) simply adjusts the ML estimator with the correction factor $\exp(-0.805/n)$. Since the bias of the ML estimator is undoubtedly dependent on both n and the rest of the design parameters (see Chapter 2), use of this estimator in only special designs would be appropriate. Fisher (1922) described an estimator ($\hat{p}_F$) which is based on the total number of negative responders ($\sum r_i$). Fisher showed his estimator to be 87.7% efficient, consequently only computational convenience could justify general use of this method. Thomas (1942) described an estimator ($\hat{p}_T$) which is the geometric mean of the estimators obtained by using the approximations $\exp(-px) \approx (1-px)$ and $\exp(-px) \approx 1/(1+px)$ in the derivative of the log-likelihood function. Finally, Johnson and Brown (1961) described an estimator ($\hat{p}_J$) which, like Fisher's estimator, uses only the total number of negative responders and is also 87.7% efficient. For similar reasons, as advanced before, the last two estimators should not generally be used when adequate computer facilities are available. Loyer (1981) used the following (small) designs for the comparison of the statistical methods described : $m=3$, $n=3,5$ and 10 , $c=2,4$

and 10, and $x_2=0.01$. As he stated, these nine experimental designs should cover virtually all $m=3$ serial dilution settings found in the literature and should be appropriate for $p \approx 160$. The comparisons have been made for $\varphi=10(10)300$. The designs with $c=4$ and $c=10$ would have been obtained with (2.1)-(2.4) using $\varphi_1=10$, $\varphi_2=300$, $d=1$, $P_1=0.3344$, $P_2=0.7604$, and $\varphi_1=10$, $\varphi_2=300$, $d=2$, $P_1=0.0042$, $P_2=0.9467$, respectively. The design with $c=2$ could not be obtained with (2.1)-(2.4) since $d=3-\text{ENTIER}(\log(30)/\log(2))$ is less than 1. Loyer (1981) finds that no single estimator performs best across all designs. He states that, in general, $\hat{\varphi}_T$ performed best at "high" dilution factors ($c=10$), $\hat{\varphi}_J$ performed best at "intermediate" dilution factors ($c=4$) and $\hat{\varphi}_C$ performed best at "low" dilution factors ($c=2$). As the considered estimators seem to be most useful for the very small designs mentioned just above, they are not considered in this Chapter. The main interest here pertains to LDA where designs are usually larger.

Fazekas de St. Groth (1982) advanced several arguments for the use of ML methods instead of MC methods. He proposed the ML estimation procedure as set out in detail in Finney (1978).

Loyer and Hamilton (1984) paid attention to interval estimation procedures commonly applied to small (standard) serial dilution experiments and based on exact probabilities. Furthermore they offer an alternative method for the calculation of two-sided confidence intervals.

Taswell (1987) repeats the conclusions of his 1981 paper. Criticism (see Fazekas de St. Groth (1982)) and other reactions (see Miller (1982) and Lefkovits and Waldmann (1984)) are discussed and suggestions are made for the Monte Carlo comparison of estimation methods in dilution analysis (see Chapter 7).

After introducing in Section 5.2 the notation to be used, Section 5.3 is devoted to the ML method and Section 5.4. to the MC method. Finally, in Section 5.5, some remarks on (one-sided) confidence intervals are given.

5.2 Notation

First some notation will be (re)introduced. The design parameters in dilution experiments are denoted by m (i.e. the number of groups of replicate cultures), x_i (i.e. the (mean) number of cells tested in replicate culture group i) and n_i (i.e. the number of replicate cultures for group i), $i=1, \dots, m$. Furthermore let φ denote the unknown frequency. The biometrical model can be formulated as follows :

$$(5.1) \quad \begin{array}{cccc} Y_{11} & Y_{12} & \cdots & Y_{1n_1} \\ Y_{21} & Y_{22} & \cdots & Y_{2n_2} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ Y_{m1} & Y_{m2} & \cdots & Y_{mn_m} \end{array}$$

where the Y_{ij} are independent Bernoulli-distributed random variables with $P(Y_{ij} = 0) = 1 - P(Y_{ij} = 1) = \exp(-\varphi x_i)$, $i=1, \dots, m$, $j=1, \dots, n_i$ (cf. Does, Strijbosch and Buurman (1988)). A negative response for a replicate culture is thus denoted by zero. From grounds of sufficiency it follows that the relevant observations from a limiting dilution assay consist of the

independent binomial random variables R_i which are defined by $R_i = \sum_{j=1}^{n_i} (1 - Y_{ij})$, i.e. $R_i \sim \text{Bin}(n_i, \exp(-\varphi x_i))$, $i=1, \dots, m$. The vector $(R_1, \dots, R_m)$ will be denoted by $\underline{R}$. Furthermore let $n_i - R_i$ be denoted by Q_i and $(Q_1, \dots, Q_m)$ by $\underline{Q}$.

5.3 The maximum likelihood (ML) method

5.3.1 Introduction

One of the most important and most used methods in dilution analysis is the ML estimation procedure. The corresponding estimator is often called the "most probable number" as introduced by McCrady (1915). The method of ML together with other principles of estimation, is presented with the required rigor in texts on statistical theory (e.g. Kendall and Stuart (1973)). Some of its main properties are asymptotic sufficiency, efficiency and

consistency. It is clear that this method in dilution series is advocated by many authors such as Mather (1949), Peto (1953), Fazekas de St. Groth (1982) and Buurman, Daemen, Groenewegen et al. (1983). One of the shortcomings of the ML method is that it leads to biased estimators. The positive bias of the ML estimator is well known by researchers in the field (see Eisenhart and Wilson (1943), Thomas and Woodward (1955), McCarthy, Thomas and Delaney (1958)). Special attention has been given in the literature to the bias of the ML estimator in small designs like the design with small n and $m=3$ which is commonly used in serial dilution assays : Johnson and Brown (1961) gave mathematical approximations expressing the bias in the serial dilution problem as being related to the number of replicates; Salama, Koch and Tolley (1978) calculated exact biases of about 10% for $m=3$ decimal solutions ($c=10$) with $n=10$ replicates per culture group.

5.3.2 The ML estimation procedure

The ML estimator $\hat{\varphi}_{ML}$ is determined as the value of φ that maximizes

$$(5.2) \quad \log L(\varphi) = \sum_{i=1}^m \sum_{j=1}^{n_i} \{ -(1-Y_{ij})\varphi x_i + Y_{ij} \log(1-\exp(-\varphi x_i)) \}$$

or

$$(5.3) \quad \log L(\varphi) = \sum_{i=1}^m \{ -R_i \varphi x_i + Q_i \log(1-\exp(-\varphi x_i)) \},$$

where $L(\varphi)$ denotes the relevant part of the likelihood function. The value of φ that maximizes $\log L(\varphi)$ is usually calculated by Newton's method of iterative approximation (see e.g. Daellenbach, George and McNickle (1983), p. 576) :

$$(5.4) \quad \hat{\varphi}_{k+1} = \hat{\varphi}_k - \frac{\partial \log L(\varphi) / \partial \varphi}{\partial^2 \log L(\varphi) / \partial \varphi^2} \bigg|_{\hat{\varphi}_k}.$$

where

$$(5.5) \quad \partial \log L(\varphi) / \partial \varphi = \sum_{i=1}^m \{ -R_i x_i + Q_i x_i \exp(-\varphi x_i) / (1 - \exp(-\varphi x_i)) \},$$

and

$$(5.6) \quad \partial^2 \log L(\varphi) / \partial \varphi^2 = \sum_{i=1}^m \{ -Q_i x_i^2 \exp(-\varphi x_i) / (1 - \exp(-\varphi x_i))^2 \}.$$

The formula for the Fisher information (3.1) is found as follows. Let $p_i = \exp(-\varphi x_i)$, $i=1, \dots, m$, and $\log L(\varphi) = l(\varphi) = \sum_{i=1}^m l_i(\varphi)$. Moreover, let ∂ denote the partial derivation with respect to φ . Then $l_i(\varphi) = R_i \log(p_i) + (n_i - R_i) \log(1 - p_i)$ and $l'_i(\varphi) = (R_i - n_i p_i) p'_i / \{p_i(1 - p_i)\}$. Thus $I(\varphi) = \text{var}_{\varphi}(\sum_{i=1}^m l'_i(\varphi)) = \sum_{i=1}^m \text{var}_{\varphi}(l'_i(\varphi)) = \sum_{i=1}^m n_i p_i (1 - p_i) (p'_i)^2 / \{p_i^2 (1 - p_i)^2\} = \sum_{i=1}^m n_i x_i^2 / (\exp(\varphi x_i) - 1)$.

A serious problem might seem to arise with the ML estimation procedure. If all R_i are 0, which event occurs with an (extremely) small but undeniably positive probability, then $\hat{\varphi}_{ML} = \infty$, and hence $E \hat{\varphi}_{ML} = \infty$, thus leading to an infinite, rather than an asymptotically negligible, bias. This undesirable phenomenon can be eliminated by the following simple modification which will be tacitly assumed hereafter: if all R_i indeed happen to be 0, then modify $\hat{\varphi}_{ML}$ by replacing an arbitrary Y_{ij} , say one corresponding with the smallest x_i , by 0. In Subsection 5.3.3 it is demonstrated by way of example for the case $m=1$ that this simple modification suffices to reduce the bias from infinity to the desired order n_1^{-1} .

Evaluated at $\hat{\varphi}_{ML}$, formula (5.6) provides an approximation of $-I(\varphi)$ where $I(\varphi)$ denotes the Fisher information (see Section 3.2). An estimator of the variance of $\hat{\varphi}_{ML}$ is given by

$$(5.7) \quad S^2(\hat{\varphi}_{ML}) = \frac{-1}{\partial^2 \log L(\varphi) / \partial \varphi^2} \Big|_{\hat{\varphi}_{ML}}.$$

5.3.3 Reducing the bias of the ML estimator from infinity to the desired order

In this Subsection it will be shown for the case $m=1$ that the modified ML estimator (see Subsection 5.3.2) has a finite bias. Let R be a binomial distributed variable with parameters n and p where $p = \exp(-\varphi)$ with $0 < p < 1$. Clearly, for φ the ML estimator $\hat{\varphi} = -\log \hat{p} = -\log(R/n)$ will have $E \hat{\varphi} = \infty$. However, let $X = \max(1, R)$ and set $\tilde{\varphi} = -\log(X/n)$. Then $E(\tilde{\varphi} - \varphi) = O(n^{-1})$ as desired.

To prove this statement observe that

$$(5.8) \quad E(\tilde{\varphi} - \varphi) = -E\{\log(X/n) - \log p\} = -E\{\log(1+Z)\} \\ = -E\{Z - Z^2(1+\Theta Z)^{-2}/2\},$$

for some $\Theta = \Theta_Z \in [0,1]$, with $Z = (X-np)/np$. It is immediate that $EZ = (np)^{-1}P(R=0) = O(n^{-1})$. To deal with the second term on the right-hand side of (5.8), Schwarz' inequality and the fact that $E(X-np)^4 = O(n^2)$ can be applied to show that this term is $O(n^{-1}\{E(1+\Theta Z)^{-4}\}^{1/2})$. According to a well-known inequality for binomial probabilities (see e.g. Okamoto (1958)), $P(A) \leq \exp(-np^2/2)$ where A is the event that $X \leq np/2$. Now the maximum of $(1+\Theta Z)^{-4}$ on A is reached for $X = 1$ and equals $(np)^4$, uniformly in $\Theta = \Theta_Z$. Hence $E(1+\Theta Z)^{-4}I_A = O((np)^4 \exp(-np^2/2)) = o(1)$. On A^c , clearly $(1+\Theta Z)^{-4}$ is bound and hence $E(1+\Theta Z)^{-4} = O(1)$, which implies that the second term on the right-hand side of (5.8) is $O(n^{-1})$ as well.

5.4 The minimum chi-square (MC) method

5.4.1 Introduction

Large sample theory shows that the difference between ML estimators and MC estimators tends to disappear when the number of observations is increased. Furthermore, ML and MC estimators have some important properties in common such as consistency, asymptotic normality of distribution and asymptotic efficiency (see Rao (1957)). The use of MC estimators as compared to ML estimators in bioassays has been a matter of discussion (see Berkson (1980), Fazekas de St. Groth (1982), Mantel (1985)). Berkson (1980) states that the use of MC estimators is preferable in cases where it is not certain that the ML estimation procedure is best. However, in the discussion of that paper this point of view is criticized. A serious objection to the MC method when applied in dilution assays is that the number of expected frequencies is necessarily small for some culture groups, since the experimental design is constructed in order to obtain sufficiently informative data for a range of parameter values (see Chapter 2). This can lead to unstable values of χ^2 (cf. (5.9)). The amount of information about parameter values in dilution

assays can increase in different ways. Designs with a large number of dilutions (m) and small numbers of replicates (n) which are useful when a parameter has to be estimated if not much a priori information is available and a provisional estimate is required, or designs with small m and large n which are useful when much a priori information is available and an accurate estimate is required, can and will be constructed in dilution analysis. As Mantel (1985) remarked, the ML estimator is consistent no matter how information increases, whereas for the MC estimator to be consistent it is essential that the information increases by getting replicate numbers increasingly large. This property of MC estimators suggests that they are not generally applicable in dilution analysis like ML estimators are. However, little is known about bias and efficiency of ML and MC estimators when applied to data from dilution experiments with small or moderate size.

