CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone 31 - 20 592 9333, telex 12571 (mactr nl),
telefax 31 - 20 592 4199

CWI is the nationally funded Dutch institute for research in Mathematics and Computer Science.

Stochastic integrals and goodness-of-fit tests

A.J. Koning

1991 Mathematics Subject Classification: 62G10, 60H05, 60F15.
ISBN 90 6196 428 8
NUGI-code: 811

Preface

I was first introduced to the subject of this book by Wim Albers in 1986. He showed me an early version of Akritas (1988), in which an apparently completely new χ^2-test in the random censoring model was described. In the classical i.i.d. model the χ^2-test statistic reduced to a simple functional of a process which was an intriguing transformation of the empirical distribution function.

Shortly after that, I learned more about this process from several other seemingly independent sources [in the end all sources trace back to Aalen (1976)]. In Aki (1986) the process was attributed to the innovation approach in Khmaladze (1981), and in Shorack and Wellner (1986) two chapters were dedicated to "the basic martingale".

Provided that T is not too large, the basic martingale behaves on the interval $[0, T]$ as a lineair transformation of some empirical process. The recognition of this fact led to Einmahl and Koning (1992), in which several results for the empirical process were transferred to stochastic integrals with respect to the basic martingale. A main tool was a strong approximation on the interval $[0, T]$.

The work on that paper made the usefulness of a more general strong approximation clear to me, so I started investigating the interval $[0, \infty)$. A major difficulty is that on this interval the basic martingale no longer behaves as a linear transformation of an empirical process. As Theorem 2 on page 26 shows, this difficulty can be taken care of. In fact, Theorem 2 goes somewhat further than a strong approximation, since it also presents an exponential inequality governing the approximation. The combination of an approximation and an exponential inequality is referred to as a KMT-type inequality.

The major theme of this book is the statistical application of the KMT-type inequality Theorem 2, especially in the areas of testing, estimation, deviations, and efficiencies. In Chapter 1 an overview is presented. Some basic tools are described in Chapter 2. The theory is developed in Chapters 3 and 4, applied to the problem of testing exponentiality in Chapter 5, and supported by computer simulations in Chapter 6.

ALEX J. KONING

May 1993

Contents

Chapter 1

An overview

This chapter is meant to provide an overview of the remainder of the book. Although some technical elements are outlined in section 1.9, the interested reader should consult the other chapters for full detail.

1.1 Introduction

The complexity of the finite sample distribution of a test statistic often makes approximation necessary. In most cases approximations are motivated by weak convergence arguments. However, such arguments fail to give insight into the accuracy of the approximation. Typically, the approximations are poor in the tail of the distribution, the region which is of importance in determining critical regions and probability values.

This book discusses the use of the approximation theory of Komlós, Major and Tusnády (1975) to construct approximations for test statistics themselves, rather than for their distribution functions. The accuracy of the approximations is indicated by accompanying probability inequalities.

These probability inequalities imply weak convergence. In fact, they have consequences which go far beyond weak convergence, such as deviation results [see section 1.2]. Thus, one may view the inequalities as refinements of the implied weak convergence results.

The test statistics considered mainly occur in the random censoring model, and are of the goodness-of-fit type. Section 1.2 pays some attention to goodness-of-fit tests in the i.i.d. model.

In section 1.2 the approximation of empirical processes is described. Sections 1.3 and 1.4 introduce the notion of KMT-type inequalities and the random censoring model, respectively. Section 1.5 addresses testing the simple null hypothesis. The transition to the composite null hypothesis, the subject of section 1.8, is made in sections 1.6 and 1.7.

1.2 KMT-inequalities for empirical processes

Many tests for assessing the fit of a cumulative distribution function $F(t)$ to an i.i.d. sample are based based on the empirical distribution function F_n of the sample. A vast majority of these so-called EDF tests have test statistics which can be expressed in terms of the empirical process

$$U_n(t) = n^{1/2}\{F_n(t) - F(t)\}. \tag{1.1}$$

Obviously, empirical process theory has direct implications for this class of EDF tests.

If the sample indeed originates from the distribution F, then it is noticed in Doob (1949) that as the sample size n tends to infinity, the empirical process seems to behave more and more as a familiar Gaussian process, Brownian bridge [with time $F(t)$]. This observation has provided the inspiration for the so-called Empirical Central Limit Theorem [ECLT], which tells us that the empirical process converges in distribution to a Brownian bridge. The ECLT appears as Theorem 1 on page 19 of this book.

Although the ECLT is a very useful device, it only describes the asymptotic behavior of the empirical process. Hence, the question arises how well a Brownian bridge approximation to the empirical process performs for finite samples. This question was answered by Komlós, Major and Tusnády (1975) by giving their famous probability inequality, which appears as Inequality 2 on page 19.

Within the class of empirical process EDF tests there is the important subclass of tests based on statistics of the form $T(U_n)$, where T is a Lipschitz functional [see Definition 2 on page 27]. As examples of members of this subclass we mention Kolmogorov-Smirnov, Cramér-von Mises and χ^2 tests. Since according to the probability integral transformation we may take U_n and B_n to be equal to $\tilde{U}_n \circ F$ and $\tilde{B}_n \circ F$, Inequality 2 implies that

$$P(|T(U_n) - T(B_n)| > c_T n^{-1/2}(c_2 \log n + x)) \leq c_3 \exp\{-c_4 x\}. \tag{1.2}$$

Hence, if we know the distribution of $T(B_n)$ then we may use the KMT-inequality to bound probability values $P(T(U_n) > x)$.

The KMT-inequality implies that the difference between $P(T(U_n) > x)$ and $P(T(B_n) > x)$ soon becomes neglible if n tends to infinity, provided that x is kept fixed. However, if x tends to infinity and n is kept fixed, $P(T(B_n) > x)$ typically behaves as $\exp\{-ax^2\}$ for some $a > 0$, whereas the difference between $P(T(U_n) > x)$ and $P(T(B_n) > x)$ as indicated by the KMT-inequality behaves as $\exp\{-bx\}$ for some $b > 0$. This difference will eventually become much larger than $P(T(B_n) > x)$. Thus, the usefulness of the KMT-inequality for bounding probability values depends on whether we let n or x tend to infinity.

This immediately raises the question what will happen if we let n and x simultaneously tend to infinity. In other words: what can be said about the difference between $P(T(U_n) > x_n)$ and $P(T(B_n) > x_n)$ as n tends to infinity, and $P(T(B_n) > x_n)$ behaves as $\exp\{-a(x_n)^2\}$? This question is addressed in Inglot and Ledwina (1990) and Inglot, Kallenberg and Ledwina (1989).

It turns out that in case $x_n \to \infty$ and $x_n = o(n^{1/2})$ the difference becomes neglible with respect to $P(T(B_n) > x_n)$. Observe that in this situation $P(T(U_n) > x_n)$ also behaves as $\exp\{-a(x_n)^2\}$. In mathematical statistics this is known as a deviation result. Deviation results are important in the evaluation of statistical tests, and are classified according to the rate at which x_n is allowed to converge to infinity. A moderate deviation result allows $x_n = \mathcal{O}((\log n)^{1/2})$, and is required in the computation of weak intermediate efficiency [see Kallenberg (1983)]. A Cramér type deviation result allows $x_n = o(n^{1/6})$, and is required in the computation of intermediate efficiency [see Kallenberg (1983)]. A Chernoff type deviation result allows $x_n = \mathcal{O}(n^{1/2})$, and is required in the computation of Bahadur efficiency [see Bahadur (1960)]. Observe that the deviation result implied bv the KMT-inequality just falls short of being Chernoff type.

For the interesting examples of test statistics $T(U_n)$ the deviation result just described was already known before. However, the earlier work on deviations of $T(U_n)$ consists of separate results. A clear advantage of the KMT-inequality is that it enables a unified approach.

1.3 KMT-type inequalities

The KMT approach to the EDF statistics sketched above may be extended to statistics based on other stochastic processes. In particular, we are interested in statistics of processes $Q_n(t)$ for which there exist a sequence of identically distributed approximating Gaussian processes $W_n(t)$ such that the KMT-type inequality

$$P(\sup_{t\in[0,\infty)} |Q_n(t) - W_n(t)| > n^{-\gamma}(c_2 \log n + x)^\tau) \leq c_3 \exp\{-c_4 x\} \tag{1.3}$$

holds for $0 < \gamma \leq 1/2$ and $\tau > 0$.

By setting x equal to $\log n$, it is easily seen that the supremum over t of $Q_n(t) - W_n(t)$ remains bounded in probability, even after multiplication by $n^\gamma(\log n)^\tau$. This implies that Q_n converges weakly to a Gaussian process having the same distribution as W_1.

Although Q_n is asymptotically Gaussian, the asymptotic distribution of the statistic $T(Q_n)$ is in general non-normal. Asymptotic Pitman efficacy can be defined for non-normal test statistics [see Rothe (1981)], but has the undesirable property of depending on the size of the test. Moreover, its computation is in most cases extremely troublesome. This makes it rather impractible to use asymptotic Pitman efficacy to evaluate test statistics based on Q_n.

From the proof of Theorem 3 on page 43 we may infer that inequality (1.3) yields a deviation result for $T(Q_n)$ allowing $x_n = o(n^{\gamma/(2\tau-1)})$, provided that

$P(T(Q_n) > x)$ behaves as $\exp\{-ax^2\}$ for n tending to ∞. Deviation results shed light on the "tail" behavior of the approximation. Hence, we may view the number $\gamma/(2\tau - 1)$ as an indicator of tail accuracy. It should be as high as possible.

The implications of KMT-type inequalities for intermediate efficiency and for the equivalence of limiting approximate Bahadur efficiency and limiting Pitman efficiency are discussed in Kallenberg and Koning (1993) [see also Appendix A]. Here the ratio γ/τ is of importance.

In later sections in this chapter KMT-type inequalities for processes occurring in the random censoring model are described. The distribution of the approximating Gaussian process is in some cases omitted from the description, since it can be quite complex. Thus, our main focus in this chapter will be the accuracy of the approximation, as indicated by γ and τ.

1.4 Censoring

In the analysis of failure time data we are interested in the time passing between a first event and a second event. Sometimes a third event makes the observation of the second event impossible. In this situation we only know a lower bound to the actual failure time. It follows that we have two types of observations, which we call "uncensored" or "censored", depending on whether the second event is observed.

Techniques for the analysis of failure time data should distinguish between these two types of observations. Hence, tests based on the empirical process are in general not applicable.

Nevertheless, an abundance of tests is available for failure time data. A striking feature of the leading tests is that the central role of the empirical process has been taken over by another stochastic process, which was named the basic martingale in Shorack and Wellner (1986). This process will be discussed in later sections.

In what follows we shall assume that the so-called random censoring model holds. In this model we have two independent sequences $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ of non-negative i.i.d. random variables which are observed indirectly by means of the sample $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$, defined by

$$Z_i = X_i \wedge Y_1, \qquad \delta_i = 1_{\{X_i \leq Y_i\}}.$$

The distribution of X_i is of interest.

It is convenient to represent the sample $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ by means of the empirical distribution functions

$$H_n^1(t) = n^{-1} \sum_{i=1}^{n} 1_{\{Z_i \leq t, \delta_i = 1\}}, \qquad H_{n-}(t) = n^{-1} \sum_{i=1}^{n} 1_{\{Z_i < t\}}. \tag{1.4}$$

One may show that results obtained under the random censoring model continue to hold if the random variables $Y_1, \ldots, Y_n$ are degenerated in $+\infty$. In this case the random censoring model reduces to the classical i.i.d. model.

1.5 Testing the simple null hypothesis

The fact that nH_n^1 is a counting process provides the entry point for an approach of the random censoring model, which has become quite popular over the last fifteen years. A counting process is a nondecreasing process starting at zero, making jumps of length 1 at random points in time, and is constant elsewhere. To each counting process corresponds another stochastic process, referred to as the compensator, such that the difference between the counting process and its compensator has the martingale property: the conditional expectation of the process at some time t given all events occurring before time $s < t$ is equal to the value attained by the process at time s. Observe that multiplying a martingale by a fixed constant does not affect the martingale property.

Let P_n be the probability measure generated by the pairs of random variables $(X_1, Y_1,), \ldots, (X_n, Y_n)$. Suppose there exists a cumulative distribution function F, indexed by θ belonging to some set Θ, such that $P_n(X_1 \leq t) = F(t; \theta_n)$ for some $\theta_n \in \Theta$. Typically, Θ is a very large class of distribution functions.

Now, suppose we are interested in testing the simple null hypothesis that θ_n equals θ_0, where θ_0 is some element of the set Θ.

Let $\Lambda(t; \theta)$ denote $-\log(1 - F(t; \theta))$, the cumulative hazard function belonging to $F(t; \theta)$. The counting process nH_n^1 has compensator $n \int_0^{\cdot}(1 - H_{n-}(s))d\Lambda(s; \theta_n)$. The basic martingale

$$M_n(t; \theta_0) = n^{1/2}\{H_n^1(t) - \int_0^t (1 - H_{n-}(s))d\Lambda(s; \theta_0)\}, \tag{1.5}$$

is the difference between nH_n^1 and its compensator under the null hypothesis, multiplied by $n^{-1/2}$. If the null hypothesis holds, then the basic martingale behaves approximately as a Wiener process with time

$$H^1(t; \theta_0) = \int_0^\infty P_n(Y_1 > s)dF(s; \theta_0). \tag{1.6}$$

Although this can be shown by Rebolledo's martingale central limit theorem, empirical process techniques [outlined briefly in section 1.9] are capable of providing a KMT-type probability inequality [with $\gamma = 1/2$ and $\tau = 2$]. Recall that KMT-type probability inequalities are refinements of central limit theorems.

Now that we are able to establish asymptotic properties of the $M_n(t; \theta_0)$ without exploiting its martingale character, we may consider the martingale concepts discussed above merely as a natural way to arrive at a process which under the null hypothesis reflects the randomness in the sample. From a goodness-of-fit

0.0035,	0.0086,	0.0132,	0.0467,	0.0490*,	0.0669,	0.0802,	0.0830,	0.1086,
0.1173,	0.1181*,	0.1518*,	0.1653*,	0.1900*,	0.2029*,	0.2037,	0.2109*,	0.2324,
0.2342,	0.2394,	0.2447*,	0.2596*,	0.2979*,	0.3681,	0.4055,	0.5316*.	

Table 1.1: Data given in Woolson (1981). A suffix * denotes that the observation was censored by the close of the study.

perspective the essential feature of $M_n(t;\theta_0)$ is its functional dependence on the null hypothesis value of the parameter [just like the functional dependence on μ_0 of the classical one-sample t-test]. Thus, if the null hypothesis does not hold there is no guarantee that the null hypothesis behavior will still apply. Actually, drastically different behavior is what we are hoping for, since this will facilitate distinguishing between the null and the alternative hypothesis.

Observe that $M_n(t;\theta_0)$ shares a time-transformation property with the empirical process: if we transform the data by applying a function ξ which is increasing almost everywhere with respect to $F(t;\theta_0)$, then the "new" basic martingale equals $M_n(\xi(t);\theta_0)$. Thus this type of transformation only has the effect of altering the time-scale.

There is one special transformation which stands out because of its simplifying effect. If the distribution $F(t;\theta_0)$ is continuous, then transforming by applying the cumulative hazard function $\Lambda(t;\theta_0)$ yields data which are under the null hypothesis standard exponentially distributed. This is convenient, since for testing standard exponentiality the basic martingale simplifies to

$$M_n(t;\theta_0) = n^{1/2}\{H_n^1(t) - \int_0^t (1 - H_{n-}(s))ds\}, \tag{1.7}$$

which is a spline function. Changes in slope relate to failure times, censored or uncensored. If a change in slope coincides with a jump, then the corresponding failure time is uncensored.

The transformation by means of the cumulative hazard function is exploited in Woolson (1981), where data obtained from a random sample of 26 psychiatric patients are analyzed. The random sample was drawn from all psychiatric patients who were first admitted to University of Iowa Hospitals during the years 1935-1948.

The cumulative hazard function was computed according to mortality tables for the State of Iowa. Observe that if there is no difference in mortality between the psychiatric patients and other residents of Iowa, then [and only then] the transformed data should have a standard exponential distribution, and hence the basic martingale takes the form (1.7). In Figure 1.1 this stochastic process is depicted.

Now the question arises whether Figure 1.1 corresponds to what we expect from the basic martingale under the null hypothesis. Informally, we could inspect

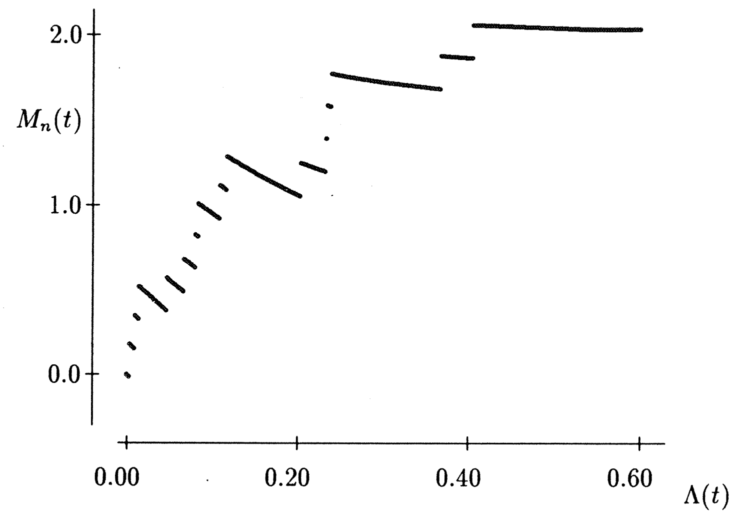

Figure 1.1: Basic martingale computed from [transformed] Woolson data.

Figure 1.1, having the properties of a Wiener process in mind. For instance, there is some drift away from the X-axis noticable in the figure, which could be a sign that we are not dealing with a process close to a Wiener process. In other words: there is some reason to doubt the null hypothesis.

If we are not content with such loose findings, then a formal approach which uses the basic martingale to construct test statistics is more appropriate. In Woolson (1981) the test proposed by Breslow (1975) is applied. Fundamentally, the test of Breslow uses $M_n(\infty;\theta_0)$ as test statistic. Under the null hypothesis $M_n(\cdot;\theta_0)$ behaves approximately as a Wiener process with time $H^1(\cdot;\theta_0)$ [see (1.6)], and thus $M_n(\infty;\theta_0)$ behaves approximately as a normal random variable with mean 0 and variance $H^1(\infty;\theta_0)$. Because the distribution of Y_1 is unknown, it becomes necessary to estimate $H^1(\infty;\theta_0)$. This is typically done using the estimator

$$\int_0^\infty (1 - H_{n-}(s))d\Lambda(s;\theta_0) = \int_0^\infty (1 - H_{n-}(s))ds, \tag{1.8}$$

which is the null hypothesis compensator of $H_n^1(t)$ evaluated in ∞. In case of the Woolson data the estimator (1.8) takes the value 0.1794. As roughly can be inferred from Figure 1.7, $M_n(\infty;\theta_0)$ itself takes the value 2.033. Thus, the test based on $M_n(\infty;\theta_0)$ and estimated variance yields an asymptotic probability value 0.0000, which leads to rejection of the null hypothesis. Hence, the mortality of psychiatric patients differs from the mortality of other inhabitants of the State of Iowa.

In Chapter 3 a broad class of tests based on sublinear Lipschitz functionals of stochastic integral with respect to the basic martingale is studied. Within this class the test statistics are of the form $T(Q_n)$, where $Q_n(t)$ is defined as

$$Q_n(t) = \int_0^t L_n(s) dM_n(s;\theta_0), \tag{1.9}$$

$L_n(t)$ is a random element of $D[0,\infty)$ satisfying certain properties, and T is sublinear [see Definition 1 on page 20] as well as Lipschitz.

An example of a sublinear Lipschitz functional is the projection in infinity $T_R(\xi) = \xi(\infty)$. Tests based on statistics of the form $T_R(Q_n)$ are called generalized rank tests. Special cases are the aforementioned test of Breslow, and the generalized rank tests of Fleming and Harrington, having weight process

$$L_n(t) = (1 - F(t;\theta_0))^\rho \tag{1.10}$$

[see Fleming, O'Fallon, O'Brien and Harrington (1980), Fleming and Harrington (1981), Harrington and Fleming (1982), Fleming, Harrington and O'Sullivan (1987)]. Generalized rank test statistics are asymptotically normal.

The supremum functional $T_S(\xi) = \sup_{t\in[0,\infty)} \xi(t)$ is another example of a sublinear Lipschitz functional. Tests of the form $T_S(Q_n)$ are called supremum type tests. Special cases are the test studied by Aki (1986), and the supremum type tests of Fleming and Harrington, also having weight process (1.10). Supremum type test statistics have a non-normal asymptotic distribution.

The general KMT-inequality given in Theorem 2 on page 26 for an appropriately centered version of the stochastic integral $Q_n(t)$ and its specialization for the null hypothesis [with $\gamma = 1/6$ and $\tau = 2$] enable the computation of various kinds of efficacies. These efficacies all coincide, and are maximized by generalized rank and supremum type tests based on a weight process which satisfies

$$L_n(t) \to_{P_0} \psi_a(t;\theta_0) \tag{1.11}$$

at a sufficient rate. Here P_0 is the probability measure corresponding to the null hypothesis, $\psi_a(t;\theta_0)$ is the score function [the evaluation in θ_0 of the partial derivative of $\log \lambda(t;\theta)$ in the direction in which the alternative approaches the null hypothesis], and $\lambda(t;\theta)$ is the derivative of $\Lambda(t;\theta)$. The generalized rank tests based on this type of weight process are asymptotically most powerful. The corresponding supremum type tests do not perform much less, since they have efficiency 1 [in the sense of approximate Bahadur efficiency, limiting approximate Bahadur efficiency, weak asymptotic intermediate efficiency, and - under extra conditions on $L_n(t)$ - asymptotic intermediate efficiency] with respect to the asymptotically most powerful generalized rank tests.

This implies that the statistics $T_R(M_n(\cdot;\theta_0))$ and $T_S(M_n(\cdot;\theta_0))$ are recommended for testing against alternatives of the proportional hazard type, and statistics $T_R(\int_0^{\cdot}(1 - F(s;\theta_0))dM_n(s;\theta_0))$ and $T_S(\int_0^{\cdot}(1 - F(s;\theta_0))dM_n(s;\theta_0))$ for testing against logistic shift alternatives.

In this respect the application of the test of Breslow to the Woolson data is certainly appropriate, since there are psychological motivations for assuming a proportional hazard model for psychiatric patients. This model states that the hazard of a psychiatric patient is proportional to the baseline hazard, in our case the hazard of an inhabitant of the State of Iowa.

1.6 Adjusting the null hypothesis

The fact that for the Woolson data $T_R(M_n(\cdot;\theta_0))$ leads to rejection of the null hypothesis could be interpreted as an indication that a proportional hazard model is indeed more appropriate. But does it fit the data?

Recall that in the previous section it was convenient to apply the cumulative hazard function to the data, since in this way the general problem was reduced to the problem of testing standard exponentiality. Under the proportional hazard model it is convenient to apply the *baseline* cumulative hazard to the data, since the transformed data follow an exponential distribution with unknown mean, say υ_n. The actual hazard experienced by a psychiatric patient is equal to the baseline hazard divided by υ_n. Hence, the general problem of testing the proportional hazard model is now reduced to the problem of testing exponentiality.

Observe that the mean of the exponential distribution is not known. Our null hypothesis gives an incomplete specification of the distribution of the data, and is henceforth composite. The unknown mean is a nuisance parameter.

Let us return to a more abstract level: since the simple null hypothesis was rejected, we adjusted the null hypothesis by adding an extra nuisance parameter to our model. Section 4 hypothesized that $P_n(X_1 \leq t)$ could be written as $F(t;\theta_n)$. From now on we write this probability as $F(t;\upsilon_n,\theta_n)$ for some (υ_n,θ_n) in the set $\Upsilon \times \Theta$. Already anticipating the necessity of the inclusion of more nuisance parameters to make the model fit the data, we assume that Υ is a subset of r-dimensional Euclidean space.

Let P_{υ_0} denote the probability measure which corresponds to the point (υ_0,θ_0). Observe that P_{υ_0} "belongs" to the null hypothesis that θ_n equals θ_0.

1.7 M-estimation

Under P_{υ_0} the basic martingale takes the form

$$M_n(t;\upsilon_0,\theta_0) = n^{1/2}\{H_n^1(t) - \int_0^t (1 - H_{n-}(s))d\Lambda(s;\upsilon_0,\theta_0)\}, \qquad (1.12)$$

where $\Lambda(t;\upsilon,\theta) = -\log(1 - F(t;\upsilon,\theta))$. In practical situations we are faced with the problem that the null hypothesis does not provide any knowledge about the actual value of υ_0. Hence, we replace υ_0 by some random element $\upsilon^{(n)} \in \Upsilon$ in an effort to obtain this knowledge elsewhere. Under P_{υ_0} this random element should be an estimator of υ_0.

A convenient class of estimators to use is the class of M-estimators as proposed by Hjort (1985). Within this class estimators are obtained as solutions to the set of M-equations

$$\Phi_{ni}(\infty; \upsilon^{(n)}) = 0, \qquad i = 1, \ldots, r, \tag{1.13}$$

where $\Phi_{ni}(t;\upsilon)$ is defined as a stochastic integral with respect to the basic martingale

$$\Phi_{ni}(t;\upsilon) = \int_0^t \phi_i(s;\upsilon) dM_n(s;\upsilon,\theta_0). \tag{1.14}$$

The integrand $\phi_i(t;\upsilon)$ is a given deterministic function. M-estimators are generalizations of the maximum likelihood estimator. The latter estimator was studied in this context by Borgan (1984), and is obtained by setting $\phi_i(t;\upsilon)$ equal to the partial derivative of $\log\lambda(t;\upsilon,\theta_0)$ with respect to the i^{th} component of υ. Here $\lambda(t;\upsilon,\theta)$ denotes the hazard function, the derivative with respect to t of the cumulative hazard function $\Lambda(t;\upsilon,\theta)$. A general theory of M-estimation is worked out in subsection 4.2.1 on page 52.

Consider the exponential distribution function $F(t;\upsilon,\theta_0) = 1 - e^{-t/\upsilon}$. Since the corresponding hazard function $\lambda(t;\upsilon,\theta_0)$ equals υ, the maximum likelihood estimator $\upsilon^{(n)}$ is obtained by setting $\phi_i(t;\upsilon)$ equal to $1/\upsilon$. Multiplying both sides of (1.13) by υ yields the maximum likelihood equation

$$M_n(\infty;\upsilon^{(n)},\theta_0) = 0. \tag{1.15}$$

For the Woolson data the maximum likelihood estimator of the unknown mean of the baseline cumulative hazard transformed data attains the value 0.309. This suggests that the baseline hazard is less than a third of the real hazard a psychiatric patient experiences, leading to the preliminary conclusion that the hazard of a psychiatric patient is more than three times as high as the hazard of an arbitrary resident of Iowa.

Of course, this only holds under the proportional hazard model. Before turning the preliminary conclusion into a definitive one, we should assess the validity of the model. This can be done by the methods described in the next section.

The stochastic integral $M_n(t;\upsilon^{(n)},\theta_0)$, depicted in Figure 1.2, is closely related to a well-known and frequently used graphical technique for assessing exponentiality. The total time on test plot is constructed by plotting

$$\int_0^t (1 - H_{n-}(s)) d\Lambda(s;\upsilon^{(n)},\theta_0)$$

versus $H_n^1(t)$ [see Barlow and Proschan (1969), and for variants Gill (1986)]. Here $\upsilon^{(n)}$ is the maximum likelihood estimator. If the data indeed follow an exponential

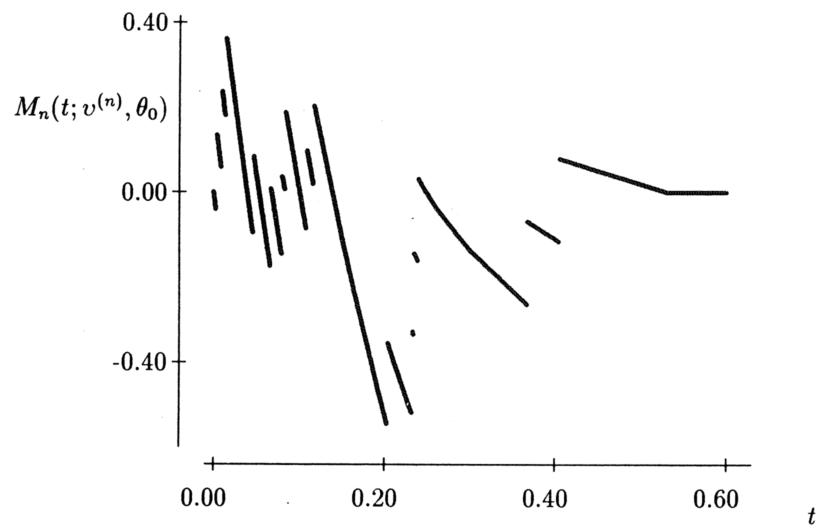

Figure 1.2: The stochastic integral involved in the construction of the maximum likelihood estimator.

distribution, then the total time on test plot should be approximately straight, making a 45 degree angle with the X-axis.

Observe that at $x = \int_0^t (1 - H_{n-}(s)) d\Lambda(s; v^{(n)}, \theta_0)$ the vertical deviation of the total time on test plot from the ideal line equals $n^{-1/2} M_n(t; v^{(n)}, \theta_0)$.

1.8 Testing the composite null hypothesis

Now that we have the M-estimators at our disposal, it seems obvious to adapt the test statistics discussed earlier to the composite null hypothesis by simply plugging in these estimators. This leads to statistics $T(Q_n(\cdot; v^{(n)}))$, where T is as before and $Q_n(t; v)$ is defined by

$$Q_n(t; v) = \int_0^t L_n(s; v) dM_n(s; v, \theta_0) \tag{1.16}$$

for some weight process $L_n(t; v)$. By exploiting the relation between M-estimators and stochastic integrals with respect to the basic martingale, the KMT-type inequalities for $Q_n(t)$ can be extended to $Q_n(t; v^{(n)})$. However, we must pay a price for the plugged-in estimator: the accuracy of the probability inequalities governing the approximation decreases. This is reflected in the values $\gamma = 1/6$ and $\tau = 3$.

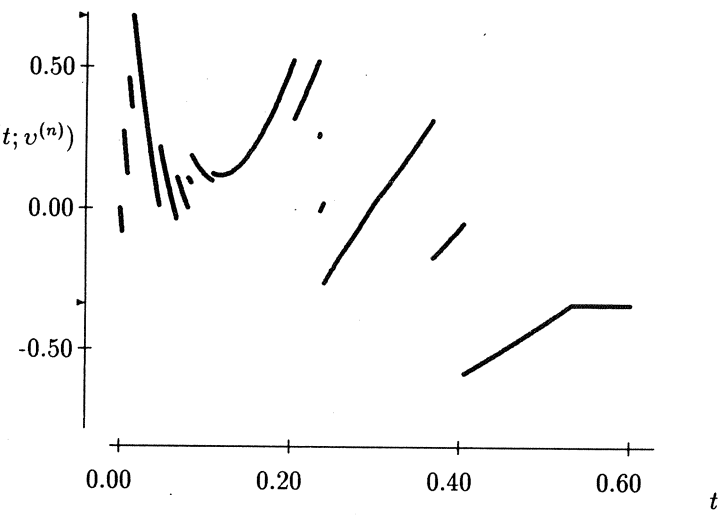

Figure 1.3: The stochastic integral leading to optimal generalized rank and supremum type tests for testing exponentiality versus logistic shift alternatives.

Again, the KMT-inequalities lead to various kinds of efficacy which all coincide and are maximized by generalized rank tests based on weight processes satisfying

$$L_n(t; v^{(n)}) \to_{P_{v_0}} \psi_{a|0}(t; v_0, \theta_0) + \sum_{i=1}^{r} c_i \phi_i(t; v_0), \tag{1.17}$$

and supremum type tests based on weight processes satisfying

$$L_n(t; v^{(n)}) \to_{P_{v_0}} \psi_{a|0}(t; v_0, \theta_0) \tag{1.18}$$

[see Theorem 13 on page 67]. Here the c_i's are constants, and $\psi_{a|0}(t; v, \theta_0)$ is the effective score function, i.e. a projected version of the score function $\psi_a(t; v, \theta_0)$, defined similar to the score function $\psi_a(t; \theta_0)$ we encountered earlier. The convergence in P_{v_0}-probability should take place at a sufficient rate.

If the estimation method is maximum likelihood then the following holds. The optimal generalized rank tests have asymptotic relative Pitman efficiency 1 with respect to the generalized likelihood ratio test, and thus are asymptotically most powerful. The optimal supremum type tests have efficiency 1 [in the sense of approximate Bahadur, limiting asymptotic relative Pitman and weak intermediate efficiency] with respect to the optimal generalized rank test.

The effective score function $\psi_{a|0}(t; v, \theta_0)$ also appears in Hjort (1990), in an investigation of χ^2-tests based on $Q_n(t; v^{(n)})$. Although the argument there lacks mathematical rigour, it underlines the universal importance of the effective score function as weight process.

As a direct consequence of (1.13), adding a term $c_i\phi_i(t;v)$ to the weight process does not alter the generalized rank test. This property, reflected in (1.17), is especially convenient when the estimation method is maximum likelihood, since in this case the difference between $\psi_{a|0}(t;v,\theta_0)$ and $\psi_a(t;v,\theta_0)$ is a linear combination of the functions $\phi_i(t;v)$. It is easier to construct a weight process converging to $\psi_a(t;v,\theta_0)$ than one converging to $\psi_{a|0}(t;v,\theta_0)$.

For given t, an optimal generalized rank test for testing the proportional hazards model versus alternatives where the ratio between the hazard and the baseline hazard function changes at point t is the statistic $M_n(t;v^{(n)},\theta_0)$, where $v^{(n)}$ is the maximum likelihood estimator. Thus, the total time on test plot is in particular useful for the detection of change points in the hazard function.

Judged by the enormous popularity of the one-sample Wilcoxon test, logistic shift alternatives are of far greater practical importance than the change point alternatives just described. For testing exponentiality versus logistic shift alternatives the weight process

$$L_n(t;v) = (1 - F(t;v,\theta_0)) - \frac{\int_0^\infty 1 - F(s;v,\theta_0)(1 - H_{n-}(s))d\Lambda(s;v,\theta_0)}{H_n^1(\infty)} \qquad (1.19)$$

leads in combination with maximum likelihood estimation to optimal generalized rank and supremum type tests [see subsection 5.3.2 on page 92]. For the Woolson data standardized versions of the corresponding test statistics attain values 0.338 and 0.711, leading to probability values 0.7355 and 0.8892 respectively. Hence, it is clear that the proportional hazards model should not be rejected, and that our earlier conclusion about the mortality of a psychiatric patient was justified.

1.9 Some technical details

In this section it is briefly outlined how the various KMT-type inequalities are derived from the original KMT-inequality, and how the sublinearity of T leads to deviation results. Of course, the remarks in this section hold only under appropriate conditions. These are given in the following chapters.

First we consider the simple null hypothesis. We may re-express $M_n(t;\theta_0)$ as follows:

$$M_n(t;\theta_0) = U_n^1(t) + \int_0^t U_{n-}(s)d\Lambda(s;\theta_0) + n^{1/2}D(t;\theta_0,\theta_n). \qquad (1.20)$$

[see page 24]. Here $U_n^1(t)$ and $U_{n-}(t)$ are empirical processes corresponding to $H_n^1(t)$ and $H_{n-}(t)$, and $D(t;\theta_0,\theta_n)$ is a deterministic function which is identical to zero if θ_n equals θ_0. Now applying the original KMT-inequality yields that a stochastic integral with respect to $M_n(t;\theta_0) - n^{1/2}D(t;\theta_0,\theta_n)$ is approximated by a Gaussian process. Under the null hypothesis the approximating process is time-transformed Wiener.

An alternative approach for obtaining KMT-type inequalities for $M_n(t;\theta_0)$ under the simple null hypothesis is given in Koning (1993), and uses partial sum process techniques. The resulting KMT-type inequality has $\gamma = 1/4$ and $\tau = 1$. The partial sum approach also has useful applications outside the random censoring model.

The transition from the simple to the null hypothesis is enabled by the two following facts. There exists an element υ_{n0} in Υ such that under P_n the random variable $n^{1/2}(\upsilon^{(n)} - \upsilon_{n0})$ is closely approximated by a random vector with components

$$\int_0^\infty \phi_i(s;\upsilon_{n0})d\{M_n(s;\upsilon_{n0},\theta_0) - n^{1/2}D(s;\upsilon_{n0},\theta_0,\upsilon_n,\theta_n)\} \tag{1.21}$$

[see Theorem 7 on page 55], where $D(t;\upsilon_{n0},\theta_0,\upsilon_n,\theta_n)$ is the obvious extension of $D(t;\theta_0,\theta_n)$ [see (4.6) on page 54]. Moreover, there exists an r-dimensional function $K_0(t;\upsilon_{n0})$ [defined by (4.16) on page 57] such that $Q_n(t;\upsilon^{(n)})$ is approximated by

$$Q_n(t;\upsilon_{n0}) + n^{1/2}(\upsilon^{(n)} - \upsilon_{n0})^T K_0(t;\upsilon_{n0}). \tag{1.22}$$

Hence, after subtracting $n^{1/2}\int_0^t L_n(s;\upsilon_{n0})dD(s;\upsilon_{n0},\theta_0,\upsilon_n,\theta_n)$ from $Q_n(t;\upsilon^{(n)})$, the resulting process is approximated by a zero mean Gaussian process. Even under P_{υ_0} the covariance function of the latter process remains intricate [see Theorem 8 on page 58].

The theory of Borell (1975) describes the tail behavior of sublinear functionals of Gaussian processes [see Inequality 5 on page 21]. It follows that there exists nonnegative constants $a, \breve{a}$ such that

$$\lim_{x\to\infty} x^{-2}\log P_0(T(Q_n) > x) = a/2, \tag{1.23}$$

$$\lim_{x\to\infty} x^{-2}\log P_{\upsilon_0}(T(Q_n(\cdot;\upsilon^{(n)})) > x) = \breve{a}/2. \tag{1.24}$$

Now the Lipschitz property of T enables us to obtain deviation results for $T(Q_n)$ [$x_n = o(n^{1/18})$ allowed] and $T(Q_n(\cdot;\upsilon^{(n)}))$ [$x_n = o(n^{1/30})$ allowed] from the KMT-type inequalities. The differences in rate are due to differences in accurateness of the corresponding inequalities.

1.10 Concluding remarks

In this chapter applications of KMT-type inequalities in the theory of goodness-of-fit tests were described. As the section on M-estimation indicates, there are other areas where KMT-type inequalities are of use as well.

In Castelle (1991) the study of stochastic integrals $\int_0^t L_n(s)dH_n^1(s)$ leads to results for some well-known two-sample generalized rank tests and for maximum likelihood estimation in the Cox regression model.

The field of nonparametric estimation in the random censoring model is covered in Burke, Csörgő and Horváth (1981), where the Nelson-Aalen estimator of the cumulative hazard function $\Lambda(t)$ and the Kaplan-Meier estimator of the distribution function $F(t)$ are investigated. Processes related to these estimators can be expressed as stochastic integrals with respect to the basic martingale. Analogues for $Q_n(t)$ of the Chibisov-O'Reilly theorem, the Lai-Wellner Glivenko-Cantelli theorem and the James law of the iterated logarithm are derived, and their implications for the Nelson-Aalen and the Kaplan-Meier estimator discussed in Einmahl and Koning (1992).

Other examples of KMT-type inequalities and their implications can be found in Csörgő and Horváth (1993).

The intricate asymptotic structure of $Q_n(t;\upsilon^{(n)})$ shows some resemblance to the asymptotic structure of the empirical process with estimated parameters [see Durbin (1973)]. In Khamaladze (1981, 1982) a transformation of this process is constructed which converges to a time-transformed Wiener process. A similar transformation can be constructed for $Q_n(t;\upsilon^{(n)})$. However, this transformation is quite complex, and involves the in general unknown distribution function $H^1(t;\upsilon_0,\theta_0)$, the expectation of $H_n^1(t)$ under P_{υ_0}. Estimating this distribution function complicates matters even further.

Chapter 2

Tail probabilities of various suprema

2.1 The empirical process

Empirical processes have been the subject of ongoing investigations for over half a century. In this chapter we encounter the most important results. For a more elaborate survey we refer to Csörgő and Révész (1981) and to Shorack and Wellner (1986).

Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space, and let $\{\tilde{Z}_i\}_{i=1}^{\infty}$ be a sequence of random variables mapping $(\Omega, \mathcal{A})$ into $([0,1], \mathcal{F})$, where $\mathcal{F} \subset \mathcal{B}[0,1]$. Let P be the probability measure induced by the sequence $\{\tilde{Z}_i\}_{i=1}^{\infty}$.

Suppose each of the random variables within this sequence has a standard uniform distribution under P. The [standard] uniform empirical process at stage n is defined by

$$\tilde{U}_n(t) = n^{-1/2} \sum_{i=1}^{n} (1_{\{\tilde{Z}_i \leq t\}} - t) \tag{2.1}$$

for $0 \leq t \leq 1$. It is easily seen that $-\inf_{t\in[0,1]} \tilde{U}_n(t)$ is equal in distribution to $\sup_{t\in[0,1]} \tilde{U}_n(t)$. Moreover, we have

$$\begin{aligned} &P(\sup_{t\in[0,1]} |\tilde{U}_n(t)| > x) \\ &\quad \leq P(\sup_{t\in[0,1]} \tilde{U}_n(t) > x) + P(-\inf_{t\in[0,1]} \tilde{U}_n(t) > x) \\ &\quad \leq 2P(\sup_{t\in[0,1]} \tilde{U}_n(t) > x) \end{aligned} \tag{2.2}$$

Hence, bounds for tail probabilities of $\sup_{t\in[0,1]} \tilde{U}_n(t)$ directly translate into bounds for tail probabilities of $-\inf_{t\in[0,1]} \tilde{U}_n(t)$ and $\sup_{t\in[0,1]} |\tilde{U}_n(t)|$.

An important early result for the tail probability of $\sup_{t\in[0,1]} \tilde{U}_n(t)$ was derived in Smirnov (1944). Surprisingly, it is known under the name Birnbaum-Tingey formula, since it was also discovered in Birnbaum and Tingey (1951). It tells us that

$$P(\sup_{t\in[0,1]} \tilde{U}_n(t) > yn^{1/2}) = \sum_{i=0}^{n(1-y)} y \binom{n}{i} (y + i/n)^{i-1}(1 - y - i/n)^{n-i} \tag{2.3}$$

for $0 < y < 1$. Applying Stirling's formula to (2.3) yields that

$$\lim_{n\to\infty} P(\sup_{t\in[0,1]} \tilde{U}_n(t) > x) = \exp\{-2x^2\} \tag{2.4}$$

for $x > 0$. The Birnbaum-Tingey formula provided the basis of the following well-known equality.

Inequality 1 (DKW-Inequality) *There exists a positive constant c_1 such that*

$$P(\sup_{t\in[0,1]} \tilde{U}_n(t) > x) \leq c_1 \exp\{-2x^2\}$$

for $x > 0$.

The DKW-inequality was first proven in Dvoretzky, Kiefer and Wolfowitz (1956). Observe that if we know a valid value of c_1, then the DKW-inequality enables us to construct a conservative version of a Kolmogorov goodness-of-fit test. The rather intriguing conjecture that c_1 could be set equal to 1 was made in Birnbaum and McCarty (1958). Equation (2.4) shows that the Birnbaum-McCarty conjecture implies that constructing a Kolmogorov goodness-of-fit test by using asymptotic methods never yields an anti-conservative test.

For over thirty years the Birnbaum-McCarty conjecture remained an open question in empirical process theory. During these years various attempts have been made to prove it, resulting in steadily decreasing permitted minimal values of c_1. As recommended values of c_1 have acted 305.2 [Devroye and Wise (1979)], 29 [Shorack and Wellner (1986)], and $2^{3/2}$ [Hu (1985)]. Recently, the Birnbaum-McCarty conjecture was proven in Massart (1990) for a slightly modified version of the DKW-inequality which restricts attention to values of x satisfying $\exp\{-2x^2\} \leq 1/2$.

