Some new aspects of dose-response multistage models with applications

Bimal Sinha
Department of Mathematics & Statistics
University of Maryland, Baltimore County

Leonid Kopylev & John Fox
National Center for Environmental Assessment, USEPA

Abstract

This paper discusses statistical inference based primarily on Self and Liang (1987), dealing with the asymptotic theory of maximum likelihood estimates and likelihood ratio tests when some parameters may lie on the boundaries. The results are widely applicable to models used in environmental risk analysis. Applications of the results to dose-response multistage models serve as illustrations. The paper considers various scenarios, and is expected to be of interest to practitioners of risk analysis.

Key Words: Dose-response models, Likelihood ratio tests, Maximum likelihood estimates, Parameters on the boundaries, Multistage Weibull model

1The interpretations and views expressed in this report are solely those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency
Executive Summary

The significance of this work for EPA methodology is that it resolves lingering questions about the accuracy and validity of confidence intervals for the benchmark dose, risk estimates, and parameters of benchmark dose models.

The landmark theoretical papers for the development of benchmark dose models acknowledged these issues but did not resolve them. Simulation studies published in 1994 and 2004 helped to explore the territory but lacked the certainty of mathematical derivation and proof.

The methods used by EPA, and others, to obtain confidence intervals for the benchmark dose make assumptions about the asymptotic distribution (i.e., Normal) of estimates of parameters of benchmark dose models, extra risk and the benchmark dose. Asymptotic approximations apply when the numbers of data (for bioassays, the numbers of test animals at each dose) are indefinitely large. Simulation studies (Bailer and Smith, 1994; Moerbeek et al., 2004) have, to some extent, resolved the question of how good is the asymptotic approximation for actual data sets. However, further work was needed to address this issue more comprehensively.

The asymptotic approximations are theoretically correct unless a parameter is on the boundary of its parameter space. Because parameters are constrained to be non-negative (e.g. multistage model) or greater than one (e.g. log-logistic model) in many dose-response models, it often happens that a parameter is estimated to be on the boundary, and so the true parameter could be on the boundary. In such a case, the asymptotic distributions of model parameters, the benchmark dose, and risks can be different from what is usually assumed.

An alternative method, the profile likelihood method, was advocated by Crump and Howe (1985) and is used by EPA’s BMDS software to find the BMDL (the lower confidence limit of the benchmark dose). The validity of this method when parameters are on the boundary was also in doubt in the past.

This report develops a theoretical approach, based on a paper by Self and Liang (1987), to resolve the issues when parameters are of the multistage model are on the boundary. For methods involving the Normal approximation, and for profile likelihood, this report establishes when these methods are correct and when they are not correct. When they are incorrect, the correct approach based on Self and Liang is provided. These conclusions
can be applied to other dose-response models (and to a wide range of other applications), not just to the multistage model that is used to illustrate the application in this report.

This report also examines the accuracy of confidence intervals reported by BMDS. Because of the difficulties connected to the parameters on the boundary, BMDS currently reports a lower confidence limit for the benchmark dose (which is asymptotically correct), but does not report confidence intervals for model parameters and risk. Without confidence bounds, estimates that BMDS software provides have restricted utility for risk assessment. This report establishes that the reported confidence interval for the benchmark dose is correct and demonstrates how to obtain confidence bounds for model parameters and extra risk which can be easily applied in BMDS software.
1. Introduction

1.1 Dose-Response Models and Risk Estimates

Quantitative risk assessment for toxic and carcinogenic chemicals relies largely upon fitting dose-response models to data from animal bioassays (epidemiological data is not discussed here). A variety of models are in use (Krewski and van Ryzin, 1981; Filipsson et al., 2003; U.S.E.P.A., 2006) to represent both quantal (dichotomous) and continuous responses. Toxicological experience and principles indicate that the response will generally be bounded and non-decreasing (Eaton and Klaassen, 2001).

Quantal response models represent the probability $P(d, \theta)$ of a quantal response, like presence or absence of a particular type of cancer, in relation to dose (d). The experimental data consist of counts of animals exposed to a chemical, the numbers exhibiting the response, and the dose levels (e.g., $n_i, x_i, d_i, i = 0, 1, ..., m$), from a bioassay (often a “long-term bioassay” lasting about 2 years) conducted with mice or rats. In a long-term (2-year) carcinogenicity assay, such as those conducted by the National Toxicology Program, there may be 3-5 dose groups including the control, each having 10-50 animals. All animals undergo pathological examination; usually, cause of death is not attributed, but presence or absence of various tumors is recorded.

Usually the maximum likelihood estimation for the parameters $\theta$ employs the binomial likelihood

$$L(\theta | x) = \prod_{i=0}^{m} \binom{n_i}{x_i} P(d_i, \theta)^{x_i} (1 - P(d_i, \theta))^{n_i - x_i}$$

The benchmark dose method (Crump, 1984; Filipsson et al., 2003; Parham and Portier, 2005) consists of estimating a lower confidence limit for the dose associated with a specified increase $\gamma$ in adverse response above the background level (i.e., increased or extra risk). This is done by fitting a dose-response model, solving for the dose (benchmark dose or "BMD") associated with the specified risk, and finding the lower confidence bound ("BMDL") for this dose (ibid.). In practice, the specified risk is typically selected to be near the lower limit of increased response that could be detected with the experimental design that generated the data, usually between a 1% and 10% increase for cancer incidence using quantal response models (Filipsson et al., 2003, Parham and Portier, 2005).
We will use "absolute risk" to refer to the probability $P(d, \theta)$ modeled by a dose response model for a quantal response. The increase in risk above background for a quantal response is quantified as "extra risk" or "additional risk" (Filipsson et al., 2003). These quantities are defined below.

**Absolute Risk**: $AR = P(d, \theta)$ \hspace{2cm} (2)

**Additional Risk**: $ADR = P(d, \theta) - P(0, \theta)$ \hspace{2cm} (3)

**Extra Risk**: $ER = \frac{P(d, \theta) - P(0, \theta)}{1 - P(0, \theta)}$ \hspace{2cm} (4)

**Benchmark Dose (BMD)**: $d^* : \gamma = \frac{P(d^*, \theta) - P(0, \theta)}{1 - P(0, \theta)}$ \hspace{2cm} (5)

### 1.2. Statistical Inference

Statistical inference for chemical risk assessment has emphasized finding confidence limits for the dose associated with a specified risk and for the risk associated with a specified dose. Parameters of the dose-response model have been of secondary concern, with one exception, the parameter remaining in linearized versions of the models in question, because this parameter governs risk at very low doses.

The review of methods for confidence limits by Crump and Howe (1985) was influential in promoting use of the profile likelihood method. An important feature of this method is invariance to transformations. Later numerical evaluations of its coverage (Bailer and Smith, 1994; Moerbeek et al., 2004) for extra risk and the benchmark dose, support its continued use, although the parametric bootstrap was similar and in some cases superior. Wald-type intervals, by comparison, were decidedly inferior.

The profile likelihood method as applied in these sources and in USEPA’s Benchmark Dose Software (USEPA, 2006) assumes that $-2\log(L(\theta|x))$ is distributed as $\chi^2_1$. This is correct, asymptotically, under certain regularity conditions (Rao, 1973; Cox and Hinkley, 1974), one of which is that the true parameters are interior to the parameter space (for more details, see: Chernoff, 1954; Feder, 1968; and Self and Liang, 1987). However, when one or more parameters (those of interest, or nuisance parameters, or both) lie on the boundary of the parameter space, the distribution of the likelihood
A ratio test statistic may not be $\chi_1^2$ and may be difficult to derive (Self and Liang, 1987).

These considerations may inconvenience inferences for parameters of the dose-response model (1), risk quantities (3 and 4), and the benchmark dose (5). While the true parameters are never known for real experiments, it often happens that maximum likelihood estimates of one or more parameters of the dose-response model are zero (or one in case of slopes; a value that is on the boundary of the parameter space when - as is usually done - parameters are constrained to be non-negative). This circumstance may adversely affect any Wald-type estimates of confidence limits and any likelihood ratio tests and related confidence limits for parameters.

1.3. Research problems

Our primary goal in this paper is to develop appropriate statistical methods to draw valid inference about the parameters $\theta$’s in the multistage model (6) given below and, more importantly, about the quantities of interest, namely, AR, ADR, ER and BMD, when some of the basic parameters may lie on their boundaries. This is a significant point that has not been addressed in the relevant literature dealing with such multistage models.

In the sequel, we discuss one commonly used model for quantal responses, called the multistage model:

$$P(d, \theta) = 1 - \exp(-\sum_{j=0}^{k} \theta_j d^j)$$  \hspace{1cm} (6)

In applications, the coefficients $\theta$’s are often constrained to be non-negative so that the dose-response function will be non-decreasing. The original motivation was a mechanistic model for carcinogenesis in which $k$ different events must occur in a cell to initiate cancer (Guess and Crump 1976, 1978). However, practical applications of the multistage model are exercises in curve-fitting (i.e., the purpose is descriptive or predictive) and there usually is no biological basis for selecting the number of ”stages” (Crump, 1996). We chose to focus on this model because it has been used often in USEPA’s IRIS program chemical risk assessments and has also figured prominently in development of statistical methods for quantitative risk assessment.

While AR, ADR and ER are direct and simple functions of $\theta$, the benchmark dose $d^*$ (5) can obviously be a complicated function of $\theta$, especially...
when \( k \) is large. To circumvent this potential difficulty, we proceed in an alternative fashion. Note that \( d^* (\theta) = \psi_\gamma (\theta) \) satisfies:

\[
\ln(1/(1 - \gamma)) = \theta_1 d^* + \theta_2 d^*^2 + \cdots + \theta_k d^*^k
\]

(7)

We propose to test \( H_{0\delta} : \psi_\gamma (\theta) = \delta \) versus \( H_{1\delta} : \psi_\gamma (\theta) \neq \delta \) for a given \( \delta > 0 \). If \( H_{0\delta} \) is accepted, we include \( \delta \) in the confidence interval for \( \psi_\gamma (\theta) \). Otherwise, we exclude it. We carry out this procedure for all possible values of \( \delta \), thus hopefully generating an interval of accepted \( \delta \)-values, all belonging to the acceptance set, and thereby leading to the lower and upper bounds of \( \delta \), which is \( \psi_\gamma (\theta) \). Since \( \psi_\gamma (\theta) \) is linear in \( \theta \), a test for \( H_0 \) versus \( H_1 \) is easily carried out! In fact, since \( \ln(1/(1 - \gamma)) = \theta_1 \delta + \theta_2 \delta^2 + \cdots + \theta_k \delta^k \) is equivalent to \( \ln(1/(1 - \gamma))/\delta = \theta_1 + \theta_2 \delta + \cdots + \theta_k \delta^{k-1} \), we can work with a re-parametrization of the parameters \((\theta_1, \theta_2, \ldots, \theta_k) \rightarrow (\theta_1^* = \theta_1 + \theta_2 \delta + \cdots + \theta_k \delta^{k-1}, \theta_2, \ldots, \theta_k)\) with \( \theta_1^* \) staying away from the boundary even if some of the \( \theta \)'s may lie on their boundaries.

The organization of the paper is as follows. We develop at length in Section 2, the core section of the paper, the necessary statistical inference in a very general multi-parameter framework with some parameters possibly lying on their boundaries, borrowing ideas from a host of key papers on this topic, notably from Self and Liang (1987). We develop both profile likelihood based inference (LRT) as well as Wald-type inference for a variety of situations with respect to boundary parameters. These results can be applied to a wide variety of models used in dose-response analysis. Applications to the specific multistage models (6) are discussed in Section 3. Some extensive simulations for linear, quadratic and cubic cases are reported in Section 4. These simulations clearly reveal that the use of Self & Liang’s procedure over the Wald procedure considerably improves the expected lengths of the confidence intervals for all the relevant parameters. An Appendix at the end contains a proof of a basic result of the paper.

2. New results on boundary value problems

Let us recall the general discussion mentioned in the previous section about the asymptotic properties of the MLEs in multi-parameter problems and the asymptotic distribution of the likelihood ratio test of a function of
such parameters. Let $X$ denote a random data set which is typically a collection of $N$ independent and identically or independent but non-identically distributed random variables and let $\theta = (\theta_1, \ldots, \theta_p)$ denote a $p$-dimensional real parameter vector which governs the distribution of $X$ through a joint density $f(X|\theta)$ or equivalently the likelihood function $L(\theta|X)$. Let us assume that $\theta \in \Omega \subset \mathbb{R}^p$ and we write $\Omega = \Omega_1 \times \cdots \Omega_p$ where it is assumed that $\theta_i \in \Omega_i$, $i = 1, \ldots, p$. We also assume that the $p$ parameters $\theta_1, \ldots, \theta_p$ are functionally independent. Keeping the specific dose-response multistage Weibull models in mind, we further assume that $\Omega_i = [\theta_{i0}, \infty)$, a half-closed interval or $\Omega_i = (\theta_{i0}, \infty)$, an open interval, where $\theta_{i0}$ is specified. In the former case, $\theta_{i0}$ is referred to as a boundary point of $\theta_i$, and in the latter case, all points of $\theta_i$ are interior points. In most dose-response models, $\theta_{i0} = 0$ or $1$ for all $i$. The statistical problem in such a set-up is to estimate the parameters $\theta$ and test suitable hypotheses about $\theta$ or some functions thereof based on the data set $X$. Since often the joint density or the likelihood function can be quite complicated as in the case of dose-response multistage Weibull models, both estimation of $\theta$ and tests about $\theta$ are carried out using suitable asymptotic theory under the assumption of a large sample size $N$.

It is well known that, under some very general conditions, the asymptotic theory of estimation based on the maximum likelihood estimates (MLEs) and the asymptotic theory of tests based on the likelihood ratio tests (LRTs) are valid and provide useful tools for meaningful statistical inference. Typically, we conclude the asymptotic normality of the MLE $\hat{\theta}_N = (\hat{\theta}_{1N}, \cdots, \hat{\theta}_{pN})$ of $\theta$ and the asymptotic chi-square distribution of $-2\ln(LRT)$ with a suitable $df$ under a null hypothesis. However, it should be noted that the validity of such results is based on a crucial assumption that the true state of nature of the unknown parameter vector $\theta$ is an interior point of $\Omega$. In many applications, including multistage dose-response Weibull models, it can often hold that some parameter spaces are half-closed and, in fact, some parameters may actually lie on their respective boundaries, thus making the standard asymptotic results on MLEs and LRTs unjustified and incorrect. Fortunately, this point has indeed been seriously addressed in the literature and a series of papers concerning this vital issue have appeared, most notably by Self and Liang (1987).

regularity conditions on the joint density of \( X \) and nature of \( \theta \) can be stated. The conditions are typical Cramer-type and are satisfied in our applications to multistage dose-response models.

Let \( \hat{\theta}_N \) denote the MLE of \( \theta \) when the likelihood function \( L(\theta|X) \) is maximized wrt \( \theta \in \Omega \) where the nature of \( \Omega \) is quite general. Recall that \( \Omega = \Omega_1 \times \cdots \Omega_p \) and also recall the nature of \( \Omega_i \) in two possible forms: open intervals and half-closed intervals. Thus, the maximization of the likelihood with respect to \( \theta \) might mean unrestricted maximization in an open set (actually product of open intervals) or a restricted maximization in a product of open and half-closed intervals, depending on which parameters can assume their boundary values. We will call these as natural restrictions on \( \theta \) and reserve the use of the term restricted maximization to the situation when the components of \( \theta \) are restricted by a null hypothesis of the form \( H_0 : \theta \in \Omega_0 \).

Assume \( \theta^* = (\theta_1^*, \ldots, \theta_p^*) \) to be the true value of the vector parameter \( \theta \) and let \( I(\theta^*) \) be the Fisher information matrix evaluated at \( \theta = \theta^* \), which is assumed to be positive definite with \( \Sigma(\theta^*) = [I(\theta^*)]^{-1} \). Note that some of the \( \theta_i^* \)'s may actually be the boundary values \( \theta_{i0} \). For ready reference, we also mention that the \((i,j)\)th element of the information matrix \( I(\theta^*) \) is computed as \( E[-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln L(\theta|X)] \) where the expectation is evaluated at \( \theta^* \). It is understood that in case of boundary points, the derivatives are computed from appropriate directions.

Our first result is concerned with the asymptotic distribution of the MLE \( \hat{\theta}_N \) of \( \theta \) under the true value \( \theta^* \) where \( \theta^* \) is either an interior point in \( \Omega \) or a combination of both interior and boundary points of \( \Omega \).

**Proposition 1.** The asymptotic distribution (as \( N \to \infty \)) of the \( p \)-vector \( \sqrt{N}(\hat{\theta}_1 - \theta_1^*, \ldots, \hat{\theta}_p - \theta_p^*) \) is the same as the (exact) distribution of the MLE \( \hat{\theta}(Z) \) of \( \theta \) based on a normal \( p \)-vector \( Z \) with mean \( \theta \) and dispersion matrix \( \Sigma(\theta^*) \) under \( \theta = 0 \). Here the MLE \( \hat{\theta}(Z) \) of \( \theta \) based on \( Z \) is computed under the assumption that the mean vector \( \theta \) of \( Z \) lies either in an open set in \( \mathbb{R}^p \) including the point 0, corresponding to the case when \((\theta_1^*, \ldots, \theta_p^*)\) is an interior point in \( \Omega \) with reference to the distribution of \( X \), or in a product of open and half-closed intervals of the type \((-\infty, \infty) \times [0, \infty) \), whose nature depends on which true parameters \( \theta_i^* \)'s are the boundary points \( \theta_{i0} \)'s of \( \theta \) with reference to the distribution of \( X \). Thus, if \( \theta_i^* = \theta_{i0} \) for some \( i \) in the distribution of \( X \), we take \( \theta_i \in [0, \infty) \) in the distribution of \( Z \) while maximizing its pdf wrt \( \theta_i \). Otherwise, we take \(-\infty < \theta_i < \infty \).
A further explanation of the interpretation of \( \theta \) in the distributions of \( X \) and \( Z \) is in order here. If the true parameter point \( \theta^* \) in the distribution of \( X \) is an interior point of \( \Omega \), \( \theta \) in the distribution of \( Z \) is also taken as an interior point of \( \mathbb{R}^p \) with center at 0. Of course, in this case, the MLE of \( \theta \) based on \( Z \) is \( Z \) itself which is \( \mathcal{N}(0, \Sigma(\theta^*)) \), and this results in the well known asymptotic normality of \( \hat{\theta}_N \) with mean at the true value \( \theta^* \) and dispersion \( \Sigma(\theta^*)/N \).

On the other hand, if the true parameter point \( \theta^* \) in the distribution of \( X \) is a boundary point of \( \Omega \) in the sense that some or all coordinates of \( \theta^* \) lie on the boundary of \( \Omega \), \( \theta \) in the distribution of \( Z \) would have an exactly similar interpretation with reference to 0 as the boundary point (thus \( \theta_i^* \) being equal to \( \theta_{i0} \) in case of being a boundary point in \( X \) would mean \( \theta_i = 0 \) in the \( Z \) distribution and similarly \( \theta_i \geq \theta_{i0} \) in the distribution of \( X \) would mean \( \theta_i \geq 0 \) in the distribution of \( Z \)). When in fact \( \theta_i^* \geq \theta_{i0} \), the maximization of the normal pdf of \( Z \) is carried out wrt \( \theta \), taking \( \theta_i \geq 0 \). Under these conditions, it is no longer true that \( Z \) is the MLE of \( \theta \) based on \( Z \), and hence the asymptotic normality of \( \hat{\theta}_N \) does not hold! Moran (1971), Chant (1974) and, most importantly, Self and Liang (1987) have clearly spelled out the MLEs of \( \theta \) based on \( Z \) under various scenarios of boundary points. Once the MLEs of \( \theta \) based on \( Z \) are identified and their joint distribution under normality of \( Z \) is derived, this would provide the asymptotic joint distribution of \( \sqrt{N}(\hat{\theta}_{1N} - \theta_{10}, \ldots, \hat{\theta}_{pN} - \theta_{p0}^*) \). Because of the nature of \( \Omega \) assumed above, it will hold that \( \hat{\theta}_{iN} \geq \theta_{i0} \), for all \( i \). Thus, in the exact and asymptotic distributions of \( (\hat{\theta}_{iN} - \theta_{i0}^*) \), it will hold that this difference can assume both positive and negative values when \( \theta_i^* \) is an interior point of \( \Omega_i = (\theta_{i0}, \infty) \), and is nonnegative when \( \theta_i^* = \theta_{i0} \), the boundary point of \( \Omega_i = [\theta_{i0}, \infty) \).

One major goal of this paper is to further develop this part when we have one, two or three boundary points, clearly explaining the joint distribution of the resultant MLEs of \( \theta \) based on \( Z \) which then readily yields the asymptotic joint distribution of the MLEs of \( \theta \) under \( X \). This knowledge is useful when one is interested in developing Wald-type inference about a smooth function of \( \theta \) based on the MLEs \( \hat{\theta}_N \) of \( \theta \).

To be specific, we establish the following results in subsection 2.1 concerning the asymptotic distribution of the MLE \( \hat{\theta}_N \) of \( \theta \), which is computed by maximizing the likelihood function \( L(\theta|X) \) wrt \( \theta \in \Omega \).