5.4.2 The MC estimation procedure

The MC estimator $\hat{\varphi}_{MC}$ is determined as the value of φ that minimizes

$$(5.9) \quad \chi^2(\varphi) = \sum_{i=1}^m \frac{(R_i - n_i \exp(-\varphi x_i))^2}{n_i \exp(-\varphi x_i) (1 - \exp(-\varphi x_i))}.$$

Since the expected value of the MC estimator is infinite, $\hat{\varphi}_{MC}$ has been adapted in the same manner as $\hat{\varphi}_{ML}$ has; see Subsection 5.3.2. Newton's method of iterative approximation can be employed to determine $\hat{\varphi}_{MC}$ in the same way as $\hat{\varphi}_{ML}$. Now

$$(5.10) \quad \hat{\varphi}_{k+1} = \hat{\varphi}_k - \frac{\partial \chi^2(\varphi) / \partial \varphi}{\partial^2 \chi^2(\varphi) / \partial \varphi^2} \bigg|_{\hat{\varphi}_k},$$

where

$$(5.11) \quad \partial \chi^2(\varphi) / \partial \varphi = \sum_{i=1}^m \frac{n_i x_i \exp(-\varphi x_i) (2R_i - n_i) + R_i^2 x_i (\exp(\varphi x_i) - 2)}{n_i (1 - \exp(-\varphi x_i))^2},$$

and

$$(5.12) \quad \partial^2_{\mathbf{x}}^2(\varphi)/\partial \varphi^2 = \sum_{i=1}^m \left[\frac{-n_i x_i^2 (\exp(-\varphi x_i) + \exp(-2\varphi x_i)) (2R_i - n_i)}{n_i (1 - \exp(-\varphi x_i))^3} + \frac{R_i^2 x_i^2 (\exp(\varphi x_i) - 3 + 4\exp(-\varphi x_i))}{n_i (1 - \exp(-\varphi x_i))^3} \right].$$

An estimator of the variance of $\hat{\varphi}_{MC}$ is given by

$$(5.13) \quad S^2(\hat{\varphi}_{MC}) = \frac{2}{\partial^2_{\mathbf{x}}^2(\varphi)/\partial \varphi^2} \Big|_{\hat{\varphi}_{MC}}.$$

5.5 Confidence intervals

As the ML and MC estimators are asymptotically normally distributed, two-sided 95% confidence intervals can be approximated by $\hat{\varphi} \pm 1.96s$ where $\hat{\varphi}$ is the point estimate and s the estimate of the standard deviation corresponding to the method used. The same procedure applies to the estimation methods described in Chapter 6 provided that s is available. In the simulation study (Chapter 7) which compares the estimation methods described in both Chapter 5 and Chapter 6, it will be investigated in which proportion the calculated confidence bounds contain the true value of the parameter.

When $r_i = n_i$, $i=1, \dots, m$, the point estimates from all methods equal 0 and estimates of the standard deviation cannot be determined. In that case a one-sided confidence interval $[0, \varphi_U]$ is more appropriate. With a confidence level of 95%, φ_U can be determined from $\prod_{i=1}^m \exp(-\varphi_U x_i n_i) = 0.05$, that is

$$(5.14) \quad \varphi_U = \log(20) / \sum_{i=1}^m n_i x_i.$$

When $r_i = 0$, $j=1, \dots, m$, a one-sided confidence interval $[\varphi_L, \infty]$ (for SDA) or $[\varphi_L, 1]$ (for LDA) is more appropriate. With a confidence level of 95%, φ_L can be determined from

$$(5.15) \quad \prod_{i=1}^m (1 - \exp(-\varphi_L x_i))^{n_i} = 0.05.$$

The value of φ_L cannot be expressed in closed form but must be determined from iterative approximation. Both one-sided confidence intervals coincide with the intervals which would have been obtained with Woodward's method based on exact probabilities (see Woodward (1957) and Loyer and Hamilton (1984)).

6 BIAS-REDUCING ESTIMATION METHODS

6.1 Introduction

In this Chapter a number of bias-reducing estimation methods related to the ML method for the evaluation of the experimental data in dilution analysis will be described : jackknife versions of the ML method, bootstrap versions of the ML method and methods to reduce the bias of the ML estimator devised by Salama, Koch and Tolley (1978). The remarks on confidence intervals given in Section 5.5 also apply to the methods described in this Chapter.

The jackknife method and the bootstrap method provide estimators of bias and variance for an extremely wide class of statistics. In this Chapter these methods are discussed and applied to dilution analysis. Although both application to the ML estimator and to the MC estimator could be performed, only application to the ML estimator will be elaborated on in this Chapter. The main reasons for this choice are the general objections as discussed in Subsection 5.4.1, the fact that a more developed theory for the ML method is available, and the restriction of available computer time for the inherently time-consuming Monte Carlo comparisons. An important part of the development and testing of the simulation program has been performed on a VAX 11/780 computer which is approximately 5 times slower than the VAX 8700 which has been used for the simulations described in Chapter 7. As an indication of the total CPU-time used it is noted that 62 hours of CPU-time were necessary for obtaining all the results of Chapter 7.

In most applications, jackknife and bootstrap estimators can be constructed in different ways. When applied to dilution analysis, three versions of jackknifed ML estimators are worthwhile to study (Section 6.2), whereas four bootstrapped ML estimators are interesting (Section 6.3). Using Taylor expansions and implicit function theorems, Salama, Koch and Tolley (1978) showed the existence of bias-reducing methods that are based on the ML method. These methods are described in Section 6.4.

6.2 The jackknife (three versions)

6.2.1 Introduction

The concept of the jackknife method was originally proposed by Quenouille (1949). The name "jackknife" is due to Tukey. Quenouille (1949) introduced a technique for reducing the bias of a serial correlation estimator based on splitting the sample into two parts. Quenouille (1956) explored the general applicability of this concept when splitting the sample into g parts of size h each. Miller (1974) gives an introduction and a historical survey. Here the "leaving-one-out" jackknife is considered where g equals the sample size and $h=1$.

Let $X_1, \dots, X_N$ be a sample of independent and identically distributed random variables. Suppose that one wishes to estimate a parameter Θ based on the sample of size N . Consider an estimator $T_N = T_N(X_1, \dots, X_N)$ of the parameter Θ . Let $T_{N-1}(X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_N)$ be the corresponding estimator based on the sample of size $(N-1)$ where observation k has been deleted. Then the k -th jackknifed pseudo-value is defined as

$$(6.1) \quad T_{Nk}^J = NT_N - (N-1)T_{N-1}(X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_N), \quad k=1, \dots, N.$$

The jackknife estimator T_N^J defined by

$$(6.2) \quad T_N^J = N^{-1} \sum_{k=1}^N T_{Nk}^J$$

in many cases reduces the bias (see Efron (1982)). The reason behind this is : if

$$(6.3) \quad E_{\Theta} T_N = \Theta + a_1 N^{-1} + a_2 N^{-2} + O(N^{-3}), \quad \text{as } N \rightarrow \infty,$$

for some constants a_1 and a_2 , then

$$(6.4) \quad E_{\Theta} T_N^J = \Theta + O(N^{-2}), \quad \text{as } N \rightarrow \infty.$$

Thus the jackknife procedure reduces the bias by eliminating the first order bias term. Tukey (1958) suggested that the jackknife method could also be

used to estimate the variance and obtain robust interval estimation. The jackknife estimator of variance is given by

$$(6.5) \quad S_{\theta}^2(T_N^J) = (N-1) \left[\sum_{k=1}^N (T_{Nk}^J)^2 - (\sum_{k=1}^N T_{Nk}^J)^2 / N \right] / N.$$

In the case of dilution series, the pseudo-values in (6.1) can be obtained in three different ways. When $n_i = n$, $i=1, \dots, m$, the biometrical model (5.1) is a matrix with columns that are independent, identically distributed (iid) random vectors. As jackknife estimators are in general determined from iid variables, the natural way to jackknife is to drop one column from (5.1) at a time (see Subsection 6.2.2). The two other versions are non-iid cases and are obtained by deleting either one row at a time or one element at a time (see Subsection 6.2.3).

6.2.2 The iid case

In general, jackknife estimators are determined from independent identically distributed (iid) random variables. In the case that the numbers of the replicates in all culture groups are equal, say n , we obtain such a jackknife method by leaving out the j -th replicate from *each* of the culture groups at a time. That is to say, in the biometrical model (5.1) the jackknife procedure is applied by leaving out one of the columns at a time. Thus, it is necessary that the individual elements of

$\{Y_{ij}\}_{i=1, \dots, m; j=1, \dots, n}$ are known. The columns of this matrix provide the $X_1, \dots, X_N$ variables in this case. This procedure produces the jackknife estimator $\hat{\varphi}_{JC}$. The bias-reducing effect of the jackknife in the iid case is a consequence of general theory of jackknife methods.

6.2.3 Two non-iid cases

Even if all $n_i = n$, the R_i will not be identically distributed, as the x_i differ. The jackknife version which has been suggested in Subsection 6.2.2 as the most natural one violates a certain symmetry inherent to the experiment, because one could permute the entries in any row in (5.1) as the Y_{ij} are iid within rows without altering the ML estimator of φ . However, this would alter the estimator $\hat{\varphi}_{JC}$ obtained by jackknifing the columns of

the matrix (5.1) which is a troubling lack of uniqueness. Another disadvantage of the procedure is that it can only be applied when there is an equal number of replicates in each culture group, i.e., $n_i = n$ for $i=1, \dots, m$.

Therefore two other jackknife versions are considered. One is based on jackknifing the rows of the matrix (5.1), i.e., the jackknife procedure is based on $R_1, \dots, R_m$ which provide the $X_1, \dots, X_N$ variables in that case. Thus the jackknife estimator $\hat{\varphi}_{J_r}$ is obtained. The other jackknife method is based on the elements of the matrix (5.1), i.e. the procedure is based on $Y_{11}, \dots, Y_{mn}$ providing the $X_1, \dots, X_N$ variables in that case. Thus the jackknife estimator $\hat{\varphi}_{J_e}$ is obtained. Resampling methods justifiable in the iid case may not work in more complex situations. In Subsection 6.2.4 it is indicated why for both methods the first order bias term of the ML estimator disappears in these non-iid cases as well (cf. Quenouille (1956)). Note that all three jackknife methods are asymptotic methods. The one based on the columns demands $n \rightarrow \infty$, the one based on the rows demands $m \rightarrow \infty$ and, finally, the one based on the elements in (5.1) demands that $\sum_{i=1}^m n_i \rightarrow \infty$.

For the determination of the 'element' jackknife estimator $\hat{\varphi}_{J_e}$ it is not necessary that the individual elements of $\{Y_{ij}\}_{i=1, \dots, m; j=1, \dots, n_i}$ be known; it is sufficient to know R . The following procedure is used. Let, for $i=1, \dots, m$, $\hat{\varphi}_{ML}(i-)$ be defined by the ML estimator obtained from the data $\{(x_1, n_1, R_1), \dots, (x_{i-1}, n_{i-1}, R_{i-1}), (x_i, n_i - 1, R_i - 1), (x_{i+1}, n_{i+1}, R_{i+1}), \dots, (x_m, n_m, R_m)\}$, and let $\hat{\varphi}_{ML}(i+)$ be the ML estimator obtained from the data $\{(x_1, n_1, R_1), \dots, (x_{i-1}, n_{i-1}, R_{i-1}), (x_i, n_i - 1, R_i), (x_{i+1}, n_{i+1}, R_{i+1}), \dots, (x_m, n_m, R_m)\}$. When $R_k = 0$, $k=1, \dots, m$, let $\hat{\varphi}_{ML}(i-)$ have an arbitrary value, and when $R_k = n_k$, $k=1, \dots, m$, let $\hat{\varphi}_{ML}(i+)$ have an arbitrary value. Then, $\hat{\varphi}_{J_e}$ and the variance estimator $S_{J_e}^2$ are calculated according to

$$(6.6) \quad \hat{\varphi}_{J_e} = N\hat{\varphi}_{ML} - (N-1)S_1/N,$$

and

$$(6.7) \quad S_{J_e}^2 = (N-1)(S_2 - S_1^2/N)/N,$$

respectively, where $N = \sum_{i=1}^m n_i$, and

$$S_k = \sum_{i=1}^m R_i (\hat{\varphi}_{ML}(i-))^{(k)} + \sum_{i=1}^m (n_i - R_i) (\hat{\varphi}_{ML}(i+))^{(k)}, \quad k=1,2.$$

6.2.4 Reducing the bias with jackknifing in the non-iid cases

Let $X_1, \dots, X_N$ be independent random variables with smooth (in Θ) density functions f_j depending on the same parameter Θ , $j=1, \dots, N$. Note that this is true in the situation of Subsection 6.2.3. Denote the first, second and third derivative of $\log f_j(\Theta)$ with respect to Θ by U_j , V_j and W_j , respectively. From the stochastic expansion of the ML estimator $\hat{\Theta}_N$ for Θ it can be shown (see Brillinger (1964), Chibisov (1973)) that

$$(6.8) \quad E_{\Theta}(\hat{\Theta}_N - \Theta) = \frac{\sum_{j=1}^N a_j}{(\sum_{j=1}^N b_j)^2} + O(N^{-3/2}), \quad \text{as } N \rightarrow \infty,$$

where

$$(6.9) \quad a_j = E_{\Theta} U_j V_j + \frac{1}{2} E_{\Theta} W_j \quad \text{and} \quad b_j = E_{\Theta} U_j^2, \quad j=1, \dots, N.$$

Note that $b_j > 0$ for $j=1, \dots, N$. Let $\hat{\Theta}_N^{(k)}$ denotes the ML estimator for Θ based on $X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_N$, then its expectation centered around Θ satisfies

$$(6.10) \quad E_{\Theta}(\hat{\Theta}_N^{(k)} - \Theta) \doteq \frac{\sum_{j \neq k} a_j}{(\sum_{j \neq k} b_j)^2},$$

where $\doteq$ denotes that higher order terms have been neglected. If $b_1, \dots, b_N$ satisfy the condition

$$(6.11) \quad \lim_{N \rightarrow \infty} \frac{\max_{1 \leq j \leq N} b_j}{\sum_{j=1}^N b_j} = 0,$$

then (6.10) can be approximated by

$$(6.12) \quad \frac{\sum_{j=1}^N a_j}{(\sum_{j=1}^N b_j)^2} \left(1 + \frac{2b_k}{\sum_{j=1}^N b_j}\right) - \frac{a_k}{(\sum_{j=1}^N b_j)^2}.$$

By neglecting higher order terms it follows from (6.10), (6.11) and (6.12) that the k-th jackknifed pseudo-value (see (6.1)) satisfies

$$(6.13) \quad E_{\Theta} T_{Nk}^J \doteq \Theta + \frac{\sum_{j=1}^N a_j}{(\sum_{j=1}^N b_j)^2} \left[1 - N \frac{2b_k}{\sum_{j=1}^N b_j} \right] + N \frac{a_k}{(\sum_{j=1}^N b_j)^2}$$

and hence it follows from (6.13) that the jackknife estimator T_N^J (see (6.2)) satisfies

$$(6.14) \quad E_{\Theta} T_N^J = E_{\Theta} (N^{-1} \sum_{k=1}^N T_{Nk}^J) \doteq \Theta$$

which shows that the first order bias term disappears.