A generalization of the DKW-inequality, which can be applied in the non-i.i.d. case and not only to the empirical process but related processes as well, can be found in Marcus and Zinn (1984).

Asymptotic expansions of the distributions of $\sup_{t\in[0,1]} \tilde{U}_n(t)$ are found in Lauwerier (1963), Penkov (1976) and Gnedenko, Korolyuk and Skorohod (1961), whereas other approximations to this distributions are contained in Stephens (1970) and Harter (1980).

2.2 The Brownian bridge approximation

In Doob (1949) it is noticed that as the sample size n tends to infinity, empirical processes seem to behave more and more like a Brownian bridge. This observation has provided the inspiration for the following result [also known as Donsker's Theorem, since it originally appeared in Donsker (1952)].

Theorem 1 (Empirical Central Limit Theorem) *The sequence of empirical processes $\{\tilde{U}_n(t)\}_{n=1}^{\infty}$ converges in P-distribution, as random elements of $D[0,1]$, to a Brownian bridge.*

As a consequence of the well-known Skorohod construction Theorem 1 implies the existence of a probability space on which we have a sequence of empirical processes and a Brownian bridge such that the supremum of the difference between each of the empirical processes and the Brownian bridge converges to zero in probability.

An alternative to the Skorohod construction [Skorohod (1956)] is the so-called Hungarian construction, given in Komlós, Major and Tusnády (1975). The Hungarian construction culminates into the following inequality.

Inequality 2 (KMT-Inequality, Brownian bridge version) *If the probability space $(\Omega, \mathcal{A}, \mathcal{P})$ is sufficiently rich, then there exists a sequence $\{\tilde{B}_n(t)\}_{n=1}^{\infty}$ of Brownian bridges such that*

$$P(\sup_{t\in[0,1]} |\tilde{U}_n(t) - \tilde{B}_n(t)| > n^{-1/2}(c_2 \log n + x)) \leq c_3 \exp\{-c_4 x\},$$

where c_2–c_4 are absolute constants.

In Bretagnolle and Massart (1989) it is shown that c_2, c_3 and c_4 may be taken as 12, 2 and 1/6, respectively.

For some applications [for instance, the derivation of a law of the iterated logarithm] it is necessary to imbed the sequence of Brownian bridges into a Kiefer process [see Kiefer (1972)]. However, imposing this extra structure on the sequence of Brownian bridges seems to worsen the rate of convergence. The most far stretching result concerning the Kiefer process approximation is given by Inequality 3. This inequality requires that the underlying probability space is the same for every $n \in \mathbb{N}$. It is still an open question whether this inequality can be improved so as to yield the same rate of convergence as does Inequality 2.

Inequality 3 (KMT-Inequality, Kiefer process version) *If the probability space $(\Omega, \mathcal{A}, \mathcal{P})$ is sufficiently rich, then there exists a Kiefer process $\tilde{K}(t,n)$ such that*

$$P(\sup_{t\in[0,1]} |\tilde{U}_n(t) - n^{-1/2}\tilde{K}(t,n)| > n^{-1/2}(c_5 \log n + x) \log n) \leq c_6 \exp\{-c_7 x\},$$

where c_5–c_7 are absolute constants.

In Mason and van Zwet (1987) a refined KMT-inequality is given, which emphasizes the behavior of the approximation near 0 and 1. However, for our purposes the original KMT-inequality suffices.

2.3 The general Gaussian process

Among the Gaussian processes, the Wiener process $W(t)$ takes first place. Results for the supremum of this process are well-known and can even be considered classical. The reflection principle yields

$$P(\sup_{t\in[0,1]} W(t) > x) = 2P(N(0,1) > x) \tag{2.5}$$

$$P(\sup_{t\in[0,1]} |W(t)| > x) = 4\sum_{k=0}^{\infty}(-1)^k P(N(0,1) > (2k+1)x) \tag{2.6}$$

Interest in the supremum of a general Gaussian process came about relatively late, at the beginning of the seventies [Fernique 1970, 1971)), Landau and Shepp (1971) and Marcus and Shepp (1971)]. The following inequality was then obtained.

Inequality 4 (Inequality of Fernique) *There exist positive constants c_8 and c_9 such that for every $x > 0$ and every separable Gaussian process $Z(t)$ satisfying $P(\sup_{t\in[0,\infty)} |Z(t)| < \infty) = 1$*

$$P(\sup_{t\in[0,\infty)} |Z(t)| > x\sqrt{\sup_{t\in[0,\infty)} \mathcal{E}\{Z(t)\}^2}) \le c_8 \exp\{-c_9 x^2\}.$$

In Borell (1975) an inequality for the tail probability of sublinear functionals of a Gaussian process is given which relates this tail probability to the norm of the sublinear functional induced by the Gaussian process, and yields a finer result than Inequality 4 when applied to the supremum of a Gaussian process.

Definition 1 *A functional T which maps $D[0,\infty)$ into $\mathbb{R}$ is said to be sublinear if $T(\xi+\zeta) \le T(\xi) + T(\zeta)$ and $T(c\xi) = cT(\xi)$ for all $c \ge 0$ and $\xi, \zeta \in D[0,\infty)$.*

Refer to Pollard (1984), page 108, for the definition of the function space $D[0,\infty)$.

The norm of a sublinear functional T induced by the Gaussian proces $Z(t)$ is defined as

$$||T||_{Z(t)} = \sup_{f\in O_{\mathcal{K}}} T(f), \tag{2.7}$$

where $O_{\mathcal{K}}$ is the unit ball in $\mathcal{K}$, the reproducing kernel Hilbert space belonging to $Z(t)$. The theory of reproducing kernels is described extensively in Aronszajn

(1950). The reproducing kernel $\mathcal{K}_W$ belonging to a standard Wiener process on [0,1] consist of functions f mapping [0,1] into $I\!R$, which are absolutely continuous and satisfy $\int_0^1 (f'(s))^2 ds < \infty$. The norm of an element of $\mathcal{K}_W$ is equal to $\int_0^1 (f'(s))^2 ds$, and hence the unit ball in $\mathcal{K}_W$ is the set $\mathcal{S}$ of functions f mapping $[0,1]$ into $I\!R$, which are absolutely continuous and satisfy $\int_0^1 (f'(s))^2 ds \leq 1$ as well as $f(0) = 0$.

The set $\mathcal{S}$ may seem familiar to statisticians and probabilists, since it also shows up in connection with the law of the iterated logarithm [see Strassen (1964)]. We shall refer to $\mathcal{S}$ as the set of Strassen functions.

Inequality 5 (Inequality of Borell) *Let $X(t)$ be a mean zero separable Gaussian process satisfying $P(\sup_{t\in[0,\infty)} |X(t)| < \infty) = 1$, and let T be a sublinear function. Assume that $P(T(X) > u||T||_X) \leq 1/2$ for some u. Then for $t \geq u$*

$$P(T(X) > t||T||_X) \leq P(N(0,1) > t - u).$$

Corollary 1 *We have*

$$\lim_{t\to\infty} t^{-2} \log P(T(X) > t) = -(||T||_X)^{-2}/2.$$

Each of the Gaussian processes to which we shall apply Inequality 5 can be considered to be a linear transformation of some standard Wiener process. Thus, a sublinear functional T of such a Gaussian process $Z(t)$ is in essence a sublinear functional $T_{Z(t)}$ of the underlying standard Wiener process. This observation leads to

$$||T||_{Z(t)} = \sup_{f\in\mathcal{S}} T_{Z(t)}(f), \tag{2.8}$$

which makes the computation of norms as in (2.7) relatively easy.

The inequality of Fernique seems to imply that it is rather the variance function than the covariance function which determines the tail behavior of the supremum of a Gaussian process. However, the covariance function is not completely irrelevant, as some finer results show. In Adler and Samorodnitsky (1987) a refinement of the inequality of Fernique is given which is based on metric entropy methods [see Dudley (1967)]. Talagrand (1988) presents a necessary and sufficient condition for

$$\lim_{x\to\infty} \frac{P(\sup_{t\in[0,\infty)} Z(t) > x)}{P(N(0,1) > x\{\sup_{t\in[0,\infty)} \mathcal{E}\{Z(t)\}^2\}^{-1/2})} = 1 \tag{2.9}$$

to hold. This condition can be verified by metric entropy methods and entails the behavior of the covariance function near points where the variance function reaches its maximum.

For a completely different approach, treating suprema of Gaussian processes from a boundary crossing perspective, refer to Durbin (1985).

Chapter 3

The simple null hypothesis

3.1 Introduction

In this chapter, which contains the material of Koning (1991), we consider the situation where $(\Omega, \mathcal{A}, \mathcal{P})$ is a probability space, $\mathcal{F}_n$ is a subset of $\mathcal{B}[0,\infty)$, and at stage n each of the independent random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ maps $(\Omega, \mathcal{A})$ into $([0,\infty), \mathcal{F}_n)$. The probability measure induced by these random variables is denoted by P_n. Each pair (X_i, Y_i) is assumed to have the same distribution.

The censoring distribution, the distribution of the censoring time Y_i, does not depend on n. Hence, there exists a cumulative distribution G such that $G(t) = P_n(Y_i \leq t)$ for each n. Defectiveness of $G(t)$ is allowed.

The failure time distribution, the distribution of X_i, is more complicated since it depends on n. This dependence is given structure in the following way: there exists a cumulative distribution function F, indexed by θ belonging to some set Θ, and a sequence of points $\{\theta_n\}_{n=1}^{\infty}$ such that $F(t;\theta_n) = P_n(X_i \leq t)$ for every $n \in I\!N$ [that is, θ_n is the actual value of θ at stage n].

Now suppose θ_0 is an element of Θ which is of special interest to us, say because we want to know whether θ_n could possibly be equal to θ_0 for every $n \in I\!N$.

If θ_n equals θ_0 [denote the probability measure corresponding to this situation by P_0, the expectation operator based on this probability measure by $\mathcal{E}_0$ and the situation itself by "under P_0"], then the basic martingale takes the form

$$M_n(t;\theta_0) = n^{1/2}\{H_n^1(t) - \int_0^t (1 - H_{n-}(s))d\Lambda(s;\theta_0)\}, \tag{3.1}$$

where $\Lambda(t;\theta) = -\log(1 - F(t;\theta))$ is the cumulative hazard function belonging to $F(t;\theta)$ and $H_n^1(t)$ and $H_{n-}(t)$ are given by (1.4) on page 4. Note that we only know that $M_n(t;\theta_0)$ is a martingale under P_0. If θ_n is arbitrary [refer to this situation as "under P_n", and denote the expectation operator belonging to P_n by $\mathcal{E}_n$] then the process $M_n(t;\theta_0)$ is in general not a martingale. Nevertheless, we also call $M_n(t;\theta_0)$ a basic martingale. Although misleading, this does not lead to difficulties since we do not encounter the original basic martingale $M_n(t;\theta_n)$ anymore.

In section 3.3 we study tests based on a special type of functional of a weighted version of a process $Q_n(t)$, which is a stochastic integral with respect to the basic martingale, that is

$$Q_n(t) = \int_0^t L_n(s) dM_n(s;\theta_0). \tag{3.2}$$

Here the weight process $L_n(t)$ is a random element of $D[0,\infty)$ which possesses certain properties.

The probability theory underlying the results in section 3.3 is presented in section 3.2, where the process $Q_n(t)$ is approximated. The common way to approach a stochastic integral with respect to the basic martingale is by martingale methods, but in this chapter an empirical process approach is followed, fundamented on an alternative representation of $M_n(t;\theta_0)$ in terms of empirical processes. A slight drawback of the empirical process approach is that we must assume that both $F(t;\theta)$ and $G(t)$ are continuous on the complete real line. However, this assumption has its rewards: the knowledge obtained by the empirical process approach is far more precise than can be obtained by using standard martingale methods such as Rebolledo's Central Limit Theorem.

Define the cumulative distribution functions $H^1(t;\theta_n)$ and $H(t;\theta_n)$ by

$$H^1(t;\theta_n) = P_n(Z_1 \le t, \delta_1 = 1) = \int_0^t (1 - G(s)) dF(s;\theta_n), \tag{3.3}$$

$$H(t;\theta_n) = P_n(Z_1 \le t) = 1 - (1 - G(t))(1 - F(t;\theta_n)), \tag{3.4}$$

and the empirical processes $U_n^1(t;\theta_n)$ and $U_{n-}(t;\theta_n)$ by

$$U_n^1(t;\theta_n) = n^{1/2}\{H_n^1(t) - H^1(t;\theta_n)\}, \tag{3.5}$$

$$U_{n-}(t;\theta_n) = n^{1/2}\{H_{n-}(t) - H(t;\theta_n)\}. \tag{3.6}$$

Then we may decompose $M_n(t;\theta_0)$ conveniently into three parts

$$M_n(t;\theta_0) = U_n^1(t;\theta_n) + \int_0^t U_{n-}(s;\theta_n) d\Lambda(s;\theta_0) + n^{1/2} D(t;\theta_0,\theta_n) \tag{3.7}$$

[compare to the decomposition given in equation (7.1.2) in Shorack and Wellner (1986)]. The first two parts involve empirical processes, and can be handled by empirical process theory. The third part is nonrandom and involves the function

$$D(t;\theta_0,\theta) = \int_0^t (1 - H(s;\theta))\{d\Lambda(s;\theta) - d\Lambda(s;\theta_0)\}. \tag{3.8}$$

Observe that if θ coincides with θ_0, then $D(t;\theta_0,\theta)$ is identical to zero. As can be expected, the function $D(t;\theta_0,\theta)$ will show up frequently in our results. Loosely speaking, it reflects the distance between the distribution functions $F(t;\theta)$ and $F(t;\theta_0)$.

Although for the empirical process approach it is needed that G is continuous on the complete real line, we may show that our approximation results remain true on $[0,t^*)$, where $t^* = \sup\{t : G(t) < 1\}$ is finite, if G is continuous only on $(-\infty, t^*)$, by appropriately modifying G on the interval $[t^*, \infty)$. Thus, our results also have implications for Type I censoring.

Moreover, the assumption of independence between X_i and Y_i may be relaxed. However, it is crucial that (3.4) still holds.

Finally, we point out to the reader that sections 3.2 and 3.3 contain results only, and that proofs are gathered in section 3.4.

3.2 Probability inequalities

In this section we present probability inequalities which concern the approximation on the halfline $[0,\infty)$ of [a centered version of] the process $Q_n(t)$ by a one-parameter Gaussian process, both under P_n as under P_0. For treatment of the former situation, the following condition is needed.

Condition 1 *There exists constants $0 < \alpha < 1/2$ and $c_\alpha < \infty$ such that*

$$\int_0^\infty (1 - F(s;\theta))^\alpha d\Lambda(s;\theta_0) < c_\alpha$$

for every $\theta \in \Theta$.

Essentially, Condition 1 relates the right tail behavior of $F(t;\theta)$ to the right tail behavior of $F(t;\theta_0)$. Note that if $(1 - F(t;\theta))/(1 - F(t;\theta_0))$ remains uniformly bounded in θ, then Condition 1 is satisfied for any $\alpha > 0$. Observe that Condition 1 implies

$$|D(t_1;\theta_0,\theta) - D(t_2;\theta_0,\theta)| \le c_{10}(1 - H(t_1 \wedge t_2;\theta))^{1-\alpha}, \qquad (3.9)$$

where $c_{10} = c_\alpha + 1/\alpha$. From (3.9) it immediately follows that $D(t;\theta_0,\theta)$ remains bounded by c_{10}.

The weight process $L_n(t)$ is assumed to satisfy Condition 2.

Condition 2 *There exists a positive non-decreasing function $q(t)$ such that*

a *$L_n(\cdot)/q(\cdot)$ is a random element in [the left continuous right limits version of] $D[0,\infty)$ endowed with the $\mathcal{J}_1$ metric.*

b *There exists a constant c_{11} not depending on θ_n such that*

$$\sup_{t\in[0,\infty)} |L_n(t)|/q(t) < c_{11}, \qquad V(L_n(\cdot)/q(\cdot)) < c_{11}$$

with P_n-probability 1 [here $V(f)$ denotes the total variation of f].

c *There exists a deterministic function $L(t;\theta_n)$ such that for every $\beta < 1/2$*

$$P_n(\sup_{t\in[0,\infty)} |L_n(t) - L(t;\theta_n)|/q(t) > c_{12}n^{-\beta}) \leq c_{13}n^{-c_{14}},$$

where c_{12}–c_{14} are positive constants not depending on θ_n.

d *There exist constants c_{15}–$c_{17} > 0$ such that for every $x > 0$*

$$P_0(\sup_{t\in[0,\infty)} |L_n(t) - L(t;\theta_0)|/q(t) > n^{-1/2}(c_{15}\log n + x)) \leq c_{16}\exp\{-c_{17}x\}.$$

Both $\sup_{t\in[0,\infty)} |L(t;\theta_n)|/q(t)$ and $V(L(\cdot;\theta_n)/q(\cdot))$ do not exceed c_{11} as a consequence of Condition 2.

Subsequent results may be viewed as bearing upon $Q_n(t)/q(t)$ rather than upon $Q_n(t)$ itself. Hence, the choice of $q(t)$ will often be inflicted by the pro, cted application of these results. Typically, one chooses $q(t) = (1-F(t;\theta_0))^{-\rho}$ for some $\rho \geq 0$.

Under Conditions 1 and 2 we have the following result.

Theorem 2 *There exists a sequence $\{W_n(t)\}_{n=1}^{\infty}$ of mean zero Gaussian processes which have covariance function*

$$\begin{aligned}\mathcal{E}_n W_n(t_1)W_n(t_2) &= H^1(t_1 \wedge t_2;\theta_n)\\ &+ \int_0^{t_1\wedge t_2} \{2D(s;\theta_0,\theta_n) - D(t_1;\theta_0,\theta_n) - D(t_2;\theta_0,\theta_n)\}d\Lambda(s;\theta_0)\\ &-D(t_1;\theta_0,\theta_n)D(t_2;\theta_0,\theta_n)\end{aligned} \tag{3.10}$$

such that for every $\beta < (1/2-\alpha)\wedge 1/6$ there exist positive constants c_{18}–c_{20} not depending on θ_n such that

$$\begin{aligned}P_n(\sup_{t\in[0,\infty)} |\{Q_n(t) - n^{1/2}\int_0^t L_n(s)dD(s;\theta_0,\theta_n)\}\\ - \int_0^t L(s;\theta_n)dW_n(s)|/q(t) > c_{18}n^{-\beta}) \leq c_{19}n^{-c_{20}}.\end{aligned} \tag{3.11}$$

Moreover, there exist positive constants c_{21}–c_{23} such that for every $x > 0$

$$\begin{aligned}P_0(\sup_{t\in[0,\infty)} |Q_n(t) - \int_0^t L(s;\theta_0)dW_n(s)|/q(t)\\ > n^{-1/6}(c_{21}\log n + x)^2) \leq c_{22}\exp\{-c_{23}x\}.\end{aligned} \tag{3.12}$$

Corollary 2 *The sequence* $\{Q_n(t)/q(t)\}_{n=1}^{\infty}$ *converges in* P_0*-distribution to* $X(t;\theta_0)/q(t)$*, where* $X(t;\theta_0)$ *is a time-transformed Wiener process with variance function* $\int_0^t (L(s;\theta_0))^2 dH^1(s;\theta_0)$.

In Theorem 2 a stochastic integral with respect to a Gaussian process appears. This integral is defined in the usual way, that is $\int_0^t L(s;\theta_n)dW_n(s)$ denotes $L(t;\theta_n)W_n(t) - \int_0^t W_n(s)dL(s;\theta_n)$.

There is a refinement of Theorem 2 worth mentioning. If the stochastic process $L_n(t)$ coincides with the function $L(t;\theta_n)$ with P_n-probability 1, then (3.11) holds for every $\beta < (1/2-\alpha)$. Moreover, the term $n^{-1/6}$ in (3.12) may be replaced by $n^{-1/2}$.

An approximation of $Q_n(t)$ by a two-parameter Gaussian process, only valid under P_0 and on some fixed closed interval, can be found in Einmahl and Koning (1992), where it is used to derive complete analogues of the Chibisov-O'Reilly theorem, the Lai-Wellner Glivenko-Cantelli theorem and the James law of the iterated logarithm.

Convergence in P_0-distribution of the sequence $\{Q_n(t)/q(t)\}_{n=1}^{\infty}$ may also be obtained by using Rebolledo's martingale central limit theorem [Shorack and Wellner (1986), p. 895], provided that the weight process $L_n(t)$ is predictable. However, such an approach does not lead to probability inequalities of the same type as (3.12), which will prove to be essential for deriving moderate deviation results for the test statistics considered in the next section.

3.3 Sublinear tests

In this section we study sublinear tests for the simple null hypothesis that θ_n equals θ_0. These tests are based on statistics of the form $T(Q_n(\cdot)/q(\cdot))$, where T is a special type functional mapping $D[0,\infty)$ into $\mathbb{R}$. We consider the behavior of these test statistics under the null hypothesis and under fixed and local alternatives, as well as efficiencies of the corresponding tests. Generalized rank and supremum type tests receive special attention, and are shown to be in some sense optimal for specific choices of the weight process.

Definition 2 *A functional* T *mapping* $D[0,\infty)$ *into* $\mathbb{R}$ *is said to be Lipschitz if there exists a constant* c_T *such that*

$$|T(\xi) - T(\zeta)| \le c_T \sup_{t\in[0,\infty)} |\xi(t) - \zeta(t)|$$

for every $\xi, \zeta \in D[0,\infty)$.

Condition 3 *The functional* T *is sublinear and Lipschitz.*

Since T is sublinear, the tail behavior of a random variable obtained by applying T to some Gaussian process ξ is described by Inequality 5. The Lipschitz property, which is borrowed from Inglot and Ledwina (1990), enables us to carry these results to some extent over to non-Gaussian processes close to ξ.

3.3.1 Deviations

Let $\mathcal{S}$ be the set of Strassen-functions and define

$$a = \{\sup_{f \in \mathcal{S}} T(\int_0^{\cdot} L(s;\theta_0) f'(H^1(s;\theta_0)) dH^1(s;\theta_0)/q(\cdot))\}^{-2}. \tag{3.13}$$

One may interpret $a^{-1/2}$ as the norm of the functional T induced by the reproducing kernel Hilbert space of the Gaussian process $X(t;\theta_0)/q(t)$, the limit in P_0-distribution of the sequence $\{Q_n(t)/q(t)\}_{n=1}^{\infty}$. Suppose that

Condition 4 *The number a is positive.*

Theorem 3 describes the tail behavior of $T(X(\cdot;\theta_0)/q(\cdot))$ [and thus the tail behavior of the asymptotic distribution of $T(Q_n(\cdot)/q(\cdot))$], and presents a moderate deviation result for $T(Q_n(\cdot)/q(\cdot))$. It holds if equation (3.12) and Conditions 3 and 4 hold.

Theorem 3 *We have*

$$\lim_{t\to\infty} t^{-2} \log P_0(T(X(\cdot;\theta_0)/q(\cdot)) > t) = -a/2. \tag{3.14}$$

Furthermore,

$$\lim_{n\to\infty} (s_n)^{-2} \log P_0(T(Q_n(\cdot)/q(\cdot)) > s_n) = -a/2 \tag{3.15}$$

for any sequence $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = o(n^{1/18})$ as $n \to \infty$.

Since $s_n = \mathcal{O}((\log n)^{1/2})$ is a special case of $s_n = o(n^{1/18})$, Theorem 3 implies that a moderate deviation result holds for $T(Q_n(\cdot)/q(\cdot))$. Moderate deviation results are important from a statistical perspective, because they play a role in evaluating the performance of a test.

As with Theorem 2, Theorem 3 can be refined in the special case where $L_n(t)$ coincides with $L(t;\theta_0)$ with P_0-probability 1. Now $s_n = o(n^{1/18})$ may be replaced by $s_n = o(n^{1/6})$, and thus we have obtained a Cramér type large deviation result. A Cramér type large deviation result should be distinguished from a Chernoff type large deviation result which allows $s_n = \mathcal{O}(n^{1/2})$.

To transform (3.14) into a moderate or large deviation result, probability inequalities of the type (3.12) are needed. Hence, Theorem 3 does not follow from Rebolledo's Central Limit Theorem.

It is possible to generalize Theorem 3 by relaxing the Lipschitz requirement in Condition 3. For example, in Inglot and Ledwina (1989) it is assumed that there exists a constant $c_T > 0$ and a weight function $\tilde{q}(t)$ belonging to some special class such that

$$|T(\xi) - T(\zeta)| \leq c_T \sup_{t\in[0,\infty)} |\xi(t) - \zeta(t)|/\tilde{q}(t) \qquad \text{for all } \xi, \zeta \in D[0,\infty).$$

However, since the functionals of our primary interest, the generalized rank and supremum type functionals, already are Lipschitz, we have preferred to present Theorem 3 in the simple version.

Since in general the distribution function $H^1(t;\theta_0)$ is involved in a, practical problems arise when the censoring distribution G is unknown. In this case it seems best to multiply the original weight process by the square root of an estimator for a. Typically, estimators for a are obtained by replacing $H^1(t;\theta_0)$ by $\int_0^t (1 - H_{n-}(s))d\Lambda(s;\theta_0)$. Of course, it should be verified whether the newly constructed weight process meets all requirements.

3.3.2 Behavior under the alternative hypothesis

Next, we consider the behavior of $T(Q_n(\cdot)/q(\cdot))$ under the alternative hypothesis. Before turning to local alternatives, we first briefly discuss the behavior under a fixed alternative. Suppose that θ is an element of Θ, not necessarily equal to θ_0, and that we have $\theta_n = \theta$ for all $n \in I\!N$ [we shall refer to this situation as "under P_θ"]. Combining equation (3.11) with Conditions 2c, and 3 yields for every $\beta > 0$

$$n^{-\beta}|T(Q_n(\cdot)/q(\cdot)) - n^{1/2}T(\int_0^{\cdot} L(s;\theta))dD(s;\theta_0,\theta)/q(\cdot))| \to_{P_\theta} 0. \tag{3.16}$$

To make treatment of the behavior under local alternatives possible we need additional notation and conditions. First we assume

Condition 5 *Θ is a convex subset of $I\!R^p$.*

The hazard function $\lambda(t;\theta)$ is defined as the derivative with respect to t of $\Lambda(t;\theta)$. Let $\psi_i(t;\theta)$ denote the first order partial derivative of $\log\lambda(t;\theta)$ with respect to the i^{th} component of θ, and $\psi_{ij}^{(1)}(t;\theta)$ the second order partial derivative of $\log\lambda(t;\theta)$ with respect to the i^{th} and j^{th} components of θ.

Condition 6 *For every $\theta \in \Theta$ and $i,j = 1,\ldots,p$ the functions $\lambda(t;\theta)$, $\psi_i(t;\theta)$ and $\psi_{ij}^{(1)}(t;\theta)$ exist. For some $\beta > 0$ there exists a constant c_{24} such that for every $\theta \in \Theta$ and $i,j = 1,\ldots,p$*

$$\int_0^\infty (\psi_i(s;\theta))^2(1 - H(s;\theta))^{(1-2\alpha)-\beta}d\Lambda(s;\theta) < c_{24},$$

$$\int_0^\infty (\psi_{ij}^{(1)}(s;\theta))^2(1 - H(s;\theta))^{2(1-2\alpha)-\beta}d\Lambda(s;\theta) < c_{24}.$$

In the description of the behavior of the test statistics under local alternatives, the p-dimensional vector function $K_a(t)$ is involved. The i^{th} element of $K_a(t)$ is defined by

$$K_{ai}(t) = \int_0^t L(s;\theta_0)\psi_i(s;\theta_0)dH^1(s;\theta_0). \tag{3.17}$$

It is easily proved that $K_{ai}(t)/q(t)$ remains uniformly bounded in t and θ_0 under Conditions 1 and 6.

Recall that $L(t;\theta)$ is the limiting function of the process $L_n(t)$ under P_θ. Let $L_i^{(1)}(t;\theta)$ denote the first order partial derivative of $L(t;\theta)$ with respect to the ith component of θ, and $L_{ij}^{(2)}(t;\theta)$ the second order partial derivative of $L(t;\theta)$ with respect to the i^{th} and j^{th} components of θ.

Condition 7 *For every $\theta \in \Theta$ and $i,j = 1,\ldots,p$ the functions $L_i^{(1)}(t;\theta)$ and $L_{ij}^{(2)}(t;\theta)$ exist. There exists a constant c_{25} such that for every $\theta \in \Theta$ and for $i,j = 1,\ldots,p$*

$$\sup_{t\in[0,\infty)} (1 - H(t;\theta))^{1/2}|L_i^{(1)}(t;\theta)|/q(t) < c_{25},$$

$$\sup_{t\in[0,\infty)} (1 - H(t;\theta))^{1-\alpha}|L_{ij}^{(2)}(t;\theta)|/q(t) < c_{25}.$$

Theorem 4 *Suppose the sequence $\{\theta_n\}_{n=1}^{\infty}$ converges to the point θ_0. Let h be the p-dimensional unit vector defined by $h = \lim_{n\to\infty}(\theta_n - \theta_0)/|\theta_n - \theta_0|$, and let σ denote $\lim_{n\to\infty} n^{1/2}|\theta_n - \theta_0|$.*

a *If $\sigma = \infty$ then $(n^{1/2}|\theta_n - \theta_0|)^{-1}T(Q_n(\cdot)/q(\cdot))$ converges in P_n-probability to $T(h^T K_a(\cdot)/q(\cdot))|$.*

b *If $\sigma < \infty$ then $\{T(Q_n(\cdot)/q(\cdot))\}_{n=1}^{\infty}$ converges in P_n-distribution to the random variable $T(\{X(\cdot;\theta_0) + \sigma h^T K_a(\cdot)\}/q(\cdot))$.*

Theorem 4 reveals three types of behavior of the sequence of test statistics $\{T(Q_n(\cdot)/q(\cdot))\}_{n=1}^{\infty}$, depending on the rate at which the alternatives converge to the null hypothesis. If the rate is faster than $n^{-1/2}$ then we have convergence in distribution to the same limit as under the null hypothesis. If the rate is of the order $n^{-1/2}$ then we also have convergence in distribution, but to a limit different from the one under the null hypothesis. If the rate is slower than $n^{-1/2}$ then the convergence in distribution is lost, since the sequence of test statistics blows up as n tends to infinity.

3.3.3 Efficiencies

Now that we have investigated the behavior of the test statistics, it is time to evaluate the corresponding tests. For assessing the performance of a test a multitude of efficiency concepts are available. A few of them are discussed.

To introduce these concept, consider two infinite sequences of tests. Each of the tests in these two sequences has the same size, that is the same probability of falsely rejecting the null hypothesis. For ease of exposition, suppose that the n^{th} test in the i^{th} sequence $[i = 1, 2]$ is based on the test statistic T_{in}, rejects the null hypothesis if $T_{in} > t_{in}$, and does not reject if $T_{in} < t_{in}$. We have no knowledge about the action taken if $T_{in} = t_{in}$. Hence, the power of the n^{th} test in the i^{th} sequence under P_θ is bounded from below by $P_\theta(T_{in} > t_{in})$ and from below by $P_\theta(T_{in} \geq t_{in})$.

The definition of asymptotic relative Pitman efficiency that follows next is adopted from Wieand (1976).

Definition 3 *For a given function $\tilde{\beta} : \Theta \to [0, 1]$, let $\overline{N}_i^{\tilde{\beta}}(\theta)$ be the largest sample size such that $P_\theta(T_{in} \geq t_{in}) < \tilde{\beta}(\theta)$. If for a unit vector $h \in I\!R^p$ there exists a constant $e_{12}^{\tilde{\beta}}(h)$ such that*

$$\liminf_{j\to\infty} \frac{\overline{N}_2^{\tilde{\beta}}(\theta_j)}{\overline{N}_1^{\tilde{\beta}}(\theta_j)} \geq e_{12}^{\tilde{\beta}}(h)$$

and

$$\limsup_{j\to\infty} \frac{\overline{N}_2^{\tilde{\beta}}(\theta_j)}{\overline{N}_1^{\tilde{\beta}}(\theta_j)} \leq e_{12}^{\tilde{\beta}}(h)$$

for any sequence $\{\theta_j\}_{j=1}^{\infty}$ tending to θ_0 while satisfying $\lim_{j\to\infty} \tilde{\beta}(\theta_j) \in (0, 1)$ and $\lim_{j\to\infty}(\theta_j - \theta_0)/|\theta_j - \theta_0| = h$, then $e_{12}^{\tilde{\beta}}(h)$ is the asymptotic relative Pitman efficiency in the direction h of the first sequence of tests with respect to the second.

By using this general definition one allows the asymptotic relative Pitman efficiency to depend on both size and power [which are kept under control] of the tests. In Rothe (1981) an alternative definition is given which involves $\liminf_{j\to\infty} \underline{N}_2^{\tilde{\beta}}(\theta_j)/\overline{N}_1^{\tilde{\beta}}(\theta_j)$ and $\limsup_{j\to\infty} \overline{N}_2^{\tilde{\beta}}(\theta_j)/\underline{N}_1^{\tilde{\beta}}(\theta_j)$, where $\underline{N}_i^{\tilde{\beta}}(\theta)$ is the smallest sample size such that $P_\theta(T_{in} > t_{in}) \geq \tilde{\beta}(\theta)$. In case there is a transformation of the test statistics available so as to obtain transformed test statistics which are asymptotically normal and have asymptotic variance 1 for any sequence $\{\theta_n\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} n^{1/2}|\theta_n - \theta_0| < \infty$, then the asymptotic relative Pitman efficiency does not depend on size and power. The asymptotic mean of these transformed test statistics divided by $\lim_{n\to\infty} n^{1/2}|\theta_n - \theta_0|$ is the square root of what is called the efficacy of the sequence of test statistics. For two such sequences the asymptotic relative Pitman efficiency is equal to the ratio of their respective efficacies. Unfortunately, since the efficacy approach can not be followed for general sublinear tests [the supremum type tests are for instance explicitly non-normal], the task of establishing asymptotic relative Pitman efficiencies becomes too formidable.

Under these circumstances approximate Bahadur efficiencies are often much easier to compute.

Definition 4 *A sequence of test statistics* $\{T_{in}\}_{n=1}^{\infty}$ *is said to be a standard sequence if the following three conditions are satisfied.*
a *The sequence* $\{T_{in}\}_{n=1}^{\infty}$ *converges in* P_0*-distribution to a random variable* T_i.
b *There exists a constant* $a_i > 0$ *such that*

$$\lim_{t\to\infty} t^{-2}\log P_0(T_i > t) = -a_i/2.$$

c *For every fixed* $\theta \in \Theta - \{\theta_0\}$ *there exists a constant* $b_i(\theta) > 0$ *such that* $|n^{-1/2}T_{in} - b_i(\theta)|$ *converges to zero in* P_θ*-probability.*
The approximate Bahadur slope of a standard sequence $\{T_{in}\}_{n=1}^{\infty}$ *is defined as* $a_i(b_i(\theta))^2$. *The approximate Bahadur efficiency of a standard sequence* $\{T_{1n}\}_{n=1}^{\infty}$ *with respect to another standard sequence* $\{T_{2n}\}_{n=1}^{\infty}$ *is defined as the ratio of their respective Bahadur slopes* $a_1(b_1(\theta))^2/a_2(b_2(\theta))^2$.

By Corollary 2 and equations (3.14) and (3.16) it immediately follows that the approximate Bahadur slope of the sequence $\{T(Q_n(\cdot)/q(\cdot))\}_{n=1}^{\infty}$ is given by $a\{T(\int_0^{\cdot} L(s;\theta)dD(s;\theta_0,\theta)/q(\cdot))\}^2$.

Approximate Bahadur efficiency has been subject to some criticism. Already in Bahadur (1960) it is advocated that conclusions should not be entirely based on approximate Bahadur slopes. In Wieand (1976) a condition is given under which the existence of the limiting [as the alternative approaches the null hypothesis] approximate Bahadur efficiency implies the existence of the limiting [as the size of the test approaches zero] asymptotic relative Pitman efficiency and the equality of the two limits. This condition obviates most of the difficulties involved in the interpretation of approximate Bahadur efficiencies, at least for θ in the vicinity of θ_0.

Definition 5 *A standard sequence* $\{T_{in}\}_{n=1}^{\infty}$ *is said to be a Wieand sequence if there is a constant* $\epsilon^* > 0$ *such that for every* $\epsilon > 0$ *and* $\delta \in (0,1)$ *there exists an integer* N *such that*

$$P_\theta(|n^{-1/2}T_{in} - b_i(\theta)| > \epsilon b_i(\theta)) < \delta$$

for every fixed θ *satisfying* $|\theta - \theta_0| < \epsilon^*$ *and* $n > N/(b_i(\theta))^2$.

Theorem 5 (Wieand (1976)) *Let* $\{T_{1n}\}_{n=1}^{\infty}$ *and* $\{T_{2n}\}_{n=1}^{\infty}$ *be two Wieand sequences. Suppose* $\lim_{\theta\to\theta_0} b_i(\theta) = 0$ *for* $i = 1,2$ *and suppose that for every p-dimensional unit vector h the limit*

$$\lim_{\theta\to\theta_0}\{a_1(b_1(\theta))^2\}/\{a_2(b_2(\theta))^2\} \tag{3.18}$$

exists if $(\theta-\theta_0)/|\theta-\theta_0|$ *tends to h. Then the limiting [as the size of the tests approaches zero] asymptotic relative Pitman efficiency exists and is equal to the limit given in (3.18).*

Theorem 5 is not valid if Rothe's definition of asymptotic relative Pitman efficiency is used. See Appendix A for more details.

Obviously, the limiting asymptotic relative Pitman efficiency does not depend on the size. Surprisingly, by letting the size of the test tend to zero, the dependence on the power has also vanished.

Theorem 6 *Let h be a p-dimensional unit vector, and let θ approach θ_0 from the direction h [that is, $\lim_{\theta \to \theta_0}(\theta - \theta_0)/|\theta - \theta_0| = h$]. If*

$$e(h) = a\{T(h^T K_a(\cdot)/q(\cdot))\}^2 \tag{3.19}$$

is not equal to zero, then $\{T(Q_n(\cdot)/q(\cdot))\}_{n=1}^{\infty}$ is a Wieand sequence with approximate Bahadur slope of the form

$$|\theta - \theta_0|^2\{e(h) + o(1)\}. \tag{3.20}$$

The Wieand approach to efficiency is based on letting both the size of the test tend to zero and the alternative tend to the null hypothesis. However, both operations are done separately. In Kallenberg (1983) a concept of efficiency is proposed based on performing both operations simultaneously. It can be considered as intermediate between the Pitman and the exact Bahadur approach.

Definition 6 *A sequence of test statistics $\{T_{in}\}_{n=1}^{\infty}$ is said to be a Kallenberg sequence if the following conditions are satisfied.*

a *There exists a constant $a_i > 0$ such that*

$$\lim_{n \to \infty} (s_n)^{-2} \log P_0(T_{in} > s_n) = -a_i/2$$

for all sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = o(n^{1/6})$ as $n \to \infty$.

b *There exists a positive function $b_i(\theta)$ such that $n^{-1/2}T_{in}/b_i(\theta_n) \to 1$ in P_n-probability for all sequences $\{\theta_n\}_{n=1}^{\infty}$ such that $\theta_n \to \theta_0$ and $n^{1/2}|\theta_n - \theta_0| \to \infty$ as n tends to infinity.*

If the sequences $\{T_{1n}\}_{n=1}^{\infty}$ and $\{T_{2n}\}_{n=1}^{\infty}$ both are Kallenberg and if the limit $\lim_{n\to\infty} a_1(b_1(\theta_n))^2/a_2(b_2(\theta_n))^2$ exists, then the asymptotic intermediate efficiency of $\{T_{1n}\}_{n=1}^{\infty}$ with respect to $\{T_{2n}\}_{n=1}^{\infty}$ is defined as this limit.

Typically, $b_i(\theta)$ behaves near θ_0 as a linear function of $|\theta - \theta_0|$, which justifies introducing the intermediate slope $\lim_{n\to\infty} a_i(b_i(\theta_n)/|\theta_n - \theta_0|)^2$. If the weight process $L_n(t)$ coincides with $L(t; \theta_0)$ with P_n-probability 1, then $T(Q_n(\cdot)/q(\cdot))\}_{n=1}^{\infty}$ is a Kallenberg sequence with intermediate slope equal to the quantity $e(h)$ defined by (3.19), as follows from Theorem 4 and the refinement of Theorem 3.

A variant of asymptotic intermediate efficiency, also proposed in Kallenberg (1983), is weak asymptotic intermediate efficiency. Here only sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = \mathcal{O}((\log n)^{1/2})$ as $n \to \infty$ are considered. Observe that the sequence $\{T(Q_n(\cdot)/q(\cdot))\}_{n=1}^{\infty}$ has weak intermediate slope $e(h)$, even if the weight process does not coincide with its limiting weight function. Hence, the weak intermediate approach yields the same picture as the Wieand approach.

The concepts in Kallenberg (1983) were proposed so as to correspond with several types of moderate and large deviation results. In the light of Theorem 3 it is tempting to propose a variant of asymptotic intermediate efficiency, which considers sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = o(n^{1/18})$ as n tends to infinity.

In the beginning of this section we assumed that the functional T is sublinear. A close look reveals that sublinearity is used in the derivation of (3.14) only. Thus, we may set up an equivalent theory for functionals other than sublinear, provided a result similar to (3.14) holds. As examples we mention the functionals occuring in Cramér-von Mises and chi-square tests [see Durbin (1973)].

However, we have preferred to restrict our attention to the class of sublinear tests, since it comprises two particularly appealing subclasses, the class of generalized rank tests, based on the generalized rank functional T_R, and the class of supremum type tests, based on the supremum functional T_S. The remainder of this section is devoted to these two subclasses. We should warn that all the attention given here to generalized rank and supremum type tests could falsely yield the impression that they are the only sublinear tests worth notice. For instance, refer to Aki and Kashiwagi (1989) for a sublinear test based on a functional other than T_R and T_S.

3.3.4 Generalized rank and supremum type tests

It is easily seen that both T_R and T_S satisfy Condition 3. Hence, our theory applies [set $q(t)$ equal to 1 for all $t \in [0,\infty)$]. An application of the Cauchy-Schwarz inequality yields

$$(\sup_{f\in S}\int_0^t L(s;\theta_0)f'(H^1(s;\theta_0))dH^1(s;\theta_0))^2 \le \int_0^t (L(s;\theta_0))^2 dH^1(s;\theta_0)$$

for every $t \in [0,\infty)$. Moreover, the right-hand side of latter inequality is achieved and bounded by $\int_0^{\infty}(L(s;\theta_0))^2 dH^1(s;\theta_0)$. Thus, defining a_R and a_S according to (3.13), with T replaced by T_R and T_S respectively, it follows that

$$a_R = a_S = \{\int_0^{\infty}(L(s;\theta_0))^2 dH^1(s;\theta_0)\}^{-1}. \tag{3.21}$$

Similarly defining $e_R(h)$ and $e_S(h)$ according to (3.19), we obtain

$$e_R(h) = a_R\{h^T K_a(\infty)\}^2, \tag{3.22}$$

$$e_S(h) = a_S\{\sup_{t\in[0,\infty)} h^T K_a(t)\}^2. \tag{3.23}$$

It should be noted that if the weight process coincides with its limiting function with P_n-probability 1, then $T_R(Q_n)$ can be written as the sum of i.i.d. random variables [see Proposition 3.1 in Aki (1986)], and hence results for this special type generalized rank test may be proven in a simpler manner. For instance, Theorem 3 is now an immediate consequence of Theorem 1 on page 549 of Feller (1971). Observe that this alternative proof also leads to a Cramér type large deviation result.

As opposed to general sublinear tests, generalized rank tests do allow us to compute asymptotic relative Pitman efficiencies. By Corollary 2 and Theorem 4b, it follows that the asymptotic power against local alternatives $\theta_n = \theta_0 + n^{-1/2}h$ of the test based on $T_R(Q_n)$ of size $\tilde{\alpha}$ equals

$$P_0(X(\infty;\theta_0) > z_{\tilde{\alpha}}(a_R)^{-1/2} - h^T K_a(\infty)),$$

where $z_{\tilde{\alpha}}$ is the $(1-\tilde{\alpha})$ quantile of the standard normal distribution. This implies that the efficacy of the sequence of test statistics $\{T_R(Q_n)\}_{n=1}^{\infty}$ is equal to $e_R(h)$.