1. The asymptotic joint distribution of \( \sqrt{N}(\hat{\theta}_{1N} - \theta_{10}, \hat{\theta}_{2N} - \theta_{20}, \ldots, \hat{\theta}_{pN} - \theta_{p0}) \),
\( \theta_i^* \) with one boundary point \( \theta_i^* = \theta_{i0}, \Omega_i = [\theta_{i0}, \infty), \theta_i^* > \theta_{i0}, \Omega_i = (\theta_{i0}, \infty), i \geq 2. \)

2. The asymptotic joint distribution of \( \sqrt{N}(\hat{\theta}_{1N} - \theta_{10}, \hat{\theta}_{2N} - \theta_{20}, \hat{\theta}_{3N} - \theta_{30}, \ldots, \hat{\theta}_{pN} - \theta_{p0}) \) with two boundary points \( \theta_i^* = \theta_{i0}, \Omega_i = [\theta_{i0}, \infty), \theta_j^* > \theta_{j0}, \Omega_j = (\theta_{j0}, \infty), j \geq 3. \)

3. The asymptotic joint distribution of \( \sqrt{N}(\hat{\theta}_{1N} - \theta_{10}, \hat{\theta}_{2N} - \theta_{20}, \hat{\theta}_{3N} - \theta_{30}, \hat{\theta}_{4N} - \theta_{40}, \ldots, \hat{\theta}_{pN} - \theta_{p0}) \) with three boundary points \( \theta_i^* = \theta_{i0}, \Omega_i = [\theta_{i0}, \infty), i = 1, 2, 3, \theta_j^* > \theta_{j0}, \Omega_j = (\theta_{j0}, \infty), j \geq 4. \)

In each case, we also describe the asymptotic distribution of \( \sum c_i \hat{\theta}_{iN} \) which can be used to draw suitable inference about \( \sum c_i \theta_i \). In particular, these results can be directly used to derive Wald-type tests of hypotheses concerning linear functions of \( \theta \), without an appeal to the alternative profile likelihood method. We will illustrate this point later by some applications.

We now describe another important aspect of the papers by Moran (1971), Chant (1974), and Self and Liang (1987) in the context of the derivation and properties of the likelihood ratio tests for \( \theta \) based on \( X \) when some parameter points may lie on the boundary. To make matters simple and easy to understand, let us consider the following two testing problems about just one component, say \( \theta_1 \), of \( \theta \). It should be noted that, as mentioned in Section 1, via a re-parametrization the testing problem about a linear function of \( \theta \) in the context of multistage dose-response Weibull models can be reduced to testing hypotheses about just one component of \( \theta \). In fact, via re-parametrization, a testing problem about any smooth function of \( \theta \) in a general multi-parameter model can be reduced to that concerning a single component of reparametrized \( \theta \).

**Problem 1.** Test the null hypothesis \( H_0 : \theta_1 = \theta_{10} + \delta \) versus \( H_1 : \theta_1 \neq \theta_{10} + \delta \) where \( \delta > 0 \). Here naturally \( \theta_1 \) is an interior point of \( \Omega_1 \).

**Problem 2.** Test the null hypothesis \( H_0 : \theta_1 = \theta_{10} \) versus \( H_1 : \theta_1 > \theta_{10} \) where \( \theta_{10} \) is the boundary point of \( \Omega_1 = [\theta_{10}, \infty) \).

Referring to the multistage dose-response Weibull models, testing problem 1 is much more relevant to us than the testing problem 2, especially after re-parametrization when we seek confidence bounds for a benchmark dose. However, generally speaking, in presence of many parameters in a model.
where some of them may in fact lie on the boundaries, testing problem 2 also makes sense and should be performed in order to reduce the number of parameters in a model not lying on the boundaries. Thus, for example, in a linear dose-response Weibull model mentioned in Section 1, we might be interested in testing $H_0: \theta_1 = 0$, a boundary point of $\theta_1$, which essentially means absence of the regression coefficient! Again, for a quadratic or cubic dose-response Weibull model as described in Section 1, we might be interested in testing if higher order regression coefficients such as $\theta_2$ and $\theta_3$ are 0, which are the boundary points of $\Omega_2 = [\theta_2, \infty)$ and $\Omega_3 = [\theta_3, \infty)$.

It is well known that when both $\Omega$, the general parameter space, and $\Omega_0$, the parameter space under a null hypothesis, are smooth in the sense of being open sets in $\mathbb{R}^p$ and in a lower dimensional subspace, respectively, the asymptotic distribution of $-2\ln(LRT)$ under the null hypothesis is central chi-square with an appropriate df. However, this result is far from being true when some parameters may lie on the boundary either under the null hypothesis or even otherwise. Thus, in the testing problem 2, which is concerned with the boundary point of $\Omega_1$, the classical asymptotic distributional result of the $LRT$ is not true. Also, even in the case of testing problem 1 which deals with an interior point of $\Omega_1$, it is quite possible that some of the nuisance parameters (unknown under the null hypothesis) do in fact lie on the boundaries, which may also affect the classical asymptotic distributional result of the $LRT$. It is precisely in the context that $\Omega$ and $\Omega_0$ may not be open sets in $\mathbb{R}^p$ that we have the following general result primarily due to Self and Liang (1987). We remark that Self and Liang’s (1987) results on the asymptotic distribution of the $LRT$ are quite general, but here we state the results keeping in mind the two null hypotheses given above under Problem 1 and Problem 2.

**Proposition 2.** (a) Consider the problem of testing the null hypothesis $H_0: \theta_1 = \theta_{10} + \delta$ versus $H_1: \theta_1 \neq \theta_{10} + \delta$, for some $\delta > 0$ when $\Omega_1 = (\theta_{10}, \infty)$, and suppose we compute the $LRT$ by maximizing $L(\theta|X)$ wrt $\theta \in \Omega$ and also under $H_0$. Write $\Omega_0 = \{\theta: \theta_1 = \theta_{10} + \delta, \theta_2, \ldots, \theta_p, unspecified\}$. Then the null distribution of the profile log likelihood based $LRT$ using $X$, namely, the null distribution of

$$-2\ln \left[ \frac{\max_{\theta \in \Omega_0} L(\theta|X)}{\max_{\theta \in \Omega} L(\theta|X)} \right]$$

is asymptotically equivalent to the distribution of the profile log likelihood.
based LRT using $Z$ under $N[\theta, \Sigma]$, namely, the distribution of

$$\min_{\theta \in \Omega^*_0} Q(\theta|Z) - \min_{\theta \in \Omega^*} Q(\theta|Z)$$

(9)

when $\theta = 0$. Here $Q(\theta|Z) = (Z - \theta)'[\Sigma]^{-1}(Z - \theta)$ is the exponent of the normal likelihood of $Z$ with mean $\theta$ and dispersion $\Sigma$, minimum under $\theta \in \Omega^*_0$ is computed when $\theta_1 = \delta$ and $\theta_2, \ldots, \theta_p$ are unspecified in $(-\infty, \infty)$, and minimum under $\theta \in \Omega^*$ is computed when all the parameters $\theta$ are unspecified in $(-\infty, \infty)$. In other words, in general, $\Omega^*$ is a translation of $\Omega$ by $\Omega_0$ so that $\Omega^*$ includes the point 0 in the parameter space of $Z$.

(b) Consider the problem of testing the null hypothesis $H_0 : \theta_1 = \theta_{10}$ versus $H_1 : \theta_1 > \theta_{10}$, and suppose we compute the LRT by maximizing $L(\theta|X)$ wrt $\theta \in \Omega$ and also under $H_0$. Write $\Omega_0 = \{\theta : \theta_{10}, \theta_2, \ldots, \theta_p \text{ unspecified}\}$. Then the null distribution of the profile log likelihood based LRT using $X$, namely, the null distribution of

$$-2\ln \left[ \frac{\max_{\theta \in \Omega_0} L(\theta|X)}{\max_{\theta \in \Omega} L(\theta|X)} \right]$$

(10)

is asymptotically equivalent to the distribution of the profile log likelihood based LRT using $Z$ under $N[\theta, \Sigma]$, namely, the distribution of

$$\min_{\theta \in \Omega^*_0} Q(\theta|Z) - \min_{\theta \in \Omega^*} Q(\theta|Z)$$

(11)

when $\theta = 0$. Here, as before, $Q(\theta|Z) = (Z - \theta)'[\Sigma]^{-1}(Z - \theta)$ is the exponent of the normal likelihood of $Z$ with mean $\theta$ and dispersion $\Sigma$, minimum under $\theta \in \Omega^*_0$ is computed when $\theta_1 = 0$ and $\theta_2, \ldots, \theta_p$ are unspecified in $(-\infty, \infty)$, and minimum under $\theta \in \Omega^*$ is computed when $\theta_1 \geq 0$ and all the other parameters $\theta$ are unspecified in $(-\infty, \infty)$. In other words, as before, $\Omega^*$ is a translation of $\Omega$ by $\Omega_0$ so that $\Omega^*$ includes the point 0 in the parameter space of $Z$.

**Remark 2.1.** It should be noted that the null hypotheses mentioned above are concerned only with $\theta_1$ and do not mention anything about the nuisance parameters $\theta_2, \ldots, \theta_p$. It is quite possible that some of the nuisance parameters may lie in a parameter space containing the boundary points. In other words, it is possible that $\theta_i \in \Omega_i = [\theta_{10}, \infty)$ for $i > 1$. When this happens, it is implied that the minimization wrt $\theta$ in the quadratic form $Q(\theta|Z)$ is done under the restriction that $\theta_i \geq 0$, implying that 0 is a boundary point wrt $\theta_i$ rather than being an interior point.
Moran (1971), Chant (1974) and, most importantly, Self and Liang (1987) discussed at length computation of the LRT based on $Z$ and its null distribution under various forms of the null hypothesis well beyond the two cases mentioned above. In general, as remarked earlier, it follows that the null distribution of the LRT based on $Z$ is central chi-square when $\Omega_0^*$ consists of interior points of the mean vector $\theta$ in the normal distribution of $Z$. However, this distribution is far from being chi-square (quite often a mixture of chi-squares) when $\Omega_0^*$ contains some boundary points of $\theta$ as in problem 2 above. Let us emphasize that, under the mapping of the parameter space of $\theta$ from the distribution of $X$ to that of $Z$, the range of values of the parameter $\theta_i$ in the distribution of $X$, namely, $\Omega_i = [\theta_{i0}, \infty)$ or $(\theta_{i0}, \infty)$ changes to $\Omega_i^* = [0, \infty)$ or $(-\infty, \infty)$ as the new range of values of $\theta_i$ in the distribution of $Z$.

A second major objective of this paper is to discuss in detail the null distribution of the LRT based on $Z$ for testing the two hypotheses about $\theta_1$ mentioned above under problem 1 and problem 2 when it is likely that there are none or some parameters lying on the boundaries, thus supplementing the earlier works of Moran (1971), Chant (1974) and Self and Liang (1987).

Referring to Proposition 1, we should note that once the MLEs of $\theta$ based on $Z$ under $\Omega^*$ and their joint distribution are derived, we can use Proposition 1 to approximate the joint as well as the marginal distributions of the actual MLEs of $\theta$ based on $X$. These distributions can then be effectively used to draw suitable Wald-type inference about a smooth function of $\theta$ such as tests for a single component or a linear function of the components of $\theta$, thus providing an alternative to the profile likelihood approach.

Likewise, referring to Proposition 2, once the LRT of $H_0$ versus $H_1$ based on $Z$ is derived and its null distribution is obtained, we can use it to get the approximate cut-off points of the LRT based on $X$. We remark that this reduction of the original inference problem based on $X$ with an arbitrary distribution to a canonical form using $Z$ which has a normal distribution, though only asymptotically valid, is a key feature of the asymptotic theory and the spirit of all the earlier works of these authors.

We are now in a position to describe the main results of this section. Based on the distributional assumption $Z \sim N_p(\theta, \Sigma)$, where $\theta$ is unknown and $\Sigma$ is positive definite known, we develop some new results for exact inference on $\theta_1$ in presence of a few nuisance parameters, allowing the possibility
that some of the nuisance parameters may lie on the boundaries.

Towards deriving the LRT of $H_0 : \theta_1 = \delta$ versus $H_1 : \theta_1 \neq \delta$ for some $\delta > 0$ based on $Z$, which corresponds to testing $H_0 : \theta_1 = \theta_{10} + \delta$ versus $H_1 : \theta_1 \neq \theta_{10} + \delta$ for some $\delta > 0$ in the original distribution of $X$, we note that the likelihood function of $Z$, namely, $L(\theta|Z)$ can be written as

$$L(\theta|Z) = K \exp[-(Z - \theta)'\Sigma^{-1}(Z - \theta)].$$

(12)

Since we assume that the $p$ parameters are functionally independent, for testing $H_0 : \theta_1 = \delta$ versus $H_1 : \theta_1 \neq \delta$ based on $Z$, it is well known that if the parameter vector $\theta$ is an interior point in $R^p$, which means the true parameter values of the parameter $\theta_1$ of interest as well as the nuisance parameters $\theta_2, \ldots, \theta_p$ are interior points, then the usual normal test based on $Z_1$ is the LRT and it provides a valid test. Recall that the dispersion matrix of $Z$ is assumed known which results in a normal test rather than a $t$-test. However, if the parameter of interest $\theta_1$ or some or all of the nuisance parameters happen to be the boundary points of their respective intervals, such a result is far from being true. It should be clearly understood that the notion of some points being on the boundary means that while true values of such parameters are unknown, we first assume that the parameter space of each such parameter is left-closed and secondly, we allow the possibility that these parameters can assume their respective left end points. Hence, parameters with values in an open interval are considered as parameters with interior points and we will call them as interior parameters.

As mentioned earlier, Self and Liang (1987), based on previous works of Chernoff (1954), Chant (1974), Moran (1971) and Shapiro (1985), developed appropriate solutions to this kind of problem for a wide variety of scenarios involving several parameters under $H_0$, and allowing some of them and also some nuisance parameters to be on the boundary. In the huge literature on mathematical statistics, Self and Liang (1987)’s paper is indeed a landmark paper with a novel contribution to this important problem particularly because it is quite common that some parameters of interest as well as some nuisance parameters in a typical multi-normal set-up can indeed lie on the boundary and also because of the fact that such situations often arise in applications.

Following essentially Self and Liang’s (1987) ideas, we derive below the likelihood ratio tests (LRTs) of $H_0$ versus $H_1$ based on $Z$, allowing one, two
and three nuisance parameters to be on the boundary. It turns out that, as one can expect, the form of the LRT becomes quite complex with the increase in the number of nuisance parameters which lie on the boundary. In the sequel, we also derive the maximum likelihood estimates of all parameters $\theta$ based on $Z$ under the condition that some of the parameters may lie on the boundary, and derive their joint and marginal distributions. As mentioned before, these results would be useful when one is interested in deriving Wald-type asymptotic inference based on $X$ about a smooth function of $\theta$, and in particular, about a linear function of these parameters which is the case for dose-response multistage Weibull models mentioned in Section 1.

Some standard results from classical multivariate analysis which are needed in the sequel are listed below. Write $Z = (Z_{(1)}, Z_{(2)}), \theta = (\theta_{(1)}, \theta_{(2)})$ where $Z_{(1)}: q \times 1, Z_{(2)}: (p-q) \times 1, \theta_{(1)}: q \times 1, \theta_{(2)}: (p-q) \times 1$, and $\Sigma_{(11)} = \text{var}(Z_{(1)}), \Sigma_{(22)} = \text{var}(Z_{(2)}), \text{and } \Sigma_{(12)} = \text{cov}(Z_{(1)}, Z_{(2)}).

1. Distribution of a quadratic form:

$$[Z - \theta]'(\Sigma)^{-1}[Z - \theta] \sim \chi^2_p$$  \hspace{1cm} (13)

2. Marginal distribution:

$$Z_1 \sim N_q[\theta_{(1)}, \Sigma_{(11)}]$$  \hspace{1cm} (14)

3. Conditional distribution of $Z_1$, given $Z_2 = z_2$:

$$Z_1|z_2 \sim N_q[\theta_1 + \Sigma_{(12)}(\Sigma_{(22)})^{-1}(z_2 - \theta_2), \Sigma_{(11,2)} = \Sigma_{(11)} - \Sigma_{(12)}(\Sigma_{(22)})^{-1}\Sigma_{(21)}]$$  \hspace{1cm} (15)

\textbf{2.1. One parameter on the boundary}

In this subsection we assume that there is one parameter, say $\theta_i$, which may lie on the boundary in the sense that $\theta_i \in \Omega_i = [\theta_{i0}, \infty)$, a half-closed interval, and carry out appropriate inference about $\theta_i$. Before we develop inferential tools, let us make a remark about testing for the existence of a boundary parameter.

Consider a general multi-parameter model based on $X$ involving $p$ parameters $\theta = (\theta_1, \ldots, \theta_p) \in \Omega$ where $\Omega = \Omega_1 \times \cdots \Omega_p$ with $\Omega_i = [\theta_{i0}, \infty)$
or \( (\theta_{i0}, \infty) \) for all \( i \). Suppose it is suspected that one parameter lies on the boundary! How do we determine which one? Here is an ad hoc approach. Assuming that \( \theta_i = \theta_{i0} \) is the point on the boundary, we can maximize the likelihood \( L(\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_p|\theta_i = \theta_{i0}, \mathbf{X}) \) wrt \( \theta_j \in \Omega_j = (\theta_{j0}, \infty) \) for all \( j \neq i \) and compute the maximum value of the likelihood, say \( L_i \). Comparing \( L_1, \ldots, L_p \) and selecting the index \( k \) such that \( L_k = \max(L_1, \ldots, L_p) \), we can conclude that \( \theta_k \) lies on the boundary. Alternatively, we can test the hypothesis \( H_{0i} : \theta_i = \theta_{i0} \) versus \( H_{1i} : \theta_i > \theta_{i0} \) by computing the LRT statistic, say \( \lambda_i \), and choose the index \( k \) for which \( \lambda_k \) is the smallest with the conclusion that \( \theta_k \) is likely to lie on the boundary. Details about such a test are given below.

Let us consider the two testing problems about \( \theta_1 \) mentioned earlier, under the assumption that there is one boundary parameter. We will derive exact tests based on \( \mathbf{Z} \) which would yield asymptotic tests based on \( \mathbf{X} \). Consider testing \( H_0 : \theta_1 = \delta \) versus \( H_1 : \theta_1 \neq \delta \) based on \( \mathbf{Z} \sim N[\boldsymbol{\theta}, \Sigma] \). We distinguish between two cases depending on whether the parameter on the boundary is the parameter of interest (\( \theta_1 \)) or a nuisance parameter, say \( \theta_2 \). In the former case, we take \( \delta = 0 \) as the boundary point of \( \Omega_1^* = [0, \infty) \) and write \( H_1 : \theta_1 > 0 \) while the other parameters are free. In the latter case, we take \( \delta > 0 \) as an interior point of \( \Omega_1^* = (0, \infty) \), and assume that one of the remaining unspecified nuisance parameters, say \( \theta_2 \), may lie on the boundary in the sense that \( \theta_2 \in [0, \infty) \). Note that, by definition, a nuisance parameter can never be known so that we cannot conclude that \( \theta_2 = 0 \)!

**Case 1:** \( \theta_1 \) is a boundary point. This has been discussed at length in Self and Liang (1987) who derived the LRT of this problem. It turns out that the LRT is based on \( Z_1^2 I[Z_1 > 0]/\sigma_{11} \) and its null distribution is a 50 : 50 mixture of \( \chi_0^2 \) and \( \chi_1^2 \) distributions. Details are omitted.

Interestingly enough, a Wald-type test based on the MLE of \( \theta_1 \), which turns out to be equivalent to the LRT described in Self and Liang (1987), can be easily developed as follows. For testing \( H_0 : \theta_1 = \theta_{10} \) versus \( H_1 : \theta_1 > \theta_{10} \) based on \( \mathbf{X} \), it seems quite reasonable to reject \( H_0 \) for large values of \( \hat{\theta}_{1N} \), the MLE of \( \theta_1 \), i.e., when \( \hat{\theta}_{1N} > c_N \) for some \( c_N \). To determine the value of \( c_N \) for a given level of significance \( \alpha \), it follows that, asymptotically, \( \alpha = P[\hat{\theta}_{1N} > c_N|H_0] = P[\sqrt{N}(\hat{\theta}_{1N} - \theta_{10}) > \sqrt{N}(c_N - \theta_{10})] \sim P[Z_1 I(Z_1 > 0) > \sqrt{N}(c_N - \theta_{10})]. \) This readily yields: \( \sqrt{N}(c_N - \theta_{10}) = \sigma_1 z_{2\alpha} \) where \( z_{2\alpha} \) is the upper \( 2\alpha \) cut-off point of a standard normal distribution. Hence we
reject $H_0$ when $\hat{\theta}_{1N} > \theta_{10} + \sigma_1 z_{2\alpha}/\sqrt{N}$. This is of course a very easy test to carry out without the need to compute a profile likelihood which is the basis of LRT. In the above, we have used the fact that in the context of the distribution of $Z$, $\hat{\theta}_1 = Z_1 I[Z_1 > 0]$ is the MLE of $\theta_1$ for $\theta_1 \geq 0$, and its distribution is a mixture of a point mass at 0 with probability 0.5 and a half-normal distribution $0.5 N[0, 1]$, and that this is indeed the asymptotic distribution of $\sqrt{N}(\hat{\theta}_{1N} - \theta_{10})/\sigma_1$. It should be noted that here $\hat{\theta}_{1N}$ is the MLE of $\theta_1$ when the likelihood $L(\theta|X)$ of $X$ is maximized wrt $\theta$ under the assumption that $\theta_1 \in \Omega_1 = [\theta_{10}, \infty)$ and $\theta_j \in (\theta_{j0}, \infty)$ for $j > 1$.