6.3 The bootstrap (four versions)

6.3.1 Introduction

The bootstrap method has been introduced by Efron (1979). Let $X_1, \dots, X_N$ be a sample of independent and identically distributed random variables with common cumulative distribution function F . Suppose again that $T_N = T_N(X_1, \dots, X_N)$ is an estimator of the parameter Θ . Furthermore, let F_N be the empirical distribution function defined by $F_N(x) = \#\{i: X_i \leq x, 1 \leq i \leq N\}/N$ and write $E_G(T_N)$ for the expectation of the statistic T_N computed based upon a random sample of size N from G and similarly for $\text{Var}_G(T_N)$, then the bootstrap bias and variance estimators for T_N are given as $[E_{F_N}(T_N) - T_N]$ and $\text{Var}_{F_N}(T_N)$, respectively. These estimators must be evaluated by drawing a large number of random samples ("bootstrap samples") from F_N . The ordinary bootstrap approach consists of drawing M bootstrap samples $(X_1^k, \dots, X_N^k)_{k=1, \dots, M}$ of size N from F_N (see Efron (1982)). Such a sample is obtained by computer generated random sampling with replacement. Although there exist more elaborate bootstrap methods (see Davison, Hinkley and Schechtman (1986)) the simple bootstrap method will be used. Repeated resampling from F_N results in an approximation $\bar{T}_N^B$ for $E_{F_N}(T_N)$:

$$(6.15) \quad \bar{T}_N^B = M^{-1} \sum_{k=1}^M T_{Nk}^B,$$

where $T_{Nk}^B = T_N(X_1^k, \dots, X_N^k)$, $k=1, \dots, M$. This estimator has a bias in estimating T_N which approximates the bias of T_N in estimating Θ . Thus, $B_N(\Theta)$ is used as an approximation of the bootstrap estimator of bias $[E_{F_N}(T_N) - T_N]$ and defined by

$$(6.16) \quad B_N(\Theta) = \bar{T}_N^B - T_N.$$

This leads to the bias-corrected "bootstrap estimator" T_N^B for Θ

$$(6.17) \quad T_N^B = T_N - B_N(\Theta) = 2T_N - \bar{T}_N^B.$$

The bootstrap estimator of variance $\text{Var}_{F_N}(T_N)$ is approximated by

$$(6.18) \quad S_{\Theta}^2(T_N^B) = [\sum_{k=1}^M (T_{Nk}^B)^2 - (\sum_{k=1}^M T_{Nk}^B)^2 / M] / (M-1).$$

As discussed in Section 6.2, there are three different ways to define a set $(X_1, \dots, X_N)$ from the biometrical model (5.1), thus resulting in the bootstrap estimators $\hat{\varphi}_{Bc}$, $\hat{\varphi}_{Br}$ and $\hat{\varphi}_{Be}$. As before, the c-variant represents the iid-case and is only applicable when all $n_i = n$, $i=1, \dots, m$. A fourth bootstrap version denoted by $\hat{\varphi}_{Bx}$ will be discussed in Subsection 6.3.3.

6.3.2 The iid case

Most of the theoretical work regarding resampling methods like the jackknife and the bootstrap thus far has been for the iid case. There seems to be no general rule to choose between the two methods for bias reduction and variance estimation. However, since examples are known for which the application of the bootstrap in more complex situations leads to inappropriate estimators (see Wu (1986)), it is interesting to compare bootstrap methods with jackknife methods. In the iid case the bootstrap samples are taken from the columns of the matrix $\{Y_{ij}\}$ which provide the $X_1, \dots, X_N$ variables in this case. In other words, if we let $\{j_t\}_{t=1, \dots, n}$ be a random sample with replacement from $\{1, \dots, n\}$, then $\{(x_{i, n}, \sum_{t=1}^n (1 - Y_{ij_t}))\}_{i=1, \dots, m}$ contains the data for the estimator corresponding with an individual bootstrap sample. Thus the bootstrap estimator $\hat{\varphi}_{Bc}$ results.

6.3.3 Three non-iid cases

The bootstrap counterpart of the 'row' jackknife version is obtained by letting $X_1, \dots, X_N$ be the rows of $\{Y_{ij}\}$ producing the estimator $\hat{\varphi}_{Br}$. Let $\{i_t\}_{t=1, \dots, m}$ be a random sample with replacement from $\{1, \dots, m\}$. Then $\{(x_{i_t, n_{i_t}, r_{i_t}})\}_{t=1, \dots, m}$ contains the data for the estimator corresponding with an individual bootstrap sample. Note that, as a result of the resampling procedure, it is possible that one or more dilutions are not included, whereas some dilutions are included more than once in the individual bootstrap samples for this bootstrap version.

The 'element' bootstrap estimator is obtained by letting $X_1, \dots, X_N$ be the elements of $\{Y_{ij}\}$ producing the estimator $\hat{\varphi}_{Be}$. Let $\{(i_t, j_t)\}_{t=1, \dots, T}$ (where $T = \sum_{i=1}^m n_i$) be a random sample with replacement from $\{(i, j)\}_{i=1, \dots, m; j=1, \dots, n_i}$ and let f_{ij} be the frequency of the combination (i, j) in that sample. Then $\{(x_{i, \sum_{j=1}^{n_i} f_{ij}, \sum_{j=1}^{n_i} f_{ij}(1-Y_{ij})})\}_{i=1, \dots, m}$ contains the data for the estimator corresponding with an individual bootstrap sample. When $\sum_{j=1}^{n_i} f_{ij} = 0$ for some i , the corresponding data $(x_i, 0, 0)$ are excluded from the computations.

A fourth bootstrap version is constructed to avoid some of the undesirable properties of the previous two versions. When each bootstrap sample is produced by joining individual samples from each row of $\{Y_{ij}\}$, a bootstrap estimator results which will be denoted by $\hat{\varphi}_{Bx}$. Let $\{j_t(i)\}_{t=1, \dots, n_i}$ be a random sample with replacement from $\{1, \dots, n_i\}$, $i=1, \dots, m$. Then $\{(x_{i, n_i, \sum_{t=1}^{n_i} (1-Y_{ij_t(i)})})\}_{i=1, \dots, m}$ contains the data for the estimator corresponding with an individual bootstrap sample. If the resampling procedure for $\hat{\varphi}_{Be}$ would be performed under the condition that the samples must contain exactly n_i elements from row i , $i=1, \dots, m$, then $\hat{\varphi}_{Be}$ and $\hat{\varphi}_{Bx}$ coincide. A major disadvantage of the x-version is that, when $(R_i = 0 \vee R_i = n_i)$, $i=1, \dots, m$, all bootstrap samples are equal, so that an estimator of the variance cannot be evaluated.

6.4 Other bias-reducing methods

As the distribution of the random variable $\underline{R}$ is specified, it is possible to investigate the stochastic expansion of $\hat{\varphi}_{ML}$ and base a corrected ML estimator on the corresponding properties. Using Taylor expansions and implicit function theorems, Salama, Koch and Tolley (1978) showed the existence of functions $H(\underline{Q}, \hat{\varphi}_{ML})$ such that the estimator $\hat{\varphi}_S$ defined by

$$(6.19) \quad \hat{\varphi}_S = \hat{\varphi}_{ML} - H(\underline{Q}, \hat{\varphi}_{ML})$$

satisfies

$$(6.20) \quad E(\hat{\varphi}_S) = \varphi + \sum_{i=1}^m O(n_i^{-2}),$$

thus removing the first order bias term. The following two alternatives for the function H are given (producing the estimators $\hat{\varphi}_{S1}$ and $\hat{\varphi}_{S2}$):

$$(6.21) \quad G_1(\underline{Q}, \hat{\varphi}_{ML}) = \frac{1}{2} \sum_{i=1}^m \left(\frac{\partial^2 \hat{\varphi}_{ML}}{\partial Q_i^2} \right) n_i e^{-\hat{\varphi}_{ML} x_i} (1 - e^{-\hat{\varphi}_{ML} x_i})$$

and

$$(6.22) \quad G_2(\underline{Q}, \hat{\varphi}_{ML}) = \frac{1}{2} \sum_{i=1}^m \left(\frac{\partial^2 \hat{\varphi}_{ML}}{\partial Q_i^2} \right) \underline{u} n_i e^{-\hat{\varphi}_{ML} x_i} (1 - e^{-\hat{\varphi}_{ML} x_i}).$$

The second order derivatives in the right-hand side of (6.22) are evaluated at $\underline{u}$ defined by $\underline{u} = E(\underline{Q})$. Unfortunately, the formula given by Salama, Koch and Tolley (1978) for $\frac{\partial^2 \hat{\varphi}_{ML}}{\partial Q_k^2}$ shows typing errors. The correct formula is :

$$(6.23) \quad \frac{\partial^2 \hat{\varphi}_{ML}}{\partial Q_k^2} = \frac{F_k^2}{D^3} \left[\sum_{i=1}^m Q_i e^{-\hat{\varphi}_{ML} x_i} F_i^2(x_i + 2F_i e^{-\hat{\varphi}_{ML} x_i}) \right] \\ - 2 \frac{F_k^3}{D^2} e^{-\hat{\varphi}_{ML} x_k},$$

where F_i and D are defined by, respectively,

$$(6.24) \quad F_i = F_i(\hat{\varphi}_{ML}) = x_i / (1 - e^{-\hat{\varphi}_{ML} x_i})$$

and

$$(6.25) \quad D = D(\underline{Q}, \hat{\varphi}_{ML}) = \sum_{j=1}^m Q_j e^{-\hat{\varphi}_{ML} x_j} F_j^2.$$

(It is noted that $\partial \hat{\varphi}_{ML} / \partial Q_i = F_i / D$, as can be established from the equation $\partial \log L(\varphi) / \partial \varphi = 0$, see (5.5).) Salama, Koch and Tolley (1978) used a special modification of the ML estimator in case all Q_i equal n_i , $i=1, \dots, m$: $\hat{\varphi}_S$ is taken as the maximum over all realizations of the experiment for which at least one of the Q_i is less than n_i . This modification is essentially the same as the modification of the ML estimator described earlier in Subsection 5.3.2.

The paper of Salama, Koch and Tolley (1978) does not provide an estimator for the standard deviation of $\hat{\varphi}_S$ based on the same principles. This, of course, is a serious drawback of the proposed method, since corresponding confidence interval estimates cannot be determined.

7 A MONTE CARLO STUDY FOR THE COMPARISON OF THE ESTIMATORS

7.1 Introduction

Research on estimators in dilution analysis requires a design method that can be used as a frame of reference. The design method proposed in Subsection 2.3.2 seems to be an acceptable candidate for use in general dilution assays and in Monte Carlo comparisons. The advantages of this design method are : it incorporates the researcher's criteria, it has suitable properties, and it can easily be used in Monte Carlo experiments aimed at a meaningful comparison of estimation methods. As a design according to this method is based on an a priori interval estimate of the unknown parameter, a means of examining an estimator's performance over a range of parameter values is automatically provided.

In order to perform a simulation study the following design parameters have to be chosen : a lower (φ_1) and an upper (φ_2) bound for the unknown parameter φ ($0 < \varphi_1 < \varphi_2$), a lower (P_1) and an upper (P_2) bound for a fraction of negative responders in a culture group to be considered sufficiently informative ($0 < P_1 < P_2 < 1$), an integer number d of groups of replicate cultures which will produce sufficiently informative data as specified by P_1 and P_2 , ($d \geq 1$), and finally, the number of replicates n_i in group i . On the basis of the parameters chosen, the number m of replicate culture groups and the number x_i of cells/organisms per replicate in group $i=1, \dots, m$, can be calculated according to the formulae (2.1)-(2.4). Though very many different combinations of these design parameters are possible and meaningful, some acceptable limitations for a Monte Carlo study can be made. All simulations are performed with an equal number of replicates $n_i = n$, $i=1, \dots, m$. With different n_i , $i=1, \dots, m$, the 'column' versions of the jackknife and bootstrap method could not be executed. Moreover, when the other design parameters are chosen carefully, there is no reason to depart from this choice. As noted by Fazekas de St. Groth (1982), biological experiments (LDA's) indicate that values of the order of 20 suffice in practice. Of

course, such a statement cannot be made without specifying the other design parameters. In particular, $n=20$ appears to suffice when P_1 is not too close to zero (e.g. 0.15) and P_2 is not too close to one (e.g. 0.70) and when $d=2$ or $d=3$. However, restriction to relatively small values of n seems to be sensible for three reasons. In the first place, it turns out that differences between estimation methods are especially interesting for these values. Secondly, the Monte Carlo experiments use a great deal of CPU-time. And, thirdly, in SDA smaller values of n are commonly used (e.g. $n=3$, or $n=6$). As for the choice of φ_1 and φ_2 , it should first be noted that only the quotient φ_2/φ_1 is relevant and not the individual values of φ_1 and φ_2 . This is both a consequence of the experimental design (according to formulae (2.1)-(2.4)) and the biological model used in dilution analysis. After consultation of researchers in the field of LDA, a reasonable choice for this quotient appears to be 10. Thus, the following experimental designs are used in the simulations : $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $d=3$ (leading to $m=7$). For the case $d=2$, the simulation program has been executed with $n=6$, 12 and 18, and, for the case $d=3$, with $n=6$ and 12. Simulation results have been obtained for 19 equidistant values of φ within the interval $[\varphi_1, \varphi_2]$ starting with $\varphi=\varphi_1$ and ending with $\varphi=\varphi_2$. The number of generated samples for each combination of $\varphi_1, \varphi_2, P_1, P_2, d, n$ and φ was 1,000. As Efron and Tibshirani (1986) indicated that a number of bootstrap replications in the range 50 to 200 is quite adequate in most situations and taking into consideration the limitations on computer time, the number M of bootstrap samples was chosen to be 100.