Since applying the Cauchy-Schwarz inequality to $h^T K_a(\infty)$ yields

$$e_R(h) \leq \int_0^{\infty} (h^T \psi(s;\theta_0))^2 dH^1(s;\theta_0), \tag{3.24}$$

where $\psi(t;\theta_0)$ is the p-dimensional vector with elements $\psi_i(t;\theta_0)$, it follows that $e_R(h)$ [and hence asymptotic relative Pitman efficiency, limiting approximate Bahadur efficiency and weak intermediate efficiency of generalized rank tests] is maximized by those tests based on weight processes with limiting weight function satisfying

$$L(t;\theta_0) \propto h^T \psi(t;\theta_0). \tag{3.25}$$

From the following lemma it follows that the upper bound for the efficacy derived in Rao (1963) coincides with the right hand side of inequality (3.24), and thus generalized rank tests based on weight processes with limiting weight function satisfying (3.25) are asymptotically most powerful.

Lemma 1 *The Fisher information matrix $I(\theta_0)$ equals the $p \times p$ matrix with elements*

$$I_{ij}(\theta_0) = \int_0^{\infty} \psi_i(s;\theta_0)\psi_j(s;\theta_0) dH^1(s;\theta_0).$$

Clearly, the use of a generalized rank test instead of a classical test [the likelihood ratio test, say] does not have to result in loss of asymptotic relative efficiency. This raises the question whether the same conclusion holds for a supremum type test. After all, in contrast to the generalized rank test which is not consistent against alternatives approaching the null hypothesis in a direction perpendicular to $K_a(\infty)$, the supremum type test has the character of an omnibus test. Needless to say, it is rather attractive to have an efficient omnibus test at our disposal.

In virtually the same way as with $e_R(h)$, we obtain that $e_S(h)$ is maximized by supremum type tests based on weight processes with limiting weight function satisfying (3.25). Furthermore, it is easily seen that these tests have efficiency 1 with respect to the optimal test in the sense of limiting asymptotic relative Pitman efficiency, limiting approximate Bahadur efficiency and weak intermediate efficiency.

This last result may impel to question the usefulness of concepts which are not able to distinguish between generalized rank tests and supremum type tests. However, as recent results on the tail behavior of the supremum of a Gaussian process show, this inability is basically a consequence of letting the size of the test tend to zero, which is the sensible thing to do if we are committed to avoiding making errors, Type I as well as Type II. From Rubin and Sethuraman (1965) it follows that minimizing the Bayes risk leads to letting the size of the test tend to zero at a rate n^{-1}. Observe that this is exactly the situation to which weak intermediate efficiency refers.

3.4 Proofs

In this section we prove the theorems presented in sections 3.2 and 3.3.

Proof of Theorem 2 As in Einmahl and Koning (1992), proof of Proposition 1 [see also Theorem 3.1 in Burke, Csörgő and Horváth (1981)], let $\tilde{U}_n$ denote the empirical process based on the uniform (0,1) random variables

$$\tilde{Z}_i = \delta_i H^1(Z_i;\theta_n) + (1-\delta_i)\{H^1(\infty;\theta_n) + H^0(Z_i;\theta_n)\},$$

where

$$H^0(t;\theta_n) = H(t;\theta_n) - H^1(t;\theta_n)$$

is the cumulative distribution function of the censored failure times under P_n. Note that

$$U_n^1(t;\theta_n) = \tilde{U}_n(H^1(t;\theta_n)),$$

$$U_{n-}(t;\theta_n) = \tilde{U}_{n-}(H^1(t;\theta_n))$$
$$+\tilde{U}_{n-}(H^1(\infty;\theta_n) + H^0(t;\theta_n)) - \tilde{U}_{n-}(H^1(\infty;\theta_n)),$$

where $\tilde{U}_{n-}(t)$ denotes the left continuous version of $\tilde{U}_n(t)$. Inequality 2 yields the existence of a sequence $\{\tilde{B}_n(t)\}_{n=1}^{\infty}$ of Brownian bridges with continuous sample paths such that for all $x > 0$

$$P_n(\sup_{t\in[0,\infty)} |\tilde{U}_n(t) - \tilde{B}_n(t)| > n^{-1/2}(c_2 \log n + x)) \le c_3 \exp\{-c_4 x\}. \tag{3.26}$$

Now define mean zero Gaussian processes $B_n^1(t)$, $B_n(t)$ and $W_n(t)$ by

$$B_n^1(t) = \tilde{B}_n(H^1(t;\theta_n)),$$

$$B_n(t) = \tilde{B}_n(H^1(t;\theta_n)) + \tilde{B}_n(H^1(\infty;\theta_n) + H^0(t;\theta_n)) - \tilde{B}_n(H^1(\infty;\theta_n)),$$

$$W_n(t) = B_n^1(t) + \int_0^t B_n(s)d\Lambda(s;\theta_0).$$

The processes $B_n^1(t)$ and $B_n(t)$ are used to approximate $U_n^1(t;\theta_n)$ and $U_{n-}(t;\theta_n)$, respectively. Thus, it follows by (3.7) that $W_n(t)$ approximates $M_n(t;\theta_0) - n^{1/2}D(t;\theta_0,\theta_n)$. Before studing the implications of this approximation, we first pay attention to the covariance structure of $W_n(t)$. Covariance calculations yield

$$\mathcal{E}_n B_n^1(t_1)B_n^1(t_2) = H^1(t_1 \wedge t_2;\theta_n) - H^1(t_1;\theta_n)H^1(t_2;\theta_n),$$

$$\mathcal{E}_n B_n^1(t_1)B_n(t_2) = H^1(t_1 \wedge t_2;\theta_n) - H^1(t_1;\theta_n)H(t_2;\theta_n),$$

$$\mathcal{E}_n B_n(t_1)B_n(t_2) = H(t_1 \wedge t_2;\theta_n) - H(t_1;\theta_n)H(t_2;\theta_n),$$

and hence (3.10). Observe that (3.10) implies

$$\begin{aligned}\mathcal{E}_n\{W_n(t_1) - W_n(t_2)\}^2 &= H^1(t_1 \vee t_2;\theta_n) - H^1(t_1 \wedge t_2;\theta_n)\\ &\quad +2\int_{t_1\wedge t_2}^{t_1\vee t_2}\{D(s;\theta_0,\theta_n) - D(t_1 \vee t_2;\theta_0,\theta_n)\}d\Lambda(s;\theta_0)\\ &\quad -(D(t_1;\theta_0,\theta_n) - D(t_2;\theta_0,\theta_n))^2, \end{aligned} \tag{3.27}$$

from which we may infer by (3.9)

$$\begin{aligned}\mathcal{E}_n\{W_n(t_1) - W_n(t_2)\}^2 &\le H^1(t_1 \vee t_2;\theta_n) - H^1(t_1 \wedge t_2;\theta_n)\\ &\quad +2c_{10}\int_{t_1\wedge t_2}^{t_1\vee t_2}(1 - H(s;\theta_n))^{1-\alpha}d\Lambda(s;\theta_0)\\ &\le c_{26}\int_{t_1\wedge t_2}^{t_1\vee t_2}(1 - H(s;\theta_n))^{\alpha}\{d\Lambda(s;\theta_n) + d\Lambda(s;\theta_0)\}, \end{aligned} \tag{3.28}$$

where $c_{26} = 1 + 2c_{10}$. To gain some probabilistic insight in the process $W_n(t)$, remark that

$$W_n(t) - \int_0^t \frac{B_n(s)}{1 - H(s;\theta_n)} dD(s;\theta_0,\theta_n)$$

is a time-transformed Wiener process with variance function $H^1(t;\theta_n)$. Moreover, the process $B_n(t)/1-H(t;\theta_n)$ is a time-transformed Wiener process with variance function $H(t;\theta_n)/(1-H(t;\theta_n))$.

Recall that $W_n(t)$ approximates $M_n(t;\theta_0) - n^{1/2}D(t;\theta_0,\theta_n)$. As a direct consequence we have that $\int_0^t L_n(s)\{dM_n(s;\theta_0) - n^{1/2}dD(s;\theta_0,\theta_n)\}$ is approximated by $\int_0^t L_n(s)dW_n(s)$. Unfortunately, the latter process is not easy to work with [it may not even be Gaussian], so we prefer replacing $\int_0^t L_n(s)dW_n(s)$ by $\int_0^t L(s;\theta_n)dW_n(s)$. To evaluate the effects of this replacement, we make use of a pure jump process $J_n(t)$, which is defined by

$$J_n(t) = W_n(x_{i,n}) \qquad \text{for } t \in I_{i,n}, \tag{3.29}$$

where $I_{i,n} = [x_{i,n}, x_{i+1,n})$ and $0 = x_{0,n} < x_{1,n} < \cdots < x_{m(n),n} = \infty$ is a grid chosen so as to satisfy

$$\int_{I_{i,n}} (1 - H(s;\theta_n))^\alpha \{d\Lambda(s;\theta_n) + d\Lambda(s;\theta_0)\}$$
$$\leq n^{-1/3} \int_0^\infty (1 - H(s;\theta_n))^\alpha \{d\Lambda(s;\theta_n) + d\Lambda(s;\theta_0)\} \tag{3.30}$$

for $i = 0, \ldots, m(n) - 1$. If the grid is chosen carefully, then there is no need for $m(n)$ to exceed $n^{1/3} + 1$. We shall assume that this is indeed the case.

It follows by (3.28) that the variance of the mean zero Gaussian process $W_n(t)$ is bounded by $c_{26}c_{10}$, whereas the variance of the mean zero Gaussian process $J_n(t) - W_n(t)$ is bounded by $c_{26}c_{10}n^{-1/3}$. Hence, by applying Inequality 4 we obtain

$$P_n(\sup_{t\in[0,\infty)} |J_n(t)| > x) \leq P_n(\sup_{t\in[0,\infty)} |W_n(t)| > x)$$
$$\leq c_8 \exp\{-c_{27}x^2\}, \tag{3.31}$$

$$P_n(\sup_{t\in[0,\infty)} |J_n(t) - W_n(t)| > xn^{-1/6}) \leq c_8 \exp\{-c_{27}x^2\}, \tag{3.32}$$

where $c_{27} = c_9/(c_{26}c_{10})$. Note that the application of Inequality 4 to $J_n(t) - W_n(t)$ is justified since this process is separable.

For any sequence $\{d_n\}_{n=1}^\infty$ of points in $(0,\infty)$ we may now write

$$\sup_{t\in[0,\infty)} |\{Q_n(t) - n^{1/2}\int_0^t L_n(s)dD(s;\theta_0,\theta_n)\}$$
$$- \int_0^t L(s;\theta_n)dW_n(s)|/q(t) \leq \sum_{i=1}^6 \Delta_{ni}, \tag{3.33}$$

where

$$\Delta_{n1} = \sup_{t\in[0,d_n]} |Q_n(t) - \int_0^t L_n(s)\{dW_n(s) + n^{1/2}dD(s;\theta_0,\theta_n)\}|/q(t),$$

$$\Delta_{n2} = \sup_{t\in[d_n,\infty)} |Q_n(t) - Q_n(d_n)|/q(t),$$

$$\Delta_{n3} = \sup_{t\in[d_n,\infty)} |\int_{d_n}^t L_n(s)dW_n(s)|/q(t),$$

$$\Delta_{n4} = n^{1/2} \sup_{t\in[d_n,\infty)} |\int_{d_n}^t L_n(s)dD(s;\theta_0,\theta_n)|/q(t),$$

$$\Delta_{n5} = \sup_{t\in[0,\infty)} |\int_0^t L_n(s)\{dW_n(s) - dJ_n(s)\}|/q(t),$$

$$\Delta_{n6} = \sup_{t\in[0,\infty)} |\int_0^t \{L_n(s) - L(s;\theta_n)\}dJ_n(s)|/q(t).$$

Observe that the terms Δ_{n5} and Δ_{n6} relate to the replacement of $\int_0^t L_n(s)dW_n(s)$ by $\int_0^t L(s;\theta_n)dW_n(s)$.

Later in this proof we will meet two specific choices of $\{d_n\}_{n=1}^{\infty}$. A first choice leads to (3.11), and a second to (3.12). Both choices have in common that $d_n \to \infty$ as $n \to \infty$. But before these choices are made, we explore the behavior of Δ_{ni}, i=1,...,6 for general sequences. Integration by parts yields with P_n-probability 1

$$\begin{aligned}
\Delta_{n1} &\le \{\sup_{t\in[0,\infty)} |L_n(t)/q(t)| + V(L_n(\cdot)/q(\cdot))\} \\
&\quad \times\{\sup_{t\in[0,d_n]} |M_n(t;\theta_0) - n^{1/2}D(t;\theta_0,\theta_n) - W_n(t)|\} \\
&\le 2c_{11}\{\sup_{t\in[0,d_n]} |U_n^1(t;\theta_n) - B_n^1(t)| \\
&\quad + \sup_{t\in[0,d_n]} |\int_0^t (U_{n-}(s;\theta_n) - B_n(s))d\Lambda(s;\theta_0)|\} \\
&\le 2c_{11}(1 + 3\Lambda(d_n;\theta_0)) \sup_{t\in[0,1]} |\tilde{U}_n(t) - \tilde{B}_n(t)|.
\end{aligned}$$

and therefore it follows by (3.26) that

$$P_n(\Delta_{n1} > 2c_{11}(1 + 3\Lambda(d_n;\theta_0))n^{-1/2}(c_2 \log n + x)) \le c_3 \exp\{-c_4 x\}. \qquad (3.34)$$

Observing that $\Delta_{n2} = 0$ if $Z_{n:n} < d_n$, where $Z_{n:n}$ denotes the largest order statistic of the sample $Z_1, \ldots, Z_n$, we obtain

$$P_n(\Delta_{n2} \ne 0) \le P_n(Z_{n:n} \ge d_n)$$

$$\leq 1 - (H(d_n;\theta_n))^n$$

$$\leq n(1 - H(d_n;\theta_n)). \tag{3.35}$$

Twice integrating by parts yields with P_n-probability 1

$$\begin{aligned}\Delta_{n3} &\leq \{V(L_n(\cdot)/q(\cdot)) + \sup_{t\in[0,\infty)} |L_n(t)|/q(t)\} \\ &\quad \times\{ \sup_{t\in[d_n,\infty)} |\int_{d_n}^t q(s)dW_n(s)|/q(t)\} \\ &\leq 4c_{11} \sup_{t\in[d_n,\infty)} |W_n(t) - W_n(d_n)|.\end{aligned}$$

Thus, applying Inequality 4 yields

$$P_n(\Delta_{n3} > 4c_{11}x\{ \sup_{t\in[d_n,\infty)} \mathcal{E}_n\{W_n(t) - W_n(d_n)\}^2\}^{1/2}) \leq c_8 \exp\{-c_9x^2\}. \tag{3.36}$$

Similarly, combining integration by parts with inequality (3.32) produces

$$P_n(\Delta_{n5} > 4c_{11}xn^{-1/6}) \leq c_8 \exp\{-c_{27}x^2\}. \tag{3.37}$$

Furthermore, we have with P_n-probability 1

$$\begin{aligned}\Delta_{n6} &\leq \{ \sup_{t\in[0,\infty)} |\int_0^t q(s)dJ_n(s)|/q(t)\} \\ &\quad \times\{\{ \sum_{i=0}^{m(n)-1} |\{L_n(x_{i+1,n}) - L(x_{i+1,n};\theta_n)\}/q(x_{i+1,n}) \\ &\qquad -\{L_n(x_{i,n}) - L(x_{i,n};\theta_n)\}/q(x_{i,n})|\} \\ &\quad + \sup_{t\in[0,\infty)} |L_n(t) - L(t;\theta_n)|/q(t)\} \\ &\leq (4n^{1/3} + 6)\{ \sup_{t\in[0,\infty)} |J_n(t)|\} \\ &\quad \times\{ \sup_{t\in[0,\infty)} |L_n(t) - L(t;\theta_n)|/q(t)\}.\end{aligned} \tag{3.38}$$

A general investigation of the behavior of Δ_{n4} is rather useless, since it depends too much on the actual situation. Thus, we content ourselves with (3.34)-(3.38).

Next we turn to the aforementioned specific choices of $\{d_n\}_{n=1}^{\infty}$. For both choices the inequalities just derived are sharpened, supplemented by an inequality for Δ_{n4}, and combined.

To prove (3.11), choose $\beta < (1/2-\alpha)\wedge 1/6$, and d_n so as to satisfy $H(d_n;\theta_n) = 1 - n^{-(\gamma+1)}$, where $\gamma = (1/2-\alpha-\beta)/\alpha$. Note that $\gamma > 0$. Since

$$\Lambda(d_n;\theta_0) < (1-H(d_n;\theta_n))^{-\alpha}\int_0^{d_n}(1-H(s;\theta_n))^{\alpha}d\Lambda(s;\theta_0)$$
$$\leq c_\alpha n^{1/2-\beta}, \tag{3.39}$$

it follows from (3.34) that

$$P_n(\Delta_{n1} > 2c_{11}(1+3c_\alpha)(1+c_2)n^{-\beta}\log n) \leq c_3 n^{-c_4}. \tag{3.40}$$

By (3.35) we immediately have

$$P_n(\Delta_{n2} \neq 0) \leq n^{-\gamma}. \tag{3.41}$$

From (3.27) we may infer for $t \geq d_n$

$$\mathcal{E}_n\{W_n(t) - W_n(d_n)\}^2 \leq c_{26}c_{10}(1-H(d_n;\theta_n))^{1-2\alpha},$$

and hence (3.36) implies

$$P_n(\Delta_{n3} > 4c_{11}\{c_{26}c_{10}n^{-(\gamma+2\beta)}\log n\}^{1/2}) \leq c_8 n^{-c_9}. \tag{3.42}$$

Integration by parts yields with P_n-probability 1

$$\Delta_{n4} \leq 2c_{11}n^{1/2}\sup_{t\in[d_n,\infty)}|D(t;\theta_0,\theta_n) - D(d_n;\theta_0,\theta_n)|,$$

which leads in combination with (3.9) to

$$P_n(\Delta_{n4} > 2c_{11}c_{10}n^{-(\beta+\gamma)}) = 0. \tag{3.43}$$

From (3.37) we obtain

$$P_n(\Delta_{n5} > 4c_{11}n^{-1/6}(\log n)^{1/2}) \leq c_8 n^{-c_{27}}. \tag{3.44}$$

Combining Condition 2c with equations (3.31) and (3.38), it follows that

$$P_n(\Delta_{n6} > 10c_{12}n^{-\beta}(\log n)^{1/2}) \leq (c_{13}+c_8)n^{-(c_{14}\wedge c_{27})} \tag{3.45}$$

[note that we have used $\beta + 1/3 < 1/2$]. Now, (3.40)–(3.45) together with (3.33) yield (3.11).

If $\theta_n = \theta_0$, then we may obtain a sharper result by making a different choice of d_n. Let $x > 0$, and choose d_n so as to satisfy $H(d_n;\theta_0) = 1 - \exp\{-x\}/n$. By noting that $\Lambda(d_n;\theta_0) \le -\log(1 - H(d_n;\theta_0)) = \log n + x$, we obtain from (3.34) for $n > 1$

$$P_0(\Delta_{n1} > 2c_{11}n^{-1/2}(5\log n + 3x)(c_2 \log n + x))$$
$$\le c_3 \exp\{-c_4 x\}. \tag{3.46}$$

Furthermore, we have by (3.35)

$$P_0(\Delta_{n2} \ne 0) \le \exp\{-x\}, \tag{3.47}$$

and since by (3.27)

$$\begin{aligned} x \sup_{t\in[d_n,\infty)} \mathcal{E}_0\{W_n(t) - W_n(d_n)\}^2 & \\ &\le x(H^1(\infty;\theta_0) - H^1(d_n;\theta_0)) \\ &\le x(1 - H(d_n;\theta_0)) \\ &\le n^{-1}, \end{aligned}$$

it follows by (3.36)

$$P_0(\Delta_{n3} > 4c_{11}n^{-1/2}) \le c_8 \exp\{-c_9 x\}. \tag{3.48}$$

As a simple consequence of the fact that $D(t;\theta_0,\theta_0)$ is identical to zero, we have

$$P_0(\Delta_{n4} \ne 0) = 0. \tag{3.49}$$

Equation (3.37) implies

$$P_0(\Delta_{n5} > 4c_{11}n^{-1/6}x) \le 2c_8 \exp\{-c_{27}x\}, \tag{3.50}$$

whereas combining Condition 2d with equations (3.31) and (3.38) yields

$$P_0(\Delta_{n6} > 10n^{-1/6}x^{1/2}(c_{15}\log n + x)) \le (c_{16} + c_8)\exp\{-(c_{17} \wedge c_{27})x\}. \tag{3.51}$$

Hence, (3.46)–(3.51) together with (3.33) yield (3.12).

We conclude the proof of Theorem 2 with the remark that if $L_n(t)$ coincides with $L(t;\theta_n)$ with P_n-probability 1, it suffices to bound the left hand side of (3.33) by $\sum_{i=1}^{4} \Delta_{ni}$. This yields the refinement of Theorem 2 mentioned at the end of section 2. □

Proof of Theorem 3 Equation (3.14) directly follows from Corollary 1. To prove (3.15), observe that as a consequence of Theorem 2 there exists a process $X_n(t;\theta_0)$, equal in P_0-distribution to $X(t;\theta_0)$, which satisfies

$$\begin{aligned} &P_0\Big(\sup_{t\in[0,\infty)} |Q_n(t) - X_n(t;\theta_0)|/q(t) > n^{-c_{28}}(c_{21}\log n + x)^{c_{29}}\Big) \\ &\leq c_{22}\exp\{-c_{23}x\}, \end{aligned} \tag{3.52}$$

with $c_{28} = 1/6$ and $c_{29} = 2$. Now, let $c_{30} = c_{28}/(2c_{29}-1)$ and choose $c_{29}^{-1} < \beta < 2$. Note that for $n \to \infty$

$$n^{c_{30}(2-\beta)}(s_n)^\beta \gg n^{c_{30}(2-\beta)} \gg \log n,$$
$$n^{c_{30}(2-\beta)}(s_n)^\beta = (n^{-c_{30}}s_n)^{\beta-2}(s_n)^2 \gg (s_n)^2,$$
$$(n^{-c_{30}}s_n)^{c_{29}\beta-1} \ll 1$$

[here $a_n \gg b_n$ denotes $b_n/a_n \to 0$], and hence (3.52) implies

$$\begin{aligned} &P_0(|T(Q_n(\cdot)/q(\cdot)) - T(X_n(\cdot;\theta_0)/q(\cdot))| > c_T n^{c_{30}(1-c_{29}\beta)}(s_n)^{c_{29}\beta}) \\ &\leq P_0\Big(\sup_{t\in[0,\infty)} |Q_n(t) - X_n(t;\theta_0)|/q(t) > n^{c_{30}(1-c_{29}\beta)}(s_n)^{c_{29}\beta}\Big) \\ &\leq c_{22}\exp\{-c_{23}(n^{c_{30}(2-\beta)}(s_n)^\beta - c_{21}\log n)\} \\ &\ll \exp\{-a(s_n)^2/2\}. \end{aligned} \tag{3.53}$$

Since $T(X_n(\cdot;\theta_0)/q(\cdot))$ and $T(X(\cdot;\theta_0)/q(\cdot))$ are equal in P_0-distribution, we may bound $P_0(T(Q_n(\cdot)/q(\cdot)) > s_n)$ from below by

$$\begin{aligned} &P_0(T(X(\cdot;\theta_0)/q(\cdot)) > s_n(1 + c_T(n^{-c_{30}}s_n)^{c_{29}\beta-1})) \\ &-P_0(T(X_n(\cdot;\theta_0)/q(\cdot)) - T(Q_n(\cdot)/q(\cdot))| > c_T n^{c_{30}(1-c_{29}\beta)}(s_n)^{c_{29}\beta}) \end{aligned}$$

and from above by

$$\begin{aligned} &P_0(T(X(\cdot;\theta_0)/q(\cdot)) > s_n(1 - c_T(n^{-c_{30}}s_n)^{c_{29}\beta-1})) \\ &+P_0(T(X_n(\cdot;\theta_0)/q(\cdot)) - T(Q_n(\cdot)/q(\cdot))| > c_T n^{c_{30}(1-c_{29}\beta)}(s_n)^{c_{29}\beta}) \end{aligned}$$

and thus it follows that (3.14) and (3.53) together yield (3.15). This concludes the proof of Theorem 3. □

The proofs of Theorem 4 and Theorem 6 make repeated use of the following lemma, which holds under Conditions 1 and 6.

Lemma 2 *Let $g(t;\theta)$ be a real valued function, $g_i^{(1)}(t;\theta)$ the first order partial derivative of $g(t;\theta)$ with respect to the i^{th} component of θ. Suppose there exists a constant c_{31} such that*

$$\sup_{t\in[0,\infty)}(1-H(t;\theta))^{\frac{1}{2}}|g(t;\theta)|<c_{31}, \tag{3.54}$$

$$\sup_{t\in[0,\infty)}(1-H(t;\theta))^{1-\alpha}|g_i^{(1)}(t;\theta)|<c_{31} \tag{3.55}$$

for every $\theta\in\Theta$. Then there exists a constant c_{32} such that

$$\sup_{t\in[0,\infty)}|\int_0^t g(s;\theta)dD(s;\theta_0,\theta)|\leq c_{32}|\theta-\theta_0|, \tag{3.56}$$

$$\sup_{t\in[0,\infty)}|\int_0^t g(s;\theta)dH^1(s;\theta)-\int_0^t g(s;\theta_0)dH^1(s;\theta_0)|\leq c_{32}|\theta-\theta_0| \tag{3.57}$$

for every $\theta\in\Theta$. Let $g_{ij}^{(2)}(t;\theta)$ the second order partial derivative of $g(t;\theta)$ with respect to the i^{th} and j^{th} components of θ. If

$$\sup_{t\in[0,\infty)}(1-H(t;\theta))^{\alpha}|g(t;\theta)|<c_{31}, \tag{3.58}$$

$$\sup_{t\in[0,\infty)}(1-H(t;\theta))^{\frac{1}{2}}|g_i^{(1)}(t;\theta)|<c_{31}, \tag{3.59}$$

$$\sup_{t\in[0,\infty)}(1-H(t;\theta))^{1-\alpha}|g_{ij}^{(2)}(t;\theta)|<c_{31} \tag{3.60}$$

then

$$\begin{aligned}\sup_{t\in[0,\infty)}|&\int_0^t g(s;\theta)dD(s;\theta_0,\theta)\\ &-\int_0^t(\theta-\theta_0)^T\psi(s;\theta_0)g(s;\theta_0)dH^1(s;\theta_0)|\\ \leq{}& c_{32}|\theta-\theta_0|^2\end{aligned} \tag{3.61}$$

for every $\theta\in\Theta$.

Corollary 3 *There exists a constant c_{32} such that*

$$\sup_{t\in[0,\infty)}|\int_0^t(1-H(s;\theta))^{-\frac{1}{2}}dD(s;\theta_0,\theta)|\leq c_{32}|\theta-\theta_0|.$$

for every $\theta\in\Theta$.

Proof of Lemma 2 For convenience, we introduce the functions

$$\Lambda_0(t;\theta) = \Lambda(t;\theta) - \Lambda(t;\theta_0),$$

$$\lambda_0(t;\theta) = \lambda(t;\theta) - \lambda(t;\theta_0).$$

Let $\Lambda_i^{(1)}(t;\theta)$ and $\lambda_i^{(1)}(t;\theta)$ respectively denote the first order partial derivatives of $\Lambda_0(t;\theta)$ and $\lambda_0(t;\theta)$ with respect to the i^{th} component of θ. Furthermore, let $\Lambda_{ij}^{(2)}(t;\theta)$ and $\lambda_{ij}^{(2)}(t;\theta)$ respectively denote the second order partial derivatives of $\Lambda_0(t;\theta)$ and $\lambda_0(t;\theta)$ with respect to the i^{th} and j^{th} components of θ. By Condition 6 we have for every $t \in [0,\infty)$

$$\begin{aligned}
&\int_0^t (1 - H(s;\theta))^{\frac{1}{2}-\alpha} |\lambda_i^{(1)}(s;\theta)| ds \\
&= \int_0^t (1 - H(s;\theta))^{\frac{1}{2}-\alpha} |\psi_i(s;\theta)| d\Lambda(s;\theta) \\
&\leq \{\int_0^\infty (1 - H(s;\theta))^\beta d\Lambda(s;\theta)\}^{1/2} \\
&\qquad \times \{\int_0^\infty (\psi_i(s;\theta))^2 (1 - H(s;\theta))^{1-2\alpha-\beta} d\Lambda(s;\theta)\}^{1/2} \\
&\leq (c_{24}/\beta)^{1/2},
\end{aligned} \tag{3.62}$$

and hence

$$\begin{aligned}
&(1 - H(t;\theta))^{\frac{1}{2}-\alpha} |\Lambda_i^{(1)}(t;\theta)| \\
&= (1 - H(t;\theta))^{\frac{1}{2}-\alpha} | \int_0^t \lambda_i^{(1)}(s;\theta) ds | \\
&\leq \int_0^t (1 - H(s;\theta))^{\frac{1}{2}-\alpha} |\lambda_i^{(1)}(s;\theta)| ds \\
&\leq (c_{24}/\beta)^{1/2}.
\end{aligned} \tag{3.63}$$

Now, we find by using the identity

$$1 - H(t;\theta) = (1 - H(t;\theta_0)) \exp\{-\Lambda_0(t;\theta)\}$$

that the first order partial derivative of $\int_0^t g(s;\theta)\lambda_0(s;\theta)(1 - H(s;\theta))ds$ with respect to the i^{th} component of θ is given by

$$\int_0^t \{\{g_i^{(1)}(s;\theta) - g(s;\theta)\Lambda_i^{(1)}(s;\theta)\}\lambda_0(s;\theta) + g(s;\theta)\lambda_i^{(1)}(s;\theta)\}(1 - H(s;\theta))ds,$$

and is bounded in t and θ, as follows from Condition 1, (3.54), (3.55), (3.62) and (3.63). By expressing $D(t;\theta_0,\theta)$ as $\int_0^t (1 - H(s;\theta))\lambda_0(s;\theta)ds$ we obtain (3.56)

As for (3.57), by writing $H^1(t;\theta) = \int_0^t (1 - H(s;\theta))\lambda(s;\theta)ds$, it follows that the first order partial derivative of $\int_0^t g(s;\theta)dH^1(s;\theta)$ with respect to the i^{th} component of θ is given by

$$\int_0^t \{g_i^{(1)}(s;\theta) + g(s;\theta)\{\psi_i(s;\theta) - \Lambda_i^{(1)}(s;\theta)\}\}(1 - H(s;\theta))\lambda(s;\theta)ds$$

and is bounded in t and θ.

Finally (3.61). For every $t \in [0,\infty)$ we have

$$\begin{aligned}
&\int_0^t (1 - H(s;\theta))^{(1-2\alpha)}|\lambda_{ij}^{(2)}(s;\theta)|ds \\
&= \int_0^t (1 - H(s;\theta))^{(1-2\alpha)}|\psi_{ij}^{(1)}(s;\theta) + \psi_i(s;\theta)\psi_j(s;\theta)|d\Lambda(s;\theta) \\
&\leq \{\int_0^\infty (1 - H(s;\theta))^\beta d\Lambda(s;\theta)\}^{1/2} \\
&\qquad \times\{\int_0^\infty (\psi_{ij}^{(1)}(s;\theta))^2(1 - H(s;\theta))^{2(1-2\alpha)-\beta}d\Lambda(s;\theta)\}^{1/2} \\
&\quad +\{\int_0^\infty (\psi_i(s;\theta))^2(1 - H(s;\theta))^{1-2\alpha-\beta}d\Lambda(s;\theta)\}^{1/2} \\
&\qquad \times\{\int_0^\infty (\psi_j(s;\theta))^2(1 - H(s;\theta))^{1-2\alpha-\beta}d\Lambda(s;\theta)\}^{1/2} \\
&\leq (c_{24}/\beta)^{1/2} + c_{24},
\end{aligned} \tag{3.64}$$

and

$$(1 - H(t;\theta))^{(1-2\alpha)}|\Lambda_{ij}^{(2)}(t;\theta)| \leq (c_{24}/\beta)^{1/2} + c_{24}. \tag{3.65}$$

Note that the first order partial derivative of $\int_0^t g(s;\theta)\lambda_0(s;\theta)(1 - H(s;\theta))ds$ with respect to the i^{th} component of θ equals $\int_0^t g(s;\theta_0)\psi_i(s;\theta_0)dH^1(s;\theta_0)$ when evaluated at $\theta = \theta_0$. Furthermore, the second order partial derivative with respect to the i^{th} and j^{th} component of θ is given by

$$\begin{aligned}
\int_0^t \{&\{g_{ij}^{(2)}(s;\theta) - g_i^{(1)}(s;\theta)\Lambda_j^{(1)}(s;\theta) - g_j^{(1)}(s;\theta)\Lambda_i^{(1)}(s;\theta) \\
&-g(s;\theta)\Lambda_{ij}^{(2)}(s;\theta) + g(s;\theta)\Lambda_i^{(1)}(s;\theta)\Lambda_j^{(1)}(s;\theta)\}\lambda_0(s;\theta) \\
&+\{g_j^{(1)}(s;\theta) - g(s;\theta)\Lambda_j^{(1)}(s;\theta)\}\lambda_i^{(1)}(s;\theta) \\
&+\{g_i^{(1)}(s;\theta) - g(s;\theta)\Lambda_i^{(1)}(s;\theta)\}\lambda_j^{(1)}(s;\theta) \\
&+g(s;\theta)\lambda_{ij}^{(2)}(s;\theta)\}(1 - H(s;\theta))ds.
\end{aligned}$$

Hence, (3.61) follows from Condition 1, Condition 6, (3.55), (3.58), (3.60) and (3.62)-(3.65). This concludes the proof of Lemma 2. □

Proof of Theorem 4 Let $\{W_n(t)\}_{n=1}^{\infty}$ be the sequence of mean zero Gaussian processes given in Theorem 2. We may write

$$\sup_{t\in[0,\infty)} |Q_n(t) - n^{1/2}|\theta_n - \theta_0|h^T K_a(t)|/q(t) \le \sum_{i=1}^{4} \Delta_{ni}, \tag{3.66}$$

where

$$\Delta_{n1} = \sup_{t\in[0,\infty)} |\{Q_n(t) - n^{1/2}\int_0^t L_n(s)dD(s;\theta_0,\theta_n)\} - \int_0^t L(s;\theta_n)dW_n(s)|/q(t),$$

$$\Delta_{n2} = \sup_{t\in[0,\infty)} |\int_0^t L(s;\theta_n)dW_n(s)|/q(t),$$

$$\Delta_{n3} = n^{1/2}\sup_{t\in[0,\infty)} |\int_0^t (L_n(s) - L(s;\theta_n))dD(s;\theta_0,\theta_n)|/q(t),$$

$$\Delta_{n4} = n^{1/2}\sup_{t\in[0,\infty)} |\int_0^t L(s;\theta_n)dD(s;\theta_0,\theta_n) - |\theta_n - \theta_0|h^T K_a(t)|/q(t).$$

By (3.11) we have for $\beta < (1/2-\alpha)\wedge 1/6$

$$P_n(\Delta_{n1} > c_{18}n^{-\beta}) \le c_{19}n^{-c_{20}}. \tag{3.67}$$

From integration by parts we obtain with P_n-probability 1

$$\begin{aligned}\Delta_{n2} &\le 2c_{11}\sup_{t\in[0,\infty)} |\int_0^t q(s)/q(t)dW_n(s)| \\ &\le 2c_{11}\sup_{t\in[0,\infty)} 2\sup_{0\le s\le t}|W_n(s)| \\ &\le 4c_{11}\sup_{t\in[0,\infty)}|W_n(t)|.\end{aligned}$$

Since by (3.28) and Condition 1 the variance function of the process $W_n(t)$ is bounded by $(c_{26})^2$, it follows by Inequality 4

$$P_n(\Delta_{n2} > 4c_{11}c_{26}x) \le c_8\exp\{-c_9x^2\}. \tag{3.68}$$

Lemma 2 yields the existence of a constant c_{33} such that

$$\Delta_{n4} \le c_{33}n^{1/2}\{|\theta_n - \theta_0|^2 + |(\theta_n - \theta_0) - |\theta_n - \theta_0|h|\} \tag{3.69}$$

and for every $\theta\in\Theta$ and $\beta < 1/2$

$$\begin{aligned}&P_n(n^\beta \sup_{t\in[0,\infty)}\int_0^t (L_n(s) - L(s;\theta_n))dD(s;\theta_0,\theta)/q(t) > c_{33}|\theta-\theta_0|) \\ &\le P_n(n^\beta \sup_{t\in[0,\infty)}\sup_{0\le s\le t}|L_n(s) - L(s;\theta_n)|/q(t) > c_{12}).\end{aligned}$$

Note that by Condition 2c this last inequality implies

$$P_n((n^{1/2}|\theta_n - \theta_0|)^{-1}\Delta_{n3} > c_{33}n^{-\beta}) \leq c_{13}n^{-c_{14}}. \tag{3.70}$$

The first part of the theorem is now easily proved by means of the inequality

$$|(n^{1/2}|\theta_n - \theta_0|)^{-1}T(Q_n(\cdot)/q(\cdot)) - T(h^T K_a(\cdot)/q(\cdot))|$$

$$\leq c_T(n^{1/2}|\theta_n - \theta_0|)^{-1}\sum_{i=1}^{4}\Delta_{ni}, \tag{3.71}$$

To prove the second part of the theorem, we first note that the supremum over t of the distance between the processes $Q_n(t)/q(t)$ and $\{\int_0^t L(s;\theta_n)dW_n(s) + n^{1/2}|\theta_n - \theta_0|h^T K_a(t)\}/q(t)$ is bounded by $\Delta_{n1} + \Delta_{n3} + \Delta_{n4}$, and hence converges to zero in P_n-probability if $n^{1/2}|\theta_n - \theta_0|$ tends to a finite limit as $n \to \infty$. Thus, it remains to show that the latter process converges in P_n-distribution to $\{X(t;\theta_0) + \sigma h^T K_a(t)\}/q(t)$, which boils down to the convergence in P_n-distribution of $\int_0^t L(s;\theta_n)dW_n(s)/q(t)$ to $X(t;\theta_0)/q(t)$.

Let $B_n(t)$ be as defined in the proof of Theorem 2. Split $\int_0^t L(s;\theta_n)dW_n(s)/q(t)$ into the two parts

$$\int_0^t L(s;\theta_n)\{dW_n(s) - \frac{B_n(s)}{1 - H(s;\theta_n)}dD(s;\theta_0,\theta_n)\}/q(t)$$

and

$$\int_0^t L(s;\theta_n)\frac{B_n(s)}{1 - H(s;\theta_n)}dD(s;\theta_0,\theta_n)/q(t).$$

The first part may be interpreted as a time-transformed Wiener process divided by $q(t)$, the time transformed Wiener process having variance function $\int_0^t (L(s;\theta_n))^2 dH^1(s;\theta_n)/(q(t))^2$. It follows from our Lemma 2 and Theorem VI.10 in Pollard (1984) that the first part converges in P_n-distribution to $X(t;\theta_0)/q(t)$. Use Lemma 2 to show that the supremum over t of the second part converges to zero in P_n-probability. This completes the proof of Theorem 4. □

Proof of Theorem 6 Let Δ_{n1}, Δ_{n2}, Δ_{n3} and Δ_{n4} be as in the proof of Theorem 4, and define $b(\theta_n)$ as $T(\int_0^{\cdot} L(s;\theta_n)dD(s;\theta_0,\theta_n)/q(\cdot))$. We may write

$$|n^{-1/2}T(Q_n(\cdot)/q(\cdot)) - b(\theta_n)| \leq c_T n^{-1/2}\{\Delta_{n1} + \Delta_{n2} + \Delta_{n3}\}. \tag{3.72}$$

Furthermore, we have

$$|b(\theta_n) - |\theta_n - \theta_0|T(h^T K_a(\cdot)/q(\cdot))| \leq c_T n^{-1/2}\Delta_{n4}. \tag{3.73}$$

Since $e(h)$ is not equal to zero, this yields the existence of positive constants ϵ^* and c_{34} such that for θ_n satisfying $|\theta_n - \theta_0| < \epsilon^*$

$$c_{34}|\theta_n - \theta_0| < b(\theta_n) < 1. \tag{3.74}$$

Now suppose $\theta \in \Theta - \{\theta_0\}$ satisfies $|\theta - \theta_0| < \epsilon^*$, and set θ_n equal to θ. Choose $\epsilon > 0$ and $\delta \in (0,1)$. By (3.67) and (3.68) it follows that there exists an integer N_1 not depending on θ such that for $n > N_1$

$$P_\theta(\Delta_{n1} > (N_1)^{1/2}\epsilon/4c_T) < \delta/4,$$
$$P_\theta(\Delta_{n2} > (N_1)^{1/2}\epsilon/4c_T) < \delta/4.$$

Hence, for $n > N_1/(b(\theta))^2$ we have

$$P_\theta(n^{-1/2}\{\Delta_{n1} + \Delta_{n2}\} > \epsilon b(\theta)/2c_T) < \delta/2, \tag{3.75}$$

since $n > N_1$ and $(N_1/n)^{1/2} < b(\theta)$. Moreover, (3.70) implies the existence of an integer $N > N_1$ not depending on θ such that for $n > N$

$$\begin{aligned} &P_\theta(\Delta_{n3} > \epsilon b(\theta)/2c_T) \\ &\leq P_\theta((n^{1/2}|\theta - \theta_0|)^{-1}\Delta_{n3} > c_{34}\epsilon) \\ &< \delta/2. \end{aligned} \tag{3.76}$$

Combining (3.72)-(3.76) now yields that $\{T(Q_n(\cdot)/q(\cdot))\}_{n=1}^\infty$ is indeed a Wieand sequence.

Finally, (3.20) immediately follows from (3.69) and (3.73). This concludes the proof of Theorem 6. □

Proof of Lemma 1 The log-likelihood of the observation (Z_1, δ_1) under P_θ is

$$\ell(\theta; Z_1, \delta_1) = \delta_1 \log \lambda(Z_1;\theta) - \Lambda(Z_1;\theta)$$

[see Borgan (1984)]. This yields that

$$\frac{\partial}{\partial\theta_i}\ell(\theta; Z_1, \delta_1) = \delta_1\psi_i(Z_1;\theta) - \int_0^{Z_1} \psi_i(s;\theta)d\Lambda(s;\theta)$$

is the score for the i^{th} component of θ. Observe the close relation between $\psi_i(t;\theta)$ and the score function. The information matrix $I(\theta)$ has elements

$$\begin{aligned} I_{ij}(\theta) &= \mathcal{E}_\theta\{\frac{\partial}{\partial\theta_i}\ell(\theta; Z_1,\delta_1)\}\{\frac{\partial}{\partial\theta_j}\ell(\theta; Z_1,\delta_1)\} \\ &= \int_0^\infty \psi_i(t;\theta)\psi_j(t;\theta)dH^1(t;\theta) \\ &\quad - \int_0^\infty \psi_i(t;\theta)\int_0^t \psi_j(s;\theta)d\Lambda(s;\theta)dH^1(t;\theta) \\ &\quad - \int_0^\infty \psi_j(t;\theta)\int_0^t \psi_i(s;\theta)d\Lambda(s;\theta)dH^1(t;\theta) \\ &\quad + \int_0^\infty \int_0^t \psi_i(s;\theta)d\Lambda(s;\theta)\int_0^t \psi_j(u;\theta)d\Lambda(u;\theta)dH(t;\theta) \\ &= \int_0^\infty \psi_i(t;\theta)\psi_j(t;\theta)dH^1(t;\theta). \end{aligned}$$

The last line follows from

$$
\begin{aligned}
&\int_0^\infty \psi_i(t;\theta) \int_0^t \psi_j(s;\theta) d\Lambda(s;\theta) dH^1(t;\theta) \\
&\qquad + \int_0^\infty \psi_j(t;\theta) \int_0^t \psi_i(s;\theta) d\Lambda(s;\theta) dH^1(t;\theta) \\
&= \int_0^\infty \{\psi_i(t;\theta) \int_0^t \psi_j(s;\theta) d\Lambda(s;\theta) \\
&\qquad + \psi_j(t;\theta) \int_0^t \psi_i(s;\theta) d\Lambda(s;\theta)\} \int_t^\infty dH(u;\theta) d\Lambda(t;\theta) \\
&= \int_0^\infty \int_0^u \{\psi_i(t;\theta) \int_0^t \psi_j(s;\theta) d\Lambda(s;\theta) \\
&\qquad + \psi_j(t;\theta) \int_0^t \psi_i(s;\theta) d\Lambda(s;\theta)\} d\Lambda(t;\theta) dH(u;\theta) \\
&= \int_0^\infty \int_0^u \psi_i(t;\theta) d\Lambda(t;\theta) \int_0^u \psi_j(s;\theta) d\Lambda(s;\theta) dH(u;\theta).
\end{aligned}
$$

This ends the proof of Lemma 1. □

Chapter 4

The composite null hypothesis

4.1 Introduction

In this chapter the results for the simple null hypothesis are extended to the composite null hypothesis, which involves estimation as well as testing. Though the lines of the previous chapter are followed rather closely, this chapter is made as self-contained as possible. As a consequence some of the remarks may seem familiar.