**Case 2.** $\theta_1$ is an interior point and $\theta_2$ is a boundary point. To derive the LRT for $H_0 : \theta_1 = \delta + \theta_{10}$ versus $H_1 : \theta_1 \neq \delta + \theta_{10}$ for some $\delta > 0$ on the basis of $X$, note that, asymptotically, this is equivalent to testing $H_0 : \theta_1 = \delta$ versus $H_1 : \theta_1 \neq \delta$ for $\delta > 0$ based on $Z$. We proceed to apply Proposition 2. It is clear from the expression of the likelihood function $L(\theta|Z)$ given in (12) that what matters is the quadratic form $Q(\theta|Z)$ given by

$$Q(\theta|Z) = (Z - \theta)' \Sigma^{-1} (Z - \theta).$$

Writing $Q(\theta|Z) = Q(\theta_1, \theta_2|Z_1, Z_2) + Q(Z_3, \ldots, Z_p|Z_1, Z_2; \theta)$ where the first part is the marginal bivariate quadratic of $(Z_1, Z_2)$ and the second part is the $(p - 2)$-dimensional conditional quadratic of $(Z_3, \ldots, Z_p)$, given $(Z_1, Z_2)$, it follows from Self and Liang (1987) that due to the interior nature of the parameters $\theta_3, \ldots, \theta_p$, the only part we need to study is the first part, and maximization of the likelihood corresponds to finding the two minimums of the first part, one under the union of null and alternative hypotheses, and the other under the null hypothesis.

To derive the maximum likelihood estimates (MLEs) of the parameters $\theta$ under the union of null and alternative parameter spaces, we can express $Q(\theta_1, \theta_2|Z_1, Z_2)$ as

$$Q(\theta_1, \theta_2|Z_1, Z_2) = Q(Z_2|\theta_1) + Q(Z_1|Z_2; \theta_1, \theta_2)$$

(17)

where $Q(Z_2|\theta_1)$ is the marginal pdf of $Z_2$ and $Q(Z_1|Z_2; \theta_1, \theta_2)$ is the conditional pdf of $Z_1$, given $Z_2$. It is clear from the discussion in Self and Liang (1987) that due to the interior nature of $\theta_1$, the minimum value of $Q(Z_1|Z_2; \theta_1, \theta_2)$ wrt $\theta_1$ is 0, irrespective of the value of $\theta_2$. Now a minimization of $Q(Z_2|\theta_1)$ wrt $\theta_2$ subject to $\theta_2 \geq 0$ readily yields:
\[ \min_{\theta_2 \geq 0} Q(Z_2 | \theta_2) = \begin{cases} 0 & \text{if } Z_2 > 0 \\ \frac{Z_2^2}{\sigma_{22}} & \text{if } Z_2 < 0. \end{cases} \] (18)

To minimize \( Q(\theta_1, \theta_2 | Z_1, Z_2) \) wrt \( \theta_2 \) under the null hypothesis when \( \theta_1 = \delta \), we write

\[ Q(\theta_1, \theta_2 | Z_1, Z_2) = Q(Z_1 | \theta_1 = \delta) + Q(Z_2 | Z_1; \theta_1 = \delta, \theta_2) \] (19)

where \( Q(Z_1 | \theta_1 = \delta) \) is the pdf of \( Z_1 \) under the null hypothesis and \( Q(Z_2 | Z_1; \theta_1 = \delta, \theta_2) \) is the pdf of \( Z_2 \), given \( Z_1 \) when \( \theta_1 = \delta \). Since the first term is independent of \( \theta_2 \), it is clear that the minimum value of \( Q(\theta_1, \theta_2 | Z_1, Z_2) \) wrt \( \theta_2 \) arises essentially from minimizing \( Q(Z_2 | Z_1; \theta_1 = \delta, \theta_2) \) wrt \( \theta_2 \). Since

\[ Q(Z_2 | Z_1; \theta_1 = \delta, \theta_2) = \frac{\left[ Z_2 - \theta_2 - \rho \frac{Z_2}{\sigma_1}(Z_1 - \delta) \right]^2}{\sigma_{22}(1 - \rho^2)} \] (20)

minimization of \( Q(Z_2 | Z_1; \theta_1 = \delta, \theta_2) \) wrt \( \theta_2 \) under the condition \( \theta_2 \geq 0 \) readily gives

\[ \min_{\theta_2 \geq 0} [Q(Z_2 | Z_1; \theta_1 = \delta, \theta_2)] = \begin{cases} 0 & \text{if } Z_{2,1} > 0 \\ \frac{Z_{2,1}^2}{\sigma_{22}(1 - \rho^2)} & \text{if } Z_{2,1} < 0 \end{cases} \] (21)

where \( Z_{2,1} = Z_2 - \rho \frac{Z_2}{\sigma_1}(Z_1 - \delta) \). Hence we get

\[ \min_{\theta_1 = \delta, \theta_2 \geq 0} Q(\theta_1, \theta_2 | Z_1, Z_2) = \begin{cases} \frac{(Z_1 - \delta)^2}{\sigma_{11}} & \text{if } Z_{2,1} > 0 \\ \frac{(Z_1 - \delta)^2}{\sigma_{11}} + \frac{Z_{2,1}^2}{\sigma_{22}(1 - \rho^2)} & \text{if } Z_{2,1} < 0 \end{cases} \] (22)

Combining (18) and (22), and taking the difference, we get

\[ -2 \ln(LRT) = W \text{(say)} \] as

\[ W = \begin{cases} (Z_1 - \delta)^2/\sigma_{11} & \text{if } Z_2 > 0, Z_{2,1} > 0 \\ (Z_1 - \delta)^2/\sigma_{11} - Z_2^2/\sigma_{22} & \text{if } Z_2 < 0, Z_{2,1} > 0 \\ Z_{1,2}^2/\sigma_{11}(1 - \rho^2) & \text{if } Z_2 < 0, Z_{2,1} < 0 \\ \frac{(Z_1 - \delta)^2}{\sigma_{11}} + \frac{Z_{2,1}^2}{\sigma_{22}(1 - \rho^2)} & \text{if } Z_2 > 0, Z_{2,1} < 0. \end{cases} \] (23)
It is easy to verify that when \( \rho = 0 \), \( W \) reduces to \( (Z_1 - \delta)^2 / \sigma_{11} \) which has a chi-square distribution with 1 d.f. under \( H_0 \), a familiar result.

To derive the null distribution of \( W \) for any given \( \rho \), we assume without loss of generality that \( \sigma_1 = \sigma_2 = 1 \) and \( \delta = 0 \). It is proved in the Appendix that the cdf of \( W \) is given by the following.

**Theorem 2.1.** The cdf \( G(w) \) of \( W \), for \( 0 < w < \infty \), is given by the sum of four parts:

(i)

\[
First \ part = \int_0^{\rho \sqrt{w}} \left[ \int_{-\sqrt{(w-v^2)(1-\rho^2)}}^{v+\sqrt{w}} N(0,1)dx \right] N(0,1)dv \tag{24}
\]

(ii)

\[
Second \ part = \int_{\rho \sqrt{w}}^{\infty} \left[ \int_{v-\sqrt{w}}^{v+\sqrt{w}} N(0,1)dx \right] N(0,1)dv \tag{25}
\]

(iii)

\[
Third \ part = \int_{-\rho \sqrt{w/(1-\rho^2)}}^{0} \left[ \int_{-\sqrt{w(1-\rho^2)}}^{v+\sqrt{w+v^2}} N(0,1)dx \right] N(0,1)dv \tag{26}
\]

(iv)

\[
Fourth \ part = \left[ \int_{-\sqrt{w(1-\rho^2)}}^{\sqrt{w(1-\rho^2)}} N(0,1)dx \right] \cdot \left[ \int_{-\rho \sqrt{w/(1-\rho^2)}}^{\infty} N(0,1)dx \right] \tag{27}
\]

The above distribution can be used to get a cut-off point of the statistic \(-2\ln(LRT)\) which can then be used to carry out the LRT based on \( X \). Quite surprisingly, our simulations carried out in the next section indicate that the above distribution does not depend on \( \rho \)!

We now discuss the application of Wald-type test for testing \( H_0 \) versus \( H_1 \). Towards this end, we note from Self and Liang (1987) that the MLEs of the parameters \( \theta_i \) when \( \theta_2 \) may be on the boundary are given by the following. Define \( Z_{i,2} = Z_i - \frac{\alpha \rho_2}{\sigma_2} Z_2 \).

**Theorem 2.2.** The MLE of \( \theta_i, \ i = 1, 3, \ldots, p \) when \( \theta_2 \) is on the boundary is given by
\[ \hat{\theta}_i = Z_i I[Z_2 > 0] + Z_{i,2} I[Z_2 < 0]. \] (28)

The marginal distribution of the MLE of \( \theta_i \), which depends on the correlation \( \rho_{i2} \) between \( Z_i \) and \( Z_2 \), is given below. This would be useful if one is interested in drawing suitable inference about just one parameter \( \theta_i \).

**Theorem 2.3.** The pdf of the MLE \( \hat{\theta}_i = U_i \) of \( \theta_i \) is given by

\[
f(u_i) = \int_0^\infty e^{-\left[\frac{x^2 + \frac{u_i^2}{\sigma_{ii}^2} - 2xu_i\rho_{i2}/\sigma_{i}}{2(1-\rho_{i2}^2)}\right]} dx 
+ e^{-\frac{u_i^2}{2\sigma_{ii}^2(1-\rho_{i2}^2)}} 2\sqrt{2\pi\sigma_{ii}(1-\rho_{i2}^2)}.
\] (29)

Proof. The pdf of \( U_i \) is derived from the cdf of \( U_i \) which is readily obtained as follows. Since \( Z_2 \) and \( Z_{i,2} \) are independent and \( P(Z_2 < 0) = 0.5 \), the second part is obvious. For the first part, we first write down the bivariate normal pdf of \( Z_2 \) and \( Z_i \), and then integrate out \( Z_2 \) over \((0, \infty)\). The pdf of \( U_i \) for the first part is then directly obtained from the cdf of the first part upon differentiation wrt \( u_i \). This completes the proof.

**Remark 2.2.** Specializing to the case of testing \( H_0 : \theta_1 = \delta + \theta_{10} \) versus \( H_1 : \theta_1 = \delta + \theta_{10} \) based on \( X \), we can easily derive a Wald-type test based on the MLE of \( \theta_1 \), namely, \( \hat{\theta}_{1N} \) which is the standard MLE of \( \theta \) based on \( X \) under \( \theta \in \Omega \) with \( \Omega_2 = [\theta_{20}, \infty) \) and \( \Omega_i = (\theta_{i0}, \infty) \) for \( i \neq 2 \). Because of the nature of \( H_0 \) and \( H_1 \), it makes sense to reject \( H_0 \) for large values of \( \hat{\theta}_{1N} \), i.e., when \( \hat{\theta}_{1N} > c_N \) for some \( c_N \). To determine the value of \( c_N \) for a given level of significance \( \alpha \), it follows that, asymptotically, \( \alpha = P[\hat{\theta}_{1N} > c_N|H_0] = P[\sqrt{N}(\hat{\theta}_{1N} - \delta - \theta_{10}) > \sqrt{N}(c_N - \delta - \theta_{10})] \sim P[Z_1 I(Z_2 > 0) + Z_{1,2} I(Z_2 < 0) > \sqrt{N}(c_N - \delta)]. \) We can now use the result of Theorem 2.3 to claim: \( \sqrt{N}(c_N - \delta) = z_{\alpha/2} \) where \( z_{\alpha/2} \) is the upper \( \alpha \) cut-off point of the distribution of \( Z_1 I(Z_2 > 0) + Z_{1,2} I(Z_2 < 0) \) given in Theorem 2.3 for \( i = 1 \). Hence we reject \( H_0 \) when \( \hat{\theta}_{1N} > \delta + \theta_{10} + z_{\alpha/2}/\sqrt{N} \). This is of course a very easy test to carry out without the need to compute a profile likelihood which is the basis of LRT. However, we should also note that we have taken a one-sided alternative as \( H_1 \). For a both-sided alternative, we may choose to reject \( H_0 \)
for large values of $|\hat{\theta}_{1N} - \delta - \theta_{10}|$ and determine the cut-off point appropriately. Note that this would provide an alternative approach to the profile likelihood method (LRT).

**Remark 2.3.** If one is interested in making inference about a linear combination of $\theta_i$’s which excludes the boundary point $\theta_2$, we note that the distribution of $U = \sum_{i \neq 2} c_i \hat{\theta}_i = \sum_{i \neq 2} c_i Z_i I[Z_2 > 0] + \sum_{i \neq 2} c_i Z_{i,2} I[Z_2 < 0]$ is again readily obtained as the sum of two parts of which the second part (when $Z_2 < 0$) is normal with mean 0 and variance obtained from the variances and covariances of the residuals $Z_{i,2}$’s, and the other part is a convolution of two normals. The latter term is obtained by writing down the bivariate normal pdf of $(\sum_{i \neq 2} c_i Z_i)$ and $Z_2$, and then integrating out $Z_2$ over $(0, \infty)$. Thus, defining $c^* = (c_1, 0, c_3, \ldots, c_p), \sigma^2 = \sum_{i \neq 2} c_i \rho_{i2} \sigma_i$ and $\Delta^2 = \sum_{i \neq 2} c_i^2 \sigma_i^2 (1 - \rho_i^2) + \sum_{i \neq j \neq 2} c_i c_j \sigma_i \sigma_j (\rho_{ij} - \rho_{i2} \rho_{j2})$, the pdf of $U$ is given by $f(u) = f_I(u) + f_{II}(u)$ where

$$f_I(u) = \frac{e^{-\frac{u^2}{2\Delta^2}}}{2\Delta \sqrt{2\pi}}$$

$$f_{II}(u) = \int_{0}^{\infty} f(x, u) dx / [2\pi \sigma^* \sqrt{1 - \rho^2}]$$

$$f(x, u) = \exp \left[ -\frac{x^2 + \frac{x^2}{2\sigma^2} - 2xu \rho^*}{2(1 - \rho^2)} \right]$$

Let us recall that the distribution of $U$ given here is precisely the asymptotic distribution of $\sqrt{N} \left[ \sum_{i \neq 2} c_i \hat{\theta}_i - \sum_{i \neq 2} c_i \theta_i \right]$ and hence can be readily used for drawing valid asymptotic inference about $\sum_{i \neq 2} c_i \theta_i$. Note that this so-called Wald-type approach avoids computing profile likelihoods for testing hypotheses about $\sum_{i \neq 2} c_i \theta_i$.

**Remark 2.4.** If, on the other hand, we are interested in making inference about a linear function of the $\theta_i$’s which also includes $\theta_2$, naturally we need to derive the distribution of $V = \sum_{i \neq 2} c_i \hat{\theta}_i + c_2 \hat{\theta}_2$. Since $\hat{\theta}_2 = Z_2 I[Z_2 > 0]$, it follows from (28) that the distribution of $V$ will again consist of two parts: one part, corresponding to $Z_2 < 0$, is just normal and is independent of $Z_2$. This is in fact the distribution of $\sum_{i \neq 2} c_i Z_{i,2}$. The other part, for $Z_2 > 0$, is obtained by first deriving the conditional distribution of $\sum_{i \neq 2} c_i Z_i$ which is normal and then convoluting it with $c_2 Z_2$. This argument leads to the pdf
of $V$ as $f(v) = f_I(v) + f_{II}^*(v)$ where $f_I(v)$ is the same as $f_I(u)$, and $f_{II}^*(v)$ is given by

$$f_{II}^*(v) = \int_0^\infty \frac{\exp\left(-\frac{(v-x(\sum_{i=1}^p c_i \rho_i^2 \sigma_i))^2}{2\Delta^2} - \frac{x^2}{2}\right)}{\Delta 2\pi} dx. \quad (31)$$

We recall that the distribution of $V$ given here is precisely the asymptotic distribution of $\sqrt{N}[\sum_{i=1}^p c_i \hat{\theta}_{iN} - \sum_{i=1}^p c_i \theta_i]$ and hence can be readily used for drawing valid asymptotic inference about $\sum_{i=1}^p c_i \theta_i$. Again, this approach avoids the computation of a profile likelihood.

2.2. Two parameters on the boundaries

We now discuss the case of two boundary parameter points and consider two cases.

Case 1. $\theta_1$, the parameter of interest, and, without loss of generality, $\theta_2$ lie on the boundaries so that $\Omega_1 = [\theta_{10}, \infty)$ and $\Omega_2 = [\theta_{20}, \infty)$ and $\Omega_i = (\theta_{i0}, \infty)$, for $i > 2$. The asymptotic joint distribution of $\sqrt{N}(\hat{\theta}_{1N} - \theta_{10}, \hat{\theta}_{2N} - \theta_{20}, \hat{\theta}_{3N} - \theta_3^*, \ldots, \hat{\theta}_{pN} - \theta_p^*)'$ for this case, given below, can be easily derived from Self and Liang (1987). This is given by the joint distribution of $W = (W_1, W_2, W_3, \ldots, W_p)$, which are functions of $Z$, under the distributional assumption $Z \sim N[0, \Sigma]$, where

$$W_1, W_2 = (Z_1, Z_2), \text{ if } Z_1 > 0, Z_2 > 0$$
$$= (Z_{12}, 0), \text{ if } Z_{12} > 0, Z_2 < 0$$
$$= (0, Z_{21}), \text{ if } Z_{21} > 0, Z_1 < 0$$
$$= (0, 0), \text{ if } Z_{12} < 0, Z_{21} < 0. \quad (32)$$

and

$$(W_3, \ldots, W_p)' = (Z_3, \ldots, Z_p)' - \Gamma \hat{\Sigma}^{-1}(Z_1 - W_1, Z_2 - W_2)' \quad (33)$$

where the matrix $\Sigma$ is partitioned as

$$\Sigma = \begin{pmatrix}
\bar{\Sigma} : 2 \times 2 & \Gamma : (p-2) \times 2 \\
\Gamma' : 2 \times (p-2) & \Sigma_{p-2} : (p-2) \times (p-2)
\end{pmatrix} \quad (34)$$

As in the case of one boundary point, it is indeed possible to develop a test for $H_0 : \theta_1 = \theta_{10}$ versus $H_1 : \theta_1 > \theta_{10}$ based on the asymptotic
marginal distribution of $\hat{\theta}_{1N}$. Since it is reasonable to reject $H_0$ for large values of $\hat{\theta}_{1N}$, i.e., when $\hat{\theta}_{1N} > c_N$ for some $c_N > \theta_{10}$ or, equivalently, when $\sqrt{N}(\hat{\theta}_{1N} - \theta_{10}) > d_N$ for some $d_N > 0$, the test can be carried out by determining the value of $d_N$ for a given level of significance $\alpha$. Note that the asymptotic null distribution of $\sqrt{N}(\hat{\theta}_{1N} - \theta_{10})$ is precisely the distribution of $W_1$ when $Z \sim N[0, \Sigma]$, which is easy to derive.

**Case 2.** Assume without any loss of generality that $\theta_2$ and $\theta_3$ lie on the boundary and our primary interest lies in the parameter $\theta_1$ which is an interior point. In the sequel, we consider both LRT and Wald-type tests for $H_0 : \theta_1 = \theta_{10} + \delta$ versus $H_1 : \theta_1 \neq \theta_{10} + \delta$ for some $\delta > 0$ based on $X$.