Many criteria could be used to compare the estimation methods. However, restriction has been made to three criteria concerning the bias, the mean square error and the realized confidence level. According to Taswell (1987), "to demonstrate convincingly in a Monte Carlo experiment that an estimator is superior for a given set of assumed parameter values and assay-design constants, it would be necessary to prove that it has a lower coefficient of variation corresponding not only to its point estimates but also to its interval estimates", the last type of coefficient of variation being defined "by taking the mean of the individual ratios of each estimate's standard error to the estimate's value". However, the last type of coefficient of variation is not at all a valid criterium, since not the magnitude of the estimate of the standard deviation is relevant but, instead, its ability to

produce correct interval estimates. The three criteria used to compare the performance of the estimation methods are the mean relative bias (MRB) determined as the mean bias relative to the true value of φ , the coefficient of variation (CV) determined as the square root of the mean square error relative to the true value of φ , and the realized confidence level (RCL) determined as the proportion of times that the calculated 95% confidence interval includes the true value of φ . Thus, the criteria are calculated as follows :

$$(7.1) \quad \text{MRB} = \text{MB}/\varphi, \text{ where MB} = \sum_{t=1}^T (\hat{\varphi}_t - \varphi)/T,$$

$$(7.2) \quad \text{CV} = (\text{MSE})^{1/2}/\varphi, \text{ where MSE} = \sum_{t=1}^T (\hat{\varphi}_t - \varphi)^2/T,$$

$$(7.3) \quad \text{RCL} = \#\{t : \varphi \in [\hat{\varphi}_t - 1.960 s_t, \hat{\varphi}_t + 1.960 s_t], 1 \leq t \leq T\}/T.$$

For Monte Carlo replicate t ($1 \leq t \leq 1,000$), $\hat{\varphi}_t$ and s_t denote the point estimate and the estimate of the standard deviation, respectively. The number T of non-missing estimates is usually 1,000. However, and in particular for the smallest designs, it can occur that a jackknifed or bootstrapped ML-estimator does not produce a meaningful value (e.g. less than zero). The lowest observed value for T has been 996 for the jackknife 'row' version used with the $(d=2, n=6)$ -design.

7.2 The simulation program

A PASCAL-program has been written to perform the Monte Carlo experiments for the comparison of the statistical methods described in Chapters 5 and 6. The program uses the NAG (FORTRAN) subroutines G05DZF and G05DYF for the generation of the Bernoulli variables $\{Y_{ij}\}$ and the bootstrap samples, respectively. The structure of the program will be described for the case $d=2$. For each of the 19 values of φ , a matrix $\{y_{ij}\}$, $i=1, \dots, 4$; $j=1, \dots, 18$ is generated 1,000 times. These numbers are used twice : one time for $n=18$ and one time for either $n=6$ or $n=12$. This concession has been made in order to curtail the required CPU-time. Depending on the value of n , the numbers r_i are determined by $r_i = n - \sum_j y_{ij}$ where $\sum_j = \sum_{j=1}^6$, for $n=6$, $\sum_j = \sum_{j=7}^{18}$, for $n=12$ and $\sum_j = \sum_{j=1}^{18}$, for $n=18$. Thus, 1,000 datasets (x_i, r_i, n) ,

$i=1, \dots, 4$, result for each combination of φ and n . For each dataset the program calculates, if possible, the weighted-mean estimate φ_0 (cf. Taswell (1981)). When all $r_i=0$ or n , this estimate cannot be determined. In that case $\varphi_0=(\varphi_1+\varphi_2)/2$ has been taken. φ_0 served as an initial estimate for the iterative determination of $\hat{\varphi}_{ML}$. $\hat{\varphi}_{ML}$ served as an initial estimate for the determination of $\hat{\varphi}_{MC}$, the jackknifed pseudo-values and the individual bootstrap estimates.

Much care has been taken to test the program. For example, the results of Salama, Koch and Tolley (1978) which have been based on the exact distributions of $\hat{\varphi}_{S1}$ and $\hat{\varphi}_{S2}$ could be reproduced quite satisfactory in the simulations.

The complete program is presented in the Appendix. However, only the statistical methods, mentioned in Subsection 7.3.5 are included. Future simulation research on dilution analysis could take advantage of this program.

7.3 Results

7.3.1 Introduction

In order to obtain a clear impression of the properties of all estimation methods in the designs used, the results of the simulation study will be presented as follows. In Subsection 7.3.2 the three jackknife versions of the ML estimator will be compared with the ML estimator which leads to a particular version as the most favourable alternative. Analogously, in Subsections 7.3.3 and 7.3.4 one bootstrap version and one method of Salama, Koch and Tolley (1978) will be chosen, respectively. Only criteria (7.1) and (7.2) are used for these comparisons. Furthermore, only the simulation results of the smallest design ($m=4$, $n=6$) will be presented. For this design differences are most explicit, whereas the properties of the compared estimators for larger designs can be described very well without presenting the numerical results obtained. Finally, in Subsection 7.3.5 the ML estimator, the MC estimator, the chosen jackknife and bootstrap alternatives and the chosen method of Salama, Koch and Tolley (1978) will be compared. All obtained numerical results are then presented including the results on realized confidence levels (see formula (7.3)).

7.3.2 Comparison of the ML estimator with the jackknife versions

The uncorrected ML estimator exhibits a positive bias. Theoretically this is due to the fact that there is a positive probability that all dilutions will produce growth so that the ML estimate becomes infinite. In practice, when m and n_i , $i=1, \dots, m$, are sufficiently large and the experimental design is chosen carefully, this situation will not occur and Monte Carlo simulations reveal the positive bias in that case. Resampling plans such as the jackknife are designed to reduce the bias. The values of the MRB and the CV obtained by the simulation program for the ML estimator and the three jackknife versions of it (J_c , J_r , and J_e) when using the smallest design are presented graphically in the Figures 7.1a and 7.1b. The MRB of the ML estimator varies from 3 to about 12% in this design. Both the J_c and the J_e estimator reduce the bias significantly to values around zero, whereas the J_r estimator shows a different behaviour. This can be explained as follows. In the first place, it is intuitively clear that J_c and J_e are capable of reducing the bias more than J_r , since the higher order bias terms of the J_c , J_r and J_e estimators are $O(N^{-r})$, $r > 1$ where $N = n, m$ and $\sum_{i=1}^m n_i$, respectively. It is gratifying to find that the Monte Carlo experiments confirm this statement. Secondly, the strong ascending behaviour of both the MRB and the CV for the J_r estimator at the end of the ϕ -range can be clarified in the following way. For all three jackknife estimators several ML estimators (the jackknifed pseudo-values) have to be evaluated. The probability of having at least one pseudo-value for which the data contains all positive responses, and hence needs to be modified according to the method described in Subsection 5.3.2, is generally larger for the J_r estimator than for the other two jackknife estimators. Furthermore, the influence of such a modification on the jackknife estimator is generally greater for the J_r version than for the other jackknife versions. For example, when using the design of Fig. 7.1, the (maximum) probabilities of the event that all responses are positive when calculating a jackknifed pseudo-value are 0.270, 0.013 and 0.011 for the J_r , J_c and J_e estimators, respectively (these figures are 0.073, 0.00007 and 0.00006, respectively, for the design of Fig. 7.5). The results obtained with the other four designs used in the simulation study can be summarized as follows. As m and n increase, the three jackknife estimators coincide increasingly, until they

Fig. 7.1a. Mean Relative Bias of the ML, Jc, Jr and Je estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=6$.

Fig. 7.1b. Coefficient of Variation of the ML, Jc, Jr and Je estimators. Design as in Fig. 7.1a.

Fig. 7.2a. Mean Relative Bias of the ML, Bc, Br, Be and Bx estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=6$.

Fig. 7.2b. Coefficient of Variation of the ML, Bc, Br, Be and Bx estimators. Design as in Fig. 7.2a.

are practically undiscernible in the largest design while still reducing the MRB over the full φ -range. Moreover, the reduction of the CV becomes significant over the full φ -range as soon as the (maximum) probability of the event that all responses are positive when calculating the jackknifed pseudo-values vanishes.

As for the jackknife versions of the ML estimator, it can be concluded from theoretical considerations, the finite sample computations and the wider applicability of the Je estimator as compared to the Jc estimator, that the Je estimator obtained by deleting one element of the matrix $\{Y_{ij}\}$ at a time is the statistical method of choice. The next best method is the Jc estimator based on deleting one column of $\{Y_{ij}\}$ at a time.

7.3.3 Comparison of the ML estimator with the bootstrap versions

Figures 7.2a and 7.2b present the simulation results for the four bootstrap versions (Bc, Br, Be and Bx) of the ML estimator obtained with the smallest design. As can be expected from the theory, the c-, r- and e-alternatives of the jackknife and the bootstrap are very strongly related and the Bx estimator is comparable to the Bc and Be estimators. The same conclusions pertain to the larger designs. On the basis of these results, the Bc, Be and Bx estimators are to be preferred. However, the Bc estimator is disregarded because of its less wide applicability. As the Be and Bx estimators are very much alike, the first is preferred because of its correspondence to the Je estimator. It is worthwhile to note here the importance of the modification of the ML estimator as suggested in Subsection 5.3.2 for the bootstrap method in dilution analysis. The probability of obtaining one or more bootstrap samples with $R_i=0$, $i=1, \dots, m$, increases strongly as the design becomes smaller. Formulae (7.4)-(7.6) give these probabilities for the Br, Bc and Be estimators, respectively.

$$(7.4) \quad P_r = 1 - \left[1 - \left[\frac{1}{m} \sum_{i=1}^m (1 - \exp(-\varphi x_i))^n \right]^m \right]^M,$$

$$(7.5) \quad P_c = 1 - \left[1 - \prod_{i=1}^m (1 - \exp(-\varphi x_i))^n \right]^M,$$

$$(7.6) \quad P_e = 1 - \left[1 - \left\{ \frac{1}{m} \sum_{i=1}^m (1 - \exp(-\varphi x_i)) \right\}^{nm} \right]^M.$$

Table 7.1 shows these probabilities for the five different designs when $\varphi = \varphi_2$. Without the modification a bootstrap estimate would not be calculable in many more cases than when the corresponding jackknife estimate is not calculable.

n =	d = 2			d = 3	
	6	12	18	6	12
Br	0.99995	0.994	0.97	0.77	0.42
Bc	0.43	0.003	0.00002	0.002	0.0000000
Be	0.69	0.014	0.0002	0.012	0.0000000

Table 7.1. The tabulated values represent the probabilities of obtaining one or more bootstrap samples with all positive responses ($R_t = 0$, $t=1, \dots, m$) when evaluating the Br, Bc and Be estimators. The designs used are the designs of Fig. 7.4-7.8, characterized by the corresponding values of the design parameters d and n. The probabilities are evaluated for $\varphi = \varphi_2$.

Figures 7.2a and 7.2b bring to the attention that several curves are not distinguishable, as will be the case in subsequent Figures. This, obviously, does not complicate the interpretation of the properties of the corresponding estimators.

7.3.4 Comparison of the ML estimator with other bias-reducing estimators

Figures 7.3a and 7.3b illustrate the MRB and CV, respectively, for the bias-corrected estimators of Salama, Koch and Tolley (1978) in the smallest design. As can be expected from the formulae given in Section 6.4, there is practically no difference between the two methods. A slight preference for the S2 estimator exists, however, on the basis of the MRB.

Fig. 7.3a. Mean Relative Bias of the ML, S1 and S2 estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=6$.

Fig. 7.3b. Coefficient of Variation of the ML, S1 and S2 estimators. Design as in Fig. 7.3a.

7.3.5 Final comparison

In this Subsection the estimation methods which are chosen to be preferable in each category (see Subsections 7.3.2-7.3.4) are compared with the ML and the MC estimators for the five experimental designs mentioned in the Introduction (7.1). The MRB and the CV are presented in the Figures 7.4a-7.8b, and (except for the S2 estimator) the RCL (formula 7.3) resulting from a separate simulation run is compared in the Figures 7.9-7.13. Note that the vertical axes in the Figures are not fixed. When comparing the five estimators on the basis of the MRB and CV, it becomes clear that in general the Je, Be and S2 estimators should be preferred. An attractive property of the jackknife and the bootstrap is that they also produce the variance estimate used to determine proper confidence bounds for φ . A major disadvantage of the S2 estimation method is the lack of an estimator for the variance. Consequently, the S2 estimator is not suitable for general use. The difference between the jackknife and the bootstrap estimators which is mainly apparent for small designs can be explained as follows. The jackknifed pseudo-values for the Jc, Jr and Je estimators are based on $n(m-1)$, $n(m-1)$ and $nm-1$ observations, respectively, whereas the corresponding ML estimators for the individual bootstrap samples are based on nm observations. On the other hand, occurrence of the situation that the modification of the ML estimation procedure has to be applied is more likely for the bootstrap than for the jackknife. It is probable that the interaction between these two phenomena chiefly determines the observed differences between the jackknife and the bootstrap estimators in dilution analysis for small designs. From a computational point of view, it is obvious that the jackknife is more attractive than the bootstrap, certainly for the larger designs. However, in very small designs, the Be estimator should be preferred over the Je estimator, since the last estimator is characterized by the undesired property of a strongly increasing CV for values of φ near φ_2 . The designs which are commonly used for LDA are sufficiently large to justify the general advocacy of the Je and Be estimators for these assays.