Again, $(\Omega, \mathcal{A}, \mathcal{P})$ is a probability space, and $\mathcal{F}_n \subset \mathcal{B}[0,\infty)$. At stage n each of the independent random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ maps $(\Omega, \mathcal{A})$ into $([0,\infty), \mathcal{F}_n)$. The probability measure induced by these random variables is denoted by P_n. Each pair (X_i, Y_i) is assumed to have the same distribution. The censoring distribution, the distribution of the censoring time Y_i, does not depend on n. Hence, there exists a cumulative distribution G such that $G(t) = P_n(Y_i \leq t)$ for each n. Defectiveness of G is allowed. The failure time distribution, the distribution of X_i, does depend on n.

However, the structure of the dependence of the failure time distribution on the sample size n is different. Now we suppose the existence of a cumulative distribution function F, indexed by (υ, θ) belonging to some set $\Upsilon \times \Theta$, and a sequence of points $\{(\upsilon_n, \theta_n)\}_{n=1}^{\infty}$ such that $F(t; \upsilon_n, \theta_n) = P_n(X_i \leq t)$.

The set Υ refers to the nuisance parameter, the parameter which value is not specified in the null hypothesis. Any information concerning these parameters must be supplied by the sample.

The set Θ refers to the parameter of interest, the parameter which value is specified in the null hypothesis. Denote the value of the parameter of interest under the null hypothesis by θ_0. Thus, at every stage n we want to test the null hypothesis that θ_n equals θ_0.

Let υ_0 be an *arbitrary* element of Υ. If (υ_n, θ_n) equals (υ_0, θ_0) [denote the probability measure corresponding to this situation by P_{υ_0}, the expectation operator belonging to this probability measure by $\mathcal{E}_{\upsilon_0}$, and the situation itself by "under P_{v_0}"] then the basic martingale takes the form

$$M_n(t; v_0, \theta_0) = n^{1/2}\{H_n^1(t) - \int_0^t (1 - H_{n-}(s))d\Lambda(s; v_0, \theta_0)\}, \tag{4.1}$$

where $\Lambda(t; v, \theta) = -\log(1 - F(t; v, \theta))$ is the cumulative hazard function belonging to $F(t; v, \theta)$. If (v_n, θ_n) is arbitrary [refer to this situation as "under P_n" and denote the expectation operator by $\mathcal{E}_n$] we also call $M_n(t; v_0, \theta_0)$ a basic martingale, despite the fact that the martingale property is in general lacking.

In section 4.3 we study tests based on a sublinear Lipschitz functional of a process $Q_n(t; v^{(n)})$, where $v^{(n)}$ is a random element of Υ and

$$Q_n(t; v) = \int_0^t L_n(s; v)dM_n(s; v, \theta_0) \tag{4.2}$$

is a stochastic integral with respect to the basic martingale. Here the weight process $L_n(t; v)$ is a stochastic process.

We restrict ourselves to the case where Υ is a convex subset of $\mathbb{R}^r$. In section 4.2 the probability theory underlying the results in section 4.3 is presented. Proofs are gathered in section 4.4.

4.2 Probability inequalities

In this section we present probability inequalities which concern the approximation on the halfline $[0, \infty)$ of [a centered version of] the process $Q_n(t; v^{(n)})$ by a one-parameter Gaussian process, both under P_n as under P_{v_0}. For treatment of the former situation, the following condition, which relates the right tail behavior of $F(t; v, \theta)$ to the right tail behavior of $F(t; v_0, \theta_0)$, is needed [compare with Condition 1].

Condition 8 *There exists constants* $0 < \alpha < 1/2$ *and* $c_\alpha < \infty$ *such that*

$$\int_0^\infty (1 - F(s; v, \theta))^\alpha d\Lambda(s; v_0, \theta_0) < c_\alpha$$

for every $v, v_0 \in \Upsilon$ *and* $\theta \in \Theta$.

4.2.1 The M-estimator $v^{(n)}$

Of course, before studying the process $Q_n(t; v^{(n)})$ we must know the properties of the random variable $v^{(n)}$. It is desirable that, at least under the null hypothesis, $v^{(n)}$ is an estimator for v_0. Hence, we let $v^{(n)}$ belong to the class of M-estimators proposed by Hjort (1985). An M-estimator $v^{(n)}$ is obtained as the solution in Υ to the equations

$$\Phi_{ni}(\infty; v^{(n)}) = 0 \qquad i = 1, \ldots, r, \tag{4.3}$$

where $\Phi_{ni}(t;v)$ is the stochastic integral

$$\Phi_{ni}(t;v) = \int_0^t \phi_i(s;v)dM_n(s;v,\theta_0), \tag{4.4}$$

and $\phi_i(t;v)$ is a given function.

The most prominent example of an M-estimator is without any doubt the maximum likelihood estimator. Since under P_{v_0} the log-likelihood of the sample $(Z_1,\delta_1),\ldots,(Z_n,\delta_n)$ is given by

$$n\int_0^\infty \log\lambda(s;v_0,\theta_0)dH_n^1(s) - n\int_0^\infty (1-H_{n-}(s))d\Lambda(s;v_0,\theta_0), \tag{4.5}$$

where $\lambda(t;v,\theta)$ is the hazard function, the derivative of the cumulative hazard function $\Lambda(t;v,\theta)$ with respect to t, the maximum likelihood estimator is obtained as M-estimator by setting $\phi_i(t;v)$ equal to $\psi_i(t;v,\theta_0)$, where $\psi_i(t;v,\theta)$ denotes the first order partial derivative of $\log\lambda(t;v,\theta)$ with respect to the i^{th} component of v [see Borgan (1984)].

The study of the behavior of $v^{(n)}$ involves the following regularity conditions on $\psi_i(t;v,\theta)$ and $\phi_i(t;v)$.

Condition 9 *There exists a constant c_ψ such that for every $i=1,\ldots,r$ and $v\in\Upsilon$ the function $\psi_i(t;v,\theta_0)$ exists, is bounded by c_ψ and has total variation bounded by c_ψ. The first order partial derivatives of $\psi_i(t;v,\theta_0)$ with respect to the components of v exist, are bounded by $(c_\psi)^2$, and have total variation bounded by $(c_\psi)^2$.*

Condition 10 *For every $i=1,\ldots,r$ and $v\in\Upsilon$ the function $\phi_i(t;v)$ is bounded by $(c_\psi)^{-2}$ and has total variation bounded by $(c_\psi)^{-2}$. The first order partial derivatives of $\phi_i(t;v)$ with respect to the components of v exist, are bounded by $(c_\psi)^{-1}$, and have total variation bounded by $(c_\psi)^{-1}$. The second order partial derivatives of $\phi_i(t;v)$ with respect to the components of v exist, are bounded by 1, and have total variation bounded by 1.*

Condition 10 is less restrictive than it might seem at first glance, since it is clear from (4.3) that multiplying $\phi_i(t;v)$ by a nonzero constant has no effect on the estimator itself. Observe that the constant may even depend on v.

However, Condition 10 does involve restrictions on the behavior of the second order partial derivatives, which may prevent the maximum likelihood estimator from falling into the framework considered here. If this is the case, one may weigh replacing the maximum likelihood estimator by a closely related M-estimator, obtained by setting $\phi_i(t;v)$ equal to a smoothed version of $\psi_i(t;v,\theta_0)$.

Before formulating two additional conditions, we first introduce the cumulative distribution functions

$$H^1(t;v_n,\theta_n) = P_n(X_1\le t,\delta_1=1),$$

$$H(t; \upsilon_n, \theta_n) = P_n(X_1 \leq t)$$

[note that $H^1(t; \upsilon_n, \theta_n)$ is possibly defective]. Moreover, we define the function $D(t; \upsilon', \theta_0, \upsilon, \theta)$ by

$$D(t; \upsilon', \theta_0, \upsilon, \theta) = \int_0^t (1 - H(s; \upsilon, \theta))\{d\Lambda(s; \upsilon, \theta) - d\Lambda(s; \upsilon', \theta_0)\}. \tag{4.6}$$

Condition 11 *There exists a function $\pi : \Upsilon \times \Theta \to \Upsilon$ such that*

$$\int_0^\infty \phi_i(s; \pi(\upsilon, \theta))dD(s; \pi(\upsilon, \theta), \theta_0, \upsilon, \theta) = 0$$

for $i = 1, \ldots, r$, $\upsilon \in \Upsilon$ and $\theta \in \Theta$.

The function $D(t; \pi(\upsilon, \theta), \theta_0, \upsilon, \theta)$ appears frequently in our results. To increase readability we shall abbreviate it by $D(t; \upsilon, \theta)$. It reflects the distance between the distribution $F(t; \upsilon, \theta)$ and the set of distributions under the null hypothesis.

It is possible to choose π so as to satisfy $\pi(\upsilon, \theta_0) = \upsilon$ for any $\upsilon \in \Upsilon$. We shall tacitly assume this is indeed the case. Hence, we may view $(\pi(\upsilon, \theta), \theta_0)$ as a projection of the point (υ, θ) on $\Upsilon \times \{\theta_0\}$. Moreover, it immediately follows that $D(t; \upsilon, \theta_0)$ equals zero for any $t \in [0, \infty)$ and $\upsilon \in \Upsilon$.

Let $\Xi_0(\upsilon, \theta)$ be the $r \times r$ matrix with elements

$$\begin{aligned} \Xi_{0ij}(\upsilon, \theta) = & \int_0^\infty \phi_i(s; \pi(\upsilon, \theta))\psi_j(s; \pi(\upsilon, \theta), \theta_0)dH^1(s; \upsilon, \theta) \\ & - \int_0^\infty \{\phi_i(s; \pi(\upsilon, \theta))\psi_j(s; \pi(\upsilon, \theta), \theta_0) \\ & \quad + \phi_{ij}^{(1)}(s; \pi(\upsilon, \theta), \theta_0)\}dD(s; \upsilon, \theta), \end{aligned}$$

where $\phi_{ij}^{(1)}(t; \upsilon)$ denotes the first order partial derivative of $\phi_i(t; \upsilon)$ with respect to the j^{th} component of υ. We shall assume that Ξ_0 satisfies Condition 12. Refer to the matrix AA^T as the square of the matrix A.

Condition 12 *There exist a constant c_{35} such that the eigenvalues of the square of the matrix $\Xi_0(\upsilon, \theta)$ all exceed $(c_{35})^2$ for every $\upsilon \in \Upsilon$ and $\theta \in \Theta$.*

The complexity of the last two conditions may seem discouraging. However, if we are only interested in what happens under the null hypothesis then we may set $\Theta = \{\theta_0\}$ and the situation clarifies considerably. Now Condition 11 holds automatically as follows from setting $\pi(\upsilon, \theta) = \upsilon$. Moreover, Condition 12 only involves $\Xi_0(\upsilon, \theta_0)$. Observe that

$$\Xi_{0ij}(\upsilon, \theta_0) = \int_0^\infty \phi_i(s; \upsilon)\psi_j(s; \upsilon, \theta_0)dH^1(s; \upsilon, \theta_0).$$

Hence, $\Xi_0(v, \theta_0)$ is of a much simpler nature than $\Xi_0(v, \theta)$.

To verify Condition 12 in the general case a two stage procedure could be succesful. First show that the eigenvalues of the square of $\Xi_0(v, \theta_0)$ are bounded from below, uniformly in v. Then try to show that the eigenvalues of the square of $\Xi_0(v, \theta) - \Xi_0(\pi(v, \theta), \theta_0)$ remain small enough to guarantee that Condition 12 holds. Note that the elements of $\Xi_0(v, \theta) - \Xi_0(\pi(v, \theta), \theta_0)$ are bounded by $4(c_\psi)^{-1} \sup_{t \in [0,\infty)} D(t; v, \theta)$. Thus this approach can be expected to work if $\sup_{t \in [0,\infty)} D(t; v, \theta)$ is small, which occurs in the neighborhood of θ_0.

Denote $\pi(v_n, \theta_n)$ by v_{n0}, $\Xi_0(v_n, \theta_n)$ by Ξ_{0n} and $\Xi_0(v_0, \theta_0)$ by Σ_{e0}, and let $\phi(t; v)$ be the r-dimensional vector with elements $\phi_i(t; v)$.

Theorem 7 *There exist positive constants c_{36}, c_{37} not depending on v_n or θ_n and a random variable S_n such that if*

$$S_n < n^{1/2}, \tag{4.7}$$

then there exists a solution $v^{(n)}$ to the equations (4.3) which satisfies

$$n^{1/2}|v^{(n)} - v_{n0}| < c_{36} S_n \tag{4.8}$$

$$\begin{aligned} n^{1/2}|\Xi_{0n}^{-1} \int_0^\infty \phi(s; v_{n0})\{dM_n(s; v_{n0}, \theta_0) - dD(s; v_n, \theta_n)\} \\ -n^{1/2}(v^{(n)} - v_{n0})| < c_{37}\{S_n\}^2 \end{aligned} \tag{4.9}$$

Furthermore, for every $\beta > \alpha$ there exist positive constants c_{38}, c_{39} not depending on v_n or θ_n, such that

$$P_n(S_n > c_{38} n^\beta (\log n)^{-1/2}) < c_{39} n^{-(\beta-\alpha)/\alpha}. \tag{4.10}$$

Moreover, if $(v_n, \theta_n) = (v_0, \theta_0)$, then there exist positive constants $c_{40} - c_{42}$ not depending on v_0 such that

$$P_{v_0}(S_n > x(c_{40} \log n + x^2)) \le c_{41} \exp\{-c_{42} x^2\} \tag{4.11}$$

for every $x > (n^{1/2} \log n)^{-1}$.

The strength of inequality (4.9) depends on the behavior of the random variable $n^{-1/2}\{S_n\}^2$. In the following corollary this behavior is characterized.

Corollary 4 *If $\alpha < 1/4$, then*

$$n^{-1/2}\{S_n\}^2 \to_{P_n} 0. \tag{4.12}$$

If $(v_n, \theta_n) = (v_0, \theta_0)$, then

$$P_{v_0}(n^{-1/2}\{S_n\}^2 > n^{-1/2} x(c_{40} \log n + x)^2) \le c_{41} \exp\{-c_{42} x\} \tag{4.13}$$

for every $x > (n^{1/2} \log n)^{-2}$.

Our approximation of the process $Q_n(t; \upsilon^{(n)})$ is based on (4.9) directly. However, (4.9) may also be used to prove asymptotic normality of the estimator $\upsilon^{(n)}$, since by the refinement of Theorem 2 we may approximate $\int_0^\infty \phi(s; \upsilon_{n0})$ $\{dM_n(s; \upsilon_{n0}, \theta_0) - dD(s; \upsilon_n, \theta_n)\}$ by the r-dimensional Gaussian random variable $\tilde{\Phi}_n$ defined by

$$\tilde{\Phi}_n = \int_0^\infty \phi(s; \upsilon_{n0}) dW_n(s; \upsilon_{n0}), \tag{4.14}$$

where $W_n(t; \upsilon_{n0})$ is the Gaussian process of Theorem 2. Under P_n the covariance function of $W_n(t; \upsilon_{n0})$ is somewhat intractable, and thus there is little use in giving an expression for the asymptotic variance of $n^{1/2}(\upsilon^{(n)} - \upsilon_{n0})$ under P_n. But again, everything becomes more transparent if (υ_n, θ_n) equals (υ_0, θ_0). Under P_{υ_0} the process $W_n(t; \upsilon_0)$ is a Wiener process with variance function $H^1(t; \upsilon_0, \theta_0)$, and hence the covariancematrix of $\tilde{\Phi}_n$ is equal to Σ_{ee}, the $r \times r$ matrix with elements

$$\Sigma_{eeij} = \int_0^\infty \phi_i(s; \upsilon_0) \phi_j(s; \upsilon_0) dH^1(s; \upsilon_0, \theta_0) \tag{4.15}$$

Hence, we arrive at the following corollary.

Corollary 5 *If $(\upsilon_n, \theta_n) = (\upsilon_0, \theta_0)$, then there exists a sequence $\{\tilde{\Phi}_n\}_{n=1}^\infty$ of r-dimensional Gaussian random vectors with zero expectation and covariance matrix Σ_{ee} such that*

$$|n^{1/2}(\upsilon^{(n)} - \upsilon_0) - \Sigma_{e0}^{-1} \tilde{\Phi}_n| = \mathcal{O}_{P_{\upsilon_0}}(n^{-1/2} (\log n)^2) \qquad \text{as } n \to \infty.$$

The convergence in P_{υ_0}-distribution of the sequence $\{n^{1/2}(\upsilon^{(n)} - \upsilon_0)\}_{n=1}^\infty$ to a Gaussian random variate with zero expectation and covariance matrix $\Sigma_{e0}^{-1} \Sigma_{ee} \Sigma_{0e}^{-1}$, where Σ_{0e} denotes the transpose of Σ_{e0}, was proved for the maximum likelihood estimator by Borgan (1984) using martingale methods. Observe that in this case $\Sigma_{e0}^{-1} \Sigma_{ee} \Sigma_{0e}^{-1}$ is equal to Σ_{ee}^{-1}. Hjort (1985) mentions that the approach of Borgan can be extended to M-estimators without difficulty. However, the results obtained in this way do not stretch as far as Corollary 4, which also conveys information about the rate of convergence.

4.2.2 The process $Q_n(t; \upsilon^{(n)})$

Now that we have gained some knowledge about the M-estimator $\upsilon^{(n)}$, it is time to direct our attention to the process $Q_n(t; \upsilon^{(n)})$. But first we make some assumptions about the weight process $L_n(t; \upsilon)$.

Condition 13 *The weight process $L_n(t; \upsilon)$ satisfies*

a *$L_n(\cdot; \upsilon)$ is a random element in [the left continuous right limits version of] $D[0, \infty]$ endowed with the $\mathcal{J}_1$ metric. There exists a constant c_{43} not depending on υ, υ_n or θ_n such that $L_n(t; \upsilon)$ is bounded by c_{43} and has total variation bounded by c_{43} with P_n-probability 1.*

b *There exists a deterministic function $L(t;v,\theta)$ such that for every $\beta < 1/2$ there exist positive constants c_{44}–c_{46} not depending on v_n or θ_n such that*

$$P_n(\sup_{t\in[0,\infty)} |L_n(t;v_{n0}) - L(t;v_n,\theta_n)| > c_{44}n^{-\beta}) < c_{45}n^{-c_{46}}.$$

In addition, there exist positive constants c_{47}–c_{49} not depending on v_0, such that for every $x > 0$

$$P_{v_0}(\sup_{t\in[0,\infty)} |L_n(t;v_0) - L(t;v_0,\theta_0)| > n^{-1/2}(c_{47}\log n + x))$$
$$< c_{48}\exp\{-c_{49}x\}.$$

c *The first order partial derivatives of $L_n(t;v)$ with respect to the components of v are random elements of $D[0,\infty)$. With P_n-probability 1 they are bounded by c_{43}, and have total variation bounded by c_{43}.*

d *For every $i = 1,\dots,r$ there exists a deterministic function $L_i^{[1]}(t;v,\theta)$ such that for every $\beta < 1/2$ there exist positive constants c_{44}–c_{46} not depending on v_n or θ_n such that*

$$P_n(\sup_{t\in[0,\infty)} |L_{ni}^{(1)}(t;v_{n0}) - L_i^{[1]}(t;v_n,\theta_n)| > c_{44}n^{-\beta}) < c_{45}n^{-c_{46}},$$

where $L_{ni}^{(1)}(t;v)$ is the first order partial derivative of $L_n(t;v)$ with respect to the i^{th} component of v.

e *With P_n-probability 1 the second order partial derivatives of $L_n(t;v)$ with respect to the components of v are random elements of $D[0,\infty)$. They are bounded by c_{43}, and have total variation bounded by c_{43}.*

Condition 13 asks of various random elements of $D[0,\infty)$ that they are uniformly bounded and have uniformly bounded total variation with P_n-probability 1. However, if needed we may replace "P_n-probability 1" by "P_n-probability tending to 1 at a fast enough rate". The rate at which $P_n(S_n < cn^{1/2})$ tends to 1 is for any $c > 0$ fast enough. In section 5.3 the relaxed version of Condition 13 is used.

To describe the effect of the M-estimator $v^{(n)}$ on the process $Q_n(t;v^{(n)})$, we introduce the r-dimensional vector function $K_0(t;v,\theta)$. The i^{th} element of $K_0(t;v,\theta)$ is defined by

$$K_{0i}(t;v,\theta) = \int_0^t L(s;v,\theta)\psi_i(s;\pi(v,\theta),\theta_0)dH^1(s;v,\theta)$$
$$+ \int_0^t \{L(s;v,\theta)\psi_i(s;\pi(v,\theta),\theta_0) + L_i^{[1]}(s;v,\theta)\}dD(s;v,\theta). \qquad (4.16)$$

Observe that $K_{0i}(t;v,\theta)$ remains bounded, uniformly in v and θ. Moreover, we have under P_{v_0}

$$K_0(t;v_0,\theta_0) = \int_0^t L(s;v_0,\theta_0)\psi(s;v_0,\theta_0)dH^1(s;v_0,\theta_0), \tag{4.17}$$

where $\psi(t;v,\theta)$ is the r-dimensional vector with components $\psi_i(t;v,\theta)$.

Theorem 8 *There exists a sequence $\{W_n(t;v_{n0})\}_{n=1}^{\infty}$ of mean zero Gaussian processes with covariance function*

$$\begin{aligned}\mathcal{E}_n W_n(t_1;v_{n0})W_n(t_2;v_{n0}) &= H^1(t_1\wedge t_2;v_n,\theta_n)\\ &-\int_0^{t_1\wedge t_2}\{D(t_1;v_n,\theta_n)+D(t_2;v_n,\theta_n)\\ &\quad -2D(s;v_n,\theta_n)\}d\Lambda(s;v_{n0},\theta_0)\\ &-D(t_1;v_n,\theta_n)D(t_2;v_n,\theta_n)\end{aligned} \tag{4.18}$$

such that for every $\beta < (1/2-2\alpha)\wedge 1/6$ we have

$$\begin{aligned}P_n(\sup_{t\in[0,\infty)} &|\{Q_n(t;v^{(n)}) - n^{1/2}\int_0^t L_n(s;v_{n0})dD(s;v_n,\theta_n)\}\\ &-\{\int_0^t L(s;v_n,\theta_n)dW_n(s;v_{n0})\\ &-(K_0(t;v_n,\theta_n))^T\Xi_{0n}^{-1}\int_0^{\infty}\phi(s;v_{n0})dW_n(s;v_{n0})\}|\\ &> c_{50}n^{-\beta}) \le c_{51}n^{-c_{52}},\end{aligned} \tag{4.19}$$

where c_{50}–c_{52} are positive constants not depending on v_n or θ_n.

If in addition $(v_n,\theta_n)=(v_0,\theta_0)$ then there exist positive constants $c_{53}-c_{55}$, not depending on (v_0,θ_0), such that

$$\begin{aligned}P_{v_0}(\sup_{t\in[0,\infty)} &|Q_n(t;v^{(n)}) - \{\int_0^t L(s;v_0,\theta_0)dW_n(s;v_0)\\ &-(K_0(t;v_0,\theta_0))^T\Sigma_{e0}^{-1}\int_0^{\infty}\phi(s;v_0)dW_n(s;v_0)\}|\\ &> n^{-1/6}(1+x)(c_{53}\log n + x)^2) \le c_{54}\exp\{-c_{55}x\}.\end{aligned} \tag{4.20}$$

Corollary 6 is based on the observation that the Gaussian process which is used in Theorem 8 to approximate $Q_n(t;v^{(n)})$ under P_{v_0} may be written as $\int_0^{\infty} L_0(s,t;v_0)dW_n(s;\theta_0)$, where the function $L_0(s,t;v_0)$ is defined by

$$L_0(s,t;v_0) = L(s;v_0,\theta_0)1_{\{s\le t\}} - (K_0(t;v_0,\theta_0))^T\Sigma_{e0}^{-1}\phi(s;v_0). \tag{4.21}$$

It seems that $L_0(s,t;v_0)$ is an adapted version of $L(s;v_0,\theta_0)$, in which the influence of the M-estimator at the stochastic integral evaluated at point t is accounted for.

Corollary 6 *The sequence $\{Q_n(t;v^{(n)})\}_{n=1}^{\infty}$ converges in P_{v_0}-distribution to a mean zero Gaussian process $\check{X}(t;v_0,\theta_0)$ with covariance function*

$$\mathcal{E}_{v_0}\check{X}(t_1;v_0,\theta_0)\check{X}(t_2;v_0,\theta_0)$$
$$= \int_0^\infty L_0(s,t_1;v_0)L_0(s,t_2;v_0)dH^1(s;v_0,\theta_0). \tag{4.22}$$

Corollary 6 implies that the variance function of $\check{X}(t;v_0,\theta_0)$ may be written as

$$\mathcal{E}_{v_0}\{\check{X}(t;v_0,\theta_0)\}^2 = \int_0^\infty (L_0(s,t;v_0))^2 dH^1(s;v_0,\theta_0). \tag{4.23}$$

Since we have

$$\int_0^\infty \int_0^\infty (L_0(s,t;v_0))^2 dH^1(s;v_0,\theta_0)dH^1(t;v_0,\theta_0) < \infty, \tag{4.24}$$

we may view the operator which assigns $\int_0^\infty L_0(s,t;v_0)f(s)dH^1(s;v_0,\theta_0)$ to any $f \in D[0,\infty)$ such that $\int_0^\infty (f(s))^2 dH^1(s;v_0,\theta_0) < \infty$ as a Hilbert-Schmidt operator. This operator is used to transform $W_n(t;v_0)$ into a Gaussian process which under P_{v_0} closely approximates $Q_n(t;v^{(n)})$. The relatively complicated form of the covariance function of $Q_n(t;v^{(n)})$ gives rise to the interesting question whether there exists another operator which can be used to transform $Q_n(t;v^{(n)})$ into a process closely approximated by $W_n(t;v_0)$ under P_{v_0}. It is possible to solve this inverse problem by following the path set out by Khamaladze (1981, 1982) in his study of the empirical process with estimated parameters.

However, the derivation of probability inequalities of the type given in Theorems 2 and 8 for the transformed process would require additional notation and, above all, quite unpleasant conditions. Therefore, we do not pursue this point any further.

The second part of Condition 13b is only needed for verifying (4.20). Thus, (4.20) and Corollary 6 remain valid if the second part of Condition 13b is dropped.

The proof of Theorem 8 makes us of Theorem 1. Just like this theorem, Theorem 8 can be sharpened in the case where $L_n(t;v_{n0})$ coincides with $L(t;v_n,\theta_n)$ with P_n-probability 1. Then, (4.19) holds for every $\beta < (1/2 - 2\alpha)$, and in (4.20) the term $n^{-1/6}$ may be replaced by $n^{-1/2}$.

Convergence in distribution of a sequence of processes strongly related to $\{Q_n(t;v^{(n)})\}_{n=1}^{\infty}$ was obtained by Hjort (1984) [see also Hjort (1990)].

4.3 Sublinear tests

As in the previous chapter we may use a functional T mapping $D[0,\infty)$ into $I\!R$ to construct a goodness-of-fit test. Applying T to the process $Q_n(t;v^{(n)})$, we obtain a statistic which can be used to test the composite null hypothesis $\theta_n = \theta_0$.

In this section we study deviations, behavior under fixed and local alternatives, and efficiencies of sublinear tests, based on test statistics $T(Q_n(t;v^{(n)}))$, where T satisfies Condition 14.

Condition 14 *The functional T is sublinear and Lipschitz.*

Observe that this condition is exactly the same as Condition 3. Thus, the remarks concerning Condition 3 also apply here.

4.3.1 Deviations

Define

$$\breve{a}(v_0) = \{\sup_{f\in\mathcal{S}} T(\int_0^\infty L_0(s,\cdot;v_0)f'(H^1(s;v_0,\theta_0))dH^1(s;v_0,\theta_0))\}^{-2} \qquad (4.25)$$

where $\mathcal{S}$ is the set of Strassen functions. One may interpret $(\breve{a}(v_0))^{-1/2}$ as the norm of the functional T, this time induced by the reproducing kernel Hilbert space of the Gaussian process $\check{X}(t;v_0,\theta_0)$, the limit in P_{v_0}-distribution of the sequence $\{Q_n(t;v^{(n)})\}_{n=1}^\infty$. Suppose that

Condition 15 $\inf_{v_0\in\Upsilon}\breve{a}(v_0)$ *is positive.*

Theorem 9 describes the tail behavior of $T(\check{X}(\cdot;v_0,\theta_0))$ [and thus the tail behavior of the asymptotic distribution of $T(Q_n(\cdot;v^{(n)}))$], and presents a deviation result for $T(Q_n(\cdot;v^{(n)}))$. It is valid whenever equation (4.20) and Conditions 14 and 15 hold.

Theorem 9 *We have*

$$\lim_{t\to\infty} t^{-2}\inf_{v_0\in\Upsilon}(\breve{a}(v_0))^{-1}\log P_{v_0}(T(\check{X}(\cdot;v_0,\theta_0))>t) = -1/2. \qquad (4.26)$$

$$\lim_{t\to\infty} t^{-2}\sup_{v_0\in\Upsilon}(\breve{a}(v_0))^{-1}\log P_{v_0}(T(\check{X}(\cdot;v_0,\theta_0))>t) = -1/2. \qquad (4.27)$$

Furthermore,

$$\lim_{n\to\infty}(s_n)^{-2}\inf_{v_0\in\Upsilon}(\breve{a}(v_0))^{-1}\log P_{v_0}(T(Q_n(\cdot;v^{(n)}))>s_n) = -1/2 \qquad (4.28)$$

$$\lim_{n\to\infty}(s_n)^{-2}\sup_{v_0\in\Upsilon}(\breve{a}(v_0))^{-1}\log P_{v_0}(T(Q_n(\cdot;v^{(n)}))>s_n) = -1/2 \qquad (4.29)$$

for any sequence $\{s_n\}_{n=1}^\infty$ such that $s_n\to\infty$ and $s_n = o(n^{1/30})$ as $n\to\infty$.

Obviously, (4.28) and (4.29) also hold for any sequence $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = \mathcal{O}((\log n)^{1/2})$. The deviation result based on the latter type of sequence is called moderate and is important from a statistical perspective, since it plays a role in evaluating the performance of a test.

As with Theorem 8, Theorem 9 can be refined in the special case where $L_n(t; v_0)$ coincides with $L(t; v_0, \theta_0)$ with P_{v_0}-probability 1. Now sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = o(n^{1/10})$ are allowed.

From a practical viewpoint it is desirable that a test for the composite null hypothesis is based on a test statistic which has under the null hypothesis an asymptotic distribution which does not depend on the actual value of the unknown parameter. Hence, the rate at which the tail of the distribution of the test statistic decays should not depend on the unknown parameter. Theorem 9 implies that in order for this latter situation to occur it is sometimes needed to multiply the original weight process by the square root of $\breve{a}(v)$. Of course, it should be verified whether the newly constructed weight process meets all requirements.

Since in general the distribution function $H^1(t; v_0, \theta_0)$ is involved in $\breve{a}(v)$, a slight problem arises when the censoring distribution G is unknown. In this case it seems best to divide the weight process by the square root of an estimator for $\breve{a}(v)$. Typically, estimators for $\breve{a}(v)$ are obtained by replacing $H^1(t; v_0, \theta_0)$ by $\int_0^t (1 - H_{n-}(s)) d\Lambda(s; v_0, \theta_0)$.

4.3.2 Behavior under the alternative hypothesis

To describe the behavior of $T(Q_n(\cdot; v^{(n)}))$ under a fixed alternative, suppose that $v_n = v \in \Upsilon$ and $\theta_n = \theta \in \Theta$ for every $n \in \mathbb{N}$ [we shall refer to this stuation as "under $P_{(v,\theta)}$"]. Then for every $\beta > (2\alpha - 1/2)^+$

$$n^{-\beta} |T(Q_n(\cdot; v^{(n)})) - n^{1/2} T(\int_0^{\cdot} L(s; v, \theta) dD(s; v, \theta))| \to_{P_{(v,\theta)}} 0. \qquad (4.30)$$

as follows from combining (4.19) with Conditions 13 and 14.

The treatment of local alternatives requires a more substantial effort. First we impose structure on Θ.

Condition 16 *Θ is a convex subset of $\mathbb{R}^p$.*

As an immediate consequence of Condition 16 we have that $\Upsilon \times \Theta$ is a subset of $\mathbb{R}^{r+p}$. Hence, it is sensible to redefine $\psi_i(t; v, \theta)$ as the first order partial derivative of $\log \lambda(t; v, \theta)$ with respect to the i^{th} component of (v, θ), $i = 1, \ldots, r + p$. Note that this redefinition is consistent with the earlier definition. Likewise, redefine $\psi_{ij}^{(1)}(t; v, \theta)$ as the second order partial derivative with respect to the i^{th} and j^{th} component of (v, θ).

Condition 17 *For every $v \in \Upsilon$, $\theta \in \Theta$ and $i,j = 1,\ldots,r+p$ the functions $\lambda(t;v,\theta)$, $\psi_i(t;v,\theta)$ and $\psi_{ij}^{(1)}(t;v,\theta)$ exist. For some $\beta > 0$ there exists a constant c_{56} such that for every $v \in \Upsilon$, $\theta \in \Theta$ and $i,j = 1,\ldots,r+p$*

$$\int_0^\infty (\psi_i(s;v,\theta))^2(1 - H(s;v,\theta))^{(1-2\alpha)-\beta} d\Lambda(s;v,\theta) < c_{56},$$

$$\int_0^\infty (\psi_{ij}^{(1)}(s;v,\theta))^2(1 - H(s;v,\theta))^{2(1-2\alpha)-\beta} d\Lambda(s;v,\theta) < c_{56}.$$

Denote the p-dimensional vector with components $\psi_{r+i}(t;v_0,\theta_0)$ by $\psi_a(t;v_0,\theta_0)$, and define

$$K_{a|0}(t;v_0) = \int_0^\infty L_0(s,t;v_0)\psi_a(s;v_0,\theta_0)dH^1(s;v_0,\theta_0) \qquad (4.31)$$

[compare with (4.21)]. Observe that under Conditions 8 and 9 the components of $K_{a|0}(t;v_0)$ remain uniformly bounded in t, v_0 and θ_0.

Let $\Lambda_i^{(1)}(t;v,\theta)$ denote the first order partial derivative of $\Lambda(t;v,\theta)$ with respect to the i^{th} component of (v,θ). Define the $r \times r$ matrix $\Xi_b(v,\theta)$ by

$$\begin{aligned}\Xi_{bij}(v,\theta) = &\int_0^\infty \phi_i(s;\pi(v,\theta))\psi_j(s;\pi(v,\theta),\theta_0)dH^1(s;v,\theta) \\ &+ \int_0^\infty \psi_i(s;\pi(v,\theta))\Lambda_j^{(1)}(s;\pi(v,\theta),\theta_0)dD(s;v,\theta), \qquad (4.32)\end{aligned}$$

Condition 18 *All eigenvalues of the square of the matrix $\Xi_b(v,\theta)$ exceed $(c_{35})^2$ for every $v \in \Upsilon$ and $\theta \in \Theta$.*

It should be noted that $\Xi_b(v,\theta_0) = \Xi_0(v,\theta_0)$ for every $v \in \Upsilon$. Since the smallest eigenvalue of the square of $\Xi_b(v,\theta)$ tends to the smallest eigenvalue of the square of $\Xi_b(v,\theta_0)$ as θ tends to θ_0, the remarks on the verifiability of Condition 12 also apply to Condition 18.

The next condition concerns $L_i^{(1)}(t;v,\theta)$, the first order partial derivative of $L(t;v,\theta)$ with respect to the i^{th} component of (v,θ), and $L_{ij}^{(2)}(t;v,\theta)$, the second order partial derivative of $L(t;v,\theta)$ with respect to the i^{th} and j^{th} components of (v,θ).

Condition 19 *For every $v \in \Upsilon$, $\theta \in \Theta$ and $i,j = 1,\ldots,r+p$ the functions $L_i^{(1)}(t;v,\theta)$ and $L_{ij}^{(2)}(t;v,\theta)$ exist. There exists a constant c_{57} such that for every $v \in \Upsilon$, $\theta \in \Theta$ and $i,j = 1,\ldots,r+p$*

$$\sup_{t\in[0,\infty)} (1 - H(t;v,\theta))^{1/2}|L_i^{(1)}(t;v,\theta)| < c_{57},$$

$$\sup_{t\in[0,\infty)} (1 - H(t;v,\theta))^{1-\alpha}|L_{ij}^{(2)}(t;v,\theta)| < c_{57}.$$

Theorem 10 *Suppose the sequence* $\{(v_n, \theta_n)\}_{n=1}^{\infty}$ *converges to the point* (v_0, θ_0). *Let h be the p-dimensional unit vector defined by* $h = \lim_{n\to\infty}(\theta_n - \theta_0)/|\theta_n - \theta_0|$, *and let* σ *denote* $\lim_{n\to\infty} n^{1/2}|\theta_n - \theta_0|$.

a *If* $\sigma = \infty$ *then* $(n^{1/2}|\theta_n - \theta_0|)^{-1}T(Q_n(\cdot; v^{(n)}))$ *converges in* P_n*-probability to* $T(h^T K_{a|0}(\cdot; v_0))$.

b *If* $\alpha < 1/4$ *and* $\sigma < \infty$ *then* $\{T(Q_n(\cdot; v^{(n)}))\}_{n=1}^{\infty}$ *converges in* P_n*-distribution to* $T(\check{X}(\cdot; v_0, \theta_0) + \sigma h^T K_{a|0}(\cdot; v_0))$.

The proof of Theorem 10b makes use of the fact that (4.19) holds for some $\beta > 0$, which provides the reason for the occurrence of the restriction $\alpha < 1/4$.

If $\alpha < 1/4$ then Theorem 10 reveals three types of behavior of the sequence of test statistics $\{T(Q_n(\cdot; v^{(n)}))\}_{n=1}^{\infty}$, depending on the rate at which the alternatives converge to the null hypothesis. If the rate is faster than $n^{-1/2}$ then we have convergence in distribution to the same limit as under the null hypothesis. If the rate is of the order $n^{-1/2}$ then we also have convergence in distribution, but to a limit different from the one under the null hypothesis. If the rate is slower than $n^{-1/2}$ then the convergence in distribution is lost, since the sequence of test statistics blows up as n tends to infinity.

4.3.3 Efficiencies

In the previous chapter a number of efficiency concepts were introduced. However, the definitions given there only apply to the simple null hypothesis. We shall now carry these definitions over to the composite null hypothesis, and reflect on their implications.

Our definition of asymptotic relative Pitman efficiency is restricted to the situation in which we have two infinite sequences of tests of the same size, and where the n^{th} test in the i^{th} sequence is based on the test statistic T_{in}, rejects the null hypothesis if $T_{in} > t_{in}$, and does not reject if $T_{in} < t_{in}$. Again, we have no knowledge about the action taken if $T_{in} = t_{in}$.

Definition 7 *For a given function* $\tilde{\beta} : \Upsilon \times \Theta \to [0, 1]$, *let* $\overline{N}_i^{\tilde{\beta}}(v, \theta)$ *be the largest sample size such that* $P_{(v,\theta)}(T_{in} \geq t_{in}) < \tilde{\beta}(v, \theta)$. *If for a unit vector* $h \in I\!R^p$ *and every* $v_0 \in \Upsilon$ *there exists a constant* $e_{12}^{\tilde{\beta}}(h; v_0)$ *such that*

$$\liminf_{j\to\infty} \frac{\overline{N}_2^{\tilde{\beta}}(v_j, \theta_j)}{\overline{N}_1^{\tilde{\beta}}(v_j, \theta_j)} \geq e_{12}^{\tilde{\beta}}(h; v_0)$$

and

$$\limsup_{j\to\infty} \frac{\overline{N}_2^{\tilde{\beta}}(v_j, \theta_j)}{\overline{N}_1^{\tilde{\beta}}(v_j, \theta_j)} \leq e_{12}^{\tilde{\beta}}(h; v_0)$$

for any sequence $\{(v_j, \theta_j)\}_{j=1}^{\infty}$ *tending to* (v_0, θ_0) *while satisfying* $\lim_{j\to\infty} \tilde{\beta}(v_j, \theta_j) \in (0,1)$ *and* $\lim_{j\to\infty}(\theta_j - \theta_0)/|\theta_j - \theta_0| = h$, *then* $e_{12}^{\tilde{\beta}}(h; v_0)$ *is the asymptotic relative Pitman efficiency of the first sequence of tests with respect to the second.*

This definition allows the asymptotic relative Pitman efficiency to depend on both size and power of the tests. It does not involve $\underline{N}_i^{\tilde{\beta}}(v, \theta)$, the smallest sample size such that $P_{(v,\theta)}(T_{in} > t_{in}) \geq \tilde{\beta}(v, \theta)$, as is propagated by Rothe (1981). An efficacy approach is possible if there is a transformation of the test statistics available such that the transformed tests statistics are asymptotically normal and have asymptotic variance 1 for any sequence $\{\theta_n\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} n^{1/2}|\theta_n - \theta_0| < \infty$. The asymptotic mean of the transformed test statistics divided by $\lim_{n\to\infty} n^{1/2}|\theta_n - \theta_0|$ is the square root of what is called the efficacy of the sequence of test statistics. For two such sequences the asymptotic relative Pitman efficiency is equal to the ratio of their respective efficacies. In absence of such a transformation, the computation of asymptotic relative Pitman efficiencies becomes too formidable for general sublinear tests.

Definition 8 *A sequence of test statistics* $\{T_{in}\}_{n=1}^{\infty}$ *is said to be a standard sequence if the following three conditions are satisfied.*
a *The sequence* $\{T_{in}\}_{n=1}^{\infty}$ *converges in* P_{v_0}*-distribution to a random variable* T_i.
b *There exists a constant* a_i *such that*

$$\lim_{t\to\infty} t^{-2} \inf_{v_0\in\Upsilon} \log P_{v_0}(T_i > t) = -a_i/2.$$

$$\lim_{t\to\infty} t^{-2} \sup_{v_0\in\Upsilon} \log P_{v_0}(T_i > t) = -a_i/2.$$

c *There exists a positive function* $b_i(v,\theta)$ *such that* $|n^{-1/2}T_{in} - b_i(v,\theta)|$ *converges to zero in* $P_{(v,\theta)}$*-probability for every* $v \in \Upsilon$ *and* $\theta \in \Theta - \{\theta_0\}$.
The approximate Bahadur slope of a standard sequence $\{T_{in}\}_{n=1}^{\infty}$ *is defined as* $a_i(b_i(v,\theta))^2$. *The approximate Bahadur efficiency of a standard sequence* $\{T_{1n}\}_{n=1}^{\infty}$ *with respect to another standard sequence* $\{T_{2n}\}_{n=1}^{\infty}$ *is defined as the ratio of their respective Bahadur slopes* $a_1(b_1(v,\theta))^2/a_2(b_2(v,\theta))^2$.

By Corollary 6 and equations (4.26), (4.27) and (4.30) it immediately follows that the approximate Bahadur slope of $\{T(Q_n(\cdot; v^{(n)}))\}_{n=1}^{\infty}$ is given by $\breve{a}(v_0)\{T(\int_0^{\cdot} L(s; v, \theta)dD(s; v, \theta))\}^2$.