As before, write $Q(Z|\theta) = Q(Z_1, Z_2, Z_3|\theta_1, \theta_2, \theta_3) + Q(Z_4, \cdots, Z_p|Z_1, Z_2, Z_3; \theta)$ and recall that in the $\theta$-space for $Z$, $\theta_2 \in [0, \infty)$, $\theta_3 \in [0, \infty)$ and $\theta_1 \in (-\infty, \infty)$ for $i \neq 2, 3$. Hence, for LRT as well as for Wald, we need to concentrate only on the first part. For the unrestricted MLEs of $\theta_2$ and $\theta_3$, we get from Self and Liang(1987):

$$
(\hat{\theta}_2, \hat{\theta}_3)' = \begin{cases} 
(Z_2, Z_3)', & \text{if } Z_2 > 0, Z_3 > 0 \\
(Z_{2,3}, 0)', & \text{if } Z_{2,3} < 0, Z_3 < 0 \\
(0, Z_{3,2})', & \text{if } Z_2 < 0, Z_{3,2} > 0 \\
(0, 0)', & \text{if } Z_{2,3} < 0, Z_{3,2} < 0.
\end{cases} 
(35)
$$

Since the minimum of $Q(Z_1|Z_2, Z_3; \theta_1, \theta_2, \theta_3)$ wrt $\theta_1$ for any given $(\theta_2, \theta_3)$ is 0 and also the minimum of $Q(Z_4, \cdots, Z_p|Z_1, Z_2, Z_3; \theta)$ wrt $\theta_4, \cdots, \theta_p$ for any given $\theta_1$, $\theta_2$ and $\theta_3$ is 0, we get

$$
\min_{\theta_1 \in \Omega} Q(Z|\theta) = Q(Z_2, Z_3|\hat{\theta}_2, \hat{\theta}_3)
(36)
$$

It is easy to show that the above minimum simplifies to

$$
\min_{\theta_1 \in \Omega} Q(Z|\theta) = \begin{cases} 
0, & \text{if } Z_2 > 0, Z_3 > 0 \\
Z_2^2/\sigma_{23}, & \text{if } Z_3 < 0, Z_{2,3} > 0 \\
Z_2^2/\sigma_{22}, & \text{if } Z_2 < 0, Z_{3,2} > 0 \\
Q(Z_2, Z_3|\theta_2 = \theta_3 = 0), & \text{if } Z_{2,3} < 0, Z_{3,2} < 0 \\
\end{cases}
(37)
$$

Once $\hat{\theta}_2$ and $\hat{\theta}_3$ are derived, the MLEs of the rest of the (interior) parameters are readily obtained from $Q(Z_1, Z_4, \cdots, Z_p|Z_2, Z_3; \theta)$ as the residuals of $Z_i$, given $(Z_2, Z_3)$, and are given by

24
\[
\hat{\theta}_i = Z_i - E[Z_i|Z_2 - \hat{\theta}_2, Z_3 - \hat{\theta}_3].
\] (38)

To derive the LRT of \( H_0 : \theta_i = \delta \) versus \( H_1 : \theta_i \neq \delta \) based on \( \mathbf{Z} \), we need to derive the restricted MLEs of \( \theta \)'s under the null hypothesis \( H_0 \). Without any loss of generality, let us assume \( \delta = 0 \), and write \( Q(Z_1, Z_2, Z_3; \theta = 0, \theta_2, \theta_3) = Q(Z_1|\theta = 0) + Q(Z_2, Z_3|Z_1; \theta_1 = 0, \theta_2, \theta_3) \). Since the first part is independent of \( \theta_2 \) and \( \theta_3 \), writing \( Y_2 = Z_{2,1} \) and \( Y_3 = Z_{3,1} \), we get from Self and Liang (1987) the following restricted MLEs of \( \theta_2 \) and \( \theta_3 \):

\[
(\hat{\theta}_2, \hat{\theta}_3)'_{null} = \begin{cases} 
(Y_2, Y_3)', & \text{if } Y_2 > 0, Y_3 > 0 \\
(Y_{2,3}, 0)', & \text{if } Y_{2,3} < 0, Y_3 < 0 \\
(0, Y_{3,2})', & \text{if } Y_2 < 0, Y_{3,2} > 0 \\
(0, 0)', & \text{if } Y_{2,3} < 0, Y_{3,2} < 0.
\end{cases}
\] (39)

Since as before the contribution from the minimization of \( Q(Z_4 \cdots, Z_p|Z_1, Z_2, Z_3; \theta) \) with respect to \( \theta_4, \cdots, \theta_p \) is 0 even under \( H_0 \) due to the interior nature of the parameters \( \theta_4, \cdots, \theta_p \), it follows that

\[
\min_{\theta \in \Omega_3} Q(\mathbf{Z} | \theta; H_0) = Q(Z_1|\theta_1 = 0) + \min_{\theta_2, \theta_3 \geq 0} Q(Z_2, Z_3|Z_1, \hat{\theta}_2null, \hat{\theta}_3null)
\] (40)

\[
= Z_1^2/\sigma_{11}, \text{ if } Z_{2,1} > 0, Z_{3,1} > 0
\]

\[
= Z_3^2/\sigma_{33} + \frac{Z_{1,3}^2}{\sigma_{11}(1 - \rho_{13}^2)}, \text{ if } Z_{3,1} < 0, Z_{2,13} > 0
\]

\[
= Z_2^2/\sigma_{22} + \frac{Z_{1,2}^2}{\sigma_{11}(1 - \rho_{12}^2)}, \text{ if } Z_{2,1} < 0, Z_{3,12} > 0
\]

\[
= Z_1^2/\sigma_{11} + (Z_{2,1}/\sigma_2, Z_{3,1}/\sigma_3) \mathbf{A}^{-1} (Z_{2,1}/\sigma_2, Z_{3,1}/\sigma_3)',
\]

\[
\text{if } Z_{2,13} < 0, Z_{3,12} < 0
\]

where the matrix \( \mathbf{A} : 2 \times 2 \) is defined as

\[
\mathbf{A} = \begin{pmatrix}
(1 - \rho_{12}^2) & \rho_{23} - \rho_{12}\rho_{13} \\
\rho_{23} - \rho_{12}\rho_{13} & 1 - \rho_{13}^2
\end{pmatrix}.
\] (41)
Combining (37) and (40), the LRT of \( H_0 : \theta_1 = 0 \) is obtained as 
\[
W = \min_{\theta \in \Omega^*} Q(Z|\theta; H_0) - \min_{\theta \in \Omega} Q(Z|\theta).
\] (42)

To simulate the null distribution of \( W \), we can take without any loss of generality \((Z_1, Z_2, Z_3) \sim N((0, 0, 0), \Sigma^*)\) where \( \Sigma^* \) has its diagonal elements as 1. Upon generating the \( j \)th iteration element as \((Z_{1j}, Z_{2j}, Z_{3j})\), we compute \((Z_{2,3j}, Z_{3,2j}, Z_{3,1j}, Z_{2,1j}, Z_{2,13j}, Z_{3,12j})\), leading to a value \( w_j \) of \( W \).

Obviously, the LRT of \( H_0 \) versus \( H_1 \) rejects \( H_0 \) when \( W \) computed as above is large. Again, quite surprisingly, it turns out from our simulation studies that the null distribution of \( W \) does not depend on the correlations between \( Z_1 \) and \((Z_2, Z_3)\).

For Wald-type inference about \( \theta_1 \), we concentrate on the distribution of the MLE of \( \theta_1 \) given in (38) under the natural restriction of two boundary points. Such a distribution can be obtained from (37) by conditioning on four disjoint subsets in \((z_2, z_3)\)-plane and deriving each component distribution and mixing them with suitable proportions. Details are omitted as this is similar to Theorem 2.1. This distribution can be used for testing \( H_0 : \theta_1 = \delta + \theta_{10} \) based on \( X \). Thus, if the alternative is \( H_1 : \theta_1 > \delta + \theta_{10} \), a reasonable test is to reject \( H_0 \) for large values of the MLE \( \hat{\theta}_{1N} \) of \( \theta_1 \), namely, when \((\hat{\theta}_{1N} - \delta - \theta_{10}) > c_N\). To determine the value of \( c_N \) for a given significance level \( \alpha \), we can use the asymptotic null distribution of \( \sqrt{N}(\hat{\theta}_{1N} - \delta - \theta_{10}) \) in presence of two boundary parameters, which is precisely the distribution mentioned above.

On the other hand, for a both-sided alternative \( H_1 : \theta_1 \neq \delta + \theta_{10} \), a reasonable test is to reject \( H_0 \) for large values of \(|\hat{\theta}_{1N} - \delta - \theta_{10}|\). To determine the cut-off point for a given significance level \( \alpha \), we can again use the asymptotic null distribution of \( \sqrt{N}(\hat{\theta}_{1N} - \delta - \theta_{10}) \) in presence of two boundary parameters, which is precisely the distribution mentioned above. Details are omitted.

**Remark 3.1.** Suppose we are interested in making a Wald-type inference about a linear function of \( \theta \), say \( \sum_{i \neq 2,3} c_i \theta_i \) which excludes the two boundary points \( \theta_2 \) and \( \theta_3 \). Naturally we would then consider the distribution of \( \sum_{i \neq 2,3} c_i \hat{\theta}_{iN} \). From (35), such a distribution can be obtained by conditioning on values of \((z_2, z_3)\) and then suitable unconditioning.
Remark 3.2. Suppose now that we are interested in making a Wald-type inference about a linear function of $\theta$, say $\sum_{i \neq 2,3} c_i \theta_i + c_2 \theta_2$ which contains one of the boundary points $\theta_2$. We would then consider the obvious statistic $\sum_{i \neq 2,3} c_i \hat{\theta}_{iN} + c_2 \hat{\theta}_{2N}$. Its asymptotic distribution can be derived from the results given above.

Remark 3.3. Let us also assume that we are interested in drawing suitable Wald-type inference about a linear function of $\theta$, which includes both the boundary parameters. Obviously, such a linear function can be written as $\eta = \sum_i c_i \theta_i$ and inference about $\eta$ is drawn by studying the asymptotic distribution of $W = \sum_i c_i \hat{\theta}_{iN}$. This distribution can also be derived from the above results.

2.3. Three points on the boundaries

In this section we deal with the case when three parameters lie on their boundaries, and as in the case of two boundary points, we consider the following two cases.

Case 1. Assume that the parameter $\theta_1$ of interest along with two other parameters, say $\theta_2$ and $\theta_3$, lie on their boundaries with $\Omega_i = [\theta_{i0}, \infty)$ for $i = 1, 2, 3$ and $\Omega_j = (\theta_{j0}, \infty)$ for $j > 3$. In such a situation, a reasonable test for the hypothesis $H_0 : \theta_1 = \theta_{10}$ versus $H_1 : \theta_1 > \theta_{10}$ can be developed on the basis of the joint asymptotic distribution of $\sqrt{N}(\hat{\theta}_{1N} - \theta_{10}, \hat{\theta}_{2N} - \theta_{20}, \hat{\theta}_{3N} - \theta_{30}, \hat{\theta}_{4N} - \theta_{40}^*, \ldots, \hat{\theta}_{pN} - \theta_{p0}^*)'$ which is derived below. Because of the nature of the null and alternative hypotheses, it makes sense to reject the null hypothesis $H_0$ for large values of the MLE of $\theta_1$, i.e., for large values of $\hat{\theta}_{1N}$ where the MLE $\hat{\theta}_N$ is computed under the assumption that there are three boundary points. The cut-off point is then obtained by considering the marginal asymptotic distribution of $\sqrt{N}(\hat{\theta}_{1N} - \theta_{10})$. This will be illustrated later in this subsection.

To derive the MLEs of $\theta$ based on $Z$ under the assumption of three boundary points corresponding to $\theta_1, \theta_2, \theta_3$, we proceed as follows. Writing

$$Q(Z|\theta) = Q(Z_1, Z_2, Z_3|\theta_1, \theta_2, \theta_3) + Q(Z_4, \ldots, Z_p | Z_1, Z_2, Z_3; \theta)$$  (43)

it follows that the MLEs of the boundary parameters $\theta_1, \theta_2, \theta_3$ are based on the first part of the quadratic and once these MLEs are obtained, the MLEs of the remaining interior parameters are simply obtained by an unrestricted
minimization of the second quadratic. The latter leads to the usual linear regressions of \( Z_4, \ldots, Z_p \), given appropriate values of \( Z_1, Z_2, Z_3 \). Details appear below.

Following a general discussion in Self and Liang (1987), let \( B_i \) denote a \( 3 \times 3 \) diagonal matrix with 1’s on the diagonal at places for which the MLEs of the three boundary parameters would be 0! The resultant vector estimates \((\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)’\) are then given by \( P_i(Z_1, Z_2, Z_3)’ \) where the projection \( P_i = I_3 - \Sigma B_i [B_i (\Sigma)^{-1} B_i] + B_i \) and \( A^+ \) denotes a Moore-Penrose inverse of \( A \). Taking the diagonal elements of \( B_i \) successively as \((0, 0, 0), (0, *, *), (*, 0, *), (0, 0, *), (*, *, 0), (0, *, 0) \) and \((*, *, *)\), we get the following theorem.

**Theorem 4.1.** The MLEs of the three boundary parameters \((\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)\) are given by

\[
(\hat{\theta}_{1N}, \hat{\theta}_{2N}, \hat{\theta}_{3N})’ = \begin{cases} 
(Z_1, Z_2, Z_3)’ & \text{if } Z_1 > 0, Z_2 > 0, Z_3 > 0 \\
(0, Z_{2,1}, Z_{3,1})’ & \text{if } Z_1 < 0, Z_{2,1} > 0, Z_{3,1} > 0 \\
(Z_{1,2}, 0, Z_{3,2})’ & \text{if } Z_2 < 0, Z_{1,2} > 0, Z_{3,2} > 0 \\
(Z_{1,3}, Z_{2,3}, 0)’ & \text{if } Z_3 < 0, Z_{1,3} > 0, Z_{2,3} > 0 \\
(0, 0, Z_{3,12})’ & \text{if } Z_{1,2} < 0, Z_{2,1} < 0, Z_{3,12} > 0 \\
(0, Z_{2,13}, 0)’ & \text{if } Z_{1,3} < 0, Z_{3,1} < 0, Z_{2,13} > 0 \\
(Z_{1,23}, 0, 0)’ & \text{if } Z_{2,3} < 0, Z_{3,2} < 0, Z_{1,23} > 0 \\
(0, 0, 0)’ & \text{if } Z_{1,23} < 0, Z_{2,13} < 0, Z_{3,12} < 0.
\end{cases}
\]

Based on the above MLEs of the three boundary parameters \( \theta_1, \theta_2 \) and \( \theta_3 \), the MLEs of the remaining interior parameters \( \theta_4, \ldots, \theta_p \), as remarked earlier, can be easily derived by computing the appropriate linear regressions.

**Theorem 4.2.** The MLEs of the interior parameters \( \hat{\theta}_4, \ldots, \hat{\theta}_p \) are given by

\[
(\hat{\theta}_{4N}, \ldots, \hat{\theta}_{pN})’ = (Z_4, \ldots, Z_p)’ - \Gamma [\tilde{\Sigma}_{3x3}]^{-1} (Z_1 - \hat{\theta}_{1N}, Z_2 - \hat{\theta}_{2N}, Z_3 - \hat{\theta}_{3N})’
\]

where \( \Gamma \) and \( \tilde{\Sigma}_{3x3} \) are defined by

\[
\Gamma = \begin{pmatrix} 
\sigma_{14} & \sigma_{24} & \sigma_{34} \\
. & . & . \\
\sigma_{1p} & \sigma_{2p} & \sigma_{3p}
\end{pmatrix}
\]
\[ \Sigma_{3 \times 3} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \] (47)

**Remark 4.1.** To draw a valid inference about a linear function \( \sum_{i>3} c_i \theta_i \) of the parameters \( \theta \), naturally we consider the estimate \( \sum_{i>3} c_i \hat{\theta}_{iN} \) and derive its asymptotic sampling distribution based on Theorem 4.1 and Theorem 4.2 along with the distributional assumption about \( Z \sim N[0, \Sigma] \).

**Remark 4.2.** Returning to the problem of testing \( H_0 : \theta_1 = \theta_{10} \) versus \( H_1 : \theta_1 > \theta_{10} \), the cut-off point of the null distribution of \( \hat{\theta}_{1N} \) can be determined from the asymptotic distribution of \( \hat{\theta}_{1N} \), using Theorem 4.1.

**Case 2.** We next assume that the parameter of interest is an interior parameter, and without any loss of generality, suppose \( \theta_1, \theta_2 \) and \( \theta_3 \) are the three boundary parameters with \( \Omega_i = [\theta_{i0}, \infty) \) for \( i = 1, 2, 3 \) and \( \Omega_j = (\theta_{j0}, \infty) \) for \( j > 3 \), and \( \theta_4 \) is the parameter of interest. The relevant testing problem about \( \theta_4 \) is then to decide between \( H_0 : \theta_4 = \theta_{40} + \delta \) versus \( H_1 : \theta_4 \neq \theta_{40} + \delta \) on the basis of \( X \) where \( \delta > 0 \) is specified. Equivalently (and asymptotically), this is the same as testing \( H_0 : \theta_4 = \delta \) versus \( H_1 : \theta_4 \neq \delta \) based on \( Z \sim N[\theta, \Sigma] \), when \( \theta_i \geq 0, i = 1, 2, 3 \) and \( -\infty < \theta_j < \infty \) for \( j > 3 \).

To derive the LRT for the above testing problem, we proceed in the usual fashion by deriving both unrestricted and restricted (under \( H_0 \)) MLEs of \( \theta \) based on \( X \) and computing the value of \( W = -2\ln(LRT(X)) \), which is indeed the profile likelihood approach. To carry out the LRT test, naturally we need to determine the cut-off point of the null distribution of \( W \). Following Self and Liang (1987) and Proposition 2, asymptotically, such a distribution is given by the distribution of the difference of two quadratic forms based on \( Z \). Theorem 4.1 and Theorem 4.2 provide the unrestricted MLEs of the parameters \( \theta \) in the distribution of \( Z \). Writing \( Q(Z|\theta) = Q(Z_1, Z_2, Z_3|\theta_1, \theta_2, \theta_3) + Q(Z_4, \cdots, Z_p|\theta) \), since the unrestricted minimum of the second term is 0, it follows that

\[
\min_{\theta \in \Omega} Q(Z|\theta) = \min_{\theta_1 \geq 0, \theta_2 \geq 0, \theta_3 \geq 0} Q(Z_1, Z_2, Z_3|\theta_1, \theta_2, \theta_3) \quad (48)
\]

\[ = Q(Z_1, Z_2, Z_3|\hat{\theta}_{1N}, \hat{\theta}_{2N}, \hat{\theta}_{3N}). \]

Because of the nature of \((\hat{\theta}_{1N}, \hat{\theta}_{2N}, \hat{\theta}_{3N})\) given in (44), it is obvious that the unrestricted minimum value of the relevant quadratic of the likelihood based
on $\mathbf{Z}$ can be divided into eight (8) disjoint sets, each set having a distinct value of the quadratic.

On the other hand, to compute the restricted maximum likelihood under the null hypothesis $H_0 : \theta_4 = \theta_{40} + \delta$ in the $\mathbf{X}$ space which is equivalent to $\theta_4 = \delta$ in the $\mathbf{Z}$ space, note that all we need to compute are the restricted MLEs of $\theta_1$, $\theta_2$, and $\theta_3$, subject to their being non-negative. This is derived as follows. We write

$$Q(Z_1, Z_2, Z_3, Z_4|\theta_1, \theta_2, \theta_3, \theta_4 = \delta) = Q(Z_4|\delta) + Q(Z_1, Z_2, Z_3|Z_4, \theta_1, \theta_2, \theta_3, \theta_4 = \delta)$$

and note from (15) that the second quadratic can be expressed as

$$Q(Z_1, Z_2, Z_3|Z_4, \theta_1, \theta_2, \theta_3, \theta_4 = \delta) = (Z_{1,4}, Z_{2,4}, Z_{3,4})\tilde{\Sigma}^{-1}(Z_{1,4}, Z_{2,4}, Z_{3,4})'$$

where the residuals $Z_{1,4}, Z_{2,4}, Z_{3,4}$ and $\tilde{\Sigma}$ are defined as

$$Z_{i,4} = Z_i - \theta_i - \rho_{i4}\sigma_i(Z_4 - \delta)/\sigma_4, \quad i = 1, 2, 3$$

$$\tilde{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{21}/\sigma_{44}$$

$$\Sigma_{11} : 3 \times 3 = \begin{pmatrix}
\sigma_{11} & \sigma_{12} & \sigma_{13} \\
\sigma_{21} & \sigma_{22} & \sigma_{23} \\
\sigma_{31} & \sigma_{32} & \sigma_{33}
\end{pmatrix}$$

$$\Sigma_{12} = (\sigma_{14}, \sigma_{24}, \sigma_{34})'$$

To determine the restricted MLEs of the parameters $\theta_1, \theta_2, \theta_3$ which must be non-negative, by minimizing $Q(Z_1, Z_2, Z_3|Z_4, \theta_1, \theta_2, \theta_3, \theta_4 = \delta)$, under the null hypothesis: $\theta_4 = \delta$, we can readily apply Theorem 4.1. Based on the first step residuals ($Z_{1,4} = Z_{1}^*, Z_{2,4} = Z_{2}^*, Z_{3,4} = Z_{3}^*$) and their dispersion matrix $\tilde{\Sigma}$, we define the second step residuals ($Z_{2,1}^*, Z_{2,1}^*, Z_{2,2}^*, Z_{2,2}^*, Z_{2,3}^*, Z_{2,3}^*, Z_{3,1}^*, Z_{3,1}^*, Z_{2,13}^*, Z_{3,12}^*$) in the usual fashion. Then, by Theorem 4.1, the restricted MLEs are given by:
\[(\hat{\theta}_{1N}, \hat{\theta}_{2N}, \hat{\theta}_{3N})'_{null} = (Z_1^*, Z_2^*, Z_3^*)', \text{ if } Z_1^* > 0, Z_2^* > 0, Z_3^* > 0 \quad (52)\]

\[(0, Z_{21}, Z_{31})', \text{ if } Z_1^* < 0, Z_{21}^* > 0, Z_{31}^* > 0 \]

\[(Z_{12}^*, 0, Z_{23}'), \text{ if } Z_2^* < 0, Z_{12}^* > 0, Z_{32}^* > 0 \]

\[(Z_{13}^*, Z_{23}^*, 0)', \text{ if } Z_3^* < 0, Z_{13}^* > 0, Z_{23}^* > 0 \]

\[(0, 0, Z_{312})', \text{ if } Z_{12}^* < 0, Z_{21}^* < 0, Z_{312}^* > 0 \]

\[(0, Z_{213}, 0)', \text{ if } Z_{13}^* < 0, Z_{31}^* < 0, Z_{213}^* > 0 \]

\[(Z_{123}^*, 0, 0)', \text{ if } Z_{23}^* < 0, Z_{32}^* < 0, Z_{123}^* > 0 \]

\[(0, 0, 0)', \text{ if } Z_{123}^* < 0, Z_{213}^* < 0, Z_{312}^* < 0.\]

This then results in the restricted quadratic form

\[\min_{\theta \in \Omega_0} Q(Z | \theta \in \Omega_0^*) \quad (53)\]

\[= \min_{\theta \in \Omega_0^*} Q(Z_1, Z_2, Z_3, Z_4 | \theta_1, \theta_2, \theta_3, \theta_4 = \delta)\]

\[= Q(Z_4 | \delta) + Q(Z_1, Z_2, Z_3 | Z_4, \hat{\theta}_{1; \text{null}}, \hat{\theta}_{2; \text{null}}, \hat{\theta}_{3; \text{null}}, \theta_4 = \delta)\]

Using (48) and (53), we readily obtain the difference \(\Delta(Z)\) between the two quadratic forms, whose distribution under \(Z \sim N[0, \Sigma]\) is essentially the asymptotic null distribution of \(-2ln(LRT)\) based on \(X\). Obviously, the distribution of \(\Delta(Z)\) would depend only on the nature of correlations among the estimates of the four parameters - one parameter of interest, namely, \(\theta_4\), and the three boundary parameters, \(\theta_1, \theta_2, \theta_3\). It may be remarked that the null distribution of \(\Delta(Z)\) may have to be obtained by simulation and it is obvious that we can take the dispersion matrix of \(Z\) as the correlation matrix, without any loss of generality. It would be interesting to check through simulations if this null distribution really depends on the correlations. Let us recall that this was not the case in the previous two instances.