Fig. 7.4a. Mean Relative Bias of the ML, Je, Be, S2 and MC estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=6$.

Fig. 7.4b. Coefficient of Variation of the ML, Je, Be, S2 and MC estimators. Design as in Fig. 7.4a.

Fig. 7.5a. Mean Relative Bias of the ML, Je, Be, S2 and MC estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=12$.

Fig. 7.5b. Coefficient of Variation of the ML, Je, Be, S2 and MC estimators. Design as in Fig. 7.5a.

Fig. 7.6a. Mean Relative Bias of the ML, Je, Be, S2 and MC estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=18$.

Fig. 7.6b. Coefficient of Variation of the ML, Je, Be, S2 and MC estimators. Design as in Fig. 7.6a.

Fig. 7.7a. Mean Relative Bias of the ML, Je, Be, S2 and MC estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=3$ (leading to $m=7$) and $n=6$.

Fig. 7.7b. Coefficient of Variation of the ML, Je, Be, S2 and MC estimators. Design as in Fig. 7.7a.

Fig. 7.8a. Mean Relative Bias of the ML, Je, Be, S2 and MC estimators. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=3$ (leading to $m=7$) and $n=12$.

Fig. 7.8b. Coefficient of Variation of the ML, Je, Be, S2 and MC estimators. Design as in Fig. 7.8a.

The simulation study of Taswell

With regard to the MC estimator, the results of Taswell (1981) will be discussed. He performed a study on the ML and MC estimators with (quoting) "Monte Carlo experiments for a wide range of values φ ($\varphi=1$ to $\varphi=10^{-6}$) and a wide variety of symmetric and asymmetric assay designs with $m=2$ to $m=5$ and $\sum_{i=1}^m n_i=48$ to 192 (symmetric and asymmetric assay designs are designs in which the doses x_i are spaced evenly or unevenly, respectively, about a dose x_i corresponding to $P_i=0.5$)". Furthermore he stated : "The efficiency of each estimator was observed to be independent of the value of φ but dependent upon the values n_i and x_i of the assay design. χ^2 minimization was clearly found to be the minimum-bias minimum-variance estimator of φ as measured by the mean square error of the estimates". The fact that the value of φ used did not influence the simulation results of Taswell (1981) is not surprising and can be understood easily from the properties of the biological model (cf. Section 7.1). On the basis of the simulation results in this Chapter it is less obvious that MC should be preferred over ML. What is clear, however, is that the properties of the MC estimator with respect to the ML estimator strongly depend on which design is used and for which values of φ the simulations are performed. Since Taswell (1981) studied only designs for which the x_i 's, $i=1, \dots, m$, are spaced evenly or unevenly about $-(\log 0.5)/\varphi$ there is an obvious explanation for his finding that depending upon the assay design ML has 3 to 15% greater bias and MSE than MC. Referring to Figures 7.4a-7.8b it is noted that MC is always better than ML for some interval in the middle of the range $[\varphi_1, \varphi_2]$, this interval narrowing when n increases. In the paper by Mantel (1985) it is argued that the MC estimation procedure minimizes the ratio of (5.9) in two ways : reducing the numerator and increasing the denominator leading to a bias toward 50% with respect to the binomial parameters involved. Thus, fitting by MC will tend to overestimate binomial parameters under 50%, but underestimate those over 50%. In LDA, this leads to overestimating φ for values smaller and underestimating φ for values larger than a certain value φ_i . The described phenomenon suggests that simulating the way Taswell (1981) did favours MC. A small Monte Carlo experiment has been performed to confirm the phenomenon. Attempting to construct a design which could be like one of the designs Taswell (1981) used, the following design parameters have been

chosen for this experiment : $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.30$, $P_2=0.70$ and $d=2$, resulting in $m=5$. Note that the middle of the informative range is 0.5. Three values of the true parameter value has been used : φ_1 , φ_2 and $\sqrt{\varphi_1\varphi_2}=0.0032$ which corresponds to the value Taswell would have used in this situation. The number of replicates are 24, 48 and 96, respectively. Table 7.2 gives the resulting Mean Relative Bias and Coefficient of Variation according to (4.1)-(4.2).

---SIMULATION---PARAMETERS---						-----MRB-----		-----CV-----	
P_1	P_2	d	m	n	φ	ML	MC	ML	MC
0.30	0.70	2	5	6	0.0010	0.026	0.156	0.391	0.409
					0.0032	0.047	0.018	0.300	0.264
					0.0100	0.080	-0.014	0.340	0.297
0.30	0.70	2	5	12	0.0010	0.025	0.105	0.274	0.291
					0.0032	0.019	0.003	0.206	0.190
					0.0100	0.036	-0.029	0.231	0.218
0.30	0.70	2	5	24	0.0010	0.012	0.054	0.197	0.204
					0.0032	0.019	0.011	0.143	0.137
					0.0100	0.022	-0.020	0.154	0.151
0.30	0.70	2	5	48	0.0010	0.007	0.029	0.133	0.136
					0.0032	0.007	0.003	0.101	0.099
					0.0100	0.008	-0.019	0.101	0.104

Table 7.2. Illustration of the phenomenon that the MC estimation procedure overestimates φ if $\varphi < \sqrt{\varphi_1\varphi_2} (=0.0032)$ and underestimates φ in the opposite case when using a design in which the values x_i , $i=1, \dots, m$, are spaced evenly on a logarithmic scale about $x = -(\log 0.5) / \sqrt{\varphi_1\varphi_2}$ where $\varphi_1=0.001$ and $\varphi_2=0.01$.

Finally, as concerns criterium (7.3), Figures 7.9-7.13 and Table 7.3 make clear that for almost all experimental situations the MC estimator of variance produces too small 95% confidence intervals. Moreover, these data clearly demonstrate that the coefficient of variation defined on the basis of the standard deviation (see Section 7.1) when used as a criterium for comparing estimation methods would wrongly favour the MC estimator.

Fig. 7.9. Realized Confidence Levels for the ML, Je, Be and MC methods. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=6$.

Fig. 7.10. Realized Confidence Levels for the ML, Je, Be and MC methods. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=12$.

Fig. 7.11. Realized Confidence Levels for the ML, Je, Be and MC methods. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=2$ (leading to $m=4$) and $n=18$.

Fig. 7.12. Realized Confidence Levels for the ML, Je, Be and MC methods. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=3$ (leading to $m=7$) and $n=6$.

Fig. 7.13. Realized Confidence Levels for the ML, Je, Be and MC methods. The design parameters are $\varphi_1=0.001$, $\varphi_2=0.01$, $P_1=0.15$, $P_2=0.70$, $d=3$ (leading to $m=7$) and $n=12$.

The ML estimator of variance produces the best confidence intervals. The Je and Be estimators of variance produce on the average 95% confidence intervals which are slightly too small. However, using percentile points of the corresponding student distribution instead of the standard normal distribution when calculating the 95% confidence bounds could correct the Je and Be confidence intervals satisfactorily.

n =	d = 2			d = 3	
	6	12	18	6	12
ML	0.9506	0.9500	0.9514	0.9510	0.9514
Je	0.9317	0.9447	0.9470	0.9404	0.9469
Be	0.9382	0.9466	0.9486	0.9444	0.9474
MC	0.9246	0.9235	0.9267	0.8981	0.9087

Table 7.3. The tabulated figures represent the mean values of RCL for the estimation methods and designs compared.

APPENDIX

A PASCAL program for the simulation of limiting dilution experiments and the comparison of estimation methods.

```

PROGRAM MonteCarloSimulation(input,output);
{=====}
CONST
  tol      = 0.00005; { stop criterium in NewtonStep }
  MaxIt    = 30;      { maximal number of NewtonSteps in ML and MC }
  maxm     = 10;      { maximal number of doses }
  maxn     = 18;      { maximal number of replicates }
  stat     = 6;       { number of statistical methods :
    1 : weighted mean (WM), used as starting
        value for the determination of the
    2 : maximum likelihood method (ML);
    3 : 'element' jackknife version of ML (Je);
    4 : 'element' bootstrap version of ML (Be);
    5 : method of Salama, Koch and Tolley (1978) (S2);
    6 : minimum chi-square (MC). }
  Vn       = 3;       { number of different values for n }
  Vphi     = 19;      { number of different values for the true phi }
  Phi1     = 0.001;   { lower bound for phi }
  Phi2     = 0.010;   { upper bound for phi }
  P1       = 0.15;    { lower bound for informative range }
  P2       = 0.70;    { upper bound for informative range }
  d        = 2;       { number of sufficiently informative doses }

  NumSim   = 1000;    { number of Monte Carlo simulations }
  NumBoot  = 100;     { number of Bootstrap samples }
  Seed     = 123456;  { seed for random number generator }
{=====}
TYPE
  Arrm_nI  = Array[1..maxm,1..maxn]      of integer;
  ArrmD    = Array[1..maxm]              of double;
  ArrmI     = Array[1..maxm]              of integer;
  ArrstatI  = Array[1..stat]              of integer;
  ArrstatB  = Array[1..stat]              of boolean;
  ArrstatD  = Array[1..stat]              of double;
  ArrVphiD  = Array[1..Vphi]              of double;
  ArrStatVnVphiI = Array[1..stat,1..Vn,1..Vphi] of integer;
  ArrStatVnVphiD = Array[1..stat,1..Vn,1..Vphi] of double;
  ArrVnI    = Array[1..Vn]                of integer;
{=====}
VAR
  n,ni,m,nm,i,j,k,del,sim,boot,NonMissing,
  f,j_start,j_end,row,mrc                : integer;
  Phi_0,Tphi,MeanPhiEst,Pneg,Phi,PhiML,SdML,
  PhiEstk,sum1,sum2,Sd,SdEstk,c,pnegS_FSi_m2,
  npqi,DS2,d1d12i_2                       : double;
  Y                                          : Arrm_nI;
  x,xrc,pnegS,FS,FS2,d2d22,npos_E,pnegS_FS2,

```

```

p_F_etc                                : ArrmD;
r,nv,rrc,nrc,zero                      : ArrmI;
NumOk,NumOkSd                          : ArrStatVnVphiI;
ok                                      : ArrStatB;
succes,AllNeg,AllPos,AllNeg0,AllPos0,n_k_is_0 : boolean;
PhiEst,SdEst                           : ArrStatD;
MeanBias,Mse,ConfLevel                 : ArrStatVnVphiD;
Phiv                                    : ArrVphiD;
nVn,jb                                 : ArrVnI;
{=====}
PROCEDURE GO5CBF( i : integer);
{sets NAG procedure GO5CAF to a repeatable initial state}
external;
{=====}
FUNCTION GO5DZF( p : double) : boolean;
{returns TRUE with probability p}
external;
{=====}
FUNCTION GO5DYF( m,n : integer) : integer;
{returns integer, uniform between m and n}
external;
{=====}
FUNCTION WM(      r      : ArrmI;
               n,m      : integer;
               x        : ArrmD;
               var PhiWM : double) : boolean;
{
  This function determines the Weighted Mean estimate of phi (PhiWM),
  provided that (r[i]<>0 and r[i]<>n) for some i. WM=false when
  (r[i]=0 or r[i]=n) for all i.
}
VAR
  SumFW,SumW,F,W,Xi : double;
  Ri,i               : integer;
BEGIN { of WM }
  SumFW := 0;
  SumW := 0;
  for i := 1 to m do
    begin
      Ri := r[i];
      if (Ri > 0) and (Ri < n) then
        begin
          Xi := x[i];
          F := (ln(n)-ln(Ri))/Xi;
          W := Ri*Xi*Xi/(1-Ri/n);
          SumFW := SumFW + F*W;
          SumW := SumW + W
        end
      end;
  end;
  if (SumW <> 0) then
    begin
      PhiWM := SumFW/SumW;
      WM := true;
    end
  end

```

```

    else WM := false
END; { of WM }
{=====}
PROCEDURE NewtonStep( var old,new    : double;
                      alfa,beta  : double;
                      var ready,ok  : boolean);
VAR rest : double;
BEGIN { of NewtonStep }
  if (beta <> 0) then
    begin
      new := old-alfa/beta;
      if (new < 0) then
        old := Phi1/100
      else
        begin
          rest := abs(old-new);
          if (rest <= tol*new) then
            ready := true
          else
            old := new
          end
        end
      end
    else
      ok := false;
END; { of NewtonStep }
{=====}
FUNCTION ML(      r,nv      : ArrmI;
                x          : ArrmD;
                m,n        : integer;
                Phi_0       : double;
                var PhiML,SdML : double;
                ij          : integer;
                var AllNeg,AllPos : boolean) : boolean;
VAR
  ok,ready          : boolean;
  old,new,Alfa,Beta,e_ox,d1,d2,Rest,d1gd2,Xi : double;
  i,j,Ri,it         : integer;
BEGIN { of ML }
  ok := true;
  old := Phi_0;
  if (ij<>0) then { adapt input to obtain pseudo-estimate
                  for jackknife estimator }
    begin
      i := abs(ij);
      if (ij<0) then { delete a zero from the data in row i }
        r[i] := r[i]-1;
        nv[i] := nv[i]-1;
      end;

      it := 1;
      ready := false;

      AllNeg := (r[1]=nv[1]);
      i := 2;