Definition 9 *A standard sequence* $\{T_{in}\}_{n=1}^{\infty}$ *is said to be a Wieand sequence if there is a constant* $\epsilon^* > 0$ *such that for every* $\epsilon > 0$ *and* $\delta \in (0,1)$ *there exists an integer* N *such that*

$$P_{(v,\theta)}(|n^{-1/2}T_{in} - b_i(v,\theta)| > \epsilon b_i(v,\theta)) < \delta$$

for every $v \in \Upsilon$, $\theta \in \Theta - \{\theta_0\}$ *satisfying* $|\theta - \theta_0| < \epsilon^*$ *and* $n > N/(b_i(v,\theta))^2$.

Since our definition of standard sequences slightly extends the one used in Wieand (1976) [it allows $P_{v_0}(T_i > t)$ to depend on v_0], an adaption of Wieand's result is in order.

Theorem 11 (Wieand (1976), adapted) *Let $\{T_{1n}\}_{n=1}^{\infty}$ and $\{T_{2n}\}_{n=1}^{\infty}$ be two Wieand sequences. Suppose $\lim_{(v,\theta)\to(v_0,\theta_0)} b_i(v,\theta) = 0$ for $i = 1,2$ and suppose that for every p-dimensional unit vector h the limit*

$$\lim_{\theta\to\theta_0} \{a_1(b_1(v,\theta))^2\}/\{a_2(b_2(v,\theta))^2\} \tag{4.33}$$

exists if $(\theta-\theta_0)/|\theta-\theta_0|$ tends to h. Then the limiting (as the size of the tests approaches zero) asymptotic relative Pitman efficiency exists and is equal to the limit given in (4.33).

Theorem 11 is not valid if variants of asymptotic relative Pitman efficiency such as in Rothe (1981) are used. See Appendix A for more details.

Theorem 12 *Let h be a p-dimensional unit vector, let (v,θ) approach (v_0,θ_0) from the direction h [that is, $h = \lim_{(v,\theta)\to(v_0,\theta_0)}(\theta-\theta_0)/|\theta-\theta_0|$], and define $e(h;v_0)$ by*

$$e(h;v_0) = \breve{a}(v_0)\{T(h^T K_{a|0}(\cdot;v_0))\}^2 \tag{4.34}$$

If $\inf_{v_0\in\Upsilon} e(h;v_0)$ is not equal to zero, then $\{T(Q_n(\cdot;v^{(n)}))\}_{n=1}^{\infty}$ is a Wieand sequence with approximate Bahadur slope of the form

$$|\theta-\theta_0|^2\{e(h;v_0)+o(1)\}. \tag{4.35}$$

Definition 10 *A sequence of test statistics $\{T_{in}\}_{n=1}^{\infty}$ is said to be a Kallenberg sequence if the following conditions are satisfied.*

a *There exists a constant a_i such that*

$$\lim_{n\to\infty}(s_n)^{-2}\inf_{v_0\in\Upsilon}\log P_{v_0}(T_{in} > s_n) = -a_i/2$$

$$\lim_{n\to\infty}(s_n)^{-2}\sup_{v_0\in\Upsilon}\log P_{v_0}(T_{in} > s_n) = -a_i/2$$

for all sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n\to\infty$ and $s_n = o(n^{1/6})$ as $n\to\infty$.

b *There exists a positive function $b_i(v,\theta)$ such that $n^{-1/2}T_{in}/b_i(v_n,\theta_n)$ converges to 1 in P_n-probability for all sequences $\{(v_n,\theta_n)\}_{n=1}^{\infty}$ tending to $\Upsilon\times\{\theta_0\}$.*

If the sequences $\{T_{1n}\}_{n=1}^{\infty}$ and $\{T_{2n}\}_{n=1}^{\infty}$ both are Kallenberg and the limit $\lim_{n\to\infty} a_1(b_1(v_n,\theta_n))^2/a_2(b_2(v_n,\theta_n))^2$ exists, then the asymptotic intermediate efficiency of $\{T_{1n}\}_{n=1}^{\infty}$ with respect to $\{T_{2n}\}_{n=1}^{\infty}$ is defined as this limit.

The typical behavior of $b_i(v,\theta)$ near $\Upsilon \times \{\theta_0\}$ is as a linear function of $|\theta_n - \theta_0|$. This brings us to call $\lim_{n\to\infty} a_i(b_i(v_n,\theta_n)/|\theta_n - \theta_0|)^2$ the intermediate slope of the Kallenberg sequence $\{T_{in}\}_{n=1}^{\infty}$.

Theorem 9 does not ensure that $\{T(Q_n(\cdot;v^{(n)}))\}_{n=1}^{\infty}$ is a Kallenberg sequence. Reason for us to turn to a variant of asymptotic intermediate efficiency, weak asymptotic intermediate efficiency. Here only sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = \mathcal{O}((\log n)^{1/2})$ as $n \to \infty$ are considered. Observe that the sequence $\{T(Q_n(\cdot;v^{(n)}))\}_{n=1}^{\infty}$ has weak intermediate slope $e(h;v_0)$. Hence, the weak intermediate approach yields the same picture as the Wieand approach. Theorem 9 shows that our tests may be evaluated using a variant of asymptotic intermediate efficiency which considers sequences $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to \infty$ and $s_n = o(n^{1/30})$ as n tends to infinity.

In the beginning of this section we assumed that the functional T is sublinear. A close look reveals that Condition 14b is used only in the derivation of (4.26) and (4.27). Thus, we may set up an equivalent theory for functionals other than sublinear, provided results similar to (4.26) and (4.27) hold.

4.3.4 Generalized rank and supremum type tests

We have already seen that both T_R and T_S satisfy Condition 14. Hence, our theory applies. Defining $a_R(v_0)$ and $a_S(v_0)$ according to (4.25), with T replaced by T_R and T_S respectively, it follows that

$$a_R(v_0) = \{\int_0^{\infty} (L_0(s,\infty;v_0))^2 dH^1(s;v_0,\theta_0)\}^{-1}, \tag{4.36}$$

$$a_S(v_0) = \{\sup_{t\in[0,\infty)} \int_0^{\infty} (L_0(s,t;v_0))^2 dH^1(s;v_0,\theta_0)\}^{-1} \tag{4.37}$$

[discover the relation with the variance function of $\check{X}(t;v_0,\theta_0)$ by comparing with equation (4.23)]. Similarly defining $e_R(h,v_0)$ and $e_S(h,v_0)$ according to (4.34), we obtain

$$e_R(h,v_0) = a_R(v_0)\{h^T K_{a|0}(\infty;v_0)\}^2, \tag{4.38}$$

$$e_S(h,v_0) = a_S(v_0)\{\sup_{t\in[0,\infty)} h^T K_{a|0}(t;v_0)\}^2. \tag{4.39}$$

As opposed to general sublinear tests, generalized rank tests do allow us to compute asymptotic Pitman efficiencies. By Corollary 6 and Theorem 10b, it follows that the asymptotic power against local alternatives $\theta_n = \theta_0 + n^{-1/2}h$ of the test based on $T_R(Q_n(\cdot;v^{(n)}))$ of size $\tilde{\alpha}$ equals

$$P_{v_0}(\check{X}(\infty;v_0,\theta_0) > z_{\tilde{\alpha}}(a_R(v_0))^{-1/2} - h^T K_{a|0}(\infty;v_0)),$$

where $z_{\tilde{\alpha}}$ is the $(1-\tilde{\alpha})$ quantile of a standard normal distribution. This implies that the efficacy of the sequence of test statistics $\{T_R(Q_n(\cdot;v^{(n)}))\}_{n=1}^{\infty}$ is equal to $e_R(h,v_0)$.

Let $I(v_0, \theta_0)$ be the Fisher information matrix, the symmetric $(r+p) \times (r+p)$ matrix with elements

$$I_{ij}(v_0, \theta_0) = \int_0^\infty \psi_i(t; v_0, \theta_0)\psi_j(t; v_0, \theta_0) dH^1(t; v_0, \theta_0). \tag{4.40}$$

Partition this matrix as follows

$$I(v_0, \theta_0) = \begin{pmatrix} I_{00} & I_{0a} \\ I_{a0} & I_{aa} \end{pmatrix}, \tag{4.41}$$

where I_{00} is an $r \times r$ matrix. Define the effective score function $\psi_{a|0}(t; v_0, \theta_0)$ by

$$\psi_{a|0}(t; v_0, \theta_0) = \psi_a(t; v_0, \theta_0) - I_{a0} I_{00}^{-1} \psi(t; v_0, \theta_0). \tag{4.42}$$

The effective score function may be interpreted as a projection of the score function $\psi_a(t; v_0, \theta_0)$. Observe that

$$\int_0^\infty \psi_a(t; v_0, \theta_0)(\psi_a(t; v_0, \theta_0))^T dH^1(t; v_0, \theta_0) = I_{aa} - I_{a0} I_{00}^{-1} I_{0a}. \tag{4.43}$$

The right-hand side of the latter equation is known as effective Fisher information.

Theorem 13 *The maximized values of $e_R(h, v_0)$ and $e_S(h, v_0)$ are both equal to $h^T\{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\}h$. Generalized rank tests based on weight processes with limiting weight function satisfying*

$$L(t; v_0, \theta_0) \propto h^T\{\psi_{a|0}(t; v_0, \theta_0)\} + \nu^T \phi(t; v_0), \tag{4.44}$$

where ν is an arbitrary r-dimensional vector, are maximizing $e_R(h, v_0)$. Supremum type tests based on maximum likelihood estimation and weight processes with limiting weight function satisfying

$$L(t; v_0, \theta_0) \propto h^T \psi_{a|0}(t; v_0, \theta_0) \tag{4.45}$$

are maximizing $e_S(h, v_0)$.

One may infer from the proof of Theorem 13 that for every supremum type test with $e_S(h, v_0) = h^T\{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\}h$ there exists a generalized rank test [based on either the same stochastic integral or a "killed" version $Q_n(t \wedge \tau; v^{(n)})$, where $0 < \tau < \infty$ fixed] with $e_R(h, v_0) = h^T\{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\}h$. The reverse is not true. If the stochastic integral underlying a generalized rank test with $e_R(h, v_0) = h^T\{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\}h$ does not converge in P_{v_0}-distribution to a Gaussian process attaining its maximum in ∞, then we have that $e_S(h, v_0)$ is less than $h^T\{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\}h$ for the supremum type test based on the same stochastic integral.

This fact makes the choice of the weight process more critical for supremum type tests than for generalized rank tests. Preferably, the construction of a weight process should lead to

$$\int_0^\infty (L_0(s,t;v_0))^2 dH^1(s;v_0,\theta_0) \leq \int_0^\infty (L_0(s,\infty;v_0))^2 dH^1(s;v_0,\theta_0).$$

That is, under maximum likelihood estimation the weight process should be based on (4.45) rather than on (4.44).

Shortly after Theorem 9 a method of standardizing a weight process was described. If we follow this method for weight processes with limiting weight function satisfying (4.45), we obtain that the standardized weight processes have limiting weight function

$$L(t;v,\theta_0) = \frac{h^T \psi_{a|0}(t;v,\theta_0)}{\sqrt{h^T\{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\} h}}. \tag{4.46}$$

As far as generalized rank tests are concerned, the choice of the M-estimator turns out to be of secondary importance. The maximized value of $e_R(h, v_0)$ does not depend on the estimation procedure. Equation (4.46) does not depend on the estimation procedure. For a weight process constructed according to (4.46) we have $a_R(v_0) = 1$, so the asymptotic behavior of the test statistics under the null hypothesis is not influenced by the estimation procedure. This implies that the quantity $e_R(h, v_0)/a_R(v_0)$, which characterizes the behavior of the test statistic under the alternative hypothesis, does not depend on the estimation procedure.

The choice of the M-estimator is far more important for supremum type tests, because of its effect on the position where the variance function $\int_0^\infty (L_0(s,t;v_0))^2 dH^1(s;v_0,\theta_0)$ is maximized. For a general M-estimation procedure this position is difficult to determine. Theorem 13 restricts itself in this respect to maximum likelihood estimation.

In Hjort (1990) χ^2-tests constructed from $Q_n(t;v^{(n)})$, where $v^{(n)}$ is the maximum likelihood estimator, are studied using martingale methods. The effective score function appearing in our (4.45) also shows up in Hjort's equation (5.7). Without rigorous mathematical support, Hjort states that as a weight process this function "is a very good choice".

The maximum likelihood procedure pays no attention to the alternative hypothesis. A modification of the maximum likelihood estimator is obtained by choosing

$$\phi(t;v) = \psi(t;v,\theta_0) - I_{0a} I_{aa}^{-1} \psi_a(t;v,\theta_0) \tag{4.47}$$

This choice sets $\int_0^\infty \phi_i(s;v)\psi_{r+j}(s;v,\theta_0) dH^1(s;v,\theta_0)$ equal to zero for every $v \in \Upsilon$, $i = 1,\ldots,r$ and $j = 1,\ldots,p$. Hence, the $r \times p$ matrix $\Xi_a(v,\theta_0)$ defined in (4.100) on page 80 is a null matrix. Since this matrix is related to the $r \times p$ matrix of partial derivatives of $\pi(v,\theta)$ with respect to the components of θ [see the proof of Lemma 3], we may say that estimation is in a sense performed "perpendicular" to the alternative hypothesis. Again the situation where the censoring distribution G is unknown calls for extra attention, since G is involved in $I_{0a}I_{aa}^{-1}$. We may try to estimate this quantity, but then the resulting function $\phi(t;v)$ is a random element of $D[0,\infty)$, and falls outside our framework.

4.4 Proofs

In this section the proofs of previously stated theorems are gathered, with exception of Theorem 11, which is proved in Appendix A.

Proof of Theorem 7 Let $\Phi_{ni}(v)$ be shorthand notation for $\Phi_{ni}(\infty;v)$. Then the i^{th} equation involved in the definition of the M-estimator $v^{(n)}$ may be written as $\Phi_{ni}(v^{(n)}) = 0$ [see equation (4.3) on page 52].

The proof is based on a second order expansion of $\Phi_{ni}(v)$ around v_{n0}. The stochastic terms in the expansion are approximated by means of deterministic counterparts. The quality of these approximations is reflected by the random variable S_n, to be introduced later. Conditional on the event $S_n < n^{1/2}$ we use the "approximated" expansion to show the existence of $v^{(n)}$, and prove (4.8) and (4.9). Subsequently, empirical process theory is used to study the behavior of the random variable S_n in some detail, yielding (4.10) and (4.11).

The first order partial derivative of $\Phi_{ni}(v)$ with respect to the j^{th} component of v is given by

$$\Phi_{nij}^{(1)}(v) = n^{-1/2}\int_0^\infty \phi_{ij}^{(1)}(s;v)dM_n(s;v,\theta_0) \\ -\int_0^\infty \phi_i(s;v)\psi_j(s;v,\theta_0)(1-H_{n-}(s))d\Lambda(s;v,\theta_0). \qquad (4.48)$$

Furthermore, the second order partial derivative of $\Phi_{ni}(v)$ with respect to the j^{th} and k^{th} components of v is given by

$$\Phi_{nijk}^{(2)}(v) = n^{-1/2}\int_0^\infty \phi_{ijk}^{(2)}(s;v)dM_n(s;v,\theta_0) \\ -\int_0^\infty \{\phi_i(s;v)\{\psi_{jk}^{(1)}(s;v,\theta_0)+\psi_j(s;v,\theta_0)\psi_k(s;v,\theta_0)\} \\ \phi_{ij}^{(1)}(s;v)\psi_k(s;v,\theta_0)\}(1-H_{n-}(s))d\Lambda(s;v,\theta_0), \qquad (4.49)$$

where $\phi_{ijk}^{(2)}(t;v)$ is the second order partial derivative of $\phi_i(t;v)$ with respect to the j^{th} and the k^{th} components of v, and $\psi_{jk}^{(1)}(t;v)$ is the first order partial derivative of $\psi_j(t;v)$ with respect to the k^{th} component of v.

Let $\Phi_n^{(1)}(v)$ denote the $r\times r$ matrix with elements $\Phi_{nij}^{(1)}(v)$, and $\Phi_{ni}^{(2)}(v)$ the $r\times r$ matrix with elements $\Phi_{nijk}^{(2)}(v)$. By making essentially the same expansion as in Borgan (1984), we obtain

$$\Phi_{ni}(v) = \Phi_{ni}(v_{n0}) + \Phi_{ni}^{(1)}(v_{n0})(v - v_{n0}) + (v - v_{n0})^T \Phi_{ni}^{(2)}(v')(v - v_{n0})/2 \quad (4.50)$$

for every $v \in \Upsilon$. Here $\Phi_{ni}^{(1)}(v_{n0})$ is the i^{th} row of $\Phi_n^{(1)}(v_{n0})$, and v' a point on the line segment between v and v_{n0}.

Now define constants c_{58}–c_{60} by $c_{58} = 4(1 + c_\alpha + 1/\alpha)$, $c_{59} = c_{35}/6c_{58}r^3$, $c_{60} = ((c_\psi)^2 c_{35}c_{59}/6r) \wedge (c_\psi c_{35}/2r^2) \wedge c_{58}$, and let V_n denote the closed ball in Υ with centre v_{n0} and radius c_{59}. Moreover, denote the empirical processes $n^{1/2}\{H_n^1(t) - H^1(t; v_n, \theta_n)\}$ and $n^{1/2}\{H_{n-}(t) - H(t; v_n, \theta_n)\}$ by $U_n^1(t; v_n, \theta_n)$ and $U_{n-}(t; v_n, \theta_n)$, respectively. The random variable S_n is defined by

$$\begin{aligned} S_n = (2/c_{60})\{ &\sup_{t\in[0,\infty)} |U_n^1(t; v_n, \theta_n)| \\ &+ \sup_{t\in[0,\infty)} \sup_{v\in V_n} \int_0^t |U_{n-}(s; v_n, \theta_n)| d\Lambda(s; v, \theta_0)\}. \end{aligned} \quad (4.51)$$

Observe that for functions f, g which are bounded by 1 and have total variation not exceeding 1, we have for every $v \in V_n$

$$\begin{aligned} |n^{-1/2} &\int_0^\infty f(s) dM_n(s; v, \theta_0) - \int_0^\infty g(s)(1 - H_{n-}(s)) d\Lambda(s; v, \theta_0) \\ &- \int_0^\infty \{f(s) + g(s)\} dD(s; v, \theta_0, v_n, \theta_n) \\ &+ \int_0^\infty g(s) dH^1(s; v_n, \theta_n)| \\ \le\ & |n^{-1/2} \int_0^\infty g(s) dU_n^1(s; v_n, \theta_n)| \\ &+ |n^{-1/2} \int_0^\infty \{f(s) + g(s)\} U_{n-}(s; v_n, \theta_n) d\Lambda(s; v, \theta_0)| \\ \le\ & c_{60} n^{-1/2} S_n, \end{aligned}$$

where the last inequality follows from integration by parts. Hence, by Condition 9, Condition 10 and Condition 11 we have

$$n^{1/2}(c_\psi)^2 |\Phi_{ni}(v_{n0})| \le c_{60} S_n, \quad (4.52)$$

$$n^{1/2} c_\psi |\Phi_{nij}^{(1)}(v_{n0}) + \Xi_{0nij}| \le c_{60} S_n, \quad (4.53)$$

and for all $v \in V_n$

$$n^{1/2} |\Phi_{nijk}^{(2)}(v) + A_{nijk}(v)| \le c_{60} S_n,$$

where $A_{nijk}(v)$ is defined by

$$A_{nijk}(v) = \int_0^\infty \{\phi_i(s;v)\{\psi_{jk}^{(1)}(s;v,\theta_0) + \psi_j(s;v,\theta_0)\psi_k(s;v,\theta_0)\}$$
$$+\phi_{ij}^{(1)}(s;v)\psi(s;v,\theta_0)\}dH^1(s;v_n,\theta_n)$$
$$-\int_0^\infty \{\phi_i(s;v)\{\psi_{jk}^{(1)}(s;v,\theta_0) + \psi_j(s;v,\theta_0)\psi_k(s;v,\theta_0)\}$$
$$+\phi_{ijk}^{(2)}(s;v) + \phi_{ij}^{(1)'}(s;v)\psi_k(s;v,\theta_0)\}dD(s;v,\theta_0,v_n,\theta_n).$$

Note that $|A_{nijk}(v)| \leq c_{58}$ for all $v \in \Upsilon$.

Assume that (4.7) holds. It follows that

$$|\Phi_{ni}(v_{n0})| < c_{35}c_{59}/6r, \tag{4.54}$$

$$|\Phi_{nij}^{(1)}(v_{n0}) + \Xi_{0nij}| < c_{35}/2r^2, \tag{4.55}$$

and for every $v \in V_n$

$$|\Phi_{nijk}^{(2)}(v)| < 2c_{58}. \tag{4.56}$$

As a consequence of (4.55) and Condition 12 we have

$$|\Phi_n^{(1)}(v_{n0})(v - v_{n0})| > c_{35}|v - v_{n0}|/2. \tag{4.57}$$

We shall now prove the existence of a solution in V_n to the equations (4.3). Let $R_n(v)$ be the r-dimensional function defined by

$$R_n(v) = v - \{\Phi_n^{(1)}(v_{n0})\}^{-1}\Phi_n(v),$$

where $\Phi_n(v)$ is the r-dimensional function with elements $\Phi_{ni}(v)$. Then it suffices to show the existence of $v^{(n)} \in V_n$ such that $v^{(n)} = R_n(v^{(n)})$.

For every $v \in V_n$ we have for some v' on the line segment between v and v_{n0}

$$|R_n(v) - v_{n0}| < 2|\Phi_n^{(1)}(v_{n0})(v - v_{n0}) - \Phi_n(v)|/c_{35}$$
$$< \sum_{i=1}^r \{2|\Phi_{ni}(v_{n0})| + \sum_{j=1}^r\sum_{k=1}^r |\Phi_{nijk}^{(2)}(v')|(|v - v_{n0}|)^2\}/c_{35}$$
$$< 2c_{59}/3, \tag{4.58}$$

and hence $R_n(v)$ maps V_n into V_n. Furthermore, for every $v, v^* \in V_n$ there exist points v', v'' on the line segment between v and v^* such that

$$|R_n(v) - R_n(v^*)| = |\{\Phi_n^{(1)}(v_{n0})\}^{-1}\{\Phi_n^{(1)}(v_{n0}) - \Phi_n^{(1)}(v')\}(v - v^*)|$$

$$\leq 2|\{\Phi_n^{(1)}(v_{n0}) - \Phi_n^{(1)}(v')\}(v - v^*)|/c_{35}$$

$$\leq 2\{\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r}|\Phi_{nijk}^{(2)}(v'')|\}(|v - v^*|)^2/c_{35}$$

$$\leq 2|v - v^*|/3. \tag{4.59}$$

Set s_0 equal to v_{n0}, and define the sequence $\{s_i\}_{i=1}^{\infty}$ recursively by $s_i = R_n(s_{i-1})$. From (4.58) and (4.59) it follows that this sequence converges to a point $v^{(n)} \in V_n$ which satisfies $v^{(n)} = R_n(v^{(n)})$.

Next we show (4.8) and (4.9). From (4.50), (4.56) and the fact that $\Phi_{ni}(v^{(n)}) = 0$ for $i = 1, \dots, r$, we may derive

$$|\Phi_n(v_{n0}) - \Xi_{0n}(v^{(n)} - v_{n0})| \leq \xi_n |v^{(n)} - v_{n0}|, \tag{4.60}$$

where

$$\xi_n = c_{58}r^3|v^{(n)} - v_{n0}| + \sum_{i=1}^{r}\sum_{j=1}^{r}|\Phi_{nij}^{(1)}(v_{n0}) + \Xi_{0nij}|. \tag{4.61}$$

Since $|\Xi_{0n}(v^{(n)} - v_{n0})| > c_{35}|v^{(n)} - v_{n0}|$ and $\xi_n \leq 3c_{35}/4$, we obtain

$$|v^{(n)} - v_{n0}| \leq 4|\Phi_n(v_{n0})|/c_{35},$$

and hence it follows from (4.52) that choosing $c_{36} = 4c_{60}r/c_{35}(c_\psi)^2$ yields (4.8). Combining (4.8), (4.53) and (4.61) gives

$$n^{1/2}\xi_n \leq \{c_{36}c_{58}r^3 + c_{60}r^2/c_\psi\}S_n, \tag{4.62}$$

which together with Condition 12, (4.8) and (4.60) leads to (4.9), with $c_{37} = c_{36}\{c_{36}c_{58}r^3 + c_{60}r^2/c_\psi\}/c_{35}$.

The remainder of this proof, the verification of (4.10) and (4.11), resembles part of the proof of Theorem 2. Let $\tilde{U}_n(t)$ be the empirical process based on the uniform (0,1) random variables $\tilde{Z}_1, \dots, \tilde{Z}_n$, defined by

$$\tilde{Z}_i = \delta_i H^1(Z_i; v_n, \theta_n) + (1 - \delta_i)\{H^1(\infty; v_n, \theta_n) + H^0(Z_i; v_n, \theta_n)\},$$

where

$$H^0(t; v_n, \theta_n) = H(t; v_n, \theta_n) - H^1(t; v_n, \theta_n)$$

is the cumulative distribution function of the censored failure times under P_n. Inequality 1 yields

$$P_n(\sup_{t\in[0,1]} |\tilde{U}_n(t)| > x) \leq 2c_1 \exp\{-2x^2\}. \tag{4.63}$$

Let $\{d_n\}_{n=1}^{\infty}$ be a sequence of points in $[0,\infty)$. Later in this proof we specify this sequence in two different ways, depending on whether we are checking (4.10) or (4.11). Since Condition 9 implies that for every $t \in [0,\infty)$ and $v \in V_n$

$$\lambda(t; v, \theta_0) \leq c_{61}\lambda(t; v_{n0}, \theta_0), \tag{4.64}$$

where $c_{61} = 1 + rc_{\psi}c_{59}$, it follows that

$$\Lambda(t; v, \theta_0) \leq c_{61}\Lambda(t; v_{n0}, \theta_0), \tag{4.65}$$

$$\begin{aligned}&\int_{d_n}^{\infty}(1 - H(s; v_n, \theta_n))d\Lambda(s; v, \theta_0)\\ &\qquad \leq c_{61}\int_{d_n}^{\infty}(1 - H(s; v_n, \theta_n))d\Lambda(s; v_{n0}, \theta_0).\end{aligned} \tag{4.66}$$

Furtherm· re, we may write

$$S_n \leq (2/c_{60})\{\Delta_{1n} + \Delta_{2n} + \Delta_{3n}\}, \tag{4.67}$$

where as consequences of the construction of $\tilde{U}_n(t)$ and (4.65) we have

$$\begin{aligned}\Delta_{1n} &= \sup_{t\in[0,\infty)} |U_n^1(t; v_{n0}, \theta_0)|\\ &\leq \sup_{t\in[0,1]} |\tilde{U}_n(t)|,\end{aligned} \tag{4.68}$$

$$\begin{aligned}\Delta_{2n} &= \sup_{t\in[0,d_n]} \sup_{v\in V_n} \int_0^t |U_{n-}(s; v_n, \theta_n)| d\Lambda(s; v, \theta_0)\\ &\leq \{\sup_{t\in[0,d_n]} |U_{n-}(s; v_n, \theta_n)|\}\{\sup_{v\in V_n} \Lambda(d_n; v, \theta_0)\}\\ &\leq 3c_{61}\Lambda(d_n; v_{n0}, \theta_0) \sup_{t\in[0,1]} |\tilde{U}_n(t)|,\end{aligned} \tag{4.69}$$

$$\Delta_{3n} = \sup_{t\in[d_n,\infty)} \sup_{v\in V_n} \int_{d_n}^t |U_{n-}(s; v_n, \theta_n)| d\Lambda(s; v, \theta_0). \tag{4.70}$$

Together with (4.68) and (4.69), inequality (4.63) yields

$$P_n(\Delta_{1n} + \Delta_{2n} > 4c_{61}\Lambda(d_n; v_{n0}, \theta_0)x) \leq 2c_1 \exp\{-2x^2\}. \tag{4.71}$$

Observing that (4.66) implies

$$\Delta_{3n} \leq n^{1/2}c_{61}\int_{d_n}^{\infty}(1 - H(s; v_n, \theta_n))d\Lambda(s; v_{n0}, \theta_0) \qquad \text{if } Z_{n:n} < d_n,$$

where $Z_{n:n}$ denotes the largest order statistic of the sample $Z_1, \dots, Z_n$, we obtain

$$P_n(\Delta_{3n} > n^{1/2} c_{61} \int_{d_n}^{\infty} (1 - H(s; v_n, \theta_n)) d\Lambda(s; v_{n0}, \theta_0))$$
$$\leq n(1 - H(d_n; v_n, \theta_n)). \tag{4.72}$$

Now choose $\beta > \alpha$, and d_n so as to satisfy $H(d_n; v_n, \theta_n) = 1 - n^{-(1+\gamma)}$, where $\gamma = (\beta - \alpha)/\alpha$. By Condition 8 we have

$$\Lambda(d_n; v_{n0}, \theta_0) < c_\alpha n^\beta,$$

$$\int_{d_n}^{\infty} (1 - H(s; v_n, \theta_n)) d\Lambda(s; v_{n0}, \theta_0) < c_\alpha n^{\beta - (\gamma + 1)},$$

it follows from (4.71) and (4.72) that

$$P_n(\Delta_{1n} + \Delta_{2n} > 4 c_\alpha c_{61} n^\beta \sqrt{(\gamma \log n)/2}) \leq 2 c_1 n^{-\gamma}, \tag{4.73}$$

$$P_n(\Delta_{3n} > c_\alpha c_{61} n^{\beta - (\gamma + 1/2)}) \leq n^{-\gamma}, \tag{4.74}$$

and hence (4.10).

Finally, assume $(v_n, \theta_n) = (v_0, \theta_0)$. Choose d_n so as to satisfy $H(d_n; v_0, \theta_0) = 1 - \exp\{-2x^2\}/n$, where $x > (n^{1/2} \log n)^{-1}$. By noting that

$$\Lambda(d_n; v_0, \theta_0) < \log n + 2x^2,$$

$$\int_{d_n}^{\infty} (1 - H(s; v_0, \theta_0)) d\Lambda(s; v_0, \theta_0) < \exp\{-2x^2\}/n$$
$$< (n^{-1/2} \log n) x,$$

we obtain from (4.71) and (4.72)

$$P_n(\Delta_{1n} + \Delta_{2n} > 4 c_{61} x (\log n + 2x^2)) < 2 c_1 \exp\{-2x^2\}, \tag{4.75}$$

$$P_{v_0}(\Delta_{3n} > c_{61} x \log n) \leq \exp\{-2x^2\}, \tag{4.76}$$

by which (4.11) readily follows. This concludes the proof of Theorem 7. □

Proof of Theorem 8 The proof is based on replacing stochastic terms in a second order expansion of $Q_n(t; v)$ by deterministic counterparts. The "approximated" expansion is applied to $Q_n(t; v^{(n)})$, and combined with (4.9) and Theorem 2.

The first order partial derivative of $Q_n(t;v)$ with respect to the i^{th} component of v is given by

$$Q_{ni}^{(1)}(t;v) = \int_0^t L_{ni}^{(1)}(s;v)dM_n(s;v,\theta_0)$$
$$-n^{1/2}\int_0^t L_n(s;v)\psi_i(s;v,\theta_0)(1-H_{n-}(s))d\Lambda(s;v,\theta_0), \qquad (4.77)$$

and the second order partial derivative of $Q_n(t;v)$ with respect to the i^{th} and j^{th} components of v is given by

$$Q_{nij}^{(2)}(t;v) = \int_0^t L_{nij}^{(2)}(s;v)dM_n(s;v,\theta_0)$$
$$-n^{1/2}\int_0^t \{L_n(s;v)\{\psi_{ij}^{(1)}(s;v,\theta_0) + \psi_i(s;v,\theta_0)\psi_j(s;v,\theta_0)\}$$
$$+L_{nj}^{(1)}(s;v)\psi_i(s;v,\theta_0)$$
$$+L_{ni}^{(1)}(s;v)\psi_j(s;v,\theta_0)\}(1-H_{n-}(s))d\Lambda(s;v,\theta_0). \qquad (4.78)$$

Here $L_{nij}^{(2)}(t;v)$ denotes the second order partial derivative of $L_n(t;v)$ with respect to the i^{th} and j^{th} components of v.

Now define

$$\tilde{Q}_{ni}^{(1)}(t;v) = n^{1/2}\int_0^t L_n(s;v)\psi_i(s;v,\theta_0)dH^1(s;v_n,\theta_n)$$
$$-n^{1/2}\int_0^t \{L_n(s;v)\psi_i(s;v,\theta_0)$$
$$+L_{ni}^{(1)}(s;v)\}dD(s;v,\theta_0,v_n,\theta_n), \qquad (4.79)$$

$$\tilde{Q}_{nij}^{(2)}(t;v) = n^{1/2}\int_0^t \{L_{nj}^{(1)}(s;v)\psi_i(s;v,\theta_0) + L_{ni}^{(1)}(s;v)\psi_j(s;v,\theta_0)$$
$$+L_n(s;v)\{\psi_{ij}^{(1)}(s;v,\theta_0)$$
$$+\psi_i(s;v,\theta_0)\psi_j(s;v,\theta_0)\}\}dH^1(s;v_n,\theta_n)$$
$$-n^{1/2}\int_0^t \{L_{nj}^{(1)}(s;v)\psi_i(s;v,\theta_0) + L_{ni}^{(1)}(s;v)\psi_j(s;v,\theta_0)$$
$$+L_n(s;v)\{\psi_{ij}^{(1)}(s;v,\theta_0) + \psi_i(s;v,\theta_0)\psi_j(s;v,\theta_0)\}$$
$$+L_{nij}^{(2)}(s;v)\}dD(s;v,\theta_0,v_n,\theta_n). \qquad (4.80)$$

By Condition 9, Condition 13 and integration by parts we obtain the existence of a constant c_{62}, not depending on v_n or θ_n, such that

$$\sup_{t\in[0,\infty)} |K_{0i}(t;v_n,\theta_n) - n^{-1/2}\tilde{Q}^{(1)}_{ni}(t;v_{n0})|$$
$$\leq c_{62}\{\sup_{t\in[0,\infty)} |L_n(t;v_{n0}) - L(t;v_n,\theta_n)|$$
$$+ \sup_{t\in[0,\infty)} |L^{(1)}_{ni}(t;v_{n0}) - L^{[1]}_i(t;v_n,\theta_n)|\}, \tag{4.81}$$

$$\sup_{t\in[0,\infty)} |Q^{(1)}_{ni}(t;v_{n0}) + \tilde{Q}^{(1)}_{ni}(t;v_{n0})| \leq c_{62}S_n, \tag{4.82}$$

and for every $v \in V_n$

$$\sup_{t\in[0,\infty)} |Q^{(2)}_{nij}(t;v) + \tilde{Q}^{(2)}_{nij}(t;v)| \leq c_{62}S_n. \tag{4.83}$$

Next we introduce the set Ω_n, consisting of all $\omega \in \Omega$ such that $S_n \leq n^{1/2}$ and that all random elements in $D[0,\infty)$ occurring in Condition 13 are bounded and have bounded variation as specified in this condition. By noting that there exists a constant c_{63}, not depending on v_n or θ_n, such that for $\omega \in \Omega_n$

$$\sup_{t\in[0,\infty)} |\tilde{Q}^{(2)}_{nij}(t;v)| \leq c_{63}n^{1/2},$$

it follows from (4.83) that for $\omega \in \Omega_n$ and $v \in V_n$

$$\sup_{t\in[0,\infty)} |Q^{(2)}_{nij}(t;v)| \leq (c_{62}+c_{63})n^{1/2},$$

and hence by (4.8)

$$\sup_{t\in[0,\infty)} |(v^{(n)} - v_{n0})^T Q^{(2)}_n(t;v)(v^{(n)} - v_{n0})| \leq c_{64}n^{-1/2}\{S_n\}^2, \tag{4.84}$$

where $Q^{(2)}_n(t;v)$ is the $r \times r$ matrix with elements $Q^{(2)}_{nij}(t;v)$ and $c_{64} = (c_{62} + c_{63})(c_{36}r)^2$.

Furthermore, letting $Q^{(1)}_n(t;v)$ denote the r-dimensional vector with elements $Q^{(1)}_{ni}(t;v)$, we obtain by (4.8), (4.81) and (4.82) for every $\omega \in \Omega_n$

$$\sup_{t\in[0,\infty)} |(v^{(n)} - v_{n0})^T\{Q^{(1)}_n(t;v_{n0}) + n^{1/2}K_0(t;v_n,\theta_n)\}|$$
$$\leq |v^{(n)} - v_{n0}|\sum_{i=1}^{r}\{\sup_{t\in[0,\infty)} |Q^{(1)}_{ni}(t;v_{n0}) + \tilde{Q}^{(1)}_{ni}(t;v_{n0})|$$
$$+ \sup_{t\in[0,\infty)} |\tilde{Q}^{(1)}_{ni}(t;v_{n0}) - n^{1/2}K_{0i}(t;v_n,\theta_n)|\}$$

$$\leq c_{36}c_{62}n^{-1/2}S_n \sum_{i=1}^{r} \{S_n$$

$$+n^{1/2} \sup_{t\in[0,\infty)} |L_n(t; v_{n0}) - L(t; v_n, \theta_n)|$$

$$+n^{1/2} \sup_{t\in[0,\infty)} |L_{ni}^{(1)}(t; v_{n0}) - L_i^{[1]}(t; v_n, \theta_n)|\}. \tag{4.85}$$

Since for every $v \in V_n$ and $t \in [0, \infty)$ we may find a point v' on the line segment between v and v_{n0} such that

$$Q_n(t; v) = Q_n(t; v_{n0}) + (v - v_{n0})^T Q_n^{(1)}(t; v_{n0})$$

$$+(v - v_{n0})^T Q_n^{(2)}(t; v')(v - v_{n0})/2, \tag{4.86}$$

it follows from (4.84) and (4.85) that for every $\omega \in \Omega_n$

$$\sup_{t\in[0,\infty)} |Q_n(t; v^{(n)}) - Q_n(t; v_{n0}) + n^{1/2}(v^{(n)} - v_{n0})^T K_0(t; v_n, \theta_n)|$$

$$\leq (c_{64} + c_{36}c_{62})n^{-1/2}S_n \sum_{i=1}^{r} \{S_n$$

$$+n^{1/2} \sup_{t\in[0,\infty)} |L_n(t; v_{n0}) - L(t; v_n, \theta_n)|$$

$$+n^{1/2} \sup_{t\in[0,\infty)} |L_{ni}^{(1)}(t; v_{n0}) - L_i^{[1]}(t; v_n, \theta_n)|\}.$$

[here we have used $v^{(n)} \in V_n$ if $\omega \in \Omega_n$]. Hence, by (4.9) and the fact that $K_0(t; v, \theta)$ remains bounded, uniformly in v and θ [say by c_{65}], we obtain

$$\sup_{t\in[0,\infty)} |Q_n(t; v^{(n)}) - Q_n(t; v_{n0}) + (K_0(t; v_n, \theta_n))^T \Xi_{0n}^{-1}$$

$$\int_0^\infty \phi(s; v_{n0})\{dM_n(s; v_{n0}, \theta_0) - dD(s; v_n, \theta_n)\}|$$

$$\leq (c_{64} + c_{36}c_{62} + c_{37}c_{65})n^{-1/2}S_n \sum_{i=1}^{r} \{S_n$$

$$+n^{1/2} \sup_{t\in[0,\infty)} |L_n(t; v_{n0}) - L(t; v_n, \theta_n)|$$

$$+n^{1/2} \sup_{t\in[0,\infty)} |L_{ni}^{(1)}(t; v_{n0}) - L_i^{[1]}(t; v_n, \theta_n)|\} \tag{4.87}$$

for every $\omega \in \Omega_n$. Now, (4.19) follows from (4.10), Condition 13 and by applying Theorem 2 to $Q_n(t; v_{n0})$ and $\int_0^\infty \phi(s; v_{n0})dM_n(s; v_{n0}, \theta_0)$ separately.

For the special case $(v_n, \theta_n) = (v_0, \theta_0)$, we may replace (4.81) by

$$\sup_{t\in[0,\infty)} |K_{0i}(t; v_n, \theta_n) - n^{1/2}\tilde{Q}_{ni}^{(1)}(t; v_{n0})|$$
$$\leq c_{62} \sup_{t\in[0,\infty)} |L_n(t; v_{n0}) - L(t; v_n, \theta_n)|, \tag{4.88}$$

and hence (4.87) simplifies to

$$\sup_{t\in[0,\infty)} |Q_n(t; v^{(n)}) - Q_n(t; v_{n0})$$
$$+(K_0(t; v_0, \theta_0))^T \Sigma_{e0}^{-1} \int_0^\infty \phi(s; v_0) dM_n(s; v_0, \theta_0)|$$
$$\leq (c_{64} + c_{36}c_{62} + c_{37}c_{65}) n^{-1/2} S_n \sum_{i=1}^{r} \{S_n$$
$$+n^{1/2} \sup_{t\in[0,\infty)} |L_n(t; v_0) - L(t; v_0, \theta_0)|\} \tag{4.89}$$

for every $\omega \in \Omega_n$. Applying Theorem 2 twice, together with Condition 13 and (4.11) yields (4.20). This concludes the proof of Theorem 8. □

Proof of Theorem 9 Equations (4.26) and (4.27) are obtained along the lines of the proof of Theorem 5.2 in Borell (1975).

To prove (4.28) and (4.29), define the mean zero Gaussian process $\check{X}_n(t; v_n, \theta_n)$ by

$$\check{X}_n(t; v_n, \theta_n) = \int_0^t L(s; v_n, \theta_n) dW_n(s; v_{n0})$$
$$-(K_0(t; v_{n0}, \theta_0))^T \Xi_{0n}^{-1} \int_0^\infty \phi(s; v_{n0}) dW_n(s; v_{n0}),$$

Observe that $\check{X}_n(t; v_0, \theta_0)$ is equal in P_{v_0}-distribution to $\check{X}(t; v_0, \theta_0)$. Moreover, by Theorem 8 we have

$$P_{v_0}(\sup_{t\in[0,\infty)} |Q_n(t; v^{(n)}) - \check{X}_n(t; v_0, \theta_0)| > n^{-c_{28}}(c_{40} \log n + x)^{c_{29}})$$
$$\leq c_{41} \exp\{-c_{42}x\}, \tag{4.90}$$

with $c_{28} = 1/6$ and $c_{29} = 3$. Let $c_{30} = c_{28}/(2c_{29} - 1)$ and choose $c_{29}^{-1} < \beta < 2$. By the same methods as used in the proof of Theorem 3 one may show

$$P_{v_0}(|T(Q_n(\cdot; v^{(n)})) - T(\check{X}_n(\cdot; v_0, \theta_0))| > c_T n^{c_{30}(1-c_{29}\beta)} (s_n)^{c_{29}\beta})$$
$$\ll \exp\{- \inf_{v_0 \in \Upsilon} \check{a}(v_0)(s_n)^2/2\}. \tag{4.91}$$

Now (4.28) and (4.29) follow from bounding $P_{v_0}(T(Q_n(\cdot; v^{(n)})) > s_n)$ between

$$P_{v_0}(T(\check{X}(\cdot; v_0, \theta_0)) > s_n(1 + c_T(n^{-c_{30}} s_n)^{c_{29}\beta - 1}))$$
$$-P_{v_0}(T(\check{X}_n(\cdot; v_0, \theta_0)) - T(Q_n(\cdot; v^{(n)}))| > c_T n^{c_{30}(1-c_{29}\beta)}(s_n)^{c_{29}\beta})$$

and

$$P_{v_0}(T(\check{X}(\cdot; v_0, \theta_0)) > s_n(1 - c_T(n^{-c_{30}} s_n)^{c_{29}\beta - 1}))$$
$$+P_{v_0}(T(\check{X}_n(\cdot; v_0, \theta_0)) - T(Q_n(\cdot; v^{(n)}))| > c_T n^{c_{30}(1-c_{29}\beta)}(s_n)^{c_{29}\beta}).$$

This concludes the proof of Theorem 9. □

To prove Theorems 10 and 12 we need the following lemma, a "composite null hypothesis" version of Lemma 2.