**Remark 4.3.** A Wald-type test of \(H_0 : \theta_4 = \theta_{40} + \delta\) versus the one-sided alternative \(H_1 : \theta_4 > \theta_{40} + \delta\) based on \(X\) can be easily derived using the unrestricted MLE \(\hat{\theta}_{4N}\) of \(\theta_4\) given in Theorem 4.2. Such a test would reject \(H_0\) for large values of \(\hat{\theta}_{4N}\), or \(\sqrt{N}(\hat{\theta}_{4N} - \theta_{40} - \delta)\) and the cut-off point of its null distribution is asymptotically computed from the distribution of the MLE of \(\theta_4\) based on \(Z\) in presence of three boundary parameters. Similarly, for testing \(H_0\) versus the both-sided alternative \(H_1 : \theta_4 \neq \theta_{40} + \delta\), one could reject \(H_0\) for large values of \(\sqrt{N}|\hat{\theta}_{4N} - \theta_{40} - \delta|\) and the cut-off point is determined.
analogously from its null distribution. From (45), this distribution, which depends on the correlation structure between \( Z_4 \) and \((Z_1, Z_2, Z_3)\), can be derived.

3. Applications of general theory to multistage Weibull models

In this section we consider some applications of the general theory developed in Section 2 in the context of multistage Weibull models and analyze some relevant data sets. In the sequel, we consider three cases: linear Weibull model, quadratic Weibull model and cubic Weibull model.

Quite generally, assume that there are \( m + 1 \) dose groups with doses \( d_0 = 0 \) (control), \( d_1, \cdots, d_m \) and \( n_i \) subjects in the \( i \)th dose group, yielding independent \( X_i \sim B[n_i, \pi_i(\theta|d_i)] \), \( i = 0, \cdots, m \) where

\[
\pi_i(\theta|d_i) = 1 - e^{-[\theta_0 + d_1 \theta_1 + \cdots + d_k \theta_k]}.
\]

(54)

Here \( \theta = (\theta_0, \cdots, \theta_k) \neq 0 \) and \( \theta_i \geq 0 \) for all \( i \), which are considered as the natural model restrictions! There are three kinds of inference problems of interest to us. First, inference about the absolute risk (AR) \( \pi(\theta|d^*) \) at a given dose \( d^* \); secondly, inference about the extra risk (ER) \( \gamma(d^*) \) at a specified dose \( d^* \), where \( \gamma(d^*) \) is defined by

\[
\gamma(d^*) = [\pi(\theta|d^*) - \pi(\theta|d_0)]/[1 - \pi(\theta|d_0)]
\]

\[
= 1 - e^{-[\theta_1 d^* + \theta_2 d^* 2 + \cdots + \theta_k d^* k]}.
\]

(55)

The third inference problem is about the benchmark dose (BMD) \( d^* \) which yields a specified relative risk \( \gamma \). From the expression for \( \pi(\theta|d^*) \), it is clear that the first two inference problems are essentially concerned with linear functions of \( \theta \), while the third problem is about drawing inference for \( d^* \) which satisfies the polynomial equation:

\[
ln\left(1/(1 - \gamma)\right) = \theta_1 d^* + \theta_2 d^* 2 + \cdots + \theta_k d^* k.
\]

(56)

Obviously, \( d^* \) is an implicit function of \( \theta \) whose solution is easy to obtain in linear, quadratic and cubic cases. We mention in passing that sometimes we may be interested in what is called an added risk at a given dose \( d \), defined by \( ADR(d) = \pi(\theta|d) - \pi(\theta|d_0) \). Since the inference on ADR is very similar to that on AR, we do not pursue it here.
Case 1: Linear Weibull model. Taking \( k = 1 \) in (54), we get
\[
\pi_i(\theta|d_i) = 1 - e^{-[\theta_0 + d_i\theta_1]} ,
\]
which readily gives
\[
AR(d*) = 1 - e^{-[\theta_0 + d*\theta_1]} 
\] (57)
\[
ER(d*) = 1 - e^{-d*\theta_1} 
\] (58)
\[
BMD = \left(\frac{1}{\theta_1}\right) \ln \left(\frac{1}{1 - \gamma}\right). 
\] (59)

Appropriate statistical inference about the above quantities follows upon computing the natural MLEs of the two parameters \( \theta_0 \) and \( \theta_1 \) based on the data \((X_0, \ldots, X_m)\), namely \( \hat{\theta}_{0N} \) and \( \hat{\theta}_{1N} \), and applying their asymptotic distributional results derived in Section 2. It is easy to derive the Fisher information matrix \( I(\theta_0, \theta_1) : 2 \times 2 \) under the assumption of independent binomial distributions of the \( X_i \)'s, and hence \( \Sigma(\theta_0, \theta_1) = I(\theta_0, \theta_1)^{-1} \), where \( I(\theta_0, \theta_1) \) is given by
\[
I(\theta_0, \theta_1) = \left( \begin{array}{cc}
\sum_{i=0}^{m} \frac{1 - \pi_i(\theta)}{\pi_i(\theta)} & \sum_{i=0}^{m} d_i \frac{1 - \pi_i(\theta)}{\pi_i(\theta)} \\
\sum_{i=0}^{m} d_i^2 \frac{1 - \pi_i(\theta)}{\pi_i(\theta)} & \sum_{i=0}^{m} d_i^2 \frac{1 - \pi_i(\theta)}{\pi_i(\theta)}
\end{array} \right). 
\] (60)

When \( \theta_0 > 0, \theta_1 > 0 \), we clearly have a regular parametric scenario. By applying the standard asymptotic theory, we then get the following result.
\[
\sqrt{N}(\hat{\theta}_{0N} - \theta_0, \hat{\theta}_{1N} - \theta_1) \to N_2[(0, 0), \Sigma(\theta_0, \theta_1)] 
\] (61)
\[
\sqrt{N}[(\hat{\theta}_{0N} + d* \hat{\theta}_{1N}) - (\theta_0 + d* \theta_1)] \to N[0, (1, d*)\Sigma(\theta_0, \theta_1)(1, d*)']. 
\] (62)

We remark that for most inference purposes, \( \theta_0 \) and \( \theta_1 \) in \( \Sigma(\theta_0, \theta_1) \) are replaced by their MLEs or their null hypotheses values, if any.

We are now in a position to discuss the inferential aspects of AR, ER and BMD. For AR\((d*)\), a point estimate is given by
\[
\hat{AR}(d*) = 1 - e^{-[\hat{\theta}_{0N} + d*\hat{\theta}_{1N}]} 
\] with its estimated asymptotic standard error (SE) as
\[
estSE(\hat{AR}(d*)) = (1 - \hat{AR}(d*)) \sqrt{[(1, d*)\Sigma(\hat{\theta})(1, d*)']} . 
\] (63)

Using the fact that, asymptotically, \( (\hat{AR}(d*) - AR(d*))/[\est.SE(\hat{AR}(d*))] \sim N[0, 1] \), it is straightforward to carry out a test for a specified value of AR\((d*)\).
as well as to construct a confidence interval of $AR(d\ast)$. Of course, it is very easy to derive an asymptotic test for $\theta_0 + d \ast \theta_1$ and also to construct a confidence interval for $\theta_0 + d \ast \theta_1$ based on (62), which would readily solve the inference problems about $AR(d\ast)$.

Similarly, a point estimate of $ER(d\ast)$ is given by $\hat{ER}(d\ast) = 1 - e^{-d\ast\hat{\theta}_{1N}}$ with its estimated asymptotic standard error (SE) as

$$\text{estSE}(\hat{ER}(d\ast)) = d \ast (1 - \hat{ER}(d\ast)) \sqrt{[\hat{\text{var}}(\hat{\theta}_{1N})]/\hat{\theta}_{1N}^2}.$$  \hspace{1cm} (64)

Using the fact that, asymptotically, $(\hat{ER}(d\ast) - ER(d\ast))/[\text{est.SE}(\hat{ER}(d\ast))] \sim N[0, 1]$, it is easy to carry out a test for a specified value of $ER(d\ast)$ as well as to construct a confidence interval of $ER(d\ast)$. Again, it is straightforward to derive an asymptotic test for $\theta_1$ and also to construct a confidence interval for $\theta_1$ based on the asymptotic distribution of $\hat{\theta}_{1N}$ which readily follows from (61). This then solves the inference problems about $ER(d\ast)$.

Finally, from (59), a point estimate of the BMD at specified extra risk level $\gamma$ is given by $\hat{BMD} = [1/\hat{\theta}_{1N}](\ln(1/(1 - \gamma)))$ with its estimated asymptotic standard error (SE) as

$$\text{est}(\text{SE}(\hat{BMD})) = \ln(1/(1 - \gamma)) \sqrt{[\hat{\text{var}}(\hat{\theta}_{1N})]/\hat{\theta}_{1N}^2}.$$  \hspace{1cm} (65)

Using the fact that, asymptotically, $(\hat{BMD} - BMD)/[\text{est.SE}(\hat{BMD})] \sim N[0, 1]$, it is straightforward to carry out a test for a specified value of BMD as well as to construct a confidence interval of BMD. Again, the asymptotic test for $\theta_1$ and its asymptotic confidence interval can be readily used to solve the inference problems about BMD.

When $\theta_0 = 0$, which may well happen in some situations, we have the case of one parameter point being on the boundary, and hence the asymptotic distributional and inferential results mentioned above are not true. Using Self and Liang (1987), we get

$$\sqrt{N}(\hat{\theta}_{0N}, \hat{\theta}_{1N} - \theta_1) \to (\tilde{\theta}_0(Z_0, Z_1), \tilde{\theta}_1(Z_0, Z_1))$$  \hspace{1cm} (66)

where $Z = (Z_0, Z_1) \sim N[0, \hat{\Sigma}]$ and $\tilde{\theta}_0(Z_0, Z_1) = 0$ if $Z_0 < 0$, = $Z_0$ if $Z_0 > 0$, and $\tilde{\theta}_1(Z_0, Z_1) = Z_{1,0}$ if $Z_0 < 0$, and = $Z_1$ if $Z_0 > 0$. Here $Z_{1,0} = Z_1 - \rho \sigma_1 Z_0/\sigma_0$ is the standard residual of $Z_1$ on $Z_0$.

A remark about $\hat{\Sigma}$ is in order here. Since $\Sigma$ is obviously a function of $\theta = (\theta_0, \theta_1)$, we can use $\hat{\Sigma}(\hat{\theta}_N)$. In case it so happens that $X_0 = 0$
so that \( \hat{\theta}_{0\,N} = 0 \), computation of \( \hat{\Sigma} \) may pose some difficulty (in terms of singularities!), and then we can take an arbitrary small value of \( \hat{\theta}_{0\,N} \).

Returning to the inference problems, although we still have the same point estimates as before, namely, \( \hat{A}\text{R}(d\star) = 1 - e\left[-\hat{\theta}_{0\,N} + d\star \hat{\theta}_1\,N\right] \), \( \hat{E}\text{R}(d\star) = 1 - e\left[-d\star \hat{\theta}_1\,N\right] \), and \( \hat{B}\text{MD} = \frac{1}{\hat{\theta}_1\,N} \ln\left(1/(1 - \gamma)\right) \), the asymptotic distributions of these estimates are not normal any more, and in fact are obtained by replacing \( \hat{\theta}_{0\,N} \) and \( \hat{\theta}_1\,N \) by \( \hat{\theta}_0(Z_1, Z_2) \) and \( \hat{\theta}_1(Z_1, Z_2) \), respectively, and deriving the resultant distributions by simulation.

We have pursued in Section 4 the inference methods for \( \hat{E}\text{R}(d\star) = 1 - e\left[-d\star \hat{\theta}_1\,N\right] \) and \( \hat{B}\text{MD} = \frac{1}{\hat{\theta}_1\,N} \ln\left(1/(1 - \gamma)\right) \) on the basis of \( \hat{E}\text{R}_1 = 1 - e\left[-d\star \hat{\theta}_1\,N\right] \) and \( \hat{B}\text{MD}_1 = \frac{1}{\hat{\theta}_1\,N} \ln\left(1/(1 - \gamma)\right) \) using some real data sets.

What we have described so far is the Wald approach! For the derivation of the LRT for hypotheses about \( \hat{E}\text{R}(d\star) \) and \( \hat{B}\text{MD} \), since these quantities involve only \( \hat{\theta}_1 \), it follows from the general theory discussed in Section 2 that irrespective of the nature of \( \hat{\theta}_0 \), whether a boundary point or not, the asymptotic distribution of \( -2\ln(LRT) \) remains as \( \chi^2 \) with 1 d.f. We have compared the Wald test with the LRT test in a few cases in Section 4.

**Case 2: Quadratic Weibull model.** Taking \( k = 2 \) in (54), we get

\[
\pi_i(\theta|d_i) = 1 - e^{-[\theta_0 + d_i \theta_1 + d_i^2 \theta_2]},
\]

which readily gives

\[
AR(d\star) = 1 - e^{-[\theta_0 + d\star \theta_1 + d\star^2 \theta_2]} \tag{67}
\]

\[
ER(d\star) = 1 - e^{-[d\star \theta_1 + d\star^2 \theta_2]} \tag{68}
\]

\[
BMD = \frac{\sqrt{(\theta_1^2 + 4\theta_2\gamma)} - \theta_1}{2\theta_2}. \tag{69}
\]

Appropriate statistical inference about the above quantities follows upon computing the natural MLEs of the three parameters \( \theta_0, \theta_1 \) and \( \theta_2 \) based on the data \( (X_0, \cdots, X_m) \), namely \( \hat{\theta}_{0\,N}, \hat{\theta}_{1\,N} \) and \( \hat{\theta}_{2\,N} \), and applying their asymptotic distributional results derived in Section 2. It is easy to derive the Fisher information matrix \( I(\theta_0, \theta_1, \theta_2) : 3 \times 3 \) under the assumption of independent binomial distributions of the \( X_i \)'s, and hence \( \Sigma(\theta_0, \theta_1, \theta_2) = I(\theta_0, \theta_1, \theta_2)^{-1} \), where \( I(\theta_0, \theta_1, \theta_2) \) is given by
When $\theta_0 > 0$, $\theta_1 > 0$, $\theta_2 > 0$, we clearly have a regular parametric scenario situation. By standard asymptotic theory, we then get the following result:

$$\sqrt{N}(\hat{\theta}_{0N} - \theta_0, \hat{\theta}_{1N} - \theta_1, \hat{\theta}_{2N} - \theta_2) \to N_3[(0, 0, 0), \Sigma(\theta)]$$  \hspace{1cm} (71)

$$\sqrt{N}[(\hat{\theta}_{0N} + d \hat{\theta}_{1N} + d^2 \hat{\theta}_{2N}) - (\theta_0 + d \theta_1 + d^2 \theta_2)] \to N[0, (1, d*, d^2*)\Sigma(\theta)(1, d*, d^2*)']$$  \hspace{1cm} (72)

$$\sqrt{N}[(\hat{\theta}_{1N} + d \hat{\theta}_{2N}) - (\theta_1 + d \theta_2)] \to N[0, (1, d*)\Sigma(\theta)(1, d*)']$$  \hspace{1cm} (73)

Here $\Sigma(\theta)$ is the $2 \times 2$ lower sub matrix of $\Sigma(\theta)$ obtained by deleting the first row and the first column.

We are now in a position to discuss the inferential aspects of AR, RR and BMD. For AR($d*$) defined in (67), a point estimate is given by

$$\hat{AR}(d*) = 1 - e^{-[\theta_{0N} + d \theta_{1N} + d^2 \theta_{2N}]}$$  \hspace{1cm} (74)

with its estimated asymptotic standard error (SE) as

$$estSE(\hat{AR}(d*)) = (1 - \hat{AR}(d*)) \cdot \sqrt{(1, d*, d^2*)\Sigma(\hat{\theta})(1, d*, d^2*)'}.$$  \hspace{1cm} (75)

Using the fact that, asymptotically, $(\hat{AR}(d*) - AR(d*))/[est.SE(\hat{AR}(d*))] \sim N[0, 1]$, it is straightforward to carry out a test for a specified value of $AR(d*)$ as well as to construct a confidence interval of $AR(d*)$. Of course, it is very easy to derive an asymptotic test for $\theta_0 + d \theta_1 + d^2 \theta_2$ and also to construct a confidence interval for $\theta_0 + d \theta_1 + d^2 \theta_2$ on the basis of $\theta_{0N} + d \theta_{1N} + d^2 \theta_{2N}$. 

36
directly from (72), which would readily solve the inference problems about $AR(d*)$.

Similarly, a point estimate of $ER(d*)$ defined in (68) is given by $\hat{ER}(d*) = 1 - e^{-(\theta_1 N + d^* \theta_2 N)}$ with its estimated asymptotic standard error (SE) as

$$
estSE(\hat{ER}(d*)) = (1 - \hat{ER}(d*)) \sqrt{(d*, d^*) \hat{\Sigma}(\hat{\theta})(d*, d^*)'}$$

where $\hat{\Sigma}(\hat{\theta})$ is the lower $2 \times 2$ submatrix of $\hat{\Sigma}(\theta)$ obtained by deleting the first row and the first column.

Using the fact that, asymptotically, $(\hat{ER}(d*) - ER(d*))/[\estSE(\hat{ER}(d*)]) \sim N[0, 1]$, it is easy to carry out a test for a specified value of $ER(d*)$ as well as to construct a confidence interval of $ER(d*)$. Again, it is straightforward to derive an asymptotic test for $\theta_1 + d^* \theta_2$ and also to construct a confidence interval for $\theta_1 + d^* \theta_2$ on the basis of $\hat{\theta}_1 N + d^* \hat{\theta}_2 N$ directly from (73), which would readily solve the inference problems about $ER(d*)$.

Finally, a point estimate of the BMD at specified extra risk level $\gamma$, defined in (69), is given by

$$BMD = \sqrt{[\hat{\theta}_1 N + 4 \hat{\theta}_2 N \gamma] - \hat{\theta}_1 N}$$

with its estimated asymptotic standard error (SE) as

$$\estSE(BMD) = \sqrt{(u(\hat{\theta}_1 N, \hat{\theta}_2 N), v(\hat{\theta}_1 N, \hat{\theta}_2 N)) \hat{\Sigma}(\hat{\theta})(u(\hat{\theta}_1 N, \hat{\theta}_2 N), v(\hat{\theta}_1 N, \hat{\theta}_2 N))'}$$

where

$$u(\theta_1, \theta_2) = [\theta_1 (\theta_1^2 + 4 \gamma \theta_2)^{-1/2} - 1]/[2 \theta_2]$$

and

$$v(\theta_1, \theta_2) = [\gamma (\theta_1^2 + 4 \gamma \theta_2)^{-1/2}/\theta_2] - \sqrt{[\theta_1^2 + 4 \theta_2^2 \gamma] - \theta_1}/2 \theta_2$$

Using the fact that, asymptotically, $(BMD - BMD)/[\estSE(BMD)] \sim N[0, 1]$, it is easy to carry out a test for a specified value of BMD as well as to construct a confidence interval of BMD.
We now consider the case when some parameters may lie on the boundaries. Since we have assumed a quadratic Weibull model, there are three possible scenarios: 

(i) \( \theta_0 = 0 \),  
(ii) \( \theta_1 = 0 \),  
(iii) \( \theta_0 = \theta_1 = 0 \).