```



```
while AllNeg and (i<=m) do
  begin
    AllNeg := (r[i]=nv[i]);
    i := i+1;
  end;
if AllNeg then
  begin
    new := 0;
    ready := true;
  end
else
  begin
    AllPos := (r[1]=0);
    i := 2;
    while AllPos and (i<=m) do
      begin
        AllPos := (r[i]=0);
        i := i+1;
      end;
    if AllPos then
      r[1] := 1; { adaptation necessary to reduce the bias of ML from
                  infinity to the desired order }
    end;
  end;
while ((not ready) and (it<=MaxIt)) do
  begin
    alfa := 0.0;
    beta := 0.0;
    for i := 1 to m do
      begin
        Ri := r[i];
        Xi := x[i];
        e_ox := exp(-old * Xi);
        d1 := (nv[i]-Ri) * Xi * e_ox;
        d2 := 1-e_ox;
        d1gd2 := d1/d2;
        alfa := alfa - Ri * Xi + d1gd2;
        beta := beta - Xi * d1gd2 / d2;
      end;
    NewtonStep(old,new,alfa,beta,ready,ok);
    it := it+1;
  end;
if (new < 0) or (not ready) then
  ok := false;
if ok then
  begin
    PhiML := new;
    if (not AllNeg) and (ij=0) then
      SdML := sqrt(-1/beta);
    end;
  end;
ML := ok
END; { of ML }
```

```

=====
FUNCTION MC(      r      : ArrmI;
                  n,m     : integer;
                  x       : ArrmD;
                  Phi_0   : double;
      var PhiMC,SdMC     : double;
                  AllPos  : boolean) : boolean;
VAR
  ok,ready        : boolean;
  i,j,Ri,it       : integer;
  Xi,old,new,alfa,beta,eox,e_ox,
  e_ox2,c1,c2,h1,h2,rest : double;
BEGIN { of MC }
  if AllPos then
    r[1] := 1; { adaptation necessary to reduce the bias of MC from
                infinity to the desired order }
  ok := true;
  old := Phi_0;
  it := 1;
  ready := false;
  while ((not ready) and (it <= MaxIt)) do
    begin
      alfa := 0.0;
      beta := 0.0;
      for i := 1 to m do
        begin
          Ri := r[i];
          Xi := x[i];
          eox := exp(old * Xi);
          e_ox := exp(-old * Xi);
          e_ox2 := exp(-2 * old * Xi);
          c1 := 1-e_ox;
          c2 := c1 * c1;
          h1 := n * Xi * (2 * Ri-n);
          h2 := Ri * Ri * Xi;
          alfa := alfa + (h1 * e_ox + h2 * (eox - 2))/(n * c2);
          beta := beta + (-Xi * h1 * (e_ox + e_ox2) +
                          Xi * h2 * (eox - 3 + 4 * e_ox))/(n * c2 * c1);
        end;
      NewtonStep(old,new,alfa,beta,ready,ok);
      it := it+1
    end; { of while }

    if (new < 0) or (not ready) then
      ok := false;

    if ok then
      begin
        PhiMC := new;
        SdMC := sqrt(2/beta);
      end;
      MC := ok
    END; { of MC }

```

```

=====
BEGIN { of MAIN PROGRAM }

nVn[1] := 6; { first number of replicates }
nVn[2] := 12; { second  ,,  ,,  }
nVn[3] := 18; { third   ,,  ,,  }
{
  For the nVn[i] replicates, columns jb[i]..(jb[i]+nVn[i]-1)
  from the matrix Y are used.
}
jb[1] := 1;
jb[2] := 7;
jb[3] := 1;

for i := 1 to Vphi do { Vphi different 'true' parameter-values }
  Phiv[i] := Phi1 + (i-1)*((Phi2-Phi1)/(Vphi-1));

G05CBF(seed);
{=====
  CALCULATION OF THE DESIGN
  =====}
c := (ln(P1)/ln(P2))**(1/d);
m := d + trunc(ln(Phi2/Phi1)/ln(c));
x[1] := sqrt(ln(P1) * ln(P2) * (c**(1-m))/(Phi1 * Phi2));
for i := 2 to m do
  x[i] := x[i-1] * c;
{=====
  INITIALIZATION
  =====}
for k := 1 to stat do
  for ni := 1 to Vn do
    for f := 1 to Vphi do
      begin
        MeanBias [k,ni,f] := 0.0;
        Mse      [k,ni,f] := 0.0;
        NumOk    [k,ni,f] := NumSim;
        NumOkSd  [k,ni,f] := NumSim;
        ConfLevel[k,ni,f] := 0.0;
      end;
  for i := 1 to m do
    zero[i] := 0;
  {=====
    START OF SIMULATIONS
    =====}
  for sim := 1 to NumSim do
    begin { of sim }
      for f := 1 to Vphi do
        begin { of f }
          Tphi := Phiv[f];
          for i := 1 to m do
            begin

```

```

pneg := exp(-Tphi*x[i]);
for j := 1 to maxn do
  if G05DZF(pneg) then
    Y[i,j] := 0 { neg. response with probability pneg }
  else Y[i,j] := 1; { pos. response with probability 1-pneg }
end;
for ni := 1 to Vn do
  begin { of ni }
    n := nVn[ni];
    j_start := jb[ni];
    j_end := j_start+n-1;
    nm := n * m;

    for i := 1 to m do
      begin
        nv[i] := n;
        r[i] := n;
        for j := j_start to j_end do
          r[i] := r[i] - Y[i,j];
        end;

        for i := 1 to stat do
          ok[i] := true;
          {=====
          DETERMINATION OF WEIGHTED-MEAN ESTIMATE (WM)
          =====}
        }
        if WM(r,n,m,x,PhiEst[1]) then
          Phi_0 := PhiEst[1]
        else
          Phi_0 := (Phi1 + Phi2)/2; {Phi_0 is initial estimate for ML}
          {=====
          DETERMINATION OF MAXIMUM LIKELIHOOD ESTIMATE (ML)
          =====}
        }
        if ML(r,nv,x,m,n,Phi_0,PhiEst[2],SdEst[2],0,AllNeg0,AllPos0)
        then
          if AllNeg0 then
            begin
              for i := 3 to stat do
                PhiEst[i] := 0
            end
          else
            begin
              Phi_0 := PhiEst[2];
              {=====
              DETERMINATION OF JACKKNIFE ESTIMATE (Je), BASED ON
              DELETING ONE ELEMENT FROM THE MATRIX Y AT A TIME
              =====}
            }
            del := 1;
            succes := true;
            sum1 := 0.0;
            sum2 := 0.0;

```

```

while ((del <= m) and succes) do
begin
  if (r[del] = n) then
    if ML(r,nv,x,m,n,Phi_0,PhiML,
          SdML,-del,AllNeg,AllPos)
    then
      begin
        sum1 := sum1 + n*PhiML;
        sum2 := sum2 + n*PhiML*PhiML;
      end
    else
      succes := false
    else
      if (r[del] = 0) then
        if ML(r,nv,x,m,n,Phi_0,PhiML,
              SdML,del,AllNeg,AllPos)
        then
          begin
            sum1 := sum1 + n*PhiML;
            sum2 := sum2 + n*PhiML*PhiML;
          end
        else
          succes := false
        else
          begin
            if ML(r,nv,x,m,n,Phi_0,PhiML,
                  SdML,-del,AllNeg,AllPos)
            then
              begin
                sum1 := sum1 + r[del]*PhiML;
                sum2 := sum2 + r[del]*PhiML*PhiML;
                if ML(r,nv,x,m,n,Phi_0,PhiML,
                      SdML,del,AllNeg,AllPos)
                then
                  begin
                    sum1 := sum1+(n-r[del])*PhiML;
                    sum2 := sum2+(n-r[del])*PhiML*PhiML;
                  end
                else succes := false
              end
            else
              succes := false
            end;
          del := del + 1
        end; { of while }
      if succes then
        begin
          PhiEst[3] := nm*PhiEst[2]- sum1*(nm-1)/nm;
          SdEst[3] := sqrt((nm-1)*abs(sum2-sum1*sum1/nm)/nm);
          if (PhiEst[3] < 0) or (PhiEst[3] > 1) then
            ok[3] := false
          end
        else
          ok[3] := false;

```

```
{=====
DETERMINATION OF BOOTSTRAP ESTIMATE (Be), BASED
ON RESAMPLING FROM THE ELEMENTS OF THE MATRIX Y
-----}
boot := 1;
succes := true;
sum1 := 0.0;
sum2 := 0.0;

while ((boot <= numboot) and succes) do
begin
  nrc := zero;
  rrc := zero;
  for i := 1 to nm do
    begin
      row := G05DYF(1,m);
      nrc[row] := nrc[row]+1;
      if (r[row]=n) then
        rrc[row] := rrc[row]+1
      else
        if (r[row]>0) then
          rrc[row] :=
            rrc[row]+1-Y[row,G05DYF(j_start,j_end)];
    end;

  xrc := x;
  k := 0;
  repeat
    k := k+1;
    n_k_is_0 := (nrc[k]=0);
  until (k=m) or n_k_is_0;
  if n_k_is_0 then
    begin
      j := k-1;
      for i := k+1 to m do
        if nrc[i]>0 then
          begin
            j := j+1;
            nrc[j] := nrc[i];
            xrc[j] := xrc[i];
            rrc[j] := rrc[i];
          end;
      mrc := j;
    end
  else
    mrc := m;

  if
    ML(rrc,nrc,xrc,mrc,n,Phi_0,PhiML,
      SdML,O,AllNeg,AllPos)
  then
    begin
      sum1 := sum1 + PhiML;
```

```

        sum2 := sum2 + PhiML*PhiML;
        boot := boot + 1
    end
else
    succes := false;
end;

if succes then
begin
    PhiEst[4] := 2*PhiEst[2]-sum1/numboot;
    SdEst[4] :=
        sqrt(abs(sum2-sum1*sum1/numboot)/(numboot-1));
    if SdEst[4]<0.01*SdEst[2] then
        SdEst[4] := 0;
    if (PhiEst[4] < 0) or (PhiEst[4] > 1) then
        ok[4] := false
    end
else
    ok[4] := false;
{=====
  DETERMINATION OF SALAMA ESTIMATE (S2)
  =====}
}
DS2 := 0.0;
for i := 1 to m do
begin
    PnegS[i]      := exp( - Phi_0 * x[i]);
    npos_E[i]     := (1 - PnegS[i]) * n;
    FS[i]         := x[i] / (1 - PnegS[i]);
    FS2[i]        := FS[i] * FS[i];
    PnegS_FS2[i]  := PnegS[i] * FS2[i];
    DS2           := DS2 + npos_E[i] * PnegS_FS2[i]
end;

for i := 1 to m do
begin
    d2d22[i] := 0.0;
    for j := 1 to m do
begin
        P_F_etc[j] := PnegS_FS2[j] *
            (x[j] + 2 * FS[j] * PnegS[j]);
        d2d22[i]   := d2d22[i] + npos_E[j] * P_F_etc[j];
    end;
    d1d12i_2      := (FS[i] / DS2) ** 2;
    pnegS_FSi_m2  := 2 * pnegS[i] * FS[i];
    d2d22[i]      := d1d12i_2 * (d2d22[i] / DS2 -
        PnegS_FSi_m2);
end;

Sum2 := 0.0;
for i := 1 to m do
begin
    npqi := n * pnegS[i] * (1 - pnegS[i]);
    Sum2 := Sum2 + d2d22[i] * npqi;
end;

```

```

end;
PhiEst[5] := Phi_0 - Sum2 / 2;

if (PhiEst[5] < 0) or (PhiEst[5] > 1) then
    ok[5] := false
end

else { (J,B,S)ML-estimates cannot be determined }
    for i := 2 to stat-1 do
        ok[i] := false;
    }
    {=====
    DETERMINATION OF MINIMUM CHI-SQUARE (MC) ESTIMATE
    -----}
}
if (not AllNeg0) then
    begin
        if MC(r,n,m,x,Phi_0,PhiEst[6],SdEst[6],AllPos0) then
            ok[6] := true
        else
            ok[6] := false;
        end;
    }
    {=====
    UPDATE TABLES
    -----}
}
for k := 2 to stat do
    if (ok[k]) then
        begin
            PhiEstk := PhiEst[k];
            MeanBias[k,ni,f] := MeanBias[k,ni,f] + PhiEstk;
            Mse[k,ni,f] := Mse[k,ni,f] + PhiEstk * PhiEstk;
            if (k>5) then {an estimate for SdS2 is not available}
                begin
                    SdEstk := SdEst[k];
                    if AllNeg0 or (SdEstk=0) then
                        NumOkSd[k,ni,f] := NumOkSd[k,ni,f] - 1
                    else
                        if (PhiEstk+1.960*SdEstk >= Tphi) and
                           (PhiEstk-1.960*SdEstk <= Tphi) then
                            ConfLevel[k,ni,f] := ConfLevel[k,ni,f]+1;
                        end
                    end
                end
            end
        else
            begin
                NumOk [k,ni,f] := NumOk [k,ni,f] - 1;
                NumOkSd[k,ni,f] := NumOkSd[k,ni,f] - 1;
            end;
        end { of ni }
    end { of f }
end; { of sim }
{=====
WRITING THE RESULTS TO STANDARD OUTPUT FILE
-----}
}