Lemma 3 *Let $g(t; v, \theta)$ be a real valued function, $g_i^{(1)}(t; v, \theta)$ the first order partial derivative with respect to the i^{th} component of (v, θ). Suppose there exists a constant c_{66} such that*

$$\sup_{t\in[0,\infty)} (1 - H(t; v, \theta))^{\frac{1}{2}} |g(t; v, \theta)| \le c_{66}, \tag{4.92}$$

$$\sup_{t\in[0,\infty)} (1 - H(t; v, \theta))^{1-\alpha} |g_i^{(1)}(t; v, \theta)| \le c_{66} \tag{4.93}$$

for every $v \in \Upsilon$ and $\theta \in \Theta$. Then there exists a constant c_{67} such that

$$\sup_{t\in[0,\infty)} |\int_0^t g(s; v, \theta) dD(s; v, \theta)| \le c_{67}|\theta - \theta_0|, \tag{4.94}$$

$$\sup_{t\in[0,\infty)} |\int_0^t g(s; v, \theta) dH^1(s; v, \theta)$$
$$- \int_0^t g(s; \pi(v, \theta), \theta_0) dH^1(s; \pi(v, \theta), \theta_0)|$$
$$\le c_{67}|\theta - \theta_0|. \tag{4.95}$$

for every $v \in \Upsilon$ and $\theta \in \Theta$. Let $g_{ij}^{(2)}(t; v, \theta)$ the second order partial derivative of $g(t; v, \theta)$ with respect to the i^{th} and j^{th} components of (v, θ). If

$$\sup_{t\in[0,\infty)} (1 - H(t; v, \theta))^{\alpha} |g(t; v, \theta)| \le c_{66}, \tag{4.96}$$

$$\sup_{t\in[0,\infty)} (1 - H(t;v,\theta))^{\frac{1}{2}} |g_i^{(1)}(t;v,\theta)| \leq c_{66}, \tag{4.97}$$

$$\sup_{t\in[0,\infty)} (1 - H(t;v,\theta))^{1-\alpha} |g_{ij}^{(2)}(t;v,\theta)| \leq c_{66} \tag{4.98}$$

then there exists an r-dimensional function $K^g_{a|0}(t;v)$ *such that*

$$\sup_{t\in[0,\infty)} |\int_0^t g(s;v,\theta)dD(s;v,\theta) - (\theta-\theta_0)^T K^g_{a|0}(t;\pi(v,\theta))| \leq c_{67}|\theta-\theta_0|^2 \tag{4.99}$$

for every $v \in \Upsilon$ *and* $\theta \in \Theta$.

Proof of Lemma 3 Let the $\Xi_a(v,\theta)$ be the $r \times p$ matrix with elements

$$\Xi_{aij}(v,\theta) = \int_0^\infty \phi_i(s;\pi(v,\theta))\psi_{r+j}(s;\pi(v,\theta),\theta_0)dH^1(s;v,\theta) + \int_0^\infty \psi_i(s;\pi(v,\theta))\Lambda^{(1)}_{r+j}(s;\pi(v,\theta),\theta_0)dD(s;v,\theta), \tag{4.100}$$

and define $K^g_{a|0}(t;v)$ by

$$K^g_{a|0}(t;v) = \int_0^\infty g(s;v,\theta_0)\{\psi_a(s;v,\theta_0) - (\{\Xi_0(v,\theta_0)\}^{-1}\Xi_a(v,\theta_0))^T\psi(s;v,\theta_0)\}dH^1(s;v,\theta_0). \tag{4.101}$$

The Implicit Function Theorem tells us that the $r\times r$ matrix of partial derivatives of $\pi(v,\theta)$ with respect to the components of v equals

$$-\{\Xi_0(v,\theta)\}^{-1}\Xi_b(v,\theta),$$

and that the $r \times p$ matrix of partial derivatives of $\pi(v,\theta)$ with respect to the components of θ equals

$$-\{\Xi_0(v,\theta)\}^{-1}\Xi_a(v,\theta).$$

Now fix $v' \in \Upsilon$ and define the function $\kappa : \Theta \to \Upsilon$ by

$$\pi(\kappa(\theta),\theta) = v'.$$

A second application of the Implicit Function Theorem yields that the $r \times p$ matrix of partial derivatives of κ with respect to the components of θ equals

$$-\{\Xi_b(\pi(v,\theta),\theta_0)\}^{-1}\Xi_a(\pi(v,\theta),\theta_0).$$

The elements of this matrix and the partial derivatives of these elements with respect to the components of θ are uniformly bounded in v', v and θ.

Let $\tilde{g}(t;v)$ be defined as $g(t;\kappa(\theta),\theta)$, let $\tilde{g}_i^{(1)}(t;v)$ be the first order partial derivative with respect to the i^{th} component of θ, and let $\tilde{g}_{ij}^{(2)}(t;v)$ denote the second order partial derivative of $\tilde{g}(t;v)$ with respect to the i^{th} and j^{th} components of θ. We have

$$\sup_{t\in[0,\infty)}(1-H(t;\pi(v,\theta),\theta_0))^{\frac{1}{2}}|\tilde{g}(t;v)|\leq c_{68},$$

$$\sup_{t\in[0,\infty)}(1-H(t;\pi(v,\theta),\theta_0))^{1-\alpha}|\tilde{g}_i^{(1)}(t;v)|\leq c_{68}$$

for some constant c_{68}. Moreover, if (4.96)-(4.98) hold, then we have

$$\sup_{t\in[0,\infty)}(1-H(t;\pi(v,\theta),\theta_0))^{\alpha}|\tilde{g}(t;v)|\leq c_{68},$$

$$\sup_{t\in[0,\infty)}(1-H(t;\pi(v,\theta),\theta_0))^{\frac{1}{2}}|\tilde{g}_i^{(1)}(t;v)|\leq c_{68},$$

$$\sup_{t\in[0,\infty)}(1-H(t;\pi(v,\theta),\theta_0))^{1-\alpha}|\tilde{g}_{ij}^{(2)}(t;v)|\leq c_{68}.$$

Hence we are in the position to apply Lemma 2. By noting that c_{68} may be chosen independently of v' our lemma follows. □

Proof of Theorem 10 Let $\{W_n(t;v_{n0})\}_{n=1}^{\infty}$ be the sequence of mean zero Gaussian processes given in Theorem 8. We may write

$$\sup_{t\in[0,\infty)}|Q_n(t;v^{(n)})-n^{1/2}|\theta_n-\theta_0|h^TK_{a|0}(t;v_0)|$$
$$\leq\Delta_{1n}+\Delta_{2n}+\Delta_{3n}+\Delta_{4n}+\Delta_{5n},$$

where

$$\Delta_{1n}=\sup_{t\in[0,\infty)}|\{Q_n(t;v^{(n)})-n^{1/2}\int_0^t L_n(s;v_{n0})dD(s;v_n,\theta_n)\}$$
$$-\{\int_0^t L(s;v_n,\theta_n)dW_n(s;v_{n0})$$
$$-(K_0(t;v_n,\theta_n))^T\Xi_{0n}^{-1}\int_0^{\infty}\phi(s;v_{n0})dW_n(s;v_{n0})\}|,$$

$$\Delta_{2n}=\sup_{t\in[0,\infty)}|\int_0^t L(s;v_n,\theta_n)dW_n(s;v_{n0})$$
$$-(K_0(t;v_n,\theta_n))^T\Xi_{0n}^{-1}\int_0^{\infty}\phi(s;v_{n0})dW_n(s;v_{n0})|,$$

$$\Delta_{3n}=n^{1/2}\sup_{t\in[0,\infty)}|\int_0^t\{L_n(s;v_{n0})-L(s;v_n,\theta_n)\}dD(s;v_n,\theta_n)|$$

$$\Delta_{4n} = n^{1/2} \sup_{t\in[0,\infty)} | \int_0^t L(s; v_n, \theta_n) dD(s; v_n, \theta_n)$$
$$-|\theta_n - \theta_0| h^T K_{a|0}(t; v_{n0})|,$$

$$\Delta_{5n} = n^{1/2}|\theta_n - \theta_0| \sup_{t\in[0,\infty)} \{h^T K_{a|0}(t; v_{n0}) - h^T K_{a|0}(t; v_0)\}.$$

By (4.19) we have for $\beta < (1/2 - 2\alpha) \wedge 1/6$

$$P_n(\Delta_{1n} > c_{50} n^{-\beta}) \leq c_{51} n^{-c_{52}}. \tag{4.102}$$

Recall that $K_{0i}(t; v, \theta)$ is bounded by c_{65}. It follows that the r components of $(\Xi_{0n}^{-1})^T K_0(t; v_n, \theta_n)$ are bounded by $c_{35}c_{65}$, and hence we obtain by integration by parts with P_n-probability 1

$$\Delta_{2n} \leq 2(c_{43} + c_{35}c_{65}r) \sup_{t\in[0,\infty)} |W_n(t; v_{n0})|$$

The variance function of the process $W_n(t; v_{n0})$ is bounded by $(1 + 4(c_\alpha + 1/\alpha))^2$, and thus Inequality 4 leads to

$$P_n(\Delta_{2n} > c_{69}x) \leq c_{37} \exp\{-c_{38}x^2\}, \tag{4.103}$$

where $c_{69} = 2(c_{43} + c_{35}c_{65}r)(1 + 4(c_\alpha + 1/\alpha))$. By observing that $K_{a|0}^g(t; v)$ equals $K_{a|0}(t; v)$ if $g(t; v, \theta)$ is equal to $L(t; v, \theta)$, and that the components of $K_{a|0}(t; v)$ remain bounded, it follows from Lemma 3 that there exists a constant c_{70} such that

$$\Delta_{4n} \leq c_{70} n^{1/2} \{|\theta_n - \theta_0|^2 + |(\theta_n - \theta_0) - (|\theta_n - \theta_0| h)|\}, \tag{4.104}$$

and for every $v \in \Upsilon$, $\theta \in \Theta$ and $\beta > 1/2$

$$P_n(n^\beta \sup_{t\in[0,\infty)} \int_0^t (L_n(s; v_{n0}) - L(s; v_n, \theta_n)) dD(s; v, \theta) > c_{70}|\theta - \theta_0|)$$
$$\leq P_n(n^\beta \sup_{t\in[0,\infty)} |L_n(t; v_{n0}) - L(t; v_n, \theta_n)| > c_{44}).$$

Note that by Condition 13 this last inequality implies

$$P_n((n^{1/2}|\theta_n - \theta_0|)^{-1} \Delta_{3n} > c_{70} n^{-\beta}) \leq c_{45} n^{-c_{46}}. \tag{4.105}$$

Finally, there exists a constant c_{71} such that

$$\Delta_{5n} \leq c_{71} n^{1/2} (|\theta_n - \theta_0|)(|v_n - v_0|). \tag{4.106}$$

The first part of the theorem is now easily proved by combining the inequality

$$|(n^{1/2}|\theta_n - \theta_0|)^{-1}T(Q_n(t;v^{(n)})) - T(h^T K_{a|0}(\cdot;v_0))|$$
$$\leq c_T(n^{1/2}|\theta_n - \theta_0|)^{-1}(\Delta_{1n} + \Delta_{2n} + \Delta_{3n} + \Delta_{4n} + \Delta_{5n}).$$

with (4.102)-(4.106).

In the proof of Theorem 9 we introduced the mean zero Gaussian process $\check{X}_n(t;v_n,\theta_n)$. Observe that

$$\sup_{t\in[0,\infty)} |Q_n(t;v^{(n)}) - \{\check{X}_n(t;v_n,\theta_n) + n^{1/2}|\theta_n - \theta_0|h^T K_{a|0}(t;v_{n0})\}|$$
$$\leq \Delta_{1n} + \Delta_{3n} + \Delta_{4n}.$$

Since the right hand side of this inequality converges to zero in P_n-probability because σ is finite, we may prove the second part of the theorem by showing that the process $\check{X}_n(t;v_n,\theta_n) + n^{1/2}|\theta_n - \theta_0|h^T K_{a|0}(t;v_{n0})$ converges in P_n-distribution to $\check{X}(t;v_0,\theta_0) + \sigma h^T K_{a|0}(t;v_0)$. This boils down to verifying the convergence in P_n-distribution of $\check{X}_n(t;v_n,\theta_n)$ to $\check{X}(t;v_0,\theta_0)$, which can be done by using Lemma 3 to check the conditions of Theorem VI.10 in Pollard (1984). This completes the proof of Theorem 10. □

Proof of Theorem 12 Let Δ_{1n}, Δ_{2n}, Δ_{3n} and Δ_{4n} be as in the proof of Theorem 10, and define $b(v_n,\theta_n)$ as $T(\int_0^{\cdot} L(s;v_n,\theta_n)dD(s;v_n,\theta_n))$. We may write

$$|n^{-1/2}T(Q_n(\cdot;v^{(n)})) - b(v_n,\theta_n)| \leq c_T n^{-1/2}\{\Delta_{1n} + \Delta_{2n} + \Delta_{3n}\}. \tag{4.107}$$

Furthermore, we have

$$|b(v_n,\theta_n) - |\theta_n - \theta_0|T(h^T K_{a|0}(\cdot;v_{n0}))| \leq c_T n^{-1/2}\Delta_{4n}. \tag{4.108}$$

Since $\inf_{v_0\in\Upsilon} e(h;v_0)$ is not equal to zero, this yields the existence of positive constants ϵ^* and c_{72} such that

$$c_{72}|\theta_n - \theta_0| < b(v_n,\theta_n) < 1 \tag{4.109}$$

for θ_n satisfying $|\theta_n - \theta_0| < \epsilon^*$.

Now suppose $\theta \in \Theta - \{\theta_0\}$ satisfies $|\theta_n - \theta_0| < \epsilon^*$, and set (v_n,θ_n) equal to (v,θ) for every $n \in \mathbb{N}$. Choose $\epsilon > 0$ and $\delta \in (0,1)$. By (4.102) and (4.103) it follows that there exists an integer N_1 not depending on v or θ such that for $n > N_1$

$$P_{(v,\theta)}(\Delta_{1n} > (N_1)^{1/2}\epsilon/4c_T) < \delta/4,$$
$$P_{(v,\theta)}(\Delta_{2n} > (N_1)^{1/2}\epsilon/4c_T) < \delta/4.$$

Hence, for $n > N_1/(b(v,\theta))^2$ we have

$$P_{(v,\theta)}(n^{-1/2}\{\Delta_{1n} + \Delta_{2n}\} > \epsilon b(v,\theta)/2c_T) < \delta/2, \tag{4.110}$$

since $n > N_1$ and $(N_1/n)^{1/2} < b(v,\theta)$. Moreover, (4.105) implies the existence of an integer $N > N_1$ not depending on v or θ such that for $n > N$

$$\begin{aligned} &P_{(v,\theta)}(\Delta_{3n} > \epsilon b(v,\theta)/2c_T) \\ &\leq P_{(v,\theta)}((n^{1/2}|\theta_n - \theta_0|)^{-1}\Delta_{3n} > c_{72}\epsilon) \\ &< \delta/2. \end{aligned} \tag{4.111}$$

Combining (4.107)-(4.111) now yields that $\{T(Q_n(\cdot; v^{(n)}))\}_{n=1}^{\infty}$ is indeed a Wieand sequence.

Finally, (4.35) immediately follows from (4.104) and (4.108). This concludes the proof of Theorem 12. □

Proof of Theorem 13 Observe that for any r-dimensional vector η and every $t \in [0,\infty) \cup \{\infty\}$

$$\begin{aligned} &\int_0^\infty \eta^T \psi(s; v_0, \theta_0) L_0(s,t; v_0) dH^1(s; v_0, \theta_0) \\ &= \eta^T K_0(t; v_0, \theta_0) - \eta^T \Sigma_{0e} \Sigma_{0e}^{-1} K_0(t; v_0, \theta_0) \\ &= 0. \end{aligned}$$

By choosing η equal to $I_{00}^{-1} I_{0a} h$ it follows that we may write $h^T K_{a|0}(t; v_0)$ as

$$\int_0^\infty h^T \{\psi_a(s; v_0, \theta_0) - I_{a0} I_{00}^{-1} \psi(s; v_0, \theta_0)\} L_0(s,t; v_0) dH^1(s; v_0, \theta_0),$$

and hence we have by the Cauchy-Schwarz inequality that $\{h^T K_{a|0}(t; v_0)\}^2$ is bounded by

$$h^T \{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\} h \int_0^\infty (L_0(s,t; v_0))^2 dH^1(s; v_0, \theta_0)$$

for every $t \in [0,\infty) \cup \{\infty\}$. This immediately yields that both $e_R(h, v_0)$ and $e_S(h, v_0)$ are bounded by $h^T \{I_{aa} - I_{a0} I_{00}^{-1} I_{0a}\} h$. It is easily established that this upper bound is achieved by generalized rank tests based on (4.44) and by supremum type tests based on maximum likelihood estimation and (4.45). □

Chapter 5

Testing exponentiality

5.1 The exponential distribution

If E_i is a random variable having an exponential distribution with mean e^{-v} then it follows that

$$P(E_i > t) = \exp\{-e^v t\}.$$

The parametrization chosen here is fairly uncommon. But, as we shall see later, for our purposes it is rather convenient. The standard exponential distribution is worth special mention. Here v equals zero.

It is easily seen that the cumulative hazard function belonging to the exponential distribution is linear with intercept 0 and slope e^v. Note that the corresponding basic martingale is a spline.

In this chapter the standard exponential distribution is used to generate samples $X_1, \ldots, X_n$ under the null hypothesis. Moreover, when investigating the effect of censoring various exponential distributions are used to generate samples $Y_1, \ldots, Y_n$. In this respect note that if X_i has a standard exponential distribution and Y_i an exponential distribution with mean e^{-v}, then the probability that X_i is censored by Y_i is given by

$$P(X_i > Y_i) = \frac{e^v}{1 + e^v}.$$

5.2 Harrington and Fleming alternatives

In the results of previous chapters the functions $\psi(t; \theta_0)$ and $\psi(t; v_0, \theta_0)$ show up regularly. Hence, the applicability of these results depend heavily on the complexity of these functions. In Harrington and Fleming (1982) a family of distributions indexed by a single location parameter is constructed such that

$$\psi(t; \theta_0) = (1 - F(t; \theta_0))^\rho \tag{5.1}$$

for some fixed $\rho \geq 0$. The distributions within this family correspond to random variables which are allowed to take values anywhere on the real line. In this section we study an extended Harrington and Fleming family, indexed by two parameters, and always containing the exponential distribution as special case.

5.2.1 Derivation

The fact that we are dealing with a location parameter enables us to write

$$\lambda(t;\theta) = \lambda(t-\theta+\theta_0;\theta_0)$$

for all $t \in \mathbb{R}$. Hence, we have

$$\psi(t;\theta_0) = \frac{\lambda'(t;\theta_0)}{\lambda(t;\theta_0)},$$

where $\lambda'(t;\theta_0)$ denotes the derivative of $\lambda(t;\theta_0)$. Thus, to find a family such that (5.1) holds, we must solve the equation

$$\frac{\lambda'(t;\theta_0)}{\lambda(t;\theta_0)} = (1-F(t;\theta_0))^\rho, \tag{5.2}$$

which is actually a second order differential equation in disguise, as can be noted by writing $1-F(t;\theta_0) = \exp\{-\Lambda(t;\theta_0)\}$. Let us for a moment restrict ourselves to positive values of ρ. Then the solution to the differential equation (5.2) is given by

$$F(t;\theta_0) = 1-(1+\rho e^t)^{-1/\rho}. \tag{5.3}$$

Note that for $\rho = 1$ we obtain the logistic distribution.

Unfortunately, the Harrington and Fleming distributions do not fit into our framework, since they put positive mass on the negative part of the real line. Nevertheless, they indicate the way to construct other families of distributions better suited for our purposes. Observe that (5.3) implies

$$1-F(t;\theta) = [1+((1-F(t;\theta_0))^{-\rho}-1)e^{\theta-\theta_0}]^{-1/\rho}. \tag{5.4}$$

Replacing $F(t;\theta_0)$ by the distribution of an exponential random variable with mean e^{-v}, we obtain from (5.4)

$$1-F(t;v,\theta) = [1+(\exp\{-\rho e^v t\}-1)e^{\theta-\theta_0}]^{-1/\rho} \tag{5.5}$$

Now let us derive the Harrington and Fleming distribution for ρ equal to zero. In this case the solution to (5.2) is given by

$$F(t;\theta_0) = 1-\exp\{-e^t\},$$

which implies

$$1 - F(t;\theta) = (1 - F(t;\theta_0))^{\exp\{\theta-\theta_0\}},$$

and hence the failure time is an exponential random variable with mean $e^{\theta_0-\theta-v}$. Of course, this distribution can be considered to be a special case of the distribution $F(t;v,\theta)$, so no adjustments of (5.5) are needed to include the case $\rho = 0$. We shall refer to the family of distributions $F(t;v,\theta)$ as the family of Harrington and Fleming alternatives to the exponential distribution. We point out that for small values of ρ, the Harrington and Fleming alternatives closely resemble the exponential distribution. In fact, for ρ equal to zero we cannot speak of alternatives anymore.

By rewriting (5.4) in the intriguing form

$$\frac{(1 - F(t;\theta))^{-\rho} - 1}{(1 - F(t;\theta_0))^{-\rho} - 1} = e^{\theta-\theta_0}$$

it follows that

$$1 - F(t;\theta_0) = [1 + ((1 - F(t;\theta))^{-\rho} - 1)e^{\theta_0-\theta}]^{-1/\rho}$$

by which the inverse of $F(t;v,\theta)$ is readily calculated. When applied to uniform random variables this inverse generates random variables having distribution $F(t;v,\theta)$. However, we shall generate latter random variables in an alternative though strongly related way. Suppose E_i is a standard exponential random variable. Then

$$(\rho e^{v})^{-1}\log(1 + (e^{\rho E_i} - 1)e^{\theta_0-\theta}) \tag{5.6}$$

is a random variable with distribution $F(t;v,\theta)$.

5.2.2 Maximum likelihood estimation

Testing whether the random variables $X_1,\ldots,X_n$ follow an exponential distribution with unknown mean involves estimation of this mean under the null-hypothesis. Since we have

$$\Lambda(t;v,\theta_0) = e^{v}t,$$

it follows that

$$\psi_1(t;v,\theta_0) = 1,$$

and hence the maximum likelihood estimator $v^{(n)}$ can be found by solving the likelihood equation

$$H_n^1(\infty) - \exp\{v^{(n)} - v_0\}\int_0^\infty (1 - H_{n-}(t))d\Lambda(s;v_0,\theta_0) = 0$$

(here v_0 is some arbitrary element of Υ), which yields

$$v^{(n)} = v_0 + \log H_n^1(\infty) - \log\int_0^\infty (1 - H_{n-}(t))d\Lambda(s;v_0,\theta_0).$$

Condition 11 on page 54 requires the existence of a certain function $\pi(v,\theta)$, related to the M-estimator. This function indeed exists and is given by

$$\pi(v,\theta) = v_0 + \log H^1(\infty; v,\theta) - \log \int_0^\infty (1 - H(s;v,\theta))d\Lambda(s;v_0,\theta_0).$$

The description of the behavior of $v^{(n)}$ involves the 1×1 matrices $\Xi_0(v,\theta)$, $\Xi_a(v,\theta)$ and $\Xi_b(v,\theta)$. For the maximum likelihood estimator both $\Xi_0(v,\theta_0)$ and $\Xi_b(v,\theta_0)$ have $H^1(\infty;v,\theta_0)$ as single element. Hence for Conditions 12 and 18 to hold it is necessary that $H^1(\infty;v,\theta_0)$, the proportion uncensored observations under the null hypothesis, remains bounded away from zero. Observe that since in our framework the censoring distribution does not depend on v_n or θ_n, the presence of censoring leads to the requirement that e^{-v} should not exceed a certain limit. However, it is possible [at the cost of additional complexity] to extend our framework by allowing the censoring distribution to depend on the parameters of the failure time distribution. This extension, which involves a straightforward generalization of Lemma 3, enables us to consider types of censoring in which the censoring distribution is linked to the failure time distribution.

If the proportion of uncensored observations remains indeed bounded away from zero, then Lemma 3 may be used to show that Conditions 12 and 18 hold if Θ is appropriately chosen. The single element of $\Xi_a(v,\theta_0)$ is equal to $\mu_\rho(v)$, where $\mu_\beta(v)$ is defined by

$$\mu_\beta(v) = \int_0^\infty (1 - F(s;v,\theta_0))^\beta dH^1(s;v,\theta_0).$$

It should be stressed that the form of the function $\psi_1(t;v,\theta_0)$ depends on the actual parametrization of the failure times. For instance, parametrizing the exponential distribution in the common way [using the reciprocal of the mean as parameter] leads to a different function $\psi_1(t;v,\theta_0)$, and hence to a different maximum likelihood estimator.

5.2.3 Modified maximum likelihood estimation

In the preceding chapter a modification of the maximum likelihood estimator was introduced, which was constructed "perpendicular" to the null hypothesis. In the situation considered here, this estimator is based on the function

$$\phi_1(t;v) = \frac{\mu_\rho(v)}{\mu_{2\rho}(v)}(1 - F(s;v,\theta_0))^\rho - 1.$$

Now $\Xi_0(v,\theta_0)$ and $\Xi_a(v,\theta_0)$ have as single element $\mu_0(v) - (\mu_\rho(v))^2/\mu_{2\rho}(v)$ and zero, respectively. The matrix $\Xi_b(v,\theta_0)$ is equal to $\Xi_0(v,\theta_0)$. Observe that due to the involvement of the censoring distribution G in $\mu_\rho(v)$ and $\mu_{2\rho}(v)$ the modified maximum likelihood estimator can only be constructed if G is known. If there is no censoring present, then $\mu_\beta(v) = (1+\beta)^{-1}$, and hence $\phi_1(t;v)$ takes the relatively easy form

$$\phi_1(t;v) = \frac{1+2\rho}{1+\rho}\exp\{-\rho e^v t\} - 1,$$

while the single elements of the matrices $\Xi_0(v,\theta_0)$ and $\Xi_a(v,\theta_0)$ equal $\rho^2/(2\rho+1)$.

Without knowledge of the censoring distribution, we can not apply the modified maximum likelihood estimation procedure to the Woolson data. Thus, to illustrate this procedure we introduce a second data set.

In Proschan (1963) a sample of 15 failure times of airconditioning equipment in aircraft is given. In this data set the phenomenon of censoring is absent. Nevertheless, when applied to the Proschan data the modified maximum likelihood procedure fails to yield an estimate for $v^{(n)}$.

The estimator $v^{(n)}$ should satisfy $\Phi_{n1}(\infty; v^{(n)}) = 0$ by equation (4.3) on page 52. In Figure 5.1 $\Phi_{n1}(\infty; v)$ is displayed as a function of v. As one can see it takes positive values only, leaving the modified maximum likelihood estimator undefined. Of course, it is possible to extend the definition of an M-estimator so as to cover this case also. E.g. rather than setting $\Phi_{n1}(\infty; v^{(n)})$ equal to zero we could minimize the absolute value of $\Phi_{n1}(\infty; v^{(n)})$. For the Proschan data this extended procedure yields -5.3474 as an estimate for v_n. In contrast, the maximum likelihood estimate is -4.7980.

Without question the phenomenon observed here is undesirable, but understanding may enable us to avoid it. To arrive at an explanation view the basic martingale $M_n(t; v, \theta_0)$ as the difference between the "counting process" part $n^{1/2}H_n^1(t)$ and the "compensator" part $\exp\{v - v_0\}\int_0^t(1 - H_{n-}(s))d\Lambda(s; v_0, \theta_0)$. Note that the latter part can alternatively be represented as $\exp\{v - v_{n0}\}\int_0^t(1 - H_{n-}(s))d\Lambda(s; v_{n0}, \theta_0)$. As v tends to $-\infty$ the process $M_n(t; v, \theta_0)$ behaves more and more like the counting process part. The compensator part becomes dominant as v tends to $+\infty$.

Now observe that for fixed t the function $\phi_1(t;v)$ changes sign from positive to negative as v travels from $-\infty$ to $+\infty$. As a consequence $\Phi_{n1}(\infty; v)$ is positive for values of v close to either $-\infty$ or $+\infty$. Thus, in order for $\Phi_{n1}(\infty; v)$ to have a zero, there must be a deep enough "dip" somewhere in its graph. As we can see from Figure 5.1 the dip exists, but is not deep enough.

74,	57,	48,	29,	502,
12,	70,	21,	29,	386,
59,	27,	153,	26,	326.

Table 5.1: Data given in Proschan (1963).

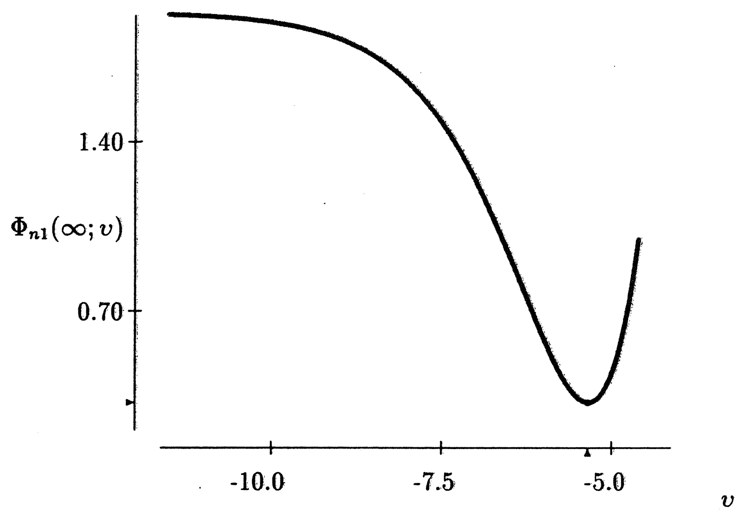

Figure 5.1: $\Phi_{n1}(\infty; v)$ for modified maximum likelihood estimation procedure, applied to the Proschan data. Minimized value is 0.32467.

In case the dip was deep enough we would have had the choice between [at least] two zeroes. Our theory leads us to expect that $\Phi_{n1}(\infty; v)$ seen as a function of v has a positive derivative near $v^{(n)}$. Hence, if there are two zeroes we should choose the one to the right of the point where the dip is maximal. In this respect it is noteworthy that for the Proschan data the maximum likelihood estimate is larger than the estimate based on the generalized definition of the modified maximum likelihood estimation procedure.

In the previous chapter we concluded that the estimation procedure was of secondary importance as far as asymptotic properties were concerned. However, at this moment it seems that we should prefer a function $\phi_1(t; v)$ which does not change sign [say, which is nonnegative]. By using similar arguments as above it follows that such a function leads to $\Phi_{n1}(\infty; v)$ being positive for values of v close to $-\infty$, being negative for values of v close to $+\infty$, and hence having an odd number of zeroes.

An additional advantage of a nonnegative function $\phi_1(t; v)$ is that Condition 11 holds automatically. The integral $\int_0^\infty \phi_1(s; v')dD(s; v', \theta_0, v, \theta)$ is positive for values of v' close to $-\infty$, and negative for values of v' close to $+\infty$. This can be seen by viewing the function $D(t; v', \theta_0, v, \theta)$ as the difference between $\int_0^t (1 - H(s; v, \theta))d\Lambda(s; v, \theta)$ and its counterpart $\int_0^t (1 - H(s; v, \theta))d\Lambda(s; v', \theta_0)$.

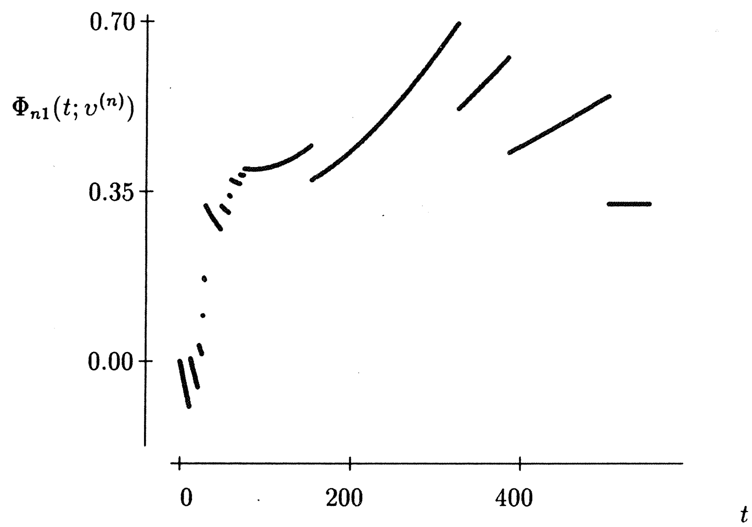

Figure 5.2: Stochastic integral related to the modified maximum likelihood estimator, constructed from the Proschan data. The value of $v^{(n)}$ is -5.3474.

5.3 Weight processes

In this section we are dealing with a sample $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ arising from a Harrington and Fleming alternative to the exponential distribution [that is, $F(t; v_n, \theta_n)$ is defined according to (5.5)], and we discuss appropriate choices of the weight process which lead to optimal generalized rank and supremum type tests for the composite null hypothesis that the sample comes from an exponential distribution with unknown mean. The choices have in common that their limiting weight functions satisfy

$$L(t; v, \theta_0) = \frac{(1 - F(t; v, \theta_0))^\rho - \mu_\rho(v)/\mu_0(v)}{\sqrt{\mu_{2\rho}(v) - (\mu_\rho(v))^2/\mu_0(v)}}. \tag{5.7}$$

(compare with (4.46)). Hence, we always have $a_R(v_0) = a_S(v_0) = 1$. Moreover, $e_R(h, v_0) = e_S(h, v_0) = \sqrt{\mu_{2\rho}(v_0) - (\mu_\rho(v_0))^2/\mu_0(v_0)}$.

5.3.1 Known censoring distribution

If the censoring distribution is known then we are in the position to make the obvious choice of using the limiting weight function $L(t; v, \theta_0)$ as weight process.

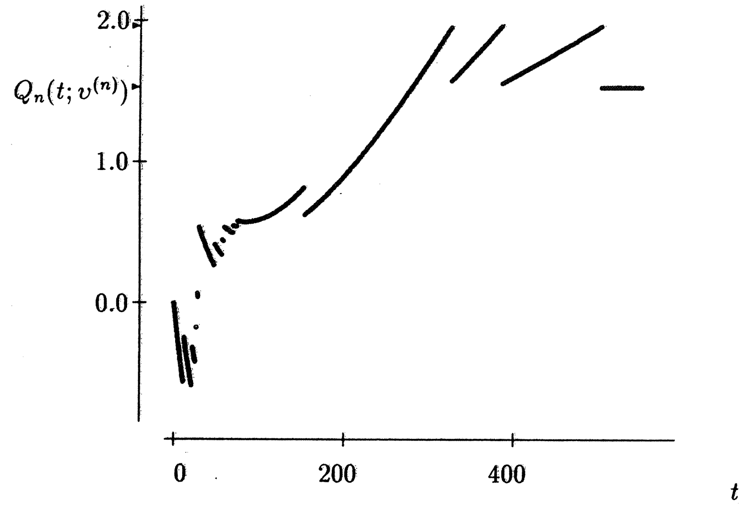

Figure 5.3: Stochastic integral based on maximum likelihood estimation and weight process for known censoring distribution, constructed from the Proschan data. The value of $v^{(n)}$ is -4.7980. The two-sided generalized rank and supremum type test statistics attain values 1.529 and 1.967, respectively. The corresponding asymptotic probability values are 0.1262 and 0.0983.

The most important case of a known censoring distribution is the absence of censoring. Here we have $\mu_\beta(v) = (1+\beta)^{-1}$, and hence setting the weight process equal to the right-hand side of (5.7) results in

$$L_n(t; v) = \frac{\sqrt{1+2\rho}}{\rho}[(1+\rho)\exp\{-\rho e^v t\} - 1]. \tag{5.8}$$

Observe that

$$L_{n1}^{(1)}(t; v) = -(1+\rho)\sqrt{1+2\rho} e^v t \exp\{-\rho e^v t\}$$

$$L_{n11}^{(2)}(t; v) = -(1+\rho)\sqrt{1+2\rho}(\rho e^v - 1)e^v t \exp\{-\rho e^v t\}$$

Moreover, using the fact that $L(t; v, \theta) = L_n(t; \pi(v, \theta))$ for every $v \in \Upsilon$ and $\theta \in \Theta$, it is easily seen that the weight process in (5.8) satisfies Condition 13.

5.3.2 Unknown censoring distribution

If G is unknown then it will not be possible to compute $\mu_\beta(v)$, and thus we are forced to estimate it. As remarked before, the common way to obtain estimators for quantities involving $H^1(t;v,\theta_0)$ is to replace this distribution function by $\int_0^t(1-H_{n-}(s))d\Lambda(s;v,\theta_0)$. This procedure yields as weight process

$$L_n(t;v)=\frac{(1-F(t;v,\theta_0))^\rho-\hat{\mu}_\rho(v)/\hat{\mu}_0(v)}{\sqrt{\hat{\mu}_{2\rho}(v)-(\hat{\mu}_\rho(v))^2/\hat{\mu}_0(v)}}, \tag{5.9}$$

where

$$\begin{aligned}\hat{\mu}_\beta(v)&=\int_0^\infty(1-F(s;v,\theta_0))^\beta(1-H_{n-}(s))d\Lambda(s;v,\theta_0)\\&=\int_0^\infty\exp\{-\beta e^v s\}(1-H_{n-}(s))e^v ds.\end{aligned} \tag{5.10}$$

To verify Condition 13 we need some insight into the random behavior of $\hat{\mu}_\beta(v)$. Lemma 4 supplies us with the necessary knowledge, by relating the difference between a generalized version of $\hat{\mu}_\beta(v)$ and its deterministic counterpart to the random variable S_n of Theorem 7 on page 55.

Lemma 4 *Let $g(t;v)$ be a real valued function, and define*

$$\hat{\mu}(v;g)=\int_0^\infty g(s;v)(1-H_{n-}(s))d\Lambda(s;v,\theta_0),$$

$$\mu(v,\theta;g)=\int_0^\infty g(s;\pi(v,\theta))(1-H(s;v,\theta))d\Lambda(s;\pi(v,\theta),\theta_0).$$

If the function $g(t;v)$ remains uniformly bounded in t and v, then there exists a constant c_{73} such that for every $\omega\in\Omega$

$$|\hat{\mu}(v_{n0};g)-\mu(v_n,\theta_n;g)|\le c_{73}n^{-1/2}S_n.$$

Corollary 7 *If for every $i=,\dots,r$ the function $g_i^{(1)}(t;v)$, the first order partial derivative of $g(t;v)$ with respect to the i^{th} component of v, remains uniformly bounded in t and v, then there exists a constant c_{73} and a function $\mu_i^{[1]}(v,\theta;g)$ such that for every $\omega\in\Omega$*

$$|\hat{\mu}_i^{(1)}(v_{n0};g)-\mu_i^{[1]}(v_n,\theta_n;g)|\le c_{73}n^{-1/2}S_n.$$

where $\hat{\mu}_i^{(1)}(v;g)$ is the first order partial derivative of $\hat{\mu}(v;g)$ with respect to the i^{th} component of v.

If in addition the function $g_{ij}^{(2)}(t;v)$, the second order partial derivative of $g(t;v)$ with respect to the i^{th} and j^{th} components of v, remains uniformly bounded in t and v, then there exists a constant c_{73} and a function $\mu_{ij}^{[2]}(v,\theta;g)$ such that for every $\omega\in\Omega$

$$|\hat{\mu}_{ij}^{(2)}(v_{n0};g)-\mu_{ij}^{[2]}(v_n,\theta_n;g)|\le c_{73}n^{-1/2}S_n.$$

where $\hat{\mu}_{ij}^{(2)}(v;g)$ is the first order partial derivative of $\hat{\mu}(v;g)$ with respect to the i^{th} and j^{th} components of v.

Proof of Lemma 4 From (4.51) we obtain

$$
\begin{aligned}
&|\hat{\mu}(v_{n0}; g) - \mu(v_n, \theta_n; g)| \\
&= |\int_0^\infty g(s; v_{n0})\{H_{n-}(s) - H(s; v_n, \theta_n)\} d\Lambda(s; v_{n0}, \theta_0)| \\
&\le |g(t; v_{n0})| n^{-1/2} \sup_{t\in[0,\infty)} |\int_0^t U_{n-}(s; v_n, \theta_n) d\Lambda(s; v_{n0}, \theta_0)| \\
&\le \sup_{t\in[0,\infty)} |g(t; v_{n0})| (c_{60}/6) n^{-1/2} S_n.
\end{aligned}
$$

□

Proof of Corollary 7 Define

$$g_i^*(t; v) = g_i^{(1)}(t; v) + g(t; v)\psi_i(t; v, \theta_0),$$

$$
\begin{aligned}
g_{ij}^*(t; v) = {} & g_{ij}^{(2)}(t; v) + g_i^{(1)}(t; v)\psi_j(t; v, \theta_0) \\
& + g_j^{(1)}(t; v)\psi_i(t; v, \theta_0) + g(t; v)\psi_{ij}^{(1)}(t; v, \theta_0).
\end{aligned}
$$

Then we have

$$\hat{\mu}_i^{(1)}(v; g) = \mu(v; g_i^*),$$

$$\hat{\mu}_{ij}^{(2)}(v; g) = \mu(v; g_{ij}^*).$$

Since both $g_i^*(t; v)$ and $g_{ij}^*(t; v)$ remain uniformly bounded in t and v, we are in the position to apply Lemma 3, which yields the desired result. □

5.4 Total time on test plots

The weight processes of the previous section were derived so as to yield generalized rank and supremum type tests which are optimal against a specific alternative. However, in some circumstances the choice for some test statistic is not based on specific optimality, but on consistency against a broad class of alternatives instead. In this respect statistics based on total time on test plots are worth mentioning.

Total time on test plots are discussed in Barlow and Proschan (1969) and in Gill (1986). According to the former paper, they are obtained by plotting $\int_0^t (1 - H_{n-}(s)) d\Lambda(s; v^{(n)}, \theta_0)$ versus $H_n^1(t)$, where $v^{(n)}$ is the maximum likelihood estimator. Under Type II censoring these plots allow to make assessments concerning the hazard function underlying the data. In particular, monotone hazard functions are easily recognized: depending on whether the hazard function is increasing or decreasing, the plot shows a convex or concave curve.

If the underlying distribution function is exponential [corresponding to a constant hazard function], then the total time on test plot should reflect a straight line through the origin, making a 45 degree angle with the X-axis. The deviations from this line are connected to the process $M_n(t; v^{(n)}, \theta_0)$. Not surprisingly, many proposals for tests for exponentiality inspired by the total time on test plot turn out to be based on functionals of $M_n(t; v^{(n)}, \theta_0)$. The supremum type test was mentioned in Barlow and Campo (1975). Observe that the generalized rank test does not make much sense, since it follows from the definition of the maximum likelihood estimator that $M_n(\infty; v^{(n)}, \theta_0)$ is zero.

The asymptotic theory for $M_n(t; v^{(n)}, \theta_0)$ is simply derived from the fact that for this process we have

$$L_0(s, t; v_0) = 1_{\{s \leq t\}} - \frac{H^1(t; v_0, \theta_0)}{H^1(\infty; v_0, \theta_0)}.$$

That is, the sequence $\{M_n(t; v^{(n)}, \theta_0)\}_{n=1}^{\infty}$ converges in P_{v_0}-distribution to a mean zero Gaussian process with covariance function

$$\frac{H^1(t_1 \wedge t_2; v_0, \theta_0)(1 - H^1(t_1 \vee t_2; v_0, \theta_0))}{H^1(\infty; v_0, \theta_0)}.$$

For a better understanding of the qualities of the total time on test plot, consider the cumulative distribution function

$$F(t; v, \theta) = \begin{cases} 1 - \exp\{-e^{v+\theta} t\} & \text{if } t \leq \tau, \\ 1 - \exp\{-e^{v+\theta}\tau - e^{v-\theta}(t - \tau) & \text{if } t > \tau, \end{cases}$$

where $\tau \in [0, \infty)$ is fixed. It follows that

$$\psi_1(t; v, \theta) = 1,$$

$$\psi_2(t; v, \theta) = \begin{cases} 1 & \text{if } t \leq \tau, \\ -1 & \text{if } t > \tau. \end{cases}$$

Now suppose that our knowledge of the censoring distribution is sufficient to determine $H^1(\tau; v, \theta_0)$ and $H^1(\infty; v, \theta_0)$. In order to find optimal generalized rank and supremum type tests, we choose a weight process recommended by Theorem 13, yielding

$$L_n(t; v) = \begin{cases} \frac{1 - H^1(\tau; v, \theta_0)}{H^1(\infty; v, \theta_0)} & \text{if } t \leq \tau, \\ -\frac{H^1(\tau; v, \theta_0)}{H^1(\infty; v, \theta_0)} & \text{if } t > \tau. \end{cases}$$

The resulting stochastic integral satisfies

$$T_R(Q_n(\cdot; v^{(n)})) = Q_n(\tau; v^{(n)}) + (Q_n(\infty; v^{(n)}) - Q_n(\tau; v^{(n)}))$$
$$= M_n(\tau; v^{(n)}, \theta_0).$$

Hence, we may view the test based on the statistic $M_n(\tau; v^{(n)}, \theta_0)$ as an optimal generalized rank test for the composite null hypothesis $\theta = 0$ in the example above. With some more effort we may show that the test based on either $\sup_{0\leq t\leq \tau} M_n(t; v^{(n)}, \theta_0)$ or $\sup_{\tau\leq t\leq\infty} M_n(t; v^{(n)}, \theta_0)$, depending on whether $\mathcal{E}_{v_0}\{M_n(\tau; v^{(n)}, \theta_0)\}^2$ is equal to $\sup_{0\leq t\leq\tau} \mathcal{E}_{v_0}\{M_n(t; v^{(n)}, \theta_0)\}^2$ or equal to $\sup_{\tau\leq t\leq\infty}$ $\mathcal{E}_{v_0}\{M_n(t; v^{(n)}, \theta_0)\}^2$, is an optimal supremum type test for the same testing problem.