When (i) \( \theta_0 = 0 \), which may well happen in some situations, we have the case of one parameter point being on the boundary, and hence the asymptotic distributional and inferential results mentioned above are not true. We should note that for statistical inference about \( ER(d^*) \) and BMD which involve the parameters \( \theta_1 \) and \( \theta_2 \), all we would need is the asymptotic joint distribution of the relevant MLEs. Using Proposition 1, quite generally we get

\[
\sqrt{N} (\hat{\theta}_{0N}, \hat{\theta}_{1N} - \theta_1, \hat{\theta}_{2N} - \theta_2) \rightarrow (\hat{\theta}_0(Z_0, Z_1, Z_2), \hat{\theta}_1(Z_0, Z_1, Z_2), \hat{\theta}_2(Z_0, Z_1, Z_2))
\]

(81)

where \( Z = (Z_0, Z_1, Z_2) \sim N[\mathbf{0}, \hat{\Sigma}] \) and from the discussion in Section 2, Theorem 2.2, we get

\[
\hat{\theta}_0(Z_0, Z_1, Z_2) = Z_0 \text{ if } Z_0 > 0, \text{ and } 0 \text{ if } Z_0 < 0 \quad (82)
\]

\[
\hat{\theta}_1(Z_0, Z_1, Z_2) = Z_1 \text{ if } Z_0 > 0, \text{ and } Z_{1,0} \text{ if } Z_0 < 0. \quad (83)
\]

\[
\hat{\theta}_2(Z_0, Z_1, Z_2) = Z_2 \text{ if } Z_0 > 0, \text{ and } Z_{2,0} \text{ if } Z_0 < 0. \quad (84)
\]

where \( Z_{1,0} \) and \( Z_{2,0} \) are the usual residual variables. A remark about \( \hat{\Sigma} \) is in order here. Since \( \Sigma \) is obviously a function of \( \theta \), we can use \( \hat{\Sigma}(\hat{\theta}_N) \). In case it so happens that \( X_0 = 0 \) so that \( \hat{\theta}_{0N} = 0 \), computation of \( \hat{\Sigma} \) may pose some difficulty (in terms of singularities!), and then we can take an arbitrary small value of \( \hat{\theta}_{0N} \).

Returning to the inference problems, although we still have the same point estimates as before, namely,

\[
AR(d^*) = 1 - e^{-[\hat{\theta}_{0N} + d^*\hat{\theta}_{1N} + d^2\hat{\theta}_{2N}]}, \quad (85)
\]

\[
ER(d^*) = 1 - e^{-(d^*\hat{\theta}_{1N} + d^2\hat{\theta}_{2N})}, \quad (86)
\]

and

\[
BMD = \left[ \sqrt{[\hat{\theta}_{1N}^2 + 4\hat{\theta}_{2N}^2\gamma]} - \hat{\theta}_{1N} \right] / 2\hat{\theta}_{2N}, \quad (87)
\]

38
the asymptotic distributions of these estimates are not normal any more, and, in fact, these are obtained by replacing \( \hat{\theta}_{0N} \), \( \hat{\theta}_{1N} \) and \( \hat{\theta}_{2N} \) by \( \hat{\theta}_0(Z_0, Z_1, Z_2) \), \( \hat{\theta}_1(Z_0, Z_1, Z_2) \) and \( \hat{\theta}_2(Z_0, Z_1, Z_2) \), respectively, and deriving the resultant distributions by simulation.

We have explicitly pursued below the inference methods for \( ER(d^*) = 1 - e^{-(d^*\theta_1 + d^*\theta_2)} \) and \( BMD = \sqrt{[\hat{\theta}_1^2 + 4\hat{\theta}_2\gamma] - \frac{\theta_1}{2\hat{\theta}_2}} \) on the basis of \( \hat{ER}(d^*) = \hat{\theta}_{1N} + d^*\hat{\theta}_{2N} \). Here are the results.

**Inference about ER.** Since \( ER(d^*) = 1 - e^{-(d^*\theta_1 + d^*\theta_2)} \), without any loss of generality, we consider equivalently the problem of drawing suitable inference about \( \tilde{ER}(d^*) = \theta_1 + d^*\theta_2 \) which is based on \( \hat{ER}(d^*) = \hat{\theta}_{1N} + d^*\hat{\theta}_{2N} \). Applying Theorem 2, we conclude that

\[
\sqrt{N}[\hat{\theta}_{1N} + d^*\hat{\theta}_{2N}] - (\theta_1 + d^*\theta_2) \sim Z_1 \quad \text{if} \quad Z_0 > 0 \quad (88)
\]

\[
\sim Z_{1.0} \quad \text{if} \quad Z_0 < 0.
\]

Since \( \hat{\Sigma}(\hat{\theta}) \) is known, it is rather easy to simulate the asymptotic distribution of \( \sqrt{N}[\hat{\theta}_{1N} + d^*\hat{\theta}_{2N}] - (\theta_1 + d^*\theta_2)] \), from which the cut-off points of this distribution can be obtained. Tests and confidence intervals about \( \tilde{ER}(d^*) \) are then routinely derived.

**Inference about BMD.** Using (81) and (87), it is easy to generate the asymptotic distribution of \( BMD(d^*) \) from which its cut-off points can be derived and used for tests and confidence intervals for \( BMD(d^*) \).

What we have described so far is the Wald approach! For the derivation of the LRT for hypotheses about \( ER(d^*) \) and BMD, since these quantities involve only \( \theta_1 \) and \( \theta_2 \) which are regular parameters, it follows from the general theory discussed in Section 2 that irrespective of the nature of \( \theta_0 \), whether a boundary point or not, the asymptotic distribution of \(-2\ln(LRT)\) remains as \( \chi^2 \) with an appropriate d.f. We have compared the Wald test with the LRT test in a few cases in Section 4.

We next consider the case (ii) \( \theta_1 = 0 \) which is still the case of one boundary parameter. Using Proposition 1, the asymptotic joint distribution of the MLEs \( \hat{\theta}_N \) is given by
\[
\sqrt{N}(\hat{\theta}_{0N} - \theta_0, \hat{\theta}_{1N} - \theta_1, \hat{\theta}_{2N} - \theta_2) \to (89)
\]

\[
\left(\hat{\theta}_0(Z_0, Z_1, Z_2), \hat{\theta}_1(Z_0, Z_1, Z_2), \hat{\theta}_2(Z_0, Z_1, Z_2)\right)
\]

where, as before, \( Z = (Z_0, Z_1, Z_2) \sim N(0, \hat{\Sigma}) \), and

\[
\hat{\theta}_0(Z_0, Z_1, Z_2) = Z_0 \text{ if } Z_1 > 0, \text{ and } = Z_{0,1} \text{ if } Z_1 < 0 \quad (90)
\]

\[
\hat{\theta}_1(Z_0, Z_1, Z_2) = Z_1 \text{ if } Z_1 > 0, \text{ and } = 0 \text{ if } Z_1 < 0 \quad (91)
\]

\[
\hat{\theta}_2(Z_0, Z_1, Z_2) = Z_2 \text{ if } Z_1 > 0, \text{ and } = Z_{2,1} \text{ if } Z_1 < 0 \quad (92)
\]

As in the previous case, a remark about \( \hat{\Sigma} \) is in order here. Since \( \Sigma \) is obviously a function of \( \theta \), we can use either \( \hat{\Sigma}(\hat{\theta}_N) \) or take \( \theta_0 = \theta_1 = 0 \) and replace \( \theta_0 \) and \( \theta_2 \) by \( \hat{\theta}_{0N} \) and \( \hat{\theta}_{2N} \), respectively.

Using the above asymptotic distribution of the MLEs of \( \theta \), we can approximate the distributions of \( \hat{ER}(d*) = 1 - e^{-(d* \hat{\theta}_{1N} + d* \hat{\theta}_{2N})} \) and \( BMD = \sqrt{[\hat{\theta}_{1N}^2 + 4\hat{\theta}_{2N}^2] - \hat{\theta}_{1N}} / 2\hat{\theta}_{2N} \) by simulations and use them to carry out tests and confidence intervals of \( ER \) and \( BMD \).

We remark that the LRT for \( ER(d*) \) which is equivalent to the linear function \( \theta_1 + d* \theta_2 \) can be carried out as follows. From the discussion in Section 2, we first consider a reparametrization from \( (\theta_0, \theta_1, \theta_2) \) to \( (\theta_0, \theta_1, \theta_1 + d* \theta_2 = \theta_2) \) and express the joint distribution in terms of these new parameters. The problem then boils down to inference about \( \theta_2 \) when \( \theta_1 \) lies on the boundary! This is precisely the set up discussed in Section 2.

Finally, we consider the case when \( (iii) \theta_0 = \theta_1 = 0 \), which involves two points on the boundary, a case discussed in Section 2. Following Proposition 1, the asymptotic distribution of the MLEs \( \hat{\theta}_N \) of \( \theta \) is given by

\[
\sqrt{N}(\hat{\theta}_{0N}, \hat{\theta}_{1N}, \hat{\theta}_{2N} - \theta_2) \to (93)
\]

where again \( Z = (Z_0, Z_1, Z_2) \sim N(0, \hat{\Sigma}) \) and \( \hat{\theta}_0(Z_0, Z_1, Z_2), \hat{\theta}_1(Z_0, Z_1, Z_2) \) and \( \hat{\theta}_2(Z_0, Z_1, Z_2) \) are as defined in (90) - (92) with obvious changes.

As remarked earlier, while using \( \hat{\Sigma} \), we can use either the natural MLEs \( \hat{\theta}_N \) or take \( \theta_0 = \theta_1 = 0 \) and replace \( \theta_2 \) by \( \hat{\theta}_{2N} \).
From the above asymptotic distribution, it is indeed possible to derive the asymptotic distribution of any function of the MLEs such as \( \hat{H}(d^*) \), \( \hat{E}(d^*) \) and \( \hat{BMD} \). Details are omitted.

**Case 3: Cubic Weibull Model.** Taking \( k = 3 \) in (54), we get

\[ \pi_i(\theta|d_i) = 1 - e^{-[\theta_0 + d_i\theta_1 + d_i^2\theta_2 + d_i^3\theta_3]} \]

which readily gives

\[ \hat{H}(d^*) = e^{-[\theta_0 + d^*\theta_1 + d^*^2\theta_2 + d^*^3\theta_3]} \]

(94)

\[ \hat{E}(d^*) = 1 - e^{-(d^*\theta_1 + d^*^2\theta_2 + d^*^3\theta_3)}. \]

(95)

Regarding \( \hat{BMD} \), we note from (56) that \( d^* \) is the solution of a cubic equation. Following Cardano (1501-1576), an explicit formula for \( d^* \) is given as follows. Let \( \ln(1/(1-\gamma)) = -u \). Define

\[ A(\theta) = \frac{3\theta_3\theta_1 - \theta_2^2}{3\theta_3^2} \]

(96)

\[ B(\theta) = \frac{9\theta_3\theta_2\theta_1 - 2\theta_2^3 - 27u\theta_3^2}{27\theta_3^2} \]

(97)

\[ v(\theta) = \frac{-\sqrt{[B^2(\theta) + \frac{4A^3(\theta)}{27}] - B(\theta)}}{2}, \quad t(\theta) = v(\theta)^{1/3} \]

(98)

\[ s(\theta) = \frac{A(\theta)}{3t(\theta)}, \quad y(\theta) = s(\theta) - t(\theta). \]

(99)

Then \( BMD d^*(\theta) \) is given by

\[ d^*(\theta) = y(\theta) - \frac{\theta_2}{3\theta_3}. \]

(100)

Appropriate statistical inference about the above quantities follows upon computing the **natural** MLEs of the three parameters \( \theta_0, \theta_1, \theta_2 \) and \( \theta_3 \) based on the data \( (X_0, \ldots, X_m) \), namely \( \hat{\theta}_0N, \hat{\theta}_1N, \hat{\theta}_2N \) and \( \hat{\theta}_3N \), and applying their asymptotic distributional results derived in Section 2. It is easy to derive the Fisher information matrix \( I(\theta_0, \theta_1, \theta_2, \theta_3) : 4 \times 4 \) under the assumption of independent binomial distributions of the \( X_i \)'s, and hence \( \Sigma(\theta_0, \theta_1, \theta_2, \theta_3) = I(\theta_0, \theta_1, \theta_2, \theta_3)^{-1} \), where \( I(\theta_0, \theta_1, \theta_2, \theta_3) \) is given by
\[ I(\theta_0, \theta_1, \theta_2, \theta_3) = \]
\[
\begin{pmatrix}
\sum_{i=0}^{m} d_i 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 2 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 3 1 - \pi_i(\theta) \\
\sum_{i=0}^{m} d_i 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 2 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 3 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 4 1 - \pi_i(\theta) \\
\sum_{i=0}^{m} d_i 1 3 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 1 3 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 3 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 4 1 - \pi_i(\theta) \\
\sum_{i=0}^{m} d_i 3 2 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 3 2 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 3 1 - \pi_i(\theta) & \sum_{i=0}^{m} d_i 5 1 - \pi_i(\theta)
\end{pmatrix}.
\]

When \( \theta_0 > 0, \theta_1 > 0, \theta_2 > 0, \theta_3 > 0 \), we clearly have a regular parametric scenario situation. By Proposition 1, we then get the following result.

\[ \sqrt{N}(\hat{\theta}_{0N} - \theta_0, \hat{\theta}_{1N} - \theta_1, \hat{\theta}_{2N} - \theta_2, \hat{\theta}_{3N} - \theta_3) \rightarrow N_4[(0, 0, 0, 0), \Sigma(\theta)] \]  
\[
\sqrt{N}[\hat{\theta}_{0N} + d \ast \hat{\theta}_{1N} + d^2 \ast \hat{\theta}_{2N} + d^3 \ast \hat{\theta}_{3N}] - (\theta_0 + d \ast \theta_1 + d^2 \ast \theta_2 + d^3 \ast \theta_3) \\
\rightarrow N[0, (1, d \ast, d^2 \ast, d^3 \ast) \Sigma(\theta)(1, d \ast, d^2 \ast, d^3 \ast)'] \]  
\[
\sqrt{N}[\hat{\theta}_{1N} + d \ast \hat{\theta}_{2N} + d^2 \ast \hat{\theta}_{3N}] - (\theta_1 + d \ast \theta_2 + d^2 \ast \theta_3) \\
\rightarrow N[0, (1, d \ast, d^2 \ast) \Sigma(\theta)(1, d \ast, d^2 \ast)'].
\]

Here \( \Sigma(\theta) \) is the \( 3 \times 3 \) lower sub matrix of \( \Sigma(\theta) \) obtained by deleting the first row and the first column.

We are now in a position to discuss the inferential aspects of AR, RR and BMD. For AR\( (d \ast) \) defined in (94), a point estimate is given by

\[ \hat{AR}(d \ast) = 1 - e^{-[\theta_{0N} + d \ast \hat{\theta}_{1N} + d^2 \ast \hat{\theta}_{2N} + d^3 \ast \hat{\theta}_{3N}]} \]  
\[ \text{estSE}(\hat{AR}(d \ast)) = (1 - AR(d \ast)) \sqrt{(1, d \ast, d^2 \ast, d^3 \ast) \Sigma(\hat{\theta})(1, d \ast, d^2 \ast, d^3 \ast)'}.
\]

Using the fact that, asymptotically, \( (\hat{AR}(d \ast) - AR(d \ast))/[\text{estSE}(\hat{AR}(d \ast))] \) \( \sim N[0, 1] \), it is straightforward to carry out a test for a specified value of
$AR(d*)$ as well as to construct a confidence interval of $AR(d*)$. Of course, it is very easy to derive an asymptotic test for $\theta_0 + d* \theta_1 + d^2* \theta_2 + d^3* \theta_3$ and also to construct a confidence interval for $\theta_0 + d* \theta_1 + d^2* \theta_2 + d^3* \theta_3$ on the basis of $\hat{\theta}_{0N} + d* \hat{\theta}_{1N} + d^2* \hat{\theta}_{2N} + d^3* \hat{\theta}_{3N}$ directly from (102), which would readily solve the inference problems about $AR(d*)$.

Similarly, a point estimate of $ER(d*)$ defined in (95) is given by $\hat{ER}(d*) = 1 - e^{-(d*\hat{\theta}_{1N} + d^2*\hat{\theta}_{2N} + d^3*\hat{\theta}_{3N})}$ with its estimated asymptotic standard error as

$$estSE(\hat{ER}(d*)) = \sqrt{(d*, d^2*, d^3*)' \tilde{\Sigma}(\hat{\theta})(d*, d^2*, d^3*)}$$ (107)

where $\tilde{\Sigma}(\hat{\theta})$ is the lower $3 \times 3$ submatrix of $\hat{\Sigma}(\theta)$ obtained by deleting the first row and the first column.

Using the fact that, asymptotically,$(\hat{ER}(d*) - ER(d*)) /[est.SE(\hat{ER}(d*))] \sim N[0, 1]$, it is easy to carry out a test for a specified value of $ER(d*)$ as well as to construct a confidence interval of $ER(d*)$. Again, it is straightforward to derive an asymptotic test for $\theta_1 + d* \theta_2 + d^2* \theta_3$ and also to construct a confidence interval for $\theta_1 + d* \theta_2 + d^2* \theta_3$ on the basis of $\hat{\theta}_{1N} + d* \hat{\theta}_{2N} + d^2* \hat{\theta}_{3N}$ directly from (104), which would readily solve the inference problems about $ER(d*)$.

Finally, a point estimate of the BMD at specified extra risk level $\gamma$, defined in (100), is given by

$$d* (\hat{\theta}_N) = y(\hat{\theta}_N) - \frac{\hat{\theta}_{2N}}{3\hat{\theta}_{3N}}.$$ (108)

The asymptotic standard error of $BMD$ can be obtained by the delta method, and it can also be readily estimated. Details are omitted. Using the fact that, asymptotically,$(BMD - BMD) /[est.SE(BMD)] \sim N[0, 1]$, it is then possible to carry out a test for a specified value of BMD as well as to construct a confidence interval of BMD.

We now consider the case when some parameters may lie on the boundaries. Since we have assumed a cubic Weibull model, there are several possible scenarios: (i) $\theta_0 = 0$, (ii) $\theta_1 = 0$, (iii) $\theta_2 = 0$, (iv) $\theta_0 = \theta_1 = 0$, (v) $\theta_0 = \theta_2 = 0$, (vi) $\theta_1 = \theta_2 = 0$, and (vii) $\theta_0 = \theta_1 = \theta_2 = 0$.

When (i) $\theta_0 = 0$, which may well happen in some situations, we have the case of one parameter point being on the boundary, and hence the asymptotic distributional and inferential results mentioned above are not true. We
should note that for statistical inference about \( ER(d^*) \) and BMD which involve the parameters \( \theta_1, \theta_2 \) and \( \theta_3 \), all we would need is the asymptotic joint distribution of the relevant MLEs. Using Proposition 1, quite generally we get

\[
\sqrt{N}(\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3) \rightarrow (\hat{\theta}_0(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_1(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_2(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_3(Z_0, Z_1, Z_2, Z_3))
\]

where \( Z = (Z_0, Z_1, Z_2, Z_3) \sim N[0, \hat{\Sigma}] \) and from the discussion in Section 2, Theorem 2.2, we get

\[
\hat{\theta}_0(Z_0, Z_1, Z_2, Z_3) = Z_0 \text{ if } Z_0 > 0, \text{ and } = 0 \text{ if } Z_0 < 0
\]

\[
\hat{\theta}_1(Z_0, Z_1, Z_2, Z_3) = Z_1 \text{ if } Z_0 > 0, \text{ and } = Z_{1,0} \text{ if } Z_0 < 0.
\]

\[
\hat{\theta}_2(Z_0, Z_1, Z_2, Z_3) = Z_2 \text{ if } Z_0 > 0, \text{ and } = Z_{2,0} \text{ if } Z_0 < 0.
\]

\[
\hat{\theta}_3(Z_0, Z_1, Z_2, Z_3) = Z_3 \text{ if } Z_0 > 0, \text{ and } = Z_{3,0} \text{ if } Z_0 < 0.
\]

where \( Z_{1,0}, Z_{2,0} \) and \( Z_{3,0} \) are the usual residual variables. A remark about \( \hat{\Sigma} \) is in order here. Since \( \Sigma \) is obviously a function of \( \theta \), we can use \( \hat{\Sigma}(\hat{\theta}_N) \). In case it so happens that \( X_0 = 0 \) so that \( \hat{\theta}_0 N = 0 \), computation of \( \hat{\Sigma} \) may pose some difficulty (in terms of singularities!), and then we can take an arbitrary small value of \( \hat{\theta}_0 N \).