```



```
for f := 1 to Vphi do
begin
  Tphi := Phiv[f];
  for ni := 1 to Vn do
    begin
      n := nVn[ni];
      for K := 2 to stat do
        begin
          NonMissing := NumOk[k,ni,f];
          if (NonMissing > 0) then
            begin
              MeanPhiEst := MeanBias[k,ni,f]/NonMissing;
              MeanBias[K,ni,f] := MeanPhiEst - Tphi;
              Mse[K,ni,f] := Mse[k,ni,f]/NonMissing -
                Tphi*(2*MeanPhiEst - Tphi);
              NonMissing := NumOkSd[k,ni,f];
              if (NonMissing > 0) then
                ConfLevel[k,ni,f] := ConfLevel[k,ni,f]/NonMissing
              else
                ConfLevel[k,ni,f] := 0.0;
            end
          else
            begin
              MeanBias[k,ni,f] := 0.0;
              Mse[k,ni,f] := 0.0;
              ConfLevel[k,ni,f] := 0.0;
            end
          end;
        end;

      WriteLn(Tphi:8:4,n:3);
      for i:= 2 to stat do
        Write(MeanBias[i,ni,f]/Tphi:11:5); { Mean Relative Bias }
      WriteLn;
      for i:= 2 to stat do
        Write(sqrt(Mse[i,ni,f])/Tphi:11:5); { Coefficient of Variation }
      WriteLn;
      ConfLevel[5,ni,f] := 0;
      for i:= 2 to stat do
        Write(ConfLevel[i,ni,f]:11:5); { Realized Confidence Level }
      WriteLn;
      for i:= 2 to stat do
        Write(NumOk[i,ni,f]:11); { number of non-failed estimations
                                of the parameter }
      WriteLn;
      NumOkSd[5,ni,f] := 0;
      for i:= 2 to stat do
        Write(NumOkSd[i,ni,f]:11);{ number of non-failed estimations
                                of the standard-error }
      WriteLn;
    end
  end
end
END.
```

REFERENCES

- Amemiya, T. (1981), Qualitative response models : a survey, *Journal of Economic Literature*, 19, 1483-1536.
- Ashton, W.D. (1972), *The Logit Transformation*, London : Griffin.
- Berkson, J. (1944), Application of the logistic function to bio-assay, *Journal of the American Statistical Association*, 39, 357-365.
- Berkson, J. (1980), Minimum chi-square, not maximum likelihood!, *Annals of Statistics*, 8, 457-487.
- Bickel, P.J. and Doksum, K.A. (1977), *Mathematical Statistics : Basic Ideas and Selected Topics*, San Francisco : Holden-Day.
- Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975), *Discrete Multivariate Analysis, Theory and Practice*, Cambridge, Mass.: MIT Press.
- Brillinger, D.R. (1964), The asymptotic behaviour of Tukey's general method of setting approximate confidence limits (the jackknife) when applied to maximum likelihood estimates, *Review of the International Statistical Institute*, 32, 202-206.
- Brown, B.Wm., Jr. (1966), Planning a quantal assay of potency, *Biometrics*, 22, 322-329.
- Buurman, W.A., Daemen, A.J.J.M., Groenewegen, G., Does, R.J.M.M., Van der Linden, C.J. and Vegt, P.A. (1983), Limiting dilution analysis of canine alloantigen-reactive T lymphocytes, *Transplantation*, 35, 363-367.

- Chibisov, D.M. (1973), An asymptotic expansion for a class of estimators containing maximum likelihood estimators, *Theory of Probability and Its Applications*, 18, 295-303.
- Cochran, W.G. (1950), Estimation of bacterial densities by means of the 'Most probable number', *Biometrics*, 6, 105-116.
- Cox, D.R. (1970), *The Analysis of Binary Data*, London : Methuen.
- Daellenbach, H.G., George, J.A. and McNickle, D.C. (1983), *Introduction to Operations Research Techniques*, Boston : Allyn and Bacon.
- Daganzo, C. (1979), *Multinomial Probit*, New York : Academic Press.
- Davison, A.C., Hinkley, D.V. and Schechtman, E. (1986), Efficient bootstrap simulation, *Biometrika*, 73, 555-566.
- Does, R.J.M.M., Strijbosch, L.W.G. and Albers, W. (1988), Using jackknife methods for estimating the parameter in dilution series, *Biometrics*, 44, 1093-1102.
- Does, R.J.M.M., Strijbosch, L.W.G. and Buurman, W.A. (1988), Dilution series : experimental design and statistical analysis, *The Proceedings of the 1-st World Congress on Mathematical Statistics and Probability*, Nauka Publishing House, Moscow.
- Doveren, R.F.C., Buurman, W.A., van der Linden, C.J., Strijbosch, L.W.G., Sproncken, E.E.M. and Kootstra, G. (1986), Analysis of cytotoxic T lymphocyte response in rejecting allografted canine kidneys, *Transplantation*, 41, 33-38.
- Efron, B. (1979), Bootstrap methods : another look at the jackknife, *Annals of Statistics*, 7, 1-26.
- Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia : SIAM.

- Efron, B. and Tibshirani, R. (1986), Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statistical Science*, 1, 54-77.
- Eisenhart, C. and Wilson, P.W. (1943), Statistical methods and control in microbiology, *Bacteriological Reviews*, 7, 57-137.
- Fazekas de St. Groth, S. (1982), The evaluation of limiting dilution assays, *Journal of Immunological Methods*, 49, R11-R23.
- Finney, D.J. (1947), The principles of biological assay, *Journal of the Royal Statistical Society, Series B Suppl.*, 2, 46-91.
- Finney, D.J. (1951), The estimation of bacterial densities from dilution series, *Journal of Hygiene*, 49, 26-35.
- Finney, D.J. (1971), *Probit Analysis*, 3rd Edition, Cambridge : University Press.
- Finney, D.J. (1978), *Statistical Method in Biological Assay*, 3rd Edition, New York : Academic Press.
- Fisher, R.A. (1922), On the mathematical foundations of theoretical statistics, *Royal Society of the London Philosophical Transactions (Series A)*, 222, 309-368.
- Fisher, R.A. and Yates, F. (1963), *Statistical Tables for Biological, Agricultural and Medical Research*, 6th Edition, Edinburgh : Oliver & Boyd.
- Haberman, S.J. (1974), *The Analysis of Frequency Data*, Chicago : University of Chicago Press.
- Haberman, S.J. (1978), *Analysis of Qualitative Data, Vol. I. Introductory Topics*, New York : Academic Press.

- Haberman, S.J. (1979), *Analysis of Qualitative Data, Vol. II. New Developments*, New York : Academic Press.
- Halvorson, H.O. and Ziegler, N.R. (1933), Application of statistics to problems in bacteriology. II. A consideration of the accuracy of dilution data obtained by using a single dilution, *Journal of Bacteriology*, 26, 331-339.
- Hamilton, M.A. (1979), Robust estimates of the ED50, *Journal of the American Statistical Association*, 74, 344-354.
- Hamilton, M.A. (1980), Inference about the ED50 using the trimmed Spearman-Kärber procedure - A Monte Carlo investigation, *Communications in Statistics - Simulation and Computation*, B9(3), 235-254.
- Hsi, B.P. (1969), The multiple sample up-and-down method in bioassay, *Journal of the American Statistical Association*, 64, 147-162.
- Johnson, E.A. and Brown, B.Wm., Jr. (1961), The Spearman estimator for serial dilution assays, *Biometrics*, 17, 79-88.
- Kendall, M.G. and Stuart, A. (1973), *The Advanced Theory of Statistics*, 3rd Edition, London : Griffin.
- Kärber, G. (1931), Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche, *Arch. Exp. Path. Pharmac.*, 162, 480-487.
- Lefkovits, I. and Waldmann, H. (1984), Limiting dilution analysis of the cells of immune system. I. The clonal basis of the immune response, *Immunology Today*, 5, 265-268.
- Loyer, M.W. (1981), Using a serial dilution experiment to estimate the density of organisms, *Unpublished Ph.D. Thesis : Montana State University*.

- Loyer, M.W. and Hamilton, M.A. (1984), Interval estimation of the density of organisms using a serial-dilution experiment, *Biometrics*, 40, 907-916.
- Mantel, N. (1985), Maximum likelihood vs. minimum chi-square, *Biometrics*, 41, 777-783.
- Maryanski, J.L., Cerottini, J.-C. and Brunner, K.T. (1980), Susceptibility of proliferating cytolytic T lymphocytes to 33258 Hoechst-modified BUdR and light treatment : quantitation of the residual precursor cells by limiting dilution analysis, *Journal of Immunology*, 124, 839-845.
- Maryanski, J.L., MacDonald, H.R. and Cerottini, J.-C. (1980), Limiting dilution analysis of alloantigen-reactive T lymphocytes. IV. High frequency of cytolytic T lymphocyte precursor cells in MLC blasts separated by velocity sedimentation, *Journal of Immunology*, 124, 42-47.
- Mather, K. (1949), The analysis of extinction time data in bioassay, *Biometrics*, 5, 127-143.
- McCarthy, J.A., Thomas, H.A., Jr. and Delaney, J.E. (1958), Evaluation of the reliability of coliform density counts, *American Journal of Public Health*, 48, 1628-1635.
- McCrary, M.H. (1915), The numerical interpretation of fermentation-tube results, *Journal of Infectious Diseases*, 17, 183-212.
- Miller, R.G. (1974), The jackknife - a review, *Biometrika*, 61, 1-15.
- Miller, R.G. (1982), Clonal analysis by limiting dilution. An overview. In *Isolation, Characterization, and Utilization of T Lymphocyte Clones*. (C.G. Fathman and F.W. Fitch, eds.), Chap. 18, pp. 219-231. New York : Academic Press.
- Okamoto, M. (1958), Some inequalities relating to the partial sum of binomial probabilities, *Annals of the Institute of Statistical Mathematics*, 10, 29-35.

- Peto, S. (1953), A dose-response equation for the invasion of micro-organisms, *Biometrics*, 9, 320-335.
- Plackett, R.L. (1974), *The Analysis of Categorical Data*, London : Griffin.
- Quenouille, M.H. (1949), Approximate tests of correlation in time series, *Journal of the Royal Statistical Society, Series B*, 11, 18-84.
- Quenouille, M.H. (1956), Notes on bias in estimation, *Biometrika*, 43, 353-360.
- Rao, C.R. (1957), Theory of the method of estimation by minimum chi-square, *Bulletin de l'Institut International de Statistique*, 35, 25-32.
- Salama, I.A., Koch, G.G. and Tolley, H.D. (1978), On the estimation of the most probable number in a serial dilution experiment, *Communications in Statistics - Theory and Methods*, A7(13), 1267-1281.
- Spearman, C. (1901), The method of the 'right and wrong cases' ('constant stimuli') without Gauss's formulae, *British Journal of Psychology*, 2, 227-242.
- Strijbosch, L.W.G., Buurman, W.A., Does, R.J.M.M., Zinken, P.H. and Groenewegen, G. (1987), Limiting dilution assays : experimental design and statistical analysis, *Journal of Immunological Methods*, 97, 133-140.
- Strijbosch, L.W.G. and Does, R.J.M.M. (1988), Comparison of bias-reducing methods for estimating the parameter in dilution series, *Communications in Statistics - Simulation and Computation*, B17(4), 1173-1190.
- Strijbosch, L.W.G., Does, R.J.M.M. and Albers, W. (1988), Design methods for some dose-response models, *Medical Informatics and Statistics Report*, 21, University of Limburg, Maastricht, Submitted for publication.

- Strijbosch, L.W.G., Does, R.J.M.M. and Buurman, W.A. (1987), Experimentele proefopzet en statistische analyse van verdunningsreeksen (in Dutch), *Kwantitatieve Methoden*, 24, 149-160.
- Strijbosch, L.W.G., Does, R.J.M.M. and Buurman, W.A. (1988), Computer aided design and evaluation of limiting and serial dilution experiments, *International Journal of Bio-Medical Computing*, 23, 279-290.
- Taswell, C. (1981), Limiting dilution assays for the determination of immunocompetent cell frequencies. I. Data analysis, *Journal of Immunology*, 126, 1614-1619.
- Taswell, C. (1984), Limiting dilution assays for the determination of immunocompetent cell frequencies III : Validity tests for the single-hit Poisson model. *Journal of Immunological Methods*, 72, 29-40.
- Taswell, C. (1987), Limiting dilution assays for the separation, characterization, and quantitation of biologically active particles and their clonal progeny. In *Cell Separation : Methods and Selected Applications*, 4, Pretlow, T.G. and Pretlow, T.P. (eds), 109-145, New York : Academic Press.
- Thomas, H.A., Jr. (1942), Bacterial densities from fermentation tube results, *Journal of the American Water Works Association*, 34, 572-576.
- Thomas, H.A., Jr. (1955), Statistical analysis of coliform data, *Sewage and Industrial Wastes*, 27, 212-222.
- Thomas, H.A., Jr. and Woodward, R.L. (1955), Estimation of coliform density by the membrane filter and the fermentation tube methods, *American Journal of Public Health*, 45, 1431-1437.
- Tsutakawa, R.K. (1972), Design of experiment for bioassay, *Journal of the American Statistical Association*, 67, 584-590.

- Tsutakawa, R.K. (1980), Selection of dose levels for estimating a percentage point of a logistic quantal response curve, *Applied Statistics*, 29, 25-33.
- Tukey, J.W. (1958), Bias and confidence in not quite large samples, abstract, *Annals of Mathematical Statistics*, 29, 614.
- Tukey, J.W. (1977), *Exploratory Data Analysis*, Reading, Mass.: Addison-Wesley.
- Vose, B.M. and Bonnard, G.D. (1983), Limiting dilution analysis of the frequency of human T cells and large granular lymphocytes proliferating in response to interleukin 2. I. The effect of lectin on the proliferative frequency and cytotoxic activity of cultured lymphoid cells, *Journal of Immunology*, 130, 687-693.
- Wetherill, G.B. (1963), Sequential estimation of quantal response curves, *Journal of the Royal Statistical Society, Series B.*, 25, 1-48.
- Woodward, R.L. (1957), How probable is the most probable number, *Journal of the American Water Works Association*, 49, 1060-1068.
- Wu, C.F.J. (1986), Jackknife, bootstrap and other resampling methods in regression analysis, *Annals of Statistics*, 14, 1261-1295.

SUMMARY

This thesis is a contribution to the theory of Limiting and Serial Dilution Assays (LDA's, SDA's) which are used in many areas of bioresearch. In general, the aim of such an assay is the estimation of the frequency of a particular cell type among other but similar cells (LDA), or the estimation of the density of particular cells or microorganisms per unit volume of solution (SDA). In both LDA and SDA the unknown parameter is commonly estimated by using the "single-hit Poisson model" with binary (positive or negative) data produced by samples taken from different dilutions. In this study research has been performed on two important statistical issues of bioassays. These are the construction of an experimental design and the choice of the estimation techniques to be used.