Chapter 6

Small sample characteristics

6.1 Simulations

The theory of the preceding chapters is asymptotic of nature. It can be expected to work if the sample size n is large, but there is no absolute guarantee that this theory is also applicable in situations where the sample size is relatively small.

In this chapter we investigate which are the limitations to our theory with respect to sample size. The investigation is conducted by means of simulation experiments in situations described in section 5.1 and subsection 5.2.1.

Within the class of goodness-of-fit tests considered in our theory, a test is determined by four items: the estimation procedure, the weight process $L_n(t;v)$, the functional T, and the testsize. In the simulations choices are made for each of these items.

- If we count "no estimation" [that is, replace $v^{(n)}$ by v_0] as an estimation procedure, then two different estimation procedures are available, the other procedure being maximum likelihood. The modified maximum likelihood procedure, introduced in section 4.3, is not included in the simulations because of the weaknesses discussed in section 5.2.

- The weight process $L_n(t;v)$ is always a linear function of $\exp\{-\rho e^v t\}$, where ρ equals 1. For administrative reasons the weight process is identified by a two-digit identification code, which determines the coefficients of the linear function. The first digit indicates whether the coefficients are deterministic [1; involve $\mu_\beta(v)$] or random [2; involve $\hat{\mu}_\beta(v)$]. The second digit marks the kind of weight process; it can be either 1 [one-sample log-rank: slope of the linear function is equal to zero], 2 [one-sample generalized Wilcoxon: intercept of the linear function is equal to zero] or 3 [weight process advocated in section 5.3].

- A description of the functionals T used is given in Table 6.1. For each test statistic an asymptotic P-value is computed. For generalized rank tests this is done by means of the standard normal distribution, while for supremum type tests the distribution of the supremum of a standard Wiener process on $[0, 1]$ is used. Observe that for a supremum type test with underlying stochastic integral not weakly converging to some time-transformed Wiener process, the asymptotic P-value is asymptotic in two respects: not only the sample size n should be large enough, the observed value of the test statistic should be large as well.

The asymptotic P-value facilitates intercomparison, since it transforms the various test statistics under study to statistics which have nearly the same null-distribution, while preserving the characteristics of the resulting test.

- Each test statistic is combined with 5 different test sizes, running from $0.5/n$ to $2.5/n$.

During the simulations, a histogram of the asymptotic P-values is constructed. This histogram, which has interval width $(5n)^{-1}$, is used afterwards for density estimation.

The estimated density under the null hypothesis serves as the basis for the construction of critical points of the P-values. If Table 6.4b tells us that the critical point of the left-sided generalized rank test is 0.0136 for size 0.020, this means that rejecting the null hypothesis if and only if the asymptotic P-value is less than or equal to 0.0136 yields a test with size 0.020 "exactly", implying that we would obtain an anticonservative test if we were to use the asymptotic P-value as an approximation to the exact P-value [that is, rejection only takes place if the asymptotic P-value is less than or equal to 0.020].

Subsequently, power is estimated by computing the left tail probability belonging to a critical point of the asymptotic P-value. This computation takes place according to the estimated density under one of the four alternative hypotheses.

The sample sizes used are 25, 50 and 100. For each of these sample sizes we construct 10000 replications of five types of censored and five types of uncensored samples. For the construction of censored samples the standard exponential distribution serves as censoring distribution.

Test	*Functional $T(\xi)$*		
	Leftsided	Rightsided	Twosided
Generalized rank	$-\xi(\infty)$	$\xi(\infty)$	$\lvert\xi(\infty)\rvert$
Supremum type	$\sup_{t\in[0,\infty)} -\xi(t)$	$\sup_{t\in[0,\infty)} \xi(t)$	$\sup_{t\in[0,\infty)} \lvert\xi(t)\rvert$

Table 6.1: Functionals used.

6.2 The simple null hypothesis

The simulated samples are drawn from a distribution $F(t; \upsilon_n, \theta_n)$ belonging to the family of Harrington and Fleming alternatives to the exponential distribution, as described in Section 5.2. The value of υ_n and ρ are always 0 and 1, respectively. Hence, we are considering logistic shift alternatives to the standard exponential distribution.

The parameter of interest θ_n varies in the following way with sample size:

$$\theta_n = c_H \sqrt{\frac{2 \log n}{n \mu_{2\rho}(\upsilon_0)}}, \tag{6.1}$$

where c_H is chosen according to the type of hypothesis under which simulation takes place. There are five types of hypotheses: H_{LL}, H_L, H_0, H_R and H_{RR}; the respective choices of c_H are -1.0, -0.5, 0.0, 0.5 and 1.0. The resulting values of θ_n are displayed in Table 6.2. The value of θ_n under H_0 coincides with θ_0.

The asymptotic power of a generalized rank test is equal to the probability that a standard normal random variable exceeds $z_{\tilde{\alpha}} - (\theta_n - \theta_0)\sqrt{n e_R(h)}$, where $z_{\tilde{\alpha}}$ is the $(1 - \tilde{\alpha})$ quantile of the standard normal distribution.

Approximating $z_{\tilde{\alpha}}$ by $\sqrt{-2 \log \tilde{\alpha}}$, we find that equation (6.1) implies that under H_{LL} and H_{RR} the power of a size $1/n$ one-sided generalized rank test based on stochastic integral 12 [for which $e_R(h) = \mu_{2\rho}(\upsilon_0)$] will tend to 0.5 as n tends to infinity. However, the power of such a test will still be far away from 0.5 in the region of the sample sizes we are considering, due to poor quality of the approximation of $z_{\tilde{\alpha}}$. Table 6.3 indicates the power we could expect from the generalized rank test based on stochastic integral 12.

Observe that the same crude approximation of $z_{\tilde{\alpha}}$ is used under similar conditions in the weak intermediate efficiency concept.

Sample size	*Type of Censoring*	H_{LL}	H_L	H_0	H_R	H_{RR}
25	None	-0.8789	-0.4395	0.0000	0.4395	0.8789
25	Exp(1)	-1.0149	-0.5075	0.0000	0.5075	1.0149
50	None	-0.6852	-0.3426	0.0000	0.3426	0.6852
50	Exp(1)	-0.7912	-0.3956	0.0000	0.3956	0.7912
100	None	-0.5257	-0.2628	0.0000	0.2628	0.5257
100	Exp(1)	-0.6070	-0.3035	0.0000	0.3035	0.6070

Table 6.2: Values of θ_n. The values of υ_n, θ_0 and ρ are always equal to 0, 0 and 1, respectively.

Size of test	One-sided test H_{LL}	H_L	H_R	H_{RR}	Two-sided test H_{LL}	H_L	H_R	H_{RR}
				Sample size 25				
0.020	0.686	0.216	0.216	0.686	0.584	0.145	0.145	0.584
0.040	0.784	0.315	0.315	0.784	0.686	0.217	0.217	0.686
0.060	0.837	0.387	0.387	0.837	0.744	0.271	0.271	0.744
0.080	0.871	0.446	0.446	0.871	0.784	0.316	0.316	0.784
0.100	0.895	0.495	0.495	0.895	0.814	0.355	0.355	0.814
				Sample size 50				
0.010	0.681	0.177	0.177	0.681	0.588	0.120	0.120	0.588
0.020	0.771	0.256	0.256	0.771	0.681	0.177	0.177	0.681
0.030	0.820	0.315	0.315	0.820	0.735	0.220	0.220	0.735
0.040	0.852	0.362	0.362	0.852	0.771	0.257	0.257	0.771
0.050	0.875	0.403	0.403	0.875	0.799	0.288	0.288	0.799
				Sample size 100				
0.005	0.677	0.145	0.145	0.677	0.590	0.099	0.099	0.590
0.010	0.761	0.209	0.209	0.761	0.677	0.145	0.145	0.677
0.015	0.806	0.257	0.257	0.806	0.727	0.180	0.180	0.727
0.020	0.837	0.296	0.296	0.837	0.761	0.209	0.209	0.761
0.025	0.859	0.329	0.329	0.859	0.786	0.235	0.235	0.786

Table 6.3: Asymptotic power of generalized rank test based on stochastic integral 12 (no estimation). Also valid for generalized rank test based on stochastic integral 13 (maximum likelihood estimation).

6.2.1 Stochastic integral 11

The first stochastic integral we encounter, type 11 [no estimation], is based on the weight process

$$L_n(t;v) = (\mu_0(v))^{-1/2}. \tag{6.2}$$

Hence, this stochastic integral yields the same tests as the basic martingale $M_n(t;v_0,\theta_0)$. The generalized rank test is the well-known test of Breslow (1975), a one-sample logrank test. The supremum type test was investigated by Aki (1986). Both tests are indicated for detecting proportional hazards alternatives. In our simulations the alternatives are of a different type, namely logistic shift. We have $a_R = a_S = 1$ and

$$e_R(h) = e_S(h) = \frac{(\mu_\rho(v_0))^2}{\mu_0(v_0)}. \tag{6.3}$$

In the absence of censoring this becomes $e_R(h) = e_S(h) = 1/4$.

To obtain some intuition about how the small sample behavior deviates from the asymptotic behavior, we first pay some attention to $M_1(\infty;v_0,\theta_0)$, the basic martingale for sample size 1, evaluated at infinity. Remark that $M_n(\infty; v_0, \theta_0)$ is an appropriately scaled sum of n independent copies of $M_1(\infty; v_0, \theta_0)$.

We may view $M_1(\infty; v_0, \theta_0)$ as a difference between two nonnegative random variables δ_1 and $\Lambda(Z_1; v_0, \theta_0)$, which do not have much in common. The former takes values 0 and 1 only [in the absence of censoring it even becomes degenerate], and hence is bounded. The latter takes values everywhere on the complete halfline, and in general has a distinct tail. Thus, for small sample sizes we may encounter asymmetric behavior of the basic martingale. Since the limiting process of the basic martingale is Gaussian the asymmetry should vanish as the sample size becomes larger.

Under H_0 both random variables have mean $H^1(\infty; v_0, \theta_0)$, and hence we expect that the mean of $M_n(\infty; v_0, \theta_0)$ is equal to zero. Tables 6.4a, 6.5a and 6.6a show that this expectation, which also could be inferred from martingale theory, becomes true.

Moreover, these tables reflect the asymmetry of the basic martingale. Under H_L and H_{LL} the mean of $M_n(\infty; v_0, \theta_0)$ deviates more from zero than under their counterparts H_R and H_{RR}. This is in part compensated by the standard deviation, which tends to inflate under H_L and H_{LL} and to deflate under H_R and H_{RR}.

If we conclude from Tables 6.4a, 6.5a and 6.6a that the generalized rank test based on $M_n(t; v_0, \theta_0)$ is more sensitive to the left than to the right, then we are drawing our conclusions too hastily. The asymmetry of the basic martingale is also reflected in the tail behavior under H_0, since the left tail of $M_n(\infty; v_0, \theta_0)$ is more outstretched than the right tail.

The adverse location of the leftsided critical points annuls the initial higher sensitivity on the left side, and even makes the leftsided generalized rank test less sensitive than the rightsided generalized rank test.

Tables 6.4, 6.5 and 6.6 also indicate that the effects of the asymmetry of the basic martingale are even worse for supremum type tests.

6.2.2 Stochastic integral 21

The stochastic integral type 21 [no estimation] is a variant of stochastic integral 11, having weight process

$$L_n(t; v) = (\hat{\mu}_0(v))^{-1/2}. \tag{6.4}$$

Hence, it is obtained by dividing the basic martingale by the square root of the variance estimator $\hat{\mu}_0(v_0) = \int_0^\infty (1 - H_{n-}(s)) d\Lambda(s; v_0, \theta_0)$. By making this division the martingale property is lost. The quantities $e_R(h)$ and $e_S(h)$ are the same as for stochastic integral 11, and are given by (6.3).

Before looking at the simulation results, we first investigate the behavior of stochastic integral 21 evaluated at ∞ for sample size 1 by viewing it as the difference between the random variables $\delta_1(\Lambda(Z_1; v_0, \theta_0))^{-1/2}$ and $(\Lambda(Z_1; v_0, \theta_0))^{1/2}$.

Both are nonnegative and take values everywhere on the complete halfline. In general, the former random variable is heavier tailed than the latter. Moreover, under H_0 the random variables differ in mean. To examplify, suppose censoring is absent. Now we have

$$P_0\left(\frac{\delta_1}{\sqrt{\Lambda(Z_1; v_0, \theta_0)}} \leq x\right) = \exp\{-x^{-2}\},$$

$$P_0\left(\sqrt{\Lambda(Z_1; v_0, \theta_0)} \leq x\right) = 1 - \exp\{-x^2\}.$$

The tail of $\delta_1(\Lambda(Z_1; v_0, \theta_0))^{-1/2}$, which vanishes at a rate x^{-2}, is quite heavy, whereas the tail of $(\Lambda(Z_1; v_0, \theta_0))^{1/2}$ is moderate [that is, comparable to the tail of a Gaussian random variable]. Moreover, under P_0 the mean of $\delta_1(\Lambda(Z_1; v_0, \theta_0))^{-1/2}$ is equal to $\pi^{1/2}$ and differs from the mean of $(\Lambda(Z_1; v_0, \theta_0))^{1/2}$, which is equal to $\pi^{1/2}/2$. It seems that the behavior of stochastic integral 21 shows asymmetry that is in some sense opposite to asymmetry in the behavior of stochastic integral 11. This is also reflected in Tables 6.7, 6.8 and 6.9.

Though $e_R(h)$ and $e_S(h)$ coincide, the supremum type tests surpass the generalized rank test in simulated power. This is not only true for stochastic integral 21, but also for stochastic integral 11, and underlines the omnibus character of the supremum type tests.

Surprisingly, Tables 6.4c and 6.7c show that the simulated power of one-sided tests based on stochastic integral 21 do not differ much from the simulated power of the corresponding tests based on stochastic integral 11. This is certainly not true for the two-sided tests. Two-sided tests based on stochastic integral 21 are relatively more sensitive to positive and less sensitive to negative values of θ_n.

6.2.3 Stochastic integral 12

The weight process underlying stochastic integral 12 is given by

$$L_n(t; v) = (\mu_\rho(v))^{-1/2} \exp\{-\rho e^v t\}. \tag{6.5}$$

The generalized rank test based on stochastic integral 12, a one-sample version of the two-sample generalized Wilcoxon test, was recommended by Harrington and Fleming (1982) as a test against the alternatives proposed in the same paper [which we have highlighted in section 5.2], to be used if there is no censoring present. Their recommendation followed from heuristic arguments. Indeed, the theory in section 3.3 shows that this test is optimal.

For stochastic integral 12 we have

$$e_R(h) = e_S(h) = \mu_{2\rho}(v_0) \tag{6.6}$$

If there is no censoring present we have $e_R(h) = e_S(h) = 1/3$. If the censoring distribution is standard exponential then $e_R(h) = e_S(h) = 1/4$.

The stochastic integral evaluated at infinity may be viewed as the scaled sum of n independent copies of the random variable $(1-F(Z_1;v_0,\theta_0))\delta_1-F(Z_1;v_0,\theta_0)$, which takes values between -1 and 1. For instance, under H_0 and in the absence of censoring this random variable is uniformly distributed on the interval $[-1,1]$.

This suggests that the behavior of the stochastic integral is very close to symmetry, an impression confirmed by the uncensored sample simulation results. The censored sample simulations show a slight difference between the left- and right-sided critical values, and a higher sensitivity to negative values of θ_n.

The simulated power of the generalized rank test is in accordance with the asymptotic power given in Table 6.3.

As we compare with the simulation results for stochastic integral 11, we see that the one-sided tests indeed yield higher power. However, the two-sided supremum type test gets the worst of it under H_L and H_{LL}.

6.2.4 Stochastic integral 22

Stochastic integral 22 is a variant of stochastic integral 21, to be used if the censoring distribution is unknown. Its weight process is

$$L_n(t;v) = (\hat{\mu}_\rho(v))^{-1/2}\exp\{-\rho e^v t\}. \tag{6.7}$$

The generalized rank test based on stochastic integral 22 appeared in Harrington and Fleming (1982). The quantities $e_R(h)$ and $e_S(h)$ are given by (6.6).

As we compare the uncensored sample results of stochastic integral 22 and stochastic integral 12, we do not spot many differences in simulated power of the one-sided tests. This in contrast to two-sided tests: those based on stochastic integral 22 are relatively more sensitive to positive and less sensitive to negative values of θ_n. Observe that we reached the same conclusion in the comparison of stochastic integral 12 with stochastic integral 11, suggesting that the use of a variance estimator does not have much effect on the power of one-sided tests, but shifts the sensitivity of two-sided tests to the right. In censored samples, the shift in sensitivity is also noticable for one-sided tests.

With respect to stochastic integral 21, the stochastic integral clearly yields higher simulated power, especially for generalized rank tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.5089	-1.1775	-0.0006	1.0270	1.8984
st.dev. $Q_n(\infty; v_0)$	1.2064	1.1066	0.9935	0.9013	0.7754
skewness $Q_n(\infty; v_0)$	-0.2450	-0.2967	-0.3690	-0.4190	-0.4382
kurtosis $Q_n(\infty; v_0)$	0.0579	0.0787	0.1239	0.2699	0.3639

a

Size	*Generalized rank test*			*Supremum type test*		
of test	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.020	0.0136	0.0340	0.0216	0.0082	0.0544	0.0154
0.040	0.0299	0.0539	0.0422	0.0209	0.0832	0.0370
0.060	0.0510	0.0741	0.0642	0.0350	0.1092	0.0599
0.080	0.0713	0.0914	0.0852	0.0509	0.1334	0.0809
0.100	0.0923	0.1094	0.1053	0.0692	0.1592	0.1045

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.579	0.175	0.194	0.572	0.550	0.157	0.063	0.332
0.040	0.690	0.253	0.278	0.673	0.643	0.215	0.129	0.464
0.060	0.763	0.325	0.349	0.740	0.699	0.262	0.189	0.561
0.080	0.806	0.377	0.395	0.782	0.737	0.300	0.235	0.620
0.100	0.837	0.426	0.440	0.815	0.766	0.330	0.278	0.668
	Supremum type test							
0.020	0.626	0.188	0.214	0.644	0.619	0.183	0.023	0.199
0.040	0.740	0.281	0.303	0.746	0.725	0.267	0.068	0.376
0.060	0.798	0.347	0.376	0.808	0.780	0.326	0.121	0.487
0.080	0.842	0.406	0.433	0.848	0.814	0.370	0.165	0.565
0.100	0.874	0.459	0.486	0.878	0.845	0.411	0.212	0.633

c

Table 6.4: Tests based on stochastic integral type 11 (no estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.6880	-1.2767	0.0009	1.1440	2.1442
st.dev. $Q_n(\infty; v_0)$	1.1637	1.0741	1.0096	0.9064	0.8265
skewness $Q_n(\infty; v_0)$	-0.1545	-0.2768	-0.3440	-0.3658	-0.3906
kurtosis $Q_n(\infty; v_0)$	0.0754	0.0255	0.0323	0.2722	0.2386

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.010	0.0054	0.0205	0.0091	0.0041	0.0290	0.0077
0.020	0.0124	0.0320	0.0195	0.0096	0.0470	0.0172
0.030	0.0205	0.0431	0.0295	0.0156	0.0592	0.0272
0.040	0.0290	0.0520	0.0411	0.0226	0.0699	0.0365
0.050	0.0385	0.0612	0.0522	0.0304	0.0818	0.0483

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.533	0.121	0.160	0.574	0.514	0.111	0.040	0.303
0.020	0.633	0.183	0.227	0.657	0.601	0.164	0.087	0.435
0.030	0.701	0.231	0.275	0.711	0.655	0.198	0.122	0.510
0.040	0.746	0.271	0.310	0.745	0.701	0.232	0.162	0.575
0.050	0.782	0.309	0.345	0.773	0.733	0.258	0.196	0.620
	Supremum type test							
0.010	0.599	0.145	0.179	0.633	0.592	0.141	0.024	0.243
0.020	0.697	0.214	0.261	0.733	0.684	0.204	0.056	0.385
0.030	0.758	0.262	0.308	0.777	0.741	0.249	0.091	0.479
0.040	0.801	0.309	0.345	0.808	0.777	0.282	0.118	0.539
0.050	0.832	0.351	0.381	0.836	0.808	0.318	0.153	0.595

c

Table 6.5: Tests based on stochastic integral type 11 (no estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.8722	-1.3656	0.0048	1.2636	2.3894
st.dev. $Q_n(\infty; v_0)$	1.1257	1.0725	1.0015	0.9377	0.8719
skewness $Q_n(\infty; v_0)$	-0.1833	-0.0575	-0.1562	-0.1560	-0.3505
kurtosis $Q_n(\infty; v_0)$	0.1006	-0.0101	0.1590	0.0286	0.0881

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.005	0.0024	0.0084	0.0038	0.0018	0.0106	0.0033
0.010	0.0061	0.0142	0.0091	0.0050	0.0208	0.0089
0.015	0.0097	0.0183	0.0137	0.0082	0.0275	0.0124
0.020	0.0153	0.0238	0.0185	0.0119	0.0319	0.0175
0.025	0.0208	0.0293	0.0236	0.0167	0.0382	0.0229

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.506	0.089	0.111	0.512	0.483	0.080	0.032	0.293
0.010	0.617	0.146	0.161	0.605	0.579	0.126	0.072	0.417
0.015	0.676	0.181	0.192	0.653	0.631	0.154	0.098	0.482
0.020	0.734	0.226	0.228	0.697	0.670	0.177	0.119	0.529
0.025	0.771	0.258	0.256	0.730	0.701	0.200	0.143	0.571
	Supremum type test							
0.005	0.562	0.104	0.120	0.576	0.559	0.101	0.024	0.271
0.010	0.685	0.172	0.193	0.692	0.672	0.162	0.057	0.424
0.015	0.747	0.213	0.233	0.741	0.712	0.188	0.079	0.480
0.020	0.790	0.251	0.258	0.764	0.754	0.219	0.104	0.541
0.025	0.825	0.288	0.291	0.793	0.785	0.246	0.128	0.589

c

Table 6.6: Tests based on stochastic integral type 11 (no estimation) and $n = 100$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-1.9871	-0.9820	0.1001	1.2910	2.5929
st.dev. $Q_n(\infty; v_0)$	0.8302	0.9106	1.0119	1.1623	1.3168
skewness $Q_n(\infty; v_0)$	0.1902	0.1855	0.2282	0.2734	0.3854
kurtosis $Q_n(\infty; v_0)$	0.1021	0.0295	0.1244	0.0630	0.1389

a

Size	*Generalized rank test*			*Supremum type test*		
of test	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.020	0.0328	0.0110	0.0178	0.0258	0.0172	0.0217
0.040	0.0543	0.0256	0.0377	0.0459	0.0378	0.0430
0.060	0.0776	0.0433	0.0563	0.0658	0.0596	0.0632
0.080	0.0989	0.0599	0.0774	0.0821	0.0849	0.0845
0.100	0.1194	0.0783	0.0966	0.1024	0.1084	0.1050

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.580	0.176	0.194	0.570	0.329	0.058	0.180	0.549
0.040	0.689	0.253	0.278	0.673	0.463	0.112	0.244	0.635
0.060	0.762	0.323	0.348	0.740	0.546	0.156	0.291	0.686
0.080	0.806	0.377	0.396	0.781	0.618	0.201	0.334	0.727
0.100	0.837	0.427	0.440	0.815	0.666	0.236	0.368	0.753
	Supremum type test							
0.020	0.624	0.185	0.213	0.632	0.446	0.093	0.172	0.573
0.040	0.737	0.280	0.300	0.737	0.585	0.162	0.236	0.663
0.060	0.801	0.353	0.370	0.799	0.665	0.215	0.278	0.714
0.080	0.838	0.404	0.433	0.845	0.722	0.267	0.318	0.751
0.100	0.873	0.459	0.478	0.874	0.764	0.307	0.352	0.781

c

Table 6.7: Tests based on stochastic integral type 21 (no estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.2409	-1.1183	0.0731	1.3365	2.6797
st.dev. $Q_n(\infty; v_0)$	0.8588	0.9193	1.0127	1.0923	1.2178
skewness $Q_n(\infty; v_0)$	0.1778	0.0907	0.0592	0.1204	0.1574
kurtosis $Q_n(\infty; v_0)$	0.0927	-0.0266	-0.0785	0.0741	0.0863

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.010	0.0145	0.0078	0.0104	0.0133	0.0104	0.0115
0.020	0.0255	0.0155	0.0209	0.0229	0.0220	0.0234
0.030	0.0360	0.0244	0.0302	0.0325	0.0312	0.0352
0.040	0.0462	0.0320	0.0397	0.0412	0.0407	0.0448
0.050	0.0566	0.0404	0.0500	0.0510	0.0515	0.0545

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.533	0.122	0.161	0.576	0.360	0.054	0.132	0.529
0.020	0.635	0.184	0.226	0.657	0.478	0.093	0.187	0.611
0.030	0.702	0.232	0.275	0.711	0.542	0.126	0.225	0.653
0.040	0.747	0.271	0.310	0.745	0.588	0.156	0.252	0.687
0.050	0.781	0.309	0.346	0.773	0.630	0.182	0.280	0.715
	Supremum type test							
0.010	0.603	0.148	0.177	0.628	0.457	0.079	0.128	0.558
0.020	0.700	0.216	0.260	0.725	0.580	0.136	0.190	0.645
0.030	0.761	0.268	0.303	0.771	0.653	0.182	0.232	0.699
0.040	0.802	0.312	0.342	0.805	0.697	0.213	0.263	0.727
0.050	0.833	0.354	0.382	0.832	0.730	0.241	0.286	0.753

c

Table 6.8: Tests based on stochastic integral type 21 (no estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.4953	-1.2375	0.0551	1.4118	2.8098
st.dev. $Q_n(\infty; v_0)$	0.8843	0.9496	1.0046	1.0812	1.1612
skewness $Q_n(\infty; v_0)$	0.0675	0.2184	0.1538	0.1735	0.0019
kurtosis $Q_n(\infty; v_0)$	0.0828	0.0387	0.1109	0.0353	0.0139

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.005	0.0064	0.0031	0.0036	0.0060	0.0039	0.0038
0.010	0.0125	0.0065	0.0083	0.0117	0.0090	0.0095
0.015	0.0177	0.0094	0.0131	0.0166	0.0135	0.0157
0.020	0.0251	0.0135	0.0169	0.0220	0.0172	0.0209
0.025	0.0317	0.0176	0.0215	0.0283	0.0223	0.0254

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.509	0.091	0.112	0.514	0.320	0.036	0.089	0.457
0.010	0.615	0.146	0.160	0.602	0.440	0.067	0.130	0.547
0.015	0.676	0.181	0.192	0.653	0.513	0.092	0.161	0.604
0.020	0.735	0.227	0.229	0.697	0.553	0.112	0.183	0.639
0.025	0.771	0.258	0.255	0.729	0.591	0.132	0.204	0.669
	Supremum type test							
0.005	0.583	0.112	0.124	0.580	0.400	0.049	0.087	0.494
0.010	0.689	0.173	0.188	0.683	0.547	0.094	0.136	0.604
0.015	0.747	0.213	0.225	0.731	0.623	0.135	0.176	0.668
0.020	0.791	0.253	0.256	0.762	0.671	0.161	0.200	0.702
0.025	0.827	0.291	0.290	0.791	0.702	0.182	0.219	0.725

c

Table 6.9: Tests based on stochastic integral type 21 (no estimation) and $n = 100$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.4705	-1.2552	-0.0056	1.2677	2.4906
st.dev. $Q_n(\infty; v_0)$	0.9437	0.9867	0.9935	0.9986	0.9411
skewness $Q_n(\infty; v_0)$	0.1673	0.0898	-0.0437	-0.0476	-0.0150
kurtosis $Q_n(\infty; v_0)$	-0.0295	-0.0596	0.0041	-0.0771	-0.1089

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.020	0.0205	0.0202	0.0213	0.0201	0.0245	0.0231
0.040	0.0409	0.0420	0.0407	0.0422	0.0484	0.0441
0.060	0.0605	0.0607	0.0619	0.0624	0.0733	0.0680
0.080	0.0827	0.0830	0.0826	0.0798	0.0993	0.0918
0.100	0.1018	0.1041	0.1023	0.1007	0.1214	0.1123

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.679	0.216	0.223	0.681	0.576	0.142	0.152	0.582
0.040	0.783	0.316	0.329	0.793	0.678	0.215	0.224	0.682
0.060	0.836	0.388	0.394	0.840	0.740	0.273	0.283	0.746
0.080	0.876	0.452	0.458	0.879	0.783	0.319	0.328	0.791
0.100	0.899	0.497	0.507	0.902	0.814	0.359	0.363	0.820
	Supremum type test							
0.020	0.634	0.194	0.217	0.675	0.546	0.139	0.135	0.557
0.040	0.746	0.299	0.314	0.780	0.648	0.207	0.206	0.660
0.060	0.800	0.365	0.391	0.840	0.716	0.266	0.264	0.727
0.080	0.834	0.413	0.454	0.875	0.758	0.316	0.307	0.771
0.100	0.866	0.466	0.498	0.894	0.786	0.348	0.343	0.802

c

Table 6.10: Tests based on stochastic integral type 12 (no estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.7531	-1.4008	0.0090	1.3952	2.7374
st.dev. $Q_n(\infty; v_0)$	0.9678	0.9949	1.0041	0.9878	0.9591
skewness $Q_n(\infty; v_0)$	0.0421	-0.0946	-0.0353	-0.0335	-0.0963
kurtosis $Q_n(\infty; v_0)$	-0.0557	-0.0788	-0.1121	-0.0061	0.0510

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.010	0.0107	0.0117	0.0125	0.0099	0.0117	0.0130
0.020	0.0187	0.0223	0.0229	0.0199	0.0237	0.0220
0.030	0.0282	0.0318	0.0313	0.0284	0.0343	0.0324
0.040	0.0367	0.0415	0.0404	0.0376	0.0430	0.0434
0.050	0.0473	0.0514	0.0498	0.0477	0.0529	0.0518

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.678	0.186	0.189	0.688	0.599	0.137	0.132	0.601
0.020	0.755	0.250	0.271	0.777	0.688	0.192	0.187	0.686
0.030	0.810	0.305	0.325	0.819	0.731	0.227	0.222	0.730
0.040	0.840	0.349	0.371	0.851	0.767	0.260	0.258	0.765
0.050	0.869	0.392	0.409	0.874	0.794	0.288	0.286	0.790
	Supremum type test							
0.010	0.628	0.164	0.172	0.661	0.569	0.131	0.121	0.578
0.020	0.729	0.237	0.253	0.759	0.644	0.173	0.166	0.652
0.030	0.776	0.287	0.306	0.805	0.698	0.212	0.208	0.707
0.040	0.810	0.335	0.343	0.830	0.742	0.248	0.242	0.747
0.050	0.841	0.376	0.380	0.855	0.766	0.273	0.265	0.771

c

Table 6.11: Tests based on stochastic integral type 12 (no estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-3.0270	-1.5210	0.0066	1.5154	2.9940
st.dev. $Q_n(\infty; v_0)$	0.9825	1.0059	0.9934	1.0063	0.9878
skewness $Q_n(\infty; v_0)$	0.0169	0.1574	0.0199	0.0238	-0.1805
kurtosis $Q_n(\infty; v_0)$	0.0282	0.0133	0.0289	0.0087	-0.0607

a

Size	*Generalized rank test*			*Supremum type test*		
of	Left-	Right-	Two-	Left-	Right-	Two-
test	Sided	Sided	Sided	Sided	Sided	Sided
0.005	0.0042	0.0055	0.0042	0.0042	0.0057	0.0043
0.010	0.0110	0.0109	0.0092	0.0101	0.0121	0.0099
0.015	0.0168	0.0160	0.0150	0.0172	0.0174	0.0159
0.020	0.0211	0.0213	0.0219	0.0230	0.0236	0.0230
0.025	0.0261	0.0257	0.0282	0.0278	0.0276	0.0286

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.660	0.135	0.155	0.676	0.565	0.090	0.090	0.563
0.010	0.777	0.223	0.220	0.759	0.670	0.141	0.141	0.655
0.015	0.823	0.276	0.265	0.801	0.732	0.182	0.183	0.718
0.020	0.845	0.303	0.302	0.832	0.777	0.222	0.221	0.760
0.025	0.864	0.334	0.332	0.853	0.805	0.253	0.248	0.788
	Supremum type test							
0.005	0.626	0.124	0.143	0.660	0.541	0.088	0.084	0.542
0.010	0.739	0.199	0.214	0.752	0.649	0.137	0.134	0.644
0.015	0.798	0.259	0.254	0.790	0.708	0.176	0.172	0.702
0.020	0.829	0.301	0.293	0.824	0.754	0.212	0.209	0.746
0.025	0.846	0.327	0.316	0.840	0.777	0.237	0.233	0.770

c

Table 6.12: Tests based on stochastic integral type 12 (no estimation) and $n =$ 100. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.6827	-1.3144	-0.0069	1.2127	2.2095
st.dev. $Q_n(\infty; v_0)$	0.9600	0.9889	0.9949	0.9891	0.9857
skewness $Q_n(\infty; v_0)$	0.1601	-0.0420	0.1245	0.1567	-0.0424
kurtosis $Q_n(\infty; v_0)$	0.0668	0.0678	-0.0579	-0.0133	-0.0032

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.020	0.0244	0.0195	0.0218	0.0269	0.0229	0.0255
0.040	0.0435	0.0401	0.0435	0.0463	0.0464	0.0498
0.060	0.0612	0.0602	0.0637	0.0658	0.0693	0.0710
0.080	0.0806	0.0779	0.0836	0.0853	0.0926	0.0921
0.100	0.0986	0.0988	0.1024	0.1062	0.1154	0.1150

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.776	0.257	0.197	0.560	0.665	0.160	0.138	0.466
0.040	0.846	0.344	0.295	0.679	0.762	0.240	0.208	0.580
0.060	0.882	0.410	0.366	0.740	0.809	0.294	0.259	0.642
0.080	0.906	0.469	0.417	0.781	0.841	0.340	0.302	0.684
0.100	0.923	0.514	0.471	0.821	0.864	0.375	0.337	0.714
	Supremum type test							
0.020	0.756	0.245	0.183	0.535	0.651	0.160	0.128	0.441
0.040	0.823	0.331	0.276	0.650	0.745	0.236	0.193	0.549
0.060	0.862	0.396	0.346	0.716	0.792	0.287	0.239	0.604
0.080	0.889	0.451	0.400	0.760	0.823	0.332	0.275	0.649
0.100	0.909	0.500	0.449	0.794	0.847	0.373	0.312	0.685

c

Table 6.13: Tests based on stochastic integral type 12 (no estimation) and $n = 25$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.9356	-1.4346	0.0160	1.3357	2.5290
st.dev. $Q_n(\infty; v_0)$	0.9731	0.9889	0.9887	0.9866	0.9940
skewness $Q_n(\infty; v_0)$	0.1730	0.0761	-0.0530	0.0011	0.0696
kurtosis $Q_n(\infty; v_0)$	-0.0120	0.0398	-0.0368	-0.0892	-0.0176

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.010	0.0133	0.0100	0.0116	0.0138	0.0115	0.0140
0.020	0.0241	0.0196	0.0233	0.0261	0.0216	0.0252
0.030	0.0346	0.0285	0.0336	0.0357	0.0314	0.0362
0.040	0.0457	0.0384	0.0435	0.0481	0.0416	0.0466
0.050	0.0547	0.0492	0.0535	0.0572	0.0539	0.0574

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.771	0.214	0.158	0.576	0.671	0.134	0.115	0.499
0.020	0.837	0.294	0.229	0.676	0.755	0.200	0.172	0.601
0.030	0.872	0.353	0.283	0.731	0.796	0.242	0.212	0.654
0.040	0.898	0.403	0.333	0.778	0.825	0.277	0.243	0.691
0.050	0.912	0.435	0.379	0.814	0.848	0.310	0.275	0.722
	Supremum type test							
0.010	0.745	0.202	0.148	0.558	0.664	0.138	0.113	0.491
0.020	0.819	0.284	0.216	0.652	0.735	0.194	0.157	0.571
0.030	0.851	0.332	0.264	0.708	0.777	0.233	0.195	0.627
0.040	0.881	0.387	0.306	0.748	0.807	0.268	0.225	0.665
0.050	0.895	0.418	0.351	0.787	0.828	0.298	0.252	0.695

c

Table 6.14: Tests based on stochastic integral type 12 (no estimation) and $n = 50$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-3.1813	-1.5650	-0.0012	1.4840	2.8180
st.dev. $Q_n(\infty; v_0)$	0.9902	1.0035	0.9990	0.9930	0.9735
skewness $Q_n(\infty; v_0)$	0.0615	-0.0067	0.0172	0.0445	0.0627
kurtosis $Q_n(\infty; v_0)$	0.0272	0.0634	-0.0628	-0.0304	-0.1104

a

Size	*Generalized rank test*			*Supremum type test*		
of	Left-	Right-	Two-	Left-	Right-	Two-
test	Sided	Sided	Sided	Sided	Sided	Sided
0.005	0.0062	0.0050	0.0055	0.0059	0.0056	0.0061
0.010	0.0108	0.0096	0.0112	0.0116	0.0098	0.0116
0.015	0.0155	0.0139	0.0157	0.0177	0.0152	0.0167
0.020	0.0210	0.0204	0.0205	0.0225	0.0206	0.0214
0.025	0.0263	0.0252	0.0249	0.0267	0.0268	0.0274

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.753	0.175	0.134	0.599	0.661	0.115	0.096	0.515
0.010	0.813	0.232	0.193	0.685	0.741	0.167	0.144	0.613
0.015	0.847	0.277	0.234	0.731	0.781	0.198	0.173	0.660
0.020	0.875	0.315	0.286	0.779	0.807	0.227	0.200	0.692
0.025	0.893	0.349	0.317	0.807	0.826	0.251	0.221	0.718
	Supremum type test							
0.005	0.725	0.161	0.128	0.575	0.646	0.115	0.091	0.498
0.010	0.796	0.227	0.176	0.655	0.723	0.159	0.130	0.581
0.015	0.838	0.279	0.225	0.714	0.762	0.193	0.161	0.633
0.020	0.862	0.308	0.262	0.751	0.788	0.219	0.186	0.668
0.025	0.876	0.335	0.299	0.785	0.813	0.247	0.213	0.700

c

Table 6.15: Tests based on stochastic integral type 12 (no estimation) and $n = 100$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.2360	-1.1624	0.0380	1.4089	2.9266
st.dev. $Q_n(\infty; v_0)$	0.8033	0.9057	1.0068	1.1480	1.2667
skewness $Q_n(\infty; v_0)$	0.3294	0.3004	0.2372	0.2789	0.3802
kurtosis $Q_n(\infty; v_0)$	0.1695	0.1496	0.2227	0.1071	0.1084

a

Size	*Generalized rank test*			*Supremum type test*		
of	Left-	Right-	Two-	Left-	Right-	Two-
test	Sided	Sided	Sided	Sided	Sided	Sided
0.020	0.0295	0.0115	0.0172	0.0316	0.0129	0.0193
0.040	0.0519	0.0306	0.0386	0.0577	0.0318	0.0425
0.060	0.0721	0.0479	0.0600	0.0790	0.0541	0.0634
0.080	0.0943	0.0695	0.0801	0.0981	0.0795	0.0876
0.100	0.1140	0.0908	0.0995	0.1192	0.1007	0.1112

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.680	0.217	0.222	0.682	0.448	0.083	0.196	0.648
0.040	0.783	0.316	0.331	0.796	0.597	0.156	0.276	0.742
0.060	0.837	0.388	0.395	0.842	0.684	0.222	0.330	0.794
0.080	0.876	0.452	0.458	0.879	0.736	0.269	0.369	0.825
0.100	0.900	0.498	0.508	0.903	0.777	0.312	0.403	0.846
	Supremum type test							
0.020	0.631	0.193	0.219	0.679	0.410	0.074	0.194	0.647
0.040	0.744	0.297	0.316	0.784	0.553	0.142	0.272	0.737
0.060	0.798	0.363	0.393	0.842	0.632	0.195	0.316	0.783
0.080	0.834	0.412	0.458	0.876	0.695	0.248	0.362	0.817
0.100	0.865	0.463	0.499	0.895	0.738	0.293	0.399	0.845

c

Table 6.16: Tests based on stochastic integral type 22 (no estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.5403	-1.3217	0.0401	1.5004	3.0692
st.dev. $Q_n(\infty; v_0)$	0.8433	0.9244	1.0094	1.0916	1.1856
skewness $Q_n(\infty; v_0)$	0.1642	0.0602	0.1406	0.1864	0.1680
kurtosis $Q_n(\infty; v_0)$	0.0072	-0.0212	-0.0447	0.0576	0.1438

a

Size of test	*Generalized rank test*			*Supremum type test*		
	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.010	0.0157	0.0073	0.0113	0.0158	0.0066	0.0108
0.020	0.0249	0.0156	0.0213	0.0280	0.0154	0.0210
0.030	0.0349	0.0240	0.0313	0.0378	0.0243	0.0312
0.040	0.0447	0.0334	0.0395	0.0483	0.0318	0.0412
0.050	0.0553	0.0426	0.0492	0.0591	0.0411	0.0514

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.682	0.188	0.190	0.691	0.511	0.094	0.170	0.663
0.020	0.756	0.251	0.269	0.777	0.614	0.146	0.226	0.736
0.030	0.808	0.303	0.324	0.820	0.682	0.187	0.269	0.778
0.040	0.842	0.351	0.373	0.852	0.719	0.216	0.296	0.800
0.050	0.869	0.392	0.409	0.875	0.754	0.250	0.327	0.821
	Supremum type test							
0.010	0.630	0.165	0.175	0.666	0.453	0.080	0.161	0.645
0.020	0.727	0.236	0.254	0.760	0.561	0.126	0.217	0.719
0.030	0.776	0.286	0.307	0.807	0.628	0.164	0.256	0.762
0.040	0.810	0.334	0.344	0.832	0.676	0.195	0.287	0.791
0.050	0.840	0.374	0.382	0.857	0.713	0.224	0.315	0.811

c

Table 6.17: Tests based on stochastic integral type 22 (no estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.8397	-1.4559	0.0280	1.5959	3.2537
st.dev. $Q_n(\infty; v_0)$	0.8787	0.9488	0.9965	1.0835	1.1513
skewness $Q_n(\infty; v_0)$	0.1150	0.2746	0.1530	0.1764	-0.0241
kurtosis $Q_n(\infty; v_0)$	0.0801	0.0891	0.0922	0.0397	-0.0521

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.005	0.0062	0.0034	0.0044	0.0066	0.0031	0.0049
0.010	0.0145	0.0079	0.0096	0.0140	0.0082	0.0097
0.015	0.0211	0.0123	0.0141	0.0225	0.0124	0.0154
0.020	0.0255	0.0169	0.0200	0.0295	0.0175	0.0202
0.025	0.0310	0.0211	0.0262	0.0346	0.0212	0.0260