Returning to the inference problems, although we still have the same point estimates as before, namely,

\[
\hat{AR}(d^*) = 1 - e^{-(\hat{\theta}_0 N + d^* \hat{\theta}_1 N + d^* \hat{\theta}_2 N + d^* \hat{\theta}_3 N)}
\]

\[
\hat{ER}(d^*) = 1 - e^{-(d^* \hat{\theta}_1 N + d^* \hat{\theta}_2 N + d^* \hat{\theta}_3 N)}
\]

and a similar expression for \( BMD \), the asymptotic distributions of these estimates are not normal any more, and, in fact, these are obtained by replacing
\hat{\theta}_0, \hat{\theta}_{1N}, \hat{\theta}_{2N} \text{ and } \hat{\theta}_{3N} \text{ by } \hat{\theta}_0(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_1(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_2(Z_0, Z_1, Z_2, Z_3) \text{ and } \hat{\theta}_3(Z_0, Z_1, Z_2, Z_3), \text{ respectively, and deriving the resultant distributions by simulation.}

We have explicitly pursued below the inference method for \( ER(d*) = 1 - e^{-(d*\hat{\theta}_1 + d*\hat{\theta}_2 + d^2*\theta_3)} \) on the basis of \( \hat{E}R(d*) = 1 - e^{-(d*\hat{\theta}_{1N} + d^2*\hat{\theta}_{2N} + d^3*\hat{\theta}_{3N}).} \) The same for \( BMD \) can be pursued with some extra effort! Here are the results.

**Inference about \( ER.** Since \( ER(d*) = 1 - e^{-(d*\theta_1 + d*\theta_2 + d^2*\theta_3)} \), without any loss of generality, we consider equivalently the problem of drawing suitable inference about \( \hat{E}R(d*) = \theta_1 + d * \theta_2 + d^2 * \theta_3 \) which is based on \( \hat{E}R(d*) = \hat{\theta}_{1N} + d * \hat{\theta}_{2N} + d^2 * \hat{\theta}_{3N}. \) Applying Theorem 2, we conclude that

\[
\sqrt{N} \left[ (\hat{\theta}_{1N} + d * \hat{\theta}_{2N} + d^2 * \hat{\theta}_{3N}) - (\theta_1 + d * \theta_2 + d^2 * \theta_3) \right] \\
\sim Z_1 + d * Z_2 + d^2 * Z_3 \quad \text{if } Z_2 > 0 \\
\sim Z_{1,0} + d * Z_{2,0} + d^2 * Z_{3,0} \quad \text{if } Z_0 < 0.
\]

Since \( \Sigma(\hat{\theta}) \) is known, it is rather easy to simulate the asymptotic distribution of \( \sqrt{N}[(\hat{\theta}_{1N} + d * \hat{\theta}_{2N} + d^2 * \hat{\theta}_{3N}) - (\theta_1 + d * \theta_2 + d^2 * \theta_3)] \), from which the cut-off points of this distribution can be obtained. Tests and confidence intervals about \( \hat{E}R(d*) \) are then routinely derived.

What we have described so far is the Wald approach! For the derivation of the LRT for hypotheses about \( \hat{E}R(d*) \) and \( BMD, \) since these quantities involve only \( \theta_1 \) and \( \theta_2 \) which are regular parameters, it follows from the general theory discussed in Section 2 that irrespective of the nature of \( \theta_0, \) whether a boundary point or not, the asymptotic distribution of \(-2\ln(LRT)\) remains as \( \chi^2 \) with an appropriate d.f. We have compared the Wald test with the LRT test in the case of \( ER \) below.

We next consider the case \((ii) \theta_1 = 0\) which is still the case of one boundary parameter. Using results from Section 2, the asymptotic joint distribution of the MLEs \( \hat{\theta}_N \) is given by

\[
\sqrt{(N)}(\theta_{0N} - \theta_0, \hat{\theta}_{1N}, \hat{\theta}_{2N} - \theta_2, \hat{\theta}_{3N} - \theta_3) \rightarrow \\
\left( \hat{\theta}_0(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_1(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_2(Z_0, Z_1, Z_2, Z_3, \hat{\theta}_3(Z_0, Z_1, Z_2, Z_3)) \right)
\]
where, as before, \( \mathbf{Z} = (Z_0, Z_1, Z_2, Z_3) \sim N[\mathbf{0}, \hat{\Sigma}] \), and, following the discussion in Section 2, we have

\[
\hat{\theta}_0(Z_0, Z_1, Z_2, Z_3) = Z_0 \text{ if } Z_1 > 0, \quad \text{and} \quad = Z_{0.1} \text{ if } Z_1 < 0 \quad (118)
\]
\[
\hat{\theta}_1(Z_0, Z_1, Z_2, Z_3) = Z_1 \text{ if } Z_1 > 0, \quad \text{and} \quad = 0 \quad \text{if } Z_1 < 0 \quad (119)
\]
\[
\hat{\theta}_2(Z_0, Z_1, Z_2, Z_3) = Z_2 \text{ if } Z_1 > 0, \quad \text{and} \quad = Z_{2.1} \text{ if } Z_1 < 0 \quad (120)
\]
\[
\hat{\theta}_3(Z_0, Z_1, Z_2, Z_3) = Z_3 \text{ if } Z_1 > 0, \quad \text{and} \quad = Z_{3.1} \text{ if } Z_1 < 0 \quad (121)
\]

As in the previous case, a remark about \( \hat{\Sigma} \) is in order here. Since \( \Sigma \) is obviously a function of \( \theta \), we can use either \( \hat{\Sigma}(\hat{\theta}_N) \) or take \( \theta_1 = 0 \) and replace \( \theta_0, \theta_2 \) and \( \theta_3 \) by \( \hat{\theta}_{0N}, \hat{\theta}_{2N} \) and \( \hat{\theta}_{3N} \), respectively.

Using the above asymptotic distribution of the MLEs of \( \theta \), we can approximate the distributions of \( ER(d^*) \) and \( BMD \) by simulation and use them to carry out tests and confidence intervals for \( ER \) and \( BMD \).

We remark that the LRT for \( ER(d^*) \) which is equivalent to the linear function \( \theta_1 + d^* \theta_2 + d^{2*} \theta_3 \) can be carried out as follows. From the discussion in Section 2, we first consider a reparametrization from \((\theta_0, \theta_1, \theta_2, \theta_3)\) to \((\theta_0, \theta_1, \theta_2, \theta_1 + d^* \theta_2 + d^{2*} \theta_3 = \theta_{3*})\) and express the joint distribution in terms of these new parameters. The problem then boils down to inference about \( \theta_{3*} \) when \( \theta_1 \) lies on the boundary! This is precisely the set up discussed in Section 2.

We next consider the case when \((iv) \ \theta_0 = \theta_1 = 0 \), which involves two points on the boundary, a case discussed in Section 2. Following Proposition 1, the asymptotic distribution of the MLEs \( \hat{\theta}_N \) of \( \theta \) is given by

\[
\sqrt{N}(\hat{\theta}_{0N}, \hat{\theta}_{1N}, \hat{\theta}_{2N} - \theta_2, \hat{\theta}_{3N} - \theta_3) \rightarrow \quad (122)
\]

\[
\left( \hat{\theta}_0(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_1(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_2(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_3(Z_0, Z_1, Z_2, Z_3) \right)
\]

where again \( \mathbf{Z} = (Z_0, Z_1, Z_2, Z_3) \sim N[\mathbf{0}, \hat{\Sigma}] \) and \( \hat{\theta}_0(Z_0, Z_1, Z_2, Z_3), \hat{\theta}_1(Z_0, Z_1, Z_2, Z_3), \)
\( \theta_2(Z_0, Z_1, Z_2, Z_3) \) and \( \theta_3(Z_0, Z_1, Z_2, Z_3) \) are defined in Section 2.

As remarked earlier, while using \( \hat{\Sigma} \), we can use either the natural MLEs \( \hat{\theta}_N \) or take \( \theta_0 = \theta_1 = 0 \) and replace \( \theta_2 \) and \( \theta_3 \) by \( \hat{\theta}_{2N} \) and \( \hat{\theta}_{3N} \), respectively.

From the above asymptotic distribution, it is indeed possible to derive the asymptotic distribution of any function of the MLEs such as \( AR(d^*) \), \( ER(d^*) \) and \( BMD \). Details are omitted.
4. Applications

4.1 Example

In this section we consider an example from National Toxicological Program (NTP, 1993). In this study, B6C3F1 mice were exposed to 1,3-Butadiene. The outcome was heart hemangiosarcomas. This data was also analyzed by Bailer and Smith (1994). The data is given in the following table (only doses up to 200ppm are given, following Bailer and Smith):

<table>
<thead>
<tr>
<th>Dose (ppm)</th>
<th>Number at risk</th>
<th>Number of tumor-bearing animals</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>6.25</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>62.5</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>200</td>
<td>50</td>
<td>21</td>
</tr>
</tbody>
</table>

A 3-stage Weibull model (see Case 3 in Section 3) fits well to this data and two parameters $\theta_0$ and $\theta_1$ are estimated to be on the boundary. Bailer and Smith (1994) considered several approaches to constructing an upper 95th confidence limit on Extra Risk. Their results for this confidence limit are tabulated below:

<table>
<thead>
<tr>
<th>Dose (ppm)</th>
<th>Likelihood</th>
<th>Nonparametric bootstrap</th>
<th>Parametric bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>1.404E-3</td>
<td>5.167E-5</td>
<td>5.168E-5</td>
</tr>
<tr>
<td>0.20</td>
<td>1.405E-4</td>
<td>5.154E-7</td>
<td>5.156E-7</td>
</tr>
<tr>
<td>0.02</td>
<td>1.405E-5</td>
<td>5.153E-9</td>
<td>5.155E-9</td>
</tr>
</tbody>
</table>

The likelihood upper bound in the table above is calculated according to a methodology (cited in Bailer and Smith, 1994) based on Crump, Guess and Deal (1977): it is the profile likelihood bound on a linear term of the multistage model using $\chi^2_1$ as an asymptotic distribution. That is the situation considered in Self and Liang (1987), Case 6: one parameter of interest and one nuisance parameter are on their boundaries. In this case, the correct asymptotic distribution is a 50:50 mixture of $\chi^2_1$ and $\chi^2_2$.

We calculated 90% confidence intervals for the extra risk using asymptotic normality (Wald) and also using asymptotic distribution derived according to Self and Liang (1987). The results are given below:
Confidence Intervals for Extra Risk at Three Doses

<table>
<thead>
<tr>
<th>Dose (ppm)</th>
<th>Wald Confidence Interval</th>
<th>Self-Liang Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>(0; 4.476E-5)</td>
<td>(2.143E-6; 4.471E-5)</td>
</tr>
<tr>
<td>0.20</td>
<td>(0; 5.027E-7)</td>
<td>(1.804E-8; 4.982E-7)</td>
</tr>
<tr>
<td>0.02</td>
<td>(0; 1.060E-8)</td>
<td>(1.770E-10; 1.063E-8)</td>
</tr>
</tbody>
</table>

It is interesting to note that the upper bound of both Wald and Self-Liang confidence intervals agree well to each other and to both nonparametric and parametric bootstrap confidence bounds. However, for all 3 doses, the Wald 90% confidence interval contains 0, but confidence interval calculated according to Self and Liang (1987) does not. Also, Self-Liang interval is shorter than Wald’s.

4.2 Simulations.

Monte-Carlo simulations were performed to compare coverage of the confidence intervals for the model parameters $\theta$, extra risk (ER) - linear function of parameters as well as benchmark dose (BMD) - linear or nonlinear, depending on the order of the model, function of parameters. We considered linear, quadratic and cubic models ($k=1,2,3$). In each case, the true value of one or two of the parameters (but not leading coefficient) was on the boundary (equal to zero) or true value of one or two parameters was not on the boundary, but their estimates often were. All simulations were performed for 4 groups (1 control group and 3 dose groups), each of 50 animals. The chosen doses were 0, 1/4, 1/2 and 1. This imitates standard rodent bioassay setup. The nonzero parameters were chosen to provide range of incidences for each dose group.

Monte-Carlo simulations were performed as follows. The model parameters define the binomial response probabilities $P(d, \theta)$ at each of the four doses. For each of 5000 realizations, these probabilities were used to generate data $X_d \sim Binom(N = 50, P(d, \theta))$. The resulting experiment data was used to estimate parameters of the model, using BMDS programs (USEPA, 2006). The information matrix was estimated. To avoid singularities, if the background rate was estimated to be 0, $\theta_0$ was taken to be 5E-4, as discussed in Section 3. In such cases, confidence interval for $\theta_0$ depends on the chosen substitution value and therefore confidence intervals for $\theta_0$ are not shown. It should be possible to derive confidence interval for $\theta_0$ via re-sampling procedure, but this is beyond the scope of this report.
Using information matrix, Wald confidence intervals for all parameters described above were computed. If a parameter estimate(s) was on the boundary, the confidence interval was computed using asymptotic distribution derived according to Self and Liang (1987) (See also section 3). The confidence intervals values were retained and procedure repeated 5000 times. For estimating coverage of confidence intervals, only those realizations were retained for which the $\chi^2$ goodness-of-fit P-value exceeded .1, conforming to the practice recommended by EPA (USEPA 2000) and for which estimate of leading coefficient was positive, so that models of the same order are compared to each other.

In the following table, the "Wald" column shows coverage for Wald confidence interval, while the "Self-Liang" column shows coverage of the procedure such that for each simulation, depending on the estimated value of the corresponding parameter being on the boundary or not, interval computed according to the asymptotic distribution derived in Self and Liang (1987) or Wald interval was used. The "Fit" column shows percent of the simulations when corresponding model fit well ($P > .1$), and the estimate of leading coefficient was positive. The desired coverage is 90%.

For linear model (Tables 1a and 1b), both Wald and Self-Liang confidence intervals have practically the same coverage and average length for wide range of the remaining parameter. The coverage is very close to the desired level.

For quadratic and cubic models (Tables 2a-2d, 3), and one or two parameters whose true or estimated value was on the boundary, the Wald confidence interval generally exceeds nominal coverage and Self-Liang confidence intervals are generally closer to the desired confidence level of 90%. Also, in almost all cases, Self-Liang intervals improve on average length compared to Wald intervals.
Table 1a. Coverage for linear model, when $\theta_0 = 0$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Wald</th>
<th>Self-Liang</th>
<th>LRT</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CI</td>
<td>Length</td>
<td>CI</td>
<td>Length</td>
</tr>
<tr>
<td>$\theta_1 = 0.5$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.211$</td>
<td>89.6</td>
<td>0.275</td>
<td>89.1</td>
</tr>
<tr>
<td>$\theta_1 = 1$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.105$</td>
<td>90.1</td>
<td>0.430</td>
<td>90.2</td>
</tr>
<tr>
<td>$\theta_1 = 2$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.053$</td>
<td>90.5</td>
<td>0.743</td>
<td>90.4</td>
</tr>
<tr>
<td>$\theta_1 = 3$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.035$</td>
<td>90.9</td>
<td>1.112</td>
<td>90.8</td>
</tr>
</tbody>
</table>

Table 1b. Coverage for linear model, when $\theta_0 = 0.02 > 0$, but $\hat{\theta}_0$ is sometimes 0.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Wald</th>
<th>Self-Liang</th>
<th>LRT</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CI</td>
<td>Length</td>
<td>CI</td>
<td>Length</td>
</tr>
<tr>
<td>$\theta_1 = 0.5$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.211$</td>
<td>89.8</td>
<td>0.304</td>
<td>89.9</td>
</tr>
<tr>
<td>$\theta_1 = 1$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.105$</td>
<td>90.3</td>
<td>0.138</td>
<td>92.1</td>
</tr>
<tr>
<td>$\theta_1 = 2$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.053$</td>
<td>90.2</td>
<td>0.767</td>
<td>90.3</td>
</tr>
<tr>
<td>$\theta_1 = 3$</td>
<td>$ER = 0.1$</td>
<td>$BD = 0.035$</td>
<td>90.8</td>
<td>1.140</td>
<td>90.7</td>
</tr>
</tbody>
</table>
Table 2a. Coverage for quadratic model, when $\theta_0 = 0$.

<table>
<thead>
<tr>
<th>Parameter True Values</th>
<th>Wald CI</th>
<th>Wald Length</th>
<th>Self-Liang CI</th>
<th>Self-Liang Length</th>
<th>LRT CI</th>
<th>LRT Length</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 = 0.05$</td>
<td>99.2</td>
<td>0.471</td>
<td>99.3</td>
<td>0.386</td>
<td></td>
<td></td>
<td>96.4</td>
</tr>
<tr>
<td>$\theta_2 = 0.5$</td>
<td>97.0</td>
<td>0.701</td>
<td>95.8</td>
<td>0.626</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>96.8</td>
<td>0.097</td>
<td>96.0</td>
<td>0.089</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 1.035$</td>
<td>96.2</td>
<td>0.239</td>
<td>96.0</td>
<td>0.224</td>
<td>94.4</td>
<td>0.226</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.05$</td>
<td>97.9</td>
<td>0.632</td>
<td>98.0</td>
<td>0.503</td>
<td></td>
<td></td>
<td>97.8</td>
</tr>
<tr>
<td>$\theta_2 = 1$</td>
<td>95.8</td>
<td>1.020</td>
<td>93.7</td>
<td>0.902</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>97.2</td>
<td>0.108</td>
<td>97.2</td>
<td>0.094</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.301$</td>
<td>95.6</td>
<td>0.182</td>
<td>97.2</td>
<td>0.158</td>
<td>93.8</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.05$</td>
<td>97.5</td>
<td>0.903</td>
<td>97.7</td>
<td>0.702</td>
<td></td>
<td></td>
<td>97.4</td>
</tr>
<tr>
<td>$\theta_2 = 2$</td>
<td>94.9</td>
<td>1.695</td>
<td>92.7</td>
<td>1.489</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>97.3</td>
<td>0.121</td>
<td>96.7</td>
<td>0.102</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.217$</td>
<td>95.7</td>
<td>0.142</td>
<td>96.7</td>
<td>0.117</td>
<td>93.9</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.05$</td>
<td>97.7</td>
<td>1.169</td>
<td>97.8</td>
<td>0.882</td>
<td></td>
<td></td>
<td>97.6</td>
</tr>
<tr>
<td>$\theta_2 = 3$</td>
<td>95.1</td>
<td>2.536</td>
<td>92.4</td>
<td>2.193</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>97.3</td>
<td>0.131</td>
<td>96.3</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.179$</td>
<td>95.0</td>
<td>0.125</td>
<td>96.3</td>
<td>0.100</td>
<td>93.5</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.8$</td>
<td>91.2</td>
<td>1.059</td>
<td>90.8</td>
<td>1.056</td>
<td></td>
<td></td>
<td>87.7</td>
</tr>
<tr>
<td>$\theta_2 = 0.5$</td>
<td>94.0</td>
<td>1.554</td>
<td>93.8</td>
<td>1.551</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>91.3</td>
<td>0.099</td>
<td>90.7</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.122$</td>
<td>93.5</td>
<td>0.131</td>
<td>90.7</td>
<td>0.136</td>
<td>92.9</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.8$</td>
<td>91.3</td>
<td>1.228</td>
<td>91.1</td>
<td>1.222</td>
<td></td>
<td></td>
<td>96.8</td>
</tr>
<tr>
<td>$\theta_2 = 1$</td>
<td>93.5</td>
<td>1.952</td>
<td>93.5</td>
<td>1.945</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>91.0</td>
<td>0.107</td>
<td>91.0</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.115$</td>
<td>89.1</td>
<td>0.118</td>
<td>91.0</td>
<td>0.121</td>
<td>92.0</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.8$</td>
<td>92.9</td>
<td>1.595</td>
<td>92.8</td>
<td>1.548</td>
<td></td>
<td></td>
<td>99.6</td>
</tr>
<tr>
<td>$\theta_2 = 2$</td>
<td>94.5</td>
<td>3.016</td>
<td>94.5</td>
<td>2.949</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>92.7</td>
<td>0.124</td>
<td>92.5</td>
<td>0.121</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.104$</td>
<td>88.9</td>
<td>0.112</td>
<td>92.5</td>
<td>0.110</td>
<td>91.3</td>
<td>0.097</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0.8$</td>
<td>96.4</td>
<td>2.045</td>
<td>96.4</td>
<td>1.899</td>
<td></td>
<td></td>
<td>99.9</td>
</tr>
<tr>
<td>$\theta_2 = 3$</td>
<td>94.6</td>
<td>4.481</td>
<td>94.6</td>
<td>4.234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>95.5</td>
<td>0.145</td>
<td>95.2</td>
<td>0.136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.097$</td>
<td>91.5</td>
<td>0.115</td>
<td>95.2</td>
<td>0.105</td>
<td>90.7</td>
<td>0.089</td>
<td></td>
</tr>
</tbody>
</table>
Table 2b. Coverage for quadratic model, when $\theta_1 = 0$.