For the setup of a design for a LDA or SDA one has to choose the number of dilutions, the number of replicates in each dilution and the (mean) number of cells/organisms per replicate. A design which has been constructed properly should take advantage of existing a priori information of the researcher and incorporate simple design parameters chosen by the researcher in such a way that the production of useless data is limited and the precision of the estimator is controlled. In Chapter 2 a useful and simple method for setting up a design on the basis of the researcher's criteria and a priori information about the unknown parameter is presented. Refinements for this design method and a program to perform these techniques are presented in Chapter 3. The methods of Chapter 2 and 3 regard the experimental design for LDA's and SDA's. The basic design method of Chapter 2 can be generalized to some other related and widely used biological models, such as the "logistic regression model" which is most common in bioassay. As the advance information on two unknown parameters have to be incorporated here, the experimental design is more complicated. The corresponding generalizations for three different cases of advance information are presented in Chapter 4.

The second main issue contained in this dissertation is the study of the commonly used estimation methods and some bias-reducing methods when applied to dilution analysis. In Chapter 5 two conventional statistical procedures (the Maximum Likelihood (ML) method and the Minimum Chi-square

(MC) method) are described. For an experiment with all positive responses the corresponding estimates are infinite. Thus, modified ML and MC estimators are proposed in order to reduce the bias of these estimators from infinity to the desired order. However, the resulting estimators are still biased. Little is known about bias and efficiency of the modified ML and MC estimators when applied to data from dilution experiments with small or moderate size. Resampling methods like the jackknife and the bootstrap provide estimators of bias and variance for an extremely wide class of statistics. In general, these methods are applied to independent identically distributed (iid) random variables. There are also several reasons to consider some jackknife and bootstrap versions based on non-iid variables. In Chapter 6 three jackknife and four bootstrap versions of the ML method are discussed and applied to dilution analysis.

Results on Monte Carlo comparison of estimators have been published by several authors. However, there is a difficulty in comparing these results because of the absence of generally accepted design methods. Some controversies emerged partly as a consequence of this kind of difficulty. Research on estimators in dilution analysis requires a design method that can be used as a frame of reference. Thus, the design method proposed in Chapter 2 is used for the simulation study conducted in this dissertation. A PASCAL-program has been written to perform the Monte Carlo experiments for the comparison of the statistical methods described in Chapters 5 and 6. The simulation results are presented and discussed in Chapter 7. The criteria used to compare the estimation methods concern the bias, the coefficient of variation (based upon the mean square error) and the realized confidence level (i.e. the proportion of times that the realized 95% confidence interval contains the true value of the parameter). The main results can be summarized as follows. The corresponding jackknife and bootstrap versions of the ML estimator are very strongly related. Generally, all resampling methods reduce the bias of the ML estimator significantly to values around zero. Provided that the design used is not too small, the most preferable method is a particular jackknife version (obtained by deleting one individual replicate culture at a time). The MC estimator is shown to be inferior to the ML estimator.

SAMENVATTING

EXPERIMENTEEL ONTWERP EN STATISTISCHE ANALYSE VAN BIOLOGISCHE VERDUNNINGSREEKSEN

In dit proefschrift wordt een bijdrage geleverd aan de theorie over biologische verdunningsreeksen. In de literatuur worden twee typen verdunningsreeksen onderscheiden, te weten de "Limiting Dilution Assay" (LDA) en de "Serial Dilution Assay" (SDA). Bij LDA gaat het er veelal om te schatten welke proportie van een verzameling gelijksoortige cellen een bepaalde eigenschap bezit, terwijl bij SDA de dichtheid van een bepaald type cel per volume eenheid wordt geschat. Zowel bij een LDA als een SDA wordt bij het schatten van de onbekende parameter meestal gebruik gemaakt van de veronderstelling dat het aantal relevante cellen per volume eenheid een Poisson verdeling heeft, en dat de aanwezigheid van één relevante cel in een volume eenheid al met zekerheid kan worden vastgesteld. Aangezien de te schatten parameter onbekend is en het tellen van cellen doorgaans onmogelijk is, dient met verschillende verdunningen gewerkt te worden ten einde onafhankelijk van de waarde van de parameter in een bepaald gebied een schatting te kunnen bepalen. In het hier beschreven onderzoek worden zowel de experimentele proefopzet als de analyse van verdunningsreeksen bestudeerd.

Voor een LDA of een SDA dient bij het ontwerp van een experiment het aantal verdunningen gekozen te worden, alsmede het aantal te herhalen bepalingen per verdunning en het (gemiddeld) aantal cellen per herhaling. Een goed ontwerp houdt rekening met de voorkennis van de onderzoeker en bevat ontwerp parameters die de statistische eigenschappen van het ontwerp, zoals de te bereiken nauwkeurigheid, bepalen. In hoofdstuk 2 wordt een dergelijk ontwerp gepresenteerd, terwijl in hoofdstuk 3 diverse verfijningen worden uitgewerkt en een computer-programma wordt beschreven dat deze technieken uitvoert. De basisprincipes uit hoofdstuk 2 kunnen worden gegeneraliseerd voor toepassing op andere biologische modellen (met twee onbekende parameters) die sterk aan LDA's en SDA's gerelateerd zijn, zoals het "logistische regressie model". Aangezien hier de voorkennis met betrekking tot *twee* onbekende parameters in het experimenteel ontwerp verwerkt moet worden, is een dergelijk ontwerp ingewikkelder. In hoofdstuk 4

worden voor drie verschillende gevallen van aanwezige voorkennis ontwerp voorstellen gedaan.

Het tweede hoofd onderwerp in dit proefschrift betreft de bestudering van de gebruikelijke schattingsmethoden en enige onzuiverheid-reducerende schattingsmethoden voor toepassing bij verdunningsreeksen. In hoofdstuk 5 worden twee conventionele schatters besproken, namelijk de "Maximum Likelihood" (ML) schatter en de "Minimum Chi-square" (MC) schatter. Bij een experiment met alleen positieve uitslagen zijn de corresponderende schattingen oneindig. Ten einde de onzuiverheid van deze schatters terug te brengen tot de gewenste orde, worden gewijzigde ML en MC schatters voorgesteld. Deze schatters vertonen echter nog steeds een onzuiverheid, die bij experimenten van kleine of bescheiden omvang bezwaarlijk kan zijn. Vanwege de brede toepasbaarheid van zgn. jackknife en bootstrap methoden kan worden geprobeerd de onzuiverheid te schatten om daarmee een schatter te construeren die minder onzuiver is. Jackknife en bootstrap methoden verschaffen ook automatisch een schatter voor de variantie. In het algemeen worden deze methoden toegepast op onafhankelijke en identiek verdeelde stochastische variabelen. De speciale toepassing op verdunningsreeksen echter suggereert ook varianten die zijn gebaseerd op onafhankelijke maar niet identiek verdeelde variabelen. In hoofdstuk 6 worden drie jackknife en vier bootstrap varianten van de ML methode besproken en toegepast op verdunningsreeksen.

Het ontbreken van algemeen geaccepteerde ontwerpmethoden voor verdunningsreeksen maakt het moeilijk om de Monte Carlo resultaten van verschillende onderzoekers te vergelijken. Sommige tegenstrijdigheden in de literatuur kunnen hieraan worden toegeschreven. Onderzoek naar de eigenschappen van schatters voor de onbekende parameter in verdunningsreeksen vraagt om een ontwerpmethode, die kan worden gebruikt als een referentiekader. De in hoofdstuk 2 beschreven methode is hiervoor geschikt en wordt gebruikt in het simulatie onderzoek, dat in hoofdstuk 7 beschreven is. Dit onderzoek vergelijkt m.b.v. een speciaal voor dit doel in PASCAL geschreven programma de volgende drie eigenschappen van de in hoofdstukken 5 en 6 beschreven schatters: de gemiddelde relatieve onzuiverheid, de variatiecoëfficiënt (afgeleid van de gemiddelde kwadratische fout) en de proportie van het aantal keren dat het gerealiseerde 95% betrouwbaarheidsinterval de werkelijke waarde van de

parameter bevat. De belangrijkste resultaten kunnen als volgt worden samengevat. De met elkaar corresponderende jackknife en bootstrap varianten hebben sterk op elkaar gelijkende eigenschappen. In het algemeen reduceren ze de onzuiverheid van de ML schatter significant tot waarden rond nul. Mits het ontwerp een voldoende aantal elementaire bepalingen bevat kan worden gesteld dat de jackknife variant die wordt verkregen door het beurtelings weglaten van één elementaire bepaling de voorkeur geniet. Van de MC schatter wordt aangetoond, dat deze niet te verkiezen is boven de ML schatter.

DANKBETUIGING

Een aantal mensen wil ik op deze plaats graag bedanken voor de rol die zij hebben gespeeld bij de voorbereidingen en de totstandkoming van het proefschrift. Dr. Wim Buurman heeft van uit zijn dagelijkse onderzoekspraktijk (een deel van) de problematiek voorgelegd aan dr. Ronald Does, die op zijn beurt mij bij het projekt betrok. De samenwerking met Wim Buurman is mede door zijn enthousiasme van doorslaggevende betekenis geweest. Bij de voorbereiding en uitvoering van de eerste computersimulaties is de inzet van Piet Zinken van groot belang geweest. Na acceptatie van het eerste artikel suggereerde Ronald Does mij om op dit onderwerp te promoveren. Via intensieve samenwerking met hem werd een onderzoeksplan opgesteld en uitgevoerd. Deze samenwerking is door mij als zeer plezierig ervaren. Met name zijn adviezen en zijn optimistische benadering hebben de voortgang van het onderzoek gunstig beïnvloed. Met veel genoegen denk ik ook terug aan de vele telefoon-gesprekken die wij samen hebben gevoerd. Van grote betekenis voor mij is geweest een gesprek met prof. D.J. Finney (destijds directeur van het "Research Centre" van het "International Statistical Institute" te Voorburg), die zeer geïnteresseerd was in het onderwerp van mijn onderzoek en mij aanmoedigde de resultaten te publiceren. Mijn promotor prof.dr. Wim Albers dank ik voor zijn nauwgezette bestudering van (eerdere versies van) het proefschrift die heeft geleid tot een aantal verbeteringen en veel waardevolle aanvullingen. Mijn ouders hebben de basis gelegd voor mijn wetenschappelijke belangstelling en studie, waarvoor ik hen veel dank verschuldigd ben. Liesbeth, zonder jouw aanmoedigingen en jouw bereidheid om gedurende lange tijd mijn aandeel in het gezin voor een groot deel over te nemen, was dit proefschrift nooit geschreven. Jij en de kinderen hebben er alles aan gedaan om alles zo optimaal mogelijk te laten verlopen, iets dat mij in het vuur van de strijd wel eens ontging. Bij het gereedmaken van het manuscript voor de drukker heb ik dankbaar gebruik gemaakt van de diensten van Jan Pijnenburg, o.a. voor het ontwerpen van de omslag, van ing. Henk Op den Brouw voor diverse adviezen op het gebied van de tekstverwerking en van Mw. Micky Schim van der Loeff voor de correctie van het Engels. De onduidelijkheden die na haar zorgvuldig en uitgebreid commentaar nog resteren komen uiteraard geheel voor mijn rekening. Tot slot bedank ik prof.dr. Frank van der Duyn Schouten en ir. Ed van den Bosch voor het in mij gestelde vertrouwen, en mijn kamergenoot ir. Anton Markink voor zijn interesse en medeleven.

CURRICULUM VITAE

Leo Strijbosch werd op 7 juni 1947 geboren te Helmond (NB). In die plaats volgde hij ook het lager onderwijs (St. Canisiusschool, 1953-1959) en het middelbaar onderwijs (St. Carolus Borromeus College, Gymnasium B, 1959-1966). In september 1966 begon hij met de studie Werktuigbouwkunde aan de Technische Hogeschool te Eindhoven, maar zette na een half jaar de studie voort aan de Onderafdeling der Wiskunde. Na het behalen van het kandidaatsdiploma in juni 1971 heeft hij zich vooral verdiept in de Informatica en de Operations Research. Het afstudeerwerk had betrekking op een onderwerp uit de Operations Research (gesloten netwerksystemen van wachtrijen). De afstudeerhoogleraar hierbij was prof.dr. J. Wessels. Het diploma wiskundig ingenieur werd behaald in februari 1975.

Van april tot oktober 1975 was hij als wetenschappelijk ambtenaar verbonden aan de vakgroep Statistiek van de Onderafdeling der Wiskunde voor de ontwikkeling van statistische programmatuur. In oktober 1975 aanvaardde hij een baan bij de Rijksuniversiteit Limburg en was daar aanvankelijk bij de capaciteitsgroep Medische Sociologie (tot augustus 1976) werkzaam, daarna bij de Dienst Informatieverwerking (tot december 1982) en tenslotte bij de capaciteitsgroep Medische Informatica en Statistiek (tot augustus 1985). In deze periode was hij met name betrokken bij diverse onderzoeksprojecten als adviserend statisticus. In de latere jaren van het dienstverband bij de Rijksuniversiteit Limburg vond een verschuiving plaats naar meer (statistiek) onderwijs en eigen onderzoek. In de jaren 1983-1985 werd in samenwerking met dr. Ronald Does en dr. Wim Buurman de basis gelegd voor het onderzoek dat heeft geleid tot de totstandkoming van dit proefschrift.

Vanaf september 1985 tot heden is hij verbonden aan de Katholieke Universiteit Brabant en als specialist wetenschappelijke toepassingen werkzaam bij de Automatiserings Unit van de Faculteit der Economische Wetenschappen. In deze periode werd het proefschrift afgerond.