<table>
<thead>
<tr>
<th>Parameter True Values</th>
<th>Wald CI</th>
<th>Wald Length</th>
<th>Self-Liang CI</th>
<th>Self-Liang Length</th>
<th>LRT CI</th>
<th>LRT Length</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_0 = 0.02$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>97.6</td>
<td>0.572</td>
<td>97.8</td>
<td>0.418</td>
<td></td>
<td></td>
<td>92.0</td>
</tr>
<tr>
<td>$\theta_2 = 0.5$</td>
<td>96.8</td>
<td>0.775</td>
<td>92.0</td>
<td>0.646</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>95.2</td>
<td>0.123</td>
<td>93.1</td>
<td>0.104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.459$</td>
<td>94.9</td>
<td>0.315</td>
<td>92.9</td>
<td>0.285</td>
<td>90.0</td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>96.4</td>
<td>0.753</td>
<td>96.5</td>
<td>0.564</td>
<td></td>
<td></td>
<td>92.9</td>
</tr>
<tr>
<td>$\theta_2 = 1$</td>
<td>95.1</td>
<td>1.128</td>
<td>90.0</td>
<td>0.955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>95.7</td>
<td>0.133</td>
<td>92.2</td>
<td>0.109</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.325$</td>
<td>94.2</td>
<td>0.229</td>
<td>91.5</td>
<td>0.189</td>
<td>89.1</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>96.5</td>
<td>1.081</td>
<td>96.6</td>
<td>0.797</td>
<td></td>
<td></td>
<td>93.0</td>
</tr>
<tr>
<td>$\theta_2 = 2$</td>
<td>94.6</td>
<td>1.896</td>
<td>90.1</td>
<td>1.593</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>96.1</td>
<td>0.147</td>
<td>92.2</td>
<td>0.117</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.230$</td>
<td>94.2</td>
<td>0.174</td>
<td>91.4</td>
<td>0.135</td>
<td>89.9</td>
<td>0.112</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>96.7</td>
<td>1.437</td>
<td>96.7</td>
<td>1.051</td>
<td></td>
<td></td>
<td>93.0</td>
</tr>
<tr>
<td>$\theta_2 = 3$</td>
<td>94.6</td>
<td>2.891</td>
<td>90.2</td>
<td>2.401</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>96.1</td>
<td>0.164</td>
<td>92.3</td>
<td>0.127</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.187$</td>
<td>94.2</td>
<td>0.156</td>
<td>91.4</td>
<td>0.117</td>
<td>89.9</td>
<td>0.094</td>
<td></td>
</tr>
</tbody>
</table>
Table 2b, continued.

<table>
<thead>
<tr>
<th>Parameter True Values</th>
<th>Wald CI</th>
<th>Length</th>
<th>Self-Liang CI</th>
<th>Length</th>
<th>LRT CI</th>
<th>Length</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 = 0$</td>
<td>99.4</td>
<td>1.299</td>
<td>99.4</td>
<td>92.1</td>
<td></td>
<td></td>
<td>88.9</td>
</tr>
<tr>
<td>$\theta_2 = 0.5$</td>
<td>100</td>
<td>1.390</td>
<td>94.7</td>
<td>1.111</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>96.7</td>
<td>0.300</td>
<td>92.4</td>
<td>0.227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.459$</td>
<td>94.4</td>
<td>0.745</td>
<td>90.1</td>
<td>0.597</td>
<td>92.0</td>
<td>Infty</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>97.1</td>
<td>1.454</td>
<td>97.1</td>
<td>1.075</td>
<td></td>
<td></td>
<td>97.4</td>
</tr>
<tr>
<td>$\theta_2 = 1$</td>
<td>97.1</td>
<td>1.721</td>
<td>92.4</td>
<td>1.429</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>95.1</td>
<td>0.276</td>
<td>91.0</td>
<td>0.210</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.325$</td>
<td>90.1</td>
<td>0.453</td>
<td>86.0</td>
<td>0.325</td>
<td>90.8</td>
<td>0.246</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>95.3</td>
<td>1.769</td>
<td>95.3</td>
<td>1.302</td>
<td></td>
<td></td>
<td>99.8</td>
</tr>
<tr>
<td>$\theta_2 = 2$</td>
<td>94.9</td>
<td>2.534</td>
<td>90.2</td>
<td>2.122</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>93.9</td>
<td>0.256</td>
<td>89.7</td>
<td>0.194</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.230$</td>
<td>89.3</td>
<td>0.288</td>
<td>85.1</td>
<td>0.201</td>
<td>88.5</td>
<td>0.152</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>95.4</td>
<td>1.439</td>
<td>95.4</td>
<td>1.049</td>
<td></td>
<td></td>
<td>93.6</td>
</tr>
<tr>
<td>$\theta_2 = 3$</td>
<td>94.7</td>
<td>2.894</td>
<td>90.1</td>
<td>2.399</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>94.4</td>
<td>0.164</td>
<td>90.0</td>
<td>0.127</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.187$</td>
<td>92.8</td>
<td>0.157</td>
<td>90.3</td>
<td>0.116</td>
<td>88.9</td>
<td>0.094</td>
<td></td>
</tr>
</tbody>
</table>
Table 2c. Coverage for quadratic model, when $\theta_0 = \theta_1 = 0$.

<table>
<thead>
<tr>
<th>Parameter True Values</th>
<th>Wald CI</th>
<th>Length</th>
<th>Self-Liang CI</th>
<th>Length</th>
<th>LRT CI</th>
<th>Length</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 = 0$</td>
<td>98.8</td>
<td>0.421</td>
<td>98.9</td>
<td>0.307</td>
<td></td>
<td></td>
<td>97.9</td>
</tr>
<tr>
<td>$\theta_2 = 0.5$</td>
<td>95.7</td>
<td>0.637</td>
<td>90.1</td>
<td>0.538</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>96.4</td>
<td>0.092</td>
<td>93.8</td>
<td>0.083</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.459$</td>
<td>96.3</td>
<td>0.238</td>
<td>93.8</td>
<td>0.226</td>
<td>92.0</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>98.1</td>
<td>0.591</td>
<td>98.3</td>
<td>0.432</td>
<td></td>
<td></td>
<td>98.1</td>
</tr>
<tr>
<td>$\theta_2 = 1$</td>
<td>95.3</td>
<td>0.962</td>
<td>90.0</td>
<td>0.819</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>97.7</td>
<td>0.106</td>
<td>94.9</td>
<td>0.088</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.324$</td>
<td>96.7</td>
<td>0.182</td>
<td>94.9</td>
<td>0.154</td>
<td>91.7</td>
<td>0.148</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>98.0</td>
<td>0.869</td>
<td>98.1</td>
<td>0.636</td>
<td></td>
<td></td>
<td>97.5</td>
</tr>
<tr>
<td>$\theta_2 = 2$</td>
<td>95.4</td>
<td>1.634</td>
<td>90.3</td>
<td>1.399</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>97.7</td>
<td>0.121</td>
<td>94.5</td>
<td>0.098</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.230$</td>
<td>96.3</td>
<td>0.143</td>
<td>94.5</td>
<td>0.114</td>
<td>91.1</td>
<td>0.102</td>
<td></td>
</tr>
<tr>
<td>$\theta_1 = 0$</td>
<td>97.7</td>
<td>1.133</td>
<td>97.8</td>
<td>0.819</td>
<td></td>
<td></td>
<td>97.4</td>
</tr>
<tr>
<td>$\theta_2 = 3$</td>
<td>94.8</td>
<td>2.459</td>
<td>90.1</td>
<td>2.089</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ER = 0.1$</td>
<td>97.4</td>
<td>0.131</td>
<td>94.2</td>
<td>0.104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BD = 0.187$</td>
<td>95.6</td>
<td>0.126</td>
<td>94.2</td>
<td>0.098</td>
<td>90.6</td>
<td>0.085</td>
<td></td>
</tr>
</tbody>
</table>
Table 2d. Coverage for quadratic model, when \( \theta_0 = \theta_1 = 0.02 > 0 \),
but \( \hat{\theta}_0 \) and/or \( \hat{\theta}_1 \) is/are sometimes 0.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Wald</th>
<th>Self-Liang</th>
<th>LRT</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CI</td>
<td>Length</td>
<td>CI</td>
<td>Length</td>
</tr>
<tr>
<td>( \theta_1 = 0.02 )</td>
<td>97.8</td>
<td>0.623</td>
<td>97.8</td>
<td>0.474</td>
<td>91.8</td>
</tr>
<tr>
<td>( \theta_2 = 0.5 )</td>
<td>97.0</td>
<td>0.813</td>
<td>93.6</td>
<td>0.692</td>
<td></td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>95.2</td>
<td>0.134</td>
<td>94.2</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>( BD = 0.439 )</td>
<td>93.8</td>
<td>0.333</td>
<td>94.0</td>
<td>0.296</td>
<td>91.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Wald</th>
<th>Self-Liang</th>
<th>LRT</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CI</td>
<td>Length</td>
<td>CI</td>
<td>Length</td>
</tr>
<tr>
<td>( \theta_1 = 0.02 )</td>
<td>96.5</td>
<td>0.778</td>
<td>96.8</td>
<td>0.590</td>
<td>93.7</td>
</tr>
<tr>
<td>( \theta_2 = 1 )</td>
<td>95.5</td>
<td>1.128</td>
<td>92.3</td>
<td>0.968</td>
<td></td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>95.4</td>
<td>0.139</td>
<td>93.9</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>( BD = 0.315 )</td>
<td>93.4</td>
<td>0.234</td>
<td>93.2</td>
<td>0.192</td>
<td>90.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Wald</th>
<th>Self-Liang</th>
<th>LRT</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CI</td>
<td>Length</td>
<td>CI</td>
<td>Length</td>
</tr>
<tr>
<td>( \theta_1 = 0.02 )</td>
<td>96.5</td>
<td>1.040</td>
<td>96.7</td>
<td>0.788</td>
<td>93.4</td>
</tr>
<tr>
<td>( \theta_2 = 2 )</td>
<td>95.1</td>
<td>1.794</td>
<td>91.0</td>
<td>1.553</td>
<td></td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>96.1</td>
<td>0.145</td>
<td>94.3</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td>( BD = 0.225 )</td>
<td>93.0</td>
<td>0.169</td>
<td>92.8</td>
<td>0.134</td>
<td>90.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Wald</th>
<th>Self-Liang</th>
<th>LRT</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CI</td>
<td>Length</td>
<td>CI</td>
<td>Length</td>
</tr>
<tr>
<td>( \theta_1 = 0.02 )</td>
<td>96.2</td>
<td>1.451</td>
<td>96.2</td>
<td>1.065</td>
<td>93.7</td>
</tr>
<tr>
<td>( \theta_2 = 3 )</td>
<td>94.5</td>
<td>2.923</td>
<td>90.8</td>
<td>2.429</td>
<td></td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>95.9</td>
<td>0.163</td>
<td>92.7</td>
<td>0.127</td>
<td></td>
</tr>
<tr>
<td>( BD = 0.184 )</td>
<td>93.5</td>
<td>0.154</td>
<td>91.5</td>
<td>0.117</td>
<td>90.8</td>
</tr>
</tbody>
</table>
Table 3. Coverage for cubic model under several scenarios, when some of the parameters or their estimates are on the boundary.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Wald CI</th>
<th>Wald Length</th>
<th>Self-Liang CI</th>
<th>Self-Liang Length</th>
<th>LRT CI</th>
<th>LRT Length</th>
<th>Fit %</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_0 = 0.1 )</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>96.5</td>
<td>98.8</td>
<td>92.7</td>
<td>79.9</td>
</tr>
<tr>
<td>( \theta_1 = 0 )</td>
<td>100</td>
<td>11.49</td>
<td>100</td>
<td>9.110</td>
<td>96.5</td>
<td>0.227</td>
<td>79.9</td>
</tr>
<tr>
<td>( \theta_2 = 0 )</td>
<td>100</td>
<td>3.268</td>
<td>100</td>
<td>1.056</td>
<td>96.5</td>
<td>0.227</td>
<td>79.9</td>
</tr>
<tr>
<td>( \theta_3 = 2 )</td>
<td>99.7</td>
<td>4.085</td>
<td>99.7</td>
<td>4.085</td>
<td>100</td>
<td>17.86</td>
<td>83.9</td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>99.4</td>
<td>0.235</td>
<td>99.4</td>
<td>0.235</td>
<td>100</td>
<td>17.86</td>
<td>83.9</td>
</tr>
<tr>
<td>( BD = 0.375 )</td>
<td>94.9</td>
<td>0.920</td>
<td>94.9</td>
<td>0.920</td>
<td>99.7</td>
<td>4.085</td>
<td>83.9</td>
</tr>
<tr>
<td>( \theta_0 = 0 )</td>
<td>100</td>
<td>20.52</td>
<td>100</td>
<td>10.45</td>
<td>100</td>
<td>20.52</td>
<td>83.9</td>
</tr>
<tr>
<td>( \theta_1 = 1 )</td>
<td>98.0</td>
<td>17.61</td>
<td>98.0</td>
<td>17.61</td>
<td>100</td>
<td>20.52</td>
<td>83.9</td>
</tr>
<tr>
<td>( \theta_2 = 0 )</td>
<td>99.4</td>
<td>0.235</td>
<td>99.4</td>
<td>0.235</td>
<td>100</td>
<td>17.86</td>
<td>83.9</td>
</tr>
<tr>
<td>( \theta_3 = 3 )</td>
<td>94.9</td>
<td>0.920</td>
<td>94.9</td>
<td>0.920</td>
<td>100</td>
<td>20.52</td>
<td>83.9</td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>100</td>
<td>17.61</td>
<td>100</td>
<td>10.45</td>
<td>100</td>
<td>20.52</td>
<td>83.9</td>
</tr>
<tr>
<td>( BD = 0.102 )</td>
<td>99.7</td>
<td>4.085</td>
<td>99.7</td>
<td>4.085</td>
<td>100</td>
<td>17.86</td>
<td>83.9</td>
</tr>
<tr>
<td>( \theta_0 = 0 )</td>
<td>100</td>
<td>20.52</td>
<td>100</td>
<td>10.45</td>
<td>100</td>
<td>20.52</td>
<td>83.9</td>
</tr>
<tr>
<td>( \theta_1 = 2 )</td>
<td>100</td>
<td>5.939</td>
<td>100</td>
<td>3.748</td>
<td>100</td>
<td>5.939</td>
<td>43.6</td>
</tr>
<tr>
<td>( \theta_2 = 0 )</td>
<td>100</td>
<td>15.72</td>
<td>100</td>
<td>9.518</td>
<td>100</td>
<td>5.939</td>
<td>43.6</td>
</tr>
<tr>
<td>( \theta_3 = 0.5 )</td>
<td>100</td>
<td>7307</td>
<td>100</td>
<td>92.6</td>
<td>100</td>
<td>5.939</td>
<td>43.6</td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>99.4</td>
<td>0.235</td>
<td>99.4</td>
<td>0.235</td>
<td>100</td>
<td>17.86</td>
<td>83.9</td>
</tr>
<tr>
<td>( BD = 0.053 )</td>
<td>100</td>
<td>7307</td>
<td>100</td>
<td>92.6</td>
<td>100</td>
<td>5.939</td>
<td>43.6</td>
</tr>
<tr>
<td>( \theta_0 = 0.2 )</td>
<td>100</td>
<td>3.313</td>
<td>100</td>
<td>1.108</td>
<td>100</td>
<td>3.313</td>
<td>79.2</td>
</tr>
<tr>
<td>( \theta_1 = 0.01 )</td>
<td>100</td>
<td>10.66</td>
<td>100</td>
<td>3.500</td>
<td>100</td>
<td>3.313</td>
<td>79.2</td>
</tr>
<tr>
<td>( \theta_2 = 0.01 )</td>
<td>100</td>
<td>7.758</td>
<td>100</td>
<td>3.152</td>
<td>100</td>
<td>3.313</td>
<td>79.2</td>
</tr>
<tr>
<td>( \theta_3 = 1 )</td>
<td>95.8</td>
<td>0.311</td>
<td>95.8</td>
<td>0.311</td>
<td>100</td>
<td>3.313</td>
<td>79.2</td>
</tr>
<tr>
<td>( ER = 0.1 )</td>
<td>100</td>
<td>7307</td>
<td>100</td>
<td>92.6</td>
<td>100</td>
<td>5.939</td>
<td>43.6</td>
</tr>
<tr>
<td>( BD = 0.462 )</td>
<td>99.9</td>
<td>1.8368</td>
<td>99.9</td>
<td>0.387</td>
<td>100</td>
<td>3.313</td>
<td>79.2</td>
</tr>
</tbody>
</table>
Acknowledgment

We thank Paul White and Ravi Subramaniam of NCEA/ORD/USEPA for encouragement. Research of the first author (Sinha) is supported under a NCEA visiting faculty fellowship program, and is thankfully acknowledged.
Here we provide a proof of Theorem 2.1. From (23), by a direct computation, we get

\[
G(w) = P[W \leq w] = P[(Z_1^2 \leq w)](Z_2 > 0, Z_{2,1} > 0) + P[(Z_1^2 - Z_2^2 \leq w)]I(Z_2 < 0, Z_{2,1} > 0) + P[(Z_{1,2}^2 \leq w(1 - \rho^2))]I(Z_2 < 0, Z_{2,1} \leq 0) + P[(Z_{1}^2 + \frac{Z_{2,1}^2}{1 - \rho^2} \leq w)]I(Z_2 > 0, Z_{2,1} \leq 0).
\]  

We now apply the method of transformation from \((Z_1, Z_2) \rightarrow (Z_{2,1} = U, Z_2 = V)\) with \((U, V) \sim N([0, 0], var(U) = 1 - \rho^2, var(V) = 1, cov(U, V) = 1 - \rho^2)\). Obviously, \(Z_1 = (V - U)/\rho\) and \(Z_{1,2} = (V - U)/\rho - \rho V\). Moreover, \(V|U = u \sim N[u, \rho^2]\), implying \((V - U)/\rho \sim N[0, 1]\). This yields the various terms in (123) as follows.

**Term 1**

\[
= P[\{(V - U)^2 \leq w\rho^2\}I(U > 0, V > 0)] = P[\{U - \rho \sqrt{w} \leq V \leq U + \rho \sqrt{w}\}I(U > 0, V > 0)]
\]

\[
= \int_{\rho \sqrt{w}}^{\rho \sqrt{w}} \left[ \int_{0}^{u + \rho \sqrt{w}} f(v|u)dv \right] f(u)du + \int_{\rho \sqrt{w}}^{\infty} \left[ \int_{u - \rho \sqrt{w}}^{u + \rho \sqrt{w}} f(v|u)dv \right] f(u)du
\]

\[
= \int_{0}^{\rho \sqrt{w}} \left[ \int_{\frac{v}{\sqrt{1 - \rho^2}}}^{v} N(0, 1)dx \right] \frac{N(0, 1)}{\rho}du + \left[ \int_{-\infty}^{\rho \sqrt{w}} N(0, 1)dx \right] \left[ \int_{\frac{\rho \sqrt{w}}{\sqrt{1 - \rho^2}}}^{\infty} N(0, 1)dx \right].
\]

**Term 2**

\[
= P[\{(V - U)^2 \leq \rho^2(w + V^2)\} \cdot I(U > 0, V < 0)] = P[U \leq V + \rho \sqrt{w + V^2}, V < 0, V + \rho \sqrt{w + V^2} > 0]
\]

\[
= \int_{-\sqrt{\frac{w \rho^2}{1 - \rho^2}}}^{0} P\left[ -\frac{v(1 - \rho^2)}{\rho} \leq N(0, 1) \leq \rho v + \sqrt{w + v^2} \right] N(0, 1)dv.
\]
Term 3 = $P \left[ \{(V - U - \rho^2 V)^2 \leq w \rho^2 (1 - \rho^2)\} \cdot I(U < 0, V < 0) \right]$ \hspace{1cm} (126)

= $P \left\{ V(1 - \rho^2) - \sqrt{w \rho^2 (1 - \rho^2)} \leq U \leq V(1 - \rho^2) + \sqrt{w \rho^2 (1 - \rho^2)} \right\} \cdot I(U < 0, V < 0)$

= $P \left\{ V(1 - \rho^2) - \sqrt{w \rho^2 (1 - \rho^2)} \leq U \leq V(1 - \rho^2) + \sqrt{w \rho^2 (1 - \rho^2)} \right\}, V < -\sqrt{\frac{w \rho^2}{(1 - \rho^2)}}$

+ $P \left\{ V(1 - \rho^2) - \sqrt{w \rho^2 (1 - \rho^2)} \leq U \leq 0, -\sqrt{\frac{w \rho^2}{(1 - \rho^2)}} < V < 0 \right\}$

= $P \left\{ -\sqrt{w(1 - \rho^2)} < N(0, 1) < \sqrt{w(1 - \rho^2)} \right\} \cdot P \left( N(0, 1) < -\sqrt{\frac{w \rho^2}{(1 - \rho^2)}} \right)$

+ $P \left\{ -\sqrt{w(1 - \rho^2)} < N(0, 1) < -V(1 - \rho^2)/\rho \right\}, -\sqrt{\frac{w \rho^2}{(1 - \rho^2)}} < V < 0 \right\}$.

Term 4 = $P \left[ \{(V - U)^2 (1 - \rho^2) + U^2 \rho^2 \leq w \rho^2 (1 - \rho^2)\} \cdot I(U < 0, V > 0) \right]$ \hspace{1cm} (127)

= $P \left[ \{(U - V(1 - \rho^2))^2 \leq (w - V^2) \rho^2 (1 - \rho^2)\}, I(U < 0, V > 0) \right]$

= $P \left[ V(1 - \rho^2) - \sqrt{(w - V^2) \rho^2 (1 - \rho^2)} \leq U < 0, 0 < V < \rho \sqrt{w} \right]$

= $\int_{0}^{\rho \sqrt{w}} P \left[ -\sqrt{(w - v^2) \rho^2 (1 - \rho^2)} \leq N(0, 1) \leq -v(1 - \rho^2)/\rho \right] \cdot N(0, 1)dv$.

The theorem follows upon combining and simplifying these four terms.
References