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.661	0.135	0.153	0.675	0.503	0.069	0.125	0.631
0.010	0.779	0.224	0.221	0.762	0.620	0.113	0.179	0.714
0.015	0.825	0.276	0.267	0.804	0.679	0.147	0.212	0.752
0.020	0.845	0.302	0.303	0.834	0.730	0.180	0.244	0.785
0.025	0.865	0.335	0.334	0.854	0.766	0.211	0.274	0.810
	Supremum type test							
0.005	0.627	0.124	0.141	0.656	0.482	0.065	0.127	0.630
0.010	0.737	0.197	0.217	0.754	0.581	0.103	0.172	0.702
0.015	0.797	0.259	0.255	0.791	0.650	0.138	0.211	0.749
0.020	0.830	0.302	0.293	0.824	0.689	0.164	0.235	0.773
0.025	0.846	0.327	0.316	0.841	0.726	0.190	0.260	0.796

c

Table 6.18: Tests based on stochastic integral type 22 (no estimation) and $n = 100$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.2593	-1.1703	0.0334	1.4153	2.9081
st.dev. $Q_n(\infty; v_0)$	0.7537	0.8685	1.0072	1.1883	1.4242
skewness $Q_n(\infty; v_0)$	0.3465	0.1826	0.3746	0.4447	0.2876
kurtosis $Q_n(\infty; v_0)$	0.3054	0.2065	0.1635	0.3483	0.2306

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.020	0.0328	0.0115	0.0183	0.0391	0.0128	0.0199
0.040	0.0533	0.0293	0.0413	0.0615	0.0321	0.0462
0.060	0.0734	0.0479	0.0620	0.0821	0.0511	0.0712
0.080	0.0919	0.0651	0.0813	0.1004	0.0745	0.0932
0.100	0.1095	0.0875	0.1011	0.1233	0.0973	0.1142

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.725	0.224	0.225	0.656	0.469	0.076	0.207	0.633
0.040	0.811	0.316	0.327	0.750	0.636	0.155	0.288	0.714
0.060	0.860	0.385	0.400	0.805	0.714	0.216	0.334	0.756
0.080	0.891	0.443	0.448	0.838	0.766	0.264	0.375	0.786
0.100	0.910	0.490	0.500	0.867	0.802	0.307	0.409	0.811
	Supremum type test							
0.020	0.694	0.209	0.224	0.653	0.412	0.062	0.202	0.629
0.040	0.779	0.292	0.323	0.748	0.586	0.137	0.282	0.713
0.060	0.826	0.359	0.386	0.794	0.673	0.195	0.337	0.757
0.080	0.857	0.409	0.446	0.832	0.727	0.240	0.374	0.785
0.100	0.884	0.463	0.491	0.860	0.764	0.280	0.404	0.805

c

Table 6.19: Tests based on stochastic integral type 22 (no estimation) and $n = 25$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.5515	-1.3127	0.0446	1.4920	3.0873
st.dev. $Q_n(\infty; v_0)$	0.7966	0.8895	0.9965	1.1267	1.3034
skewness $Q_n(\infty; v_0)$	0.3045	0.2346	0.1233	0.1841	0.2629
kurtosis $Q_n(\infty; v_0)$	0.0985	0.1021	0.0337	0.0002	0.1047

a

Size	*Generalized rank test*			*Supremum type test*		
of	Left-	Right-	Two-	Left-	Right-	Two-
test	Sided	Sided	Sided	Sided	Sided	Sided
0.010	0.0182	0.0059	0.0095	0.0199	0.0064	0.0096
0.020	0.0304	0.0139	0.0210	0.0331	0.0147	0.0223
0.030	0.0408	0.0221	0.0307	0.0462	0.0230	0.0334
0.040	0.0533	0.0313	0.0404	0.0586	0.0315	0.0438
0.050	0.0630	0.0414	0.0517	0.0694	0.0414	0.0551

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.724	0.192	0.179	0.656	0.494	0.070	0.163	0.633
0.020	0.805	0.271	0.259	0.746	0.637	0.127	0.228	0.715
0.030	0.845	0.324	0.314	0.796	0.698	0.170	0.269	0.756
0.040	0.879	0.380	0.365	0.828	0.742	0.206	0.302	0.787
0.050	0.896	0.415	0.404	0.855	0.780	0.242	0.337	0.811
	Supremum type test							
0.010	0.688	0.174	0.174	0.653	0.439	0.059	0.154	0.624
0.020	0.768	0.245	0.251	0.737	0.591	0.112	0.225	0.711
0.030	0.818	0.305	0.302	0.784	0.658	0.152	0.264	0.751
0.040	0.849	0.353	0.343	0.813	0.703	0.186	0.295	0.779
0.050	0.871	0.391	0.383	0.840	0.739	0.217	0.324	0.800

c

Table 6.20: Tests based on stochastic integral type 22 (no estimation) and $n = 50$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v_0)$	-2.8453	-1.4611	0.0189	1.6074	3.2586
st.dev. $Q_n(\infty; v_0)$	0.8423	0.9220	1.0016	1.0965	1.1892
skewness $Q_n(\infty; v_0)$	0.1730	0.1140	0.1342	0.1735	0.1851
kurtosis $Q_n(\infty; v_0)$	0.0830	0.1155	-0.0324	0.0143	-0.0617

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.005	0.0087	0.0034	0.0051	0.0091	0.0036	0.0059
0.010	0.0139	0.0071	0.0102	0.0161	0.0071	0.0101
0.015	0.0195	0.0107	0.0149	0.0231	0.0108	0.0156
0.020	0.0254	0.0165	0.0198	0.0282	0.0159	0.0201
0.025	0.0314	0.0207	0.0244	0.0329	0.0207	0.0263

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.714	0.158	0.157	0.670	0.530	0.070	0.138	0.642
0.010	0.781	0.213	0.217	0.743	0.638	0.114	0.189	0.710
0.015	0.823	0.261	0.257	0.781	0.693	0.144	0.222	0.748
0.020	0.853	0.300	0.309	0.825	0.734	0.172	0.248	0.773
0.025	0.876	0.332	0.338	0.845	0.762	0.198	0.274	0.794
	Supremum type test							
0.005	0.678	0.147	0.152	0.663	0.506	0.068	0.139	0.642
0.010	0.760	0.207	0.209	0.730	0.591	0.100	0.180	0.699
0.015	0.807	0.257	0.249	0.771	0.657	0.133	0.218	0.739
0.020	0.830	0.286	0.291	0.808	0.695	0.157	0.242	0.765
0.025	0.850	0.312	0.323	0.831	0.732	0.182	0.269	0.790

c

Table 6.21: Tests based on stochastic integral type 22 (no estimation) and $n = 100$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

Sample size	*Type of Censoring*	H_{LL}	H_L	H_0	H_R	H_{RR}
25	None	-1.7579	-0.8789	0.0000	0.8789	1.7579
25	Exp(1)	-3.0447	-1.5224	0.0000	1.5224	3.0447
50	None	-1.3703	-0.6852	0.0000	0.6852	1.3703
50	Exp(1)	-2.3735	-1.1867	0.0000	1.1867	2.3735
100	None	-1.0513	-0.5257	0.0000	0.5257	1.0513
100	Exp(1)	-1.8209	-0.9105	0.0000	0.9105	1.8209

Table 6.22: Values of θ_n. The value of v_n, v_0, θ_0 and ρ are always equal to 0, 0, 0 and 1, respectively.

6.3 The composite null hypothesis

As in the previous section, the simulated samples are drawn from a distribution $F(t; v_n, \theta_n)$ belonging to the family of Harrington and Fleming alternatives to the exponential distribution. The values of v_n and ρ are 0 and 1, respectively. However, the parameter of interest θ_n varies in a different way with sample size:

$$\theta_n = c_H \sqrt{\frac{2 \log n}{n\{\mu_{2\rho}(v_0) - (\mu_\rho(v_0))^2/\mu_0(v_0)\}}}, \tag{6.8}$$

where v_0 is equal to zero. Again c_H is chosen according to the five types of hypotheses H_{LL}, H_L, H_0, H_R and H_{RR} under which simulation takes place; the respective choices of c_H are -1.0, -0.5, 0.0, 0.5 and 1.0, resulting in the values of θ_n listed in Table 6.22. The asymptotic power of the generalized rank test based on stochastic integral 13 and maximum likelihood estimation is given in Table 6.3. For size $1/n$ the power of a one-sided version of this test will under H_{RR} and H_{LL} tend to 0.5 as n tends to infinity.

6.3.1 The maximum likelihood estimator

On the basis of Theorem 7 and its corollaries we expect that $v^{(n)}$ has a distribution which is close to Gaussian if the sample size is sufficiently large. Moreover, under H_0 the mean is close to zero and the variance should be near $(nH^1(\infty; v_0, \theta_0))^{-1}$.

As far as H_0 concerns, in Table 6.23 we see that although the mean of $v^{(n)}$ deviates from zero in uncensored samples, it tends to zero rather quickly [roughly at a rate n^{-1}]. For censored samples the deviation is unnoticable. The variance of $v^{(n)}$ is in both tables as indicated.

	H_{LL}	H_L	H_0	H_R	H_{RR}
	Sample size 25; no censoring				
mean $v^{(n)}$	-0.7436	-0.3936	0.0177	0.5029	1.0580
st.dev. $v^{(n)}$	0.1331	0.1643	0.2002	0.2496	0.3153
skewness $v^{(n)}$	0.2820	0.2259	0.2756	0.1672	0.0561
kurtosis $v^{(n)}$	0.0179	0.0488	0.1580	0.0956	-0.0528
	Sample size 50; no censoring				
mean $v^{(n)}$	-0.6023	-0.3167	0.0101	0.3782	0.7834
st.dev. $v^{(n)}$	0.1016	0.1201	0.1411	0.1694	0.2021
skewness $v^{(n)}$	0.0718	0.2544	-0.0947	0.1701	0.1833
kurtosis $v^{(n)}$	0.0658	0.0818	0.0426	0.0462	0.0115
	Sample size 100; no censoring				
mean $v^{(n)}$	-0.4773	-0.2475	0.0044	0.2824	0.5787
st.dev. $v^{(n)}$	0.0766	0.0875	0.0995	0.1143	0.1300
skewness $v^{(n)}$	0.1303	-0.0761	0.0374	0.1051	0.0959
kurtosis $v^{(n)}$	-0.0133	-0.0021	-0.0002	0.0058	0.0374
	Sample size 25; standard exponential censoring				
mean $v^{(n)}$	-1.5431	-0.8943	-0.0015	1.1522	2.5095
st.dev. $v^{(n)}$	0.2363	0.2581	0.2950	0.3486	0.3997
skewness $v^{(n)}$	-0.5300	-0.2887	-0.1889	-0.2045	-0.1894
kurtosis $v^{(n)}$	0.8844	0.5535	0.2429	0.1699	0.2315
	Sample size 50; standard exponential censoring				
mean $v^{(n)}$	-1.2828	-0.7154	0.0009	0.8723	1.8736
st.dev. $v^{(n)}$	0.1669	0.1812	0.2025	0.2336	0.2721
skewness $v^{(n)}$	-0.2753	-0.2176	-0.1274	-0.0040	-0.1408
kurtosis $v^{(n)}$	0.1905	0.2161	0.1948	0.1749	0.0635
	Sample size 100; standard exponential censoring				
mean $v^{(n)}$	-1.0317	-0.5621	0.0002	0.6539	1.3898
st.dev. $v^{(n)}$	0.1221	0.1313	0.1425	0.1586	0.1809
skewness $v^{(n)}$	-0.2834	-0.0988	0.0171	-0.0144	0.1241
kurtosis $v^{(n)}$	0.2477	0.1269	0.0046	-0.0584	0.0962

Table 6.23: Description of the behavior of the maximum likelihood estimator.

6.3.2 Stochastic integral 12

We have already encountered stochastic integral 12 in discussing the simulation results for the simple null hypothesis. The inclusion of this stochastic integral in the simulations for the composite null hypothesis is intended to show the perils of plugging in estimators and subsequently carrying on as under the simple null hypothesis.

Since stochastic integral 12 is scaled for the simple null hypothesis, we do not have that $a_R(v_0)$ and $a_S(v_0)$ are equal to 1. Instead, we have $a_R(v_0) = 4$ and $a_S(v_0)$ is even larger. As a consequence, the computed asymptotic P-values are far too conservative [see Tables 6.24b, 6.25b and 6.26b].

The magnitude of the computed asymptotic P-values raised some difficulties with the storage of the simulation results, These were solved by choosing the interval width of the histogram of the asymptotic P-values differently, which unfortunately made the precision of the simulated power decrease.

If we are using critical values as given in Tables 6.24b, 6.25b and 6.26b, then Theorem 13 on page 67 yields that the generalized rank test is optimal. However, the supremum type test is not. It is clear from Tables 6.24c, 6.25c and 6.26c that a supremum type test is outdistanced by the corresponding generalized rank test.

6.3.3 Stochastic integral 13

Stochastic integral 13 is based on the weight process for known censoring distribution proposed in section 5.3. The quantities $e_R(h, v_0)$ and $e_S(h, v_0)$ are given by

$$e_R(h, v_0) = e_S(h, v_0) = \sqrt{\mu_{2\rho}(v_0) - (\mu_\rho(v_0))^2/\mu_0(v_0)}. \tag{6.9}$$

If censoring is absent then both $e_R(h, v_0)$ and $e_S(h, v_0)$ are equal to $1/12$. If the censoring distribution is standard exponential they become $1/36$.

Since the generalized rank tests based on stochastic integrals 12 and 13 differ from each other by a multitude of $M_n(\infty; v^{(n)}, \theta_0)$, which is zero by the definition of the maximum likelihood estimator, their power should be about the same. The differences between the simulated power of generalized rank tests in Table 6.24c and Table 6.27c should be adressed to the inaccuracy of the simulated power mentioned in the discussion of stochastic integral 12. The same holds for Tables 6.25c and 6.28c, and Tables 6.26c and 6.29c.

The supremum type tests based on stochastic integral 13 perform better than those based on stochastic integral 12. The only exception is the two-sided supremum type test when applied to a situation in which θ_n is negative.

The simulated power of the generalized rank test tends to be lower than the asymptotic power given in Table 6.3, especially for the smaller test sizes. This effect is stronger for samples where censoring is present.

6.3.4 Stochastic integral 23

The weight process for unknown censoring distribution proposed in section 5.3 leads to stochastic integral 23. For this stochastic integral $e_R(h, v_0)$ and $e_S(h, v_0)$ coincide with those for stochastic integral 13, and hence are given by (6.9).

On the whole, tests based on stochastic integral 23 seem to be more sensitive to negative values, and less sensitive to positive values than the corresponding tests based on stochastic integral 13. The difference in sensitivity is stronger for two-sided tests than for one-sided tests, for supremum type tests than for generalized rank tests, and for small samples than for large samples. In rather surprising contrast to two-sided tests, the simulated power of one-sided tests in uncensored samples remains relatively unaffected by the choice between the two stochastic integrals.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-1.0809	-0.6409	-0.0898	0.5314	1.1619
st.dev. $Q_n(\infty; v^{(n)})$	0.3097	0.3921	0.4905	0.6101	0.7482
skewness $Q_n(\infty; v^{(n)})$	0.4893	0.3543	0.3055	0.3904	0.3986
kurtosis $Q_n(\infty; v^{(n)})$	0.3068	0.0904	0.2030	0.1534	0.1028

a

Size of test	*Generalized rank test*			*Supremum type test*		
	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.020	0.1583	0.1524	0.2536	0.0666	0.0896	0.1000
0.040	0.1865	0.2070	0.3105	0.0963	0.1324	0.1533
0.060	0.2065	0.2416	0.3548	0.1207	0.1689	0.1887
0.080	0.2268	0.2686	0.3862	0.1435	0.1983	0.2158
0.100	0.2436	0.2914	0.4135	0.1641	0.2253	0.2504

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.625	0.182	0.206	0.553	0.450	0.092	0.161	0.491
0.040	0.743	0.276	0.305	0.661	0.610	0.174	0.213	0.560
0.060	0.810	0.346	0.371	0.717	0.709	0.247	0.253	0.606
0.080	0.861	0.412	0.420	0.757	0.767	0.299	0.283	0.636
0.100	0.893	0.470	0.465	0.786	0.810	0.348	0.311	0.662
	Supremum type test							
0.020	0.542	0.155	0.180	0.496	0.453	0.107	0.103	0.368
0.040	0.669	0.243	0.262	0.599	0.593	0.186	0.158	0.458
0.060	0.744	0.311	0.326	0.668	0.662	0.239	0.193	0.510
0.080	0.797	0.367	0.373	0.708	0.706	0.278	0.222	0.544
0.100	0.836	0.421	0.415	0.738	0.756	0.325	0.256	0.586

c

Table 6.24: Tests based on stochastic integral type 12 (maximum likelihood estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-1.2453	-0.7006	-0.0607	0.6449	1.3887
st.dev. $Q_n(\infty; v^{(n)})$	0.3464	0.4138	0.4950	0.5834	0.6955
skewness $Q_n(\infty; v^{(n)})$	0.3151	0.3780	0.2343	0.1073	0.0958
kurtosis $Q_n(\infty; v^{(n)})$	0.0730	0.1572	0.0186	-0.1256	0.0173

a

Size of test	*Generalized rank test*			*Supremum type test*		
	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.010	0.1283	0.1230	0.2050	0.0434	0.0580	0.0641
0.020	0.1530	0.1576	0.2497	0.0662	0.0850	0.0997
0.030	0.1683	0.1819	0.2828	0.0841	0.1084	0.1244
0.040	0.1879	0.2001	0.3094	0.0992	0.1291	0.1484
0.050	0.1998	0.2148	0.3302	0.1122	0.1446	0.1698

b

Size of test	*One-sided test*				*Two-sided test*			
	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.644	0.150	0.188	0.622	0.496	0.078	0.145	0.562
0.020	0.748	0.222	0.269	0.705	0.626	0.140	0.191	0.626
0.030	0.798	0.273	0.319	0.753	0.706	0.187	0.232	0.669
0.040	0.849	0.340	0.361	0.783	0.753	0.227	0.263	0.698
0.050	0.875	0.378	0.392	0.804	0.788	0.262	0.285	0.720
	Supremum type test							
0.010	0.541	0.114	0.153	0.539	0.447	0.076	0.082	0.407
0.020	0.668	0.191	0.221	0.631	0.584	0.136	0.133	0.504
0.030	0.738	0.254	0.276	0.689	0.650	0.177	0.164	0.557
0.040	0.786	0.301	0.320	0.731	0.703	0.220	0.195	0.599
0.050	0.816	0.339	0.352	0.758	0.740	0.258	0.223	0.631

c

Table 6.25: Tests based on stochastic integral type 12 (maximum likelihood estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-1.3881	-0.7630	-0.0323	0.7286	1.5403
st.dev. $Q_n(\infty; v^{(n)})$	0.3779	0.4358	0.4986	0.5694	0.6417
skewness $Q_n(\infty; v^{(n)})$	0.1092	0.2182	0.2227	0.0705	0.0306
kurtosis $Q_n(\infty; v^{(n)})$	-0.0570	0.0015	0.0517	-0.0068	0.0099

a

Size of test	*Generalized rank test*			*Supremum type test*		
	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.005	0.1137	0.0877	0.1591	0.0339	0.0331	0.0440
0.010	0.1339	0.1142	0.1996	0.0489	0.0497	0.0672
0.015	0.1480	0.1334	0.2305	0.0588	0.0617	0.0834
0.020	0.1584	0.1471	0.2510	0.0696	0.0739	0.0985
0.025	0.1678	0.1598	0.2675	0.0798	0.0841	0.1109

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.691	0.157	0.139	0.610	0.484	0.067	0.121	0.579
0.010	0.775	0.220	0.200	0.698	0.614	0.114	0.165	0.653
0.015	0.821	0.265	0.247	0.746	0.699	0.161	0.204	0.701
0.020	0.846	0.299	0.281	0.777	0.743	0.192	0.227	0.729
0.025	0.867	0.330	0.313	0.799	0.774	0.220	0.248	0.747
	Supremum type test							
0.005	0.561	0.113	0.111	0.519	0.446	0.068	0.074	0.427
0.010	0.663	0.167	0.164	0.612	0.560	0.112	0.113	0.525
0.015	0.716	0.203	0.201	0.662	0.623	0.142	0.140	0.574
0.020	0.763	0.241	0.236	0.699	0.665	0.168	0.163	0.610
0.025	0.796	0.276	0.265	0.726	0.700	0.189	0.183	0.638

c

Table 6.26: Tests based on stochastic integral type 12 (maximum likelihood estimation) and $n = 100$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.1617	-1.2819	-0.1796	1.0628	2.3237
st.dev. $Q_n(\infty; v^{(n)})$	0.6195	0.7841	0.9811	1.2203	1.4965
skewness $Q_n(\infty; v^{(n)})$	0.4893	0.3543	0.3055	0.3903	0.3986
kurtosis $Q_n(\infty; v^{(n)})$	0.3068	0.0904	0.2030	0.1534	0.1028

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.020	0.0225	0.0198	0.0226	0.0422	0.0134	0.0222
0.040	0.0377	0.0512	0.0430	0.0699	0.0324	0.0474
0.060	0.0507	0.0806	0.0646	0.0920	0.0546	0.0722
0.080	0.0673	0.1082	0.0831	0.1180	0.0792	0.0985
0.100	0.0826	0.1360	0.1017	0.1404	0.1043	0.1246

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.623	0.181	0.205	0.552	0.451	0.093	0.161	0.492
0.040	0.745	0.278	0.305	0.660	0.613	0.175	0.214	0.560
0.060	0.809	0.345	0.372	0.718	0.710	0.248	0.254	0.607
0.080	0.862	0.414	0.420	0.757	0.766	0.298	0.283	0.636
0.100	0.894	0.470	0.465	0.785	0.809	0.347	0.311	0.661
	Supremum type test							
0.020	0.618	0.180	0.217	0.562	0.292	0.042	0.202	0.544
0.040	0.739	0.278	0.295	0.649	0.471	0.102	0.266	0.620
0.060	0.802	0.342	0.361	0.707	0.576	0.154	0.308	0.661
0.080	0.853	0.408	0.411	0.747	0.657	0.206	0.348	0.694
0.100	0.882	0.453	0.456	0.779	0.714	0.256	0.383	0.723

c

Table 6.27: Tests based on stochastic integral type 13 (maximum likelihood estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.4907	-1.4012	-0.1215	1.2897	2.7773
st.dev. $Q_n(\infty; v^{(n)})$	0.6929	0.8275	0.9900	1.1667	1.3910
skewness $Q_n(\infty; v^{(n)})$	0.3151	0.3780	0.2343	0.1073	0.0958
kurtosis $Q_n(\infty; v^{(n)})$	0.0730	0.1572	0.0186	-0.1256	0.0173

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.010	0.0117	0.0101	0.0112	0.0200	0.0061	0.0107
0.020	0.0203	0.0224	0.0215	0.0346	0.0154	0.0215
0.030	0.0274	0.0349	0.0320	0.0496	0.0250	0.0351
0.040	0.0381	0.0461	0.0419	0.0627	0.0354	0.0466
0.050	0.0461	0.0563	0.0516	0.0764	0.0444	0.0578

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.643	0.150	0.188	0.621	0.495	0.078	0.145	0.561
0.020	0.748	0.221	0.270	0.705	0.627	0.140	0.192	0.627
0.030	0.799	0.272	0.320	0.754	0.708	0.189	0.232	0.671
0.040	0.848	0.338	0.360	0.783	0.753	0.227	0.262	0.698
0.050	0.875	0.378	0.391	0.802	0.788	0.262	0.285	0.720
	Supremum type test							
0.010	0.630	0.144	0.179	0.603	0.357	0.040	0.171	0.592
0.020	0.738	0.218	0.259	0.692	0.501	0.081	0.223	0.656
0.030	0.800	0.281	0.313	0.739	0.603	0.128	0.272	0.705
0.040	0.839	0.325	0.351	0.772	0.661	0.164	0.305	0.732
0.050	0.866	0.372	0.384	0.794	0.705	0.192	0.328	0.751

c

Table 6.28: Tests based on stochastic integral type 13 (maximum likelihood estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.7763	-1.5259	-0.0645	1.4571	3.0805
st.dev. $Q_n(\infty; v^{(n)})$	0.7558	0.8716	0.9972	1.1389	1.2834
skewness $Q_n(\infty; v^{(n)})$	0.1092	0.2182	0.2227	0.0705	0.0306
kurtosis $Q_n(\infty; v^{(n)})$	-0.0570	0.0015	0.0517	-0.0068	0.0099

a

Size	*Generalized rank test*			*Supremum type test*		
of test	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.005	0.0080	0.0036	0.0049	0.0124	0.0026	0.0048
0.010	0.0134	0.0082	0.0104	0.0208	0.0058	0.0094
0.015	0.0179	0.0129	0.0163	0.0276	0.0102	0.0149
0.020	0.0226	0.0181	0.0214	0.0341	0.0138	0.0219
0.025	0.0267	0.0230	0.0260	0.0408	0.0182	0.0263

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.693	0.157	0.141	0.616	0.486	0.068	0.121	0.580
0.010	0.776	0.221	0.203	0.701	0.617	0.115	0.165	0.654
0.015	0.819	0.261	0.245	0.744	0.696	0.159	0.203	0.700
0.020	0.846	0.298	0.282	0.777	0.741	0.191	0.226	0.727
0.025	0.865	0.327	0.312	0.798	0.772	0.216	0.246	0.745
	Supremum type test							
0.005	0.674	0.146	0.139	0.605	0.387	0.043	0.135	0.600
0.010	0.760	0.212	0.194	0.681	0.500	0.074	0.177	0.662
0.015	0.803	0.253	0.242	0.733	0.579	0.101	0.214	0.704
0.020	0.832	0.288	0.271	0.759	0.651	0.133	0.248	0.739
0.025	0.856	0.317	0.304	0.785	0.686	0.154	0.266	0.755

c

Table 6.29: Tests based on stochastic integral type 13 (maximum likelihood estimation) and $n = 100$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-1.8702	-1.1822	-0.1362	0.9217	1.6082
st.dev. $Q_n(\infty; v^{(n)})$	0.7950	0.8010	0.9887	1.3607	1.7703
skewness $Q_n(\infty; v^{(n)})$	-0.3152	-0.1281	0.3725	0.7136	0.9926
kurtosis $Q_n(\infty; v^{(n)})$	-0.0514	0.1657	0.5918	1.0510	1.7083

a

Size	*Generalized rank test*			*Supremum type test*		
of test	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.020	0.0204	0.0188	0.0178	0.0353	0.0118	0.0151
0.040	0.0381	0.0464	0.0394	0.0657	0.0327	0.0407
0.060	0.0552	0.0733	0.0617	0.0945	0.0593	0.0690
0.080	0.0715	0.1030	0.0817	0.1200	0.0834	0.0958
0.100	0.0885	0.1310	0.1035	0.1449	0.1091	0.1255

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.396	0.138	0.180	0.336	0.256	0.073	0.140	0.278
0.040	0.527	0.224	0.253	0.421	0.389	0.135	0.187	0.342
0.060	0.615	0.294	0.304	0.475	0.481	0.190	0.224	0.386
0.080	0.679	0.354	0.351	0.520	0.543	0.237	0.252	0.415
0.100	0.728	0.409	0.391	0.555	0.599	0.280	0.278	0.440
	Supremum type test							
0.020	0.384	0.136	0.183	0.338	0.165	0.038	0.160	0.309
0.040	0.512	0.220	0.257	0.421	0.288	0.090	0.220	0.378
0.060	0.598	0.288	0.315	0.481	0.380	0.135	0.265	0.428
0.080	0.654	0.342	0.354	0.521	0.447	0.173	0.298	0.460
0.100	0.707	0.393	0.389	0.555	0.502	0.214	0.330	0.494

c

Table 6.30: Tests based on stochastic integral type 13 (maximum likelihood estimation) and $n = 25$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.2354	-1.3145	-0.1076	1.2216	2.2743
st.dev. $Q_n(\infty; v^{(n)})$	0.7809	0.8333	0.9853	1.2965	1.7302
skewness $Q_n(\infty; v^{(n)})$	-0.1759	0.0189	0.2438	0.3669	0.7487
kurtosis $Q_n(\infty; v^{(n)})$	0.0986	0.1111	0.2499	0.4027	0.7663

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.010	0.0136	0.0079	0.0103	0.0222	0.0053	0.0081
0.020	0.0221	0.0186	0.0207	0.0376	0.0156	0.0226
0.030	0.0324	0.0332	0.0313	0.0516	0.0239	0.0347
0.040	0.0401	0.0458	0.0422	0.0657	0.0363	0.0469
0.050	0.0493	0.0587	0.0543	0.0776	0.0479	0.0602

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.498	0.143	0.171	0.426	0.322	0.066	0.148	0.388
0.020	0.604	0.201	0.238	0.501	0.444	0.118	0.190	0.449
0.030	0.685	0.258	0.298	0.560	0.528	0.157	0.224	0.486
0.040	0.731	0.294	0.338	0.595	0.594	0.194	0.251	0.513
0.050	0.771	0.339	0.366	0.627	0.646	0.232	0.278	0.538
	Supremum type test							
0.010	0.479	0.137	0.170	0.420	0.215	0.035	0.154	0.399
0.020	0.594	0.200	0.248	0.515	0.354	0.080	0.221	0.486
0.030	0.662	0.249	0.287	0.556	0.431	0.113	0.258	0.525
0.040	0.714	0.292	0.331	0.600	0.491	0.143	0.286	0.554
0.050	0.749	0.323	0.367	0.629	0.545	0.170	0.310	0.580

c

Table 6.31: Tests based on stochastic integral type 13 (maximum likelihood estimation) and $n = 50$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.6127	-1.4708	-0.0769	1.4259	2.7542
st.dev. $Q_n(\infty; v^{(n)})$	0.7917	0.8383	1.0164	1.2443	1.5497
skewness $Q_n(\infty; v^{(n)})$	-0.0218	0.0448	0.1697	0.2699	0.2691
kurtosis $Q_n(\infty; v^{(n)})$	-0.0534	0.1373	0.1738	0.2143	0.0762

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.005	0.0050	0.0030	0.0037	0.0073	0.0021	0.0032
0.010	0.0107	0.0074	0.0076	0.0168	0.0055	0.0073
0.015	0.0147	0.0130	0.0129	0.0230	0.0097	0.0125
0.020	0.0192	0.0173	0.0190	0.0291	0.0133	0.0181
0.025	0.0229	0.0213	0.0233	0.0355	0.0185	0.0233

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.513	0.093	0.143	0.473	0.355	0.040	0.120	0.435
0.010	0.646	0.161	0.203	0.555	0.466	0.075	0.158	0.492
0.015	0.700	0.199	0.249	0.610	0.558	0.111	0.191	0.541
0.020	0.748	0.238	0.275	0.641	0.624	0.147	0.222	0.578
0.025	0.777	0.266	0.297	0.661	0.661	0.169	0.238	0.598
	Supremum type test							
0.005	0.486	0.084	0.138	0.462	0.258	0.023	0.124	0.442
0.010	0.631	0.157	0.195	0.546	0.372	0.048	0.169	0.511
0.015	0.686	0.196	0.238	0.598	0.459	0.075	0.205	0.558
0.020	0.725	0.230	0.268	0.628	0.521	0.099	0.232	0.592
0.025	0.759	0.259	0.301	0.657	0.566	0.119	0.255	0.615

c

Table 6.32: Tests based on stochastic integral type 13 (maximum likelihood estimation) and $n = 100$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.5843	-1.4571	-0.2413	0.9568	2.0603
st.dev. $Q_n(\infty; v^{(n)})$	0.8544	0.9362	1.0125	1.1183	1.2755
skewness $Q_n(\infty; v^{(n)})$	0.1440	0.0118	-0.0374	0.0973	0.2046
kurtosis $Q_n(\infty; v^{(n)})$	-0.0257	-0.0991	0.0559	0.0741	0.1778

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.020	0.0101	0.0313	0.0154	0.0187	0.0255	0.0211
0.040	0.0214	0.0636	0.0313	0.0403	0.0499	0.0423
0.060	0.0337	0.0934	0.0473	0.0609	0.0742	0.0671
0.080	0.0482	0.1215	0.0660	0.0846	0.0997	0.0887
0.100	0.0629	0.1477	0.0865	0.1092	0.1261	0.1096

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.618	0.178	0.209	0.558	0.575	0.151	0.095	0.379
0.040	0.744	0.274	0.303	0.658	0.694	0.232	0.145	0.463
0.060	0.812	0.349	0.369	0.717	0.760	0.289	0.185	0.520
0.080	0.862	0.414	0.419	0.757	0.810	0.346	0.221	0.566
0.100	0.893	0.470	0.464	0.786	0.849	0.395	0.257	0.606
	Supremum type test							
0.020	0.620	0.178	0.213	0.555	0.521	0.124	0.132	0.441
0.040	0.747	0.278	0.292	0.646	0.641	0.192	0.196	0.532
0.060	0.811	0.349	0.356	0.699	0.717	0.253	0.245	0.595
0.080	0.859	0.412	0.408	0.742	0.762	0.293	0.280	0.632
0.100	0.889	0.464	0.452	0.775	0.795	0.331	0.311	0.658

c

Table 6.33: Tests based on stochastic integral type 23 (maximum likelihood estimation) and $n = 25$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.8551	-1.5364	-0.1641	1.1944	2.5041
st.dev. $Q_n(\infty; v^{(n)})$	0.8911	0.9439	1.0071	1.0731	1.1909
skewness $Q_n(\infty; v^{(n)})$	0.0749	0.1252	-0.0094	-0.0902	-0.0735
kurtosis $Q_n(\infty; v^{(n)})$	-0.0674	-0.0068	-0.0208	-0.1120	0.0969

a

Size	*Generalized rank test*			*Supremum type test*		
of	Left-	Right-	Two-	Left-	Right-	Two-
test	Sided	Sided	Sided	Sided	Sided	Sided
0.010	0.0054	0.0156	0.0083	0.0094	0.0122	0.0105
0.020	0.0119	0.0303	0.0176	0.0202	0.0255	0.0213
0.030	0.0178	0.0443	0.0262	0.0312	0.0373	0.0320
0.040	0.0270	0.0563	0.0341	0.0431	0.0487	0.0436
0.050	0.0337	0.0672	0.0449	0.0568	0.0599	0.0562

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.635	0.145	0.187	0.620	0.599	0.124	0.090	0.457
0.020	0.749	0.221	0.268	0.704	0.710	0.189	0.137	0.549
0.030	0.801	0.274	0.322	0.754	0.762	0.233	0.171	0.597
0.040	0.849	0.340	0.361	0.783	0.794	0.268	0.196	0.630
0.050	0.874	0.377	0.392	0.803	0.829	0.311	0.229	0.668
	Supremum type test							
0.010	0.632	0.145	0.179	0.602	0.547	0.101	0.115	0.502
0.020	0.745	0.221	0.259	0.691	0.651	0.158	0.168	0.587
0.030	0.802	0.278	0.311	0.736	0.714	0.198	0.205	0.633
0.040	0.843	0.330	0.348	0.767	0.755	0.231	0.240	0.672
0.050	0.872	0.380	0.384	0.794	0.788	0.264	0.272	0.704

c

Table 6.34: Tests based on stochastic integral type 23 (maximum likelihood estimation) and $n = 50$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-3.0751	-1.6314	-0.0937	1.3753	2.8332
st.dev. $Q_n(\infty; v^{(n)})$	0.9156	0.9623	1.0027	1.0607	1.1205
skewness $Q_n(\infty; v^{(n)})$	-0.0646	0.0458	0.0515	-0.0831	-0.1074
kurtosis $Q_n(\infty; v^{(n)})$	-0.0999	-0.0676	-0.0034	-0.0067	0.0394

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.005	0.0045	0.0057	0.0053	0.0068	0.0051	0.0058
0.010	0.0085	0.0119	0.0104	0.0125	0.0097	0.0114
0.015	0.0129	0.0169	0.0146	0.0183	0.0159	0.0168
0.020	0.0164	0.0234	0.0196	0.0238	0.0199	0.0221
0.025	0.0202	0.0291	0.0247	0.0304	0.0254	0.0277

b

Size of test	*One-sided test* H_{LL}	H_L	H_R	H_{RR}	*Two-sided test* H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.692	0.155	0.139	0.612	0.618	0.114	0.093	0.519
0.010	0.774	0.219	0.203	0.700	0.712	0.168	0.133	0.602
0.015	0.823	0.266	0.241	0.742	0.756	0.203	0.157	0.642
0.020	0.848	0.300	0.283	0.776	0.792	0.235	0.183	0.676
0.025	0.867	0.329	0.314	0.799	0.818	0.262	0.207	0.705
	Supremum type test							
0.005	0.678	0.148	0.139	0.605	0.557	0.092	0.104	0.540
0.010	0.758	0.210	0.192	0.675	0.651	0.134	0.147	0.620
0.015	0.807	0.255	0.242	0.732	0.708	0.169	0.178	0.660
0.020	0.834	0.290	0.268	0.755	0.743	0.197	0.203	0.691
0.025	0.861	0.324	0.302	0.783	0.771	0.221	0.226	0.717

c

Table 6.35: Tests based on stochastic integral type 23 (maximum likelihood estimation) and $n = 100$. No censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-2.8138	-1.5424	-0.2213	0.7090	1.1824
st.dev. $Q_n(\infty; v^{(n)})$	0.9407	1.0018	1.0388	1.0960	1.2419
skewness $Q_n(\infty; v^{(n)})$	0.2262	0.0034	-0.0759	-0.0518	0.1543
kurtosis $Q_n(\infty; v^{(n)})$	-0.0464	-0.1069	-0.0652	0.0672	0.3960

a

Size of test	*Generalized rank test* Left-Sided	Right-Sided	Two-Sided	*Supremum type test* Left-Sided	Right-Sided	Two-Sided
0.020	0.0076	0.0330	0.0128	0.0134	0.0358	0.0195
0.040	0.0178	0.0583	0.0267	0.0297	0.0658	0.0417
0.060	0.0306	0.0865	0.0443	0.0502	0.0913	0.0637
0.080	0.0429	0.1139	0.0627	0.0730	0.1174	0.0901
0.100	0.0577	0.1374	0.0791	0.0947	0.1435	0.1150

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.020	0.663	0.191	0.153	0.288	0.642	0.173	0.051	0.140
0.040	0.773	0.294	0.216	0.374	0.742	0.256	0.088	0.195
0.060	0.836	0.373	0.275	0.438	0.802	0.325	0.123	0.246
0.080	0.873	0.434	0.327	0.491	0.838	0.378	0.160	0.288
0.100	0.899	0.492	0.367	0.531	0.864	0.419	0.185	0.324
	Supremum type test							
0.020	0.686	0.197	0.152	0.291	0.643	0.167	0.063	0.162
0.040	0.791	0.298	0.228	0.381	0.747	0.250	0.109	0.235
0.060	0.848	0.380	0.281	0.442	0.799	0.310	0.148	0.282
0.080	0.888	0.451	0.326	0.488	0.836	0.363	0.186	0.326
0.100	0.910	0.503	0.367	0.532	0.864	0.405	0.220	0.366

c

Table 6.36: Tests based on stochastic integral type 23 (maximum likelihood estimation) and $n = 25$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-3.1650	-1.6180	-0.1682	1.0020	1.7408
st.dev. $Q_n(\infty; v^{(n)})$	0.9699	1.0124	1.0113	1.0526	1.1820
skewness $Q_n(\infty; v^{(n)})$	0.1576	0.0563	-0.0837	-0.2078	0.1002
kurtosis $Q_n(\infty; v^{(n)})$	-0.0450	-0.0851	-0.0480	0.0441	0.0562

a

Size	*Generalized rank test*			*Supremum type test*		
of	Left-	Right-	Two-	Left-	Right-	Two-
test	Sided	Sided	Sided	Sided	Sided	Sided
0.010	0.0056	0.0171	0.0092	0.0094	0.0179	0.0125
0.020	0.0107	0.0308	0.0172	0.0180	0.0321	0.0256
0.030	0.0183	0.0442	0.0255	0.0278	0.0465	0.0359
0.040	0.0245	0.0605	0.0356	0.0398	0.0584	0.0482
0.050	0.0311	0.0733	0.0449	0.0496	0.0720	0.0589

b

Size	*One-sided test*				*Two-sided test*			
of test	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.010	0.745	0.187	0.143	0.374	0.720	0.168	0.060	0.227
0.020	0.816	0.252	0.208	0.457	0.793	0.229	0.091	0.289
0.030	0.864	0.324	0.259	0.515	0.833	0.274	0.119	0.336
0.040	0.888	0.368	0.310	0.565	0.862	0.320	0.149	0.380
0.050	0.908	0.407	0.340	0.597	0.881	0.354	0.173	0.410
	Supremum type test							
0.010	0.760	0.191	0.142	0.371	0.709	0.155	0.069	0.251
0.020	0.828	0.262	0.205	0.457	0.794	0.223	0.113	0.328
0.030	0.870	0.325	0.257	0.514	0.828	0.262	0.144	0.372
0.040	0.897	0.379	0.291	0.552	0.856	0.302	0.175	0.415
0.050	0.913	0.416	0.328	0.588	0.874	0.333	0.197	0.444

c

Table 6.37: Tests based on stochastic integral type 23 (maximum likelihood estimation) and $n = 50$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

	H_{LL}	H_L	H_0	H_R	H_{RR}
mean $Q_n(\infty; v^{(n)})$	-3.4306	-1.7217	-0.1192	1.2254	2.2003
st.dev. $Q_n(\infty; v^{(n)})$	0.9883	0.9882	1.0271	1.0378	1.1022
skewness $Q_n(\infty; v^{(n)})$	0.0873	0.0091	-0.0652	-0.1170	-0.1709
kurtosis $Q_n(\infty; v^{(n)})$	-0.0689	0.0228	-0.0293	0.0102	0.0174

a

Size of test	*Generalized rank test*			*Supremum type test*		
	Left-Sided	Right-Sided	Two-Sided	Left-Sided	Right-Sided	Two-Sided
0.005	0.0030	0.0070	0.0037	0.0041	0.0064	0.0056
0.010	0.0061	0.0129	0.0089	0.0096	0.0124	0.0111
0.015	0.0087	0.0195	0.0127	0.0142	0.0168	0.0170
0.020	0.0120	0.0258	0.0162	0.0188	0.0225	0.0212
0.025	0.0159	0.0312	0.0202	0.0238	0.0286	0.0266

b

Size of test	*One-sided test*				*Two-sided test*			
	H_{LL}	H_L	H_R	H_{RR}	H_{LL}	H_L	H_R	H_{RR}
	Generalized rank test							
0.005	0.754	0.149	0.117	0.411	0.697	0.114	0.050	0.263
0.010	0.824	0.215	0.164	0.493	0.794	0.184	0.089	0.357
0.015	0.856	0.256	0.211	0.551	0.829	0.220	0.111	0.399
0.020	0.882	0.295	0.246	0.592	0.850	0.248	0.128	0.431
0.025	0.902	0.336	0.274	0.626	0.870	0.274	0.145	0.460
	Supremum type test							
0.005	0.746	0.140	0.108	0.388	0.701	0.113	0.064	0.293
0.010	0.833	0.222	0.158	0.476	0.778	0.164	0.100	0.372
0.015	0.866	0.268	0.190	0.519	0.820	0.208	0.127	0.425
0.020	0.889	0.305	0.220	0.559	0.842	0.233	0.143	0.456
0.025	0.906	0.338	0.254	0.597	0.861	0.260	0.164	0.486

c

Table 6.38: Tests based on stochastic integral type 23 (maximum likelihood estimation) and $n = 100$. Standard exponential censoring.
a: description stochastic integral;
b: critical points asymptotic P-values;
c: simulated power generalized rank and supremum type tests.

6.4 Conclusion

Replacing $\mu_\beta(v_0)$ by $\hat{\mu}_\beta(v_0)$ in the weight process shifts the sensitivity. The direction of the shift depends on whether we are dealing with the simple [increased sensitivity to the right] or composite [increased sensitivity to the left] null hypothesis.

An exception should be made for one-sided tests. The indifference spotted here seems to suggest that the replacement has about the same effect as a monotone transformation of the test statistic.

If censoring occurs, then left-sided tests should be preferably be based on stochastic integral 12 [simple null hypothesis] or 23 [composite null hypothesis], and right-sided tests on stochastic integral 22 [simple null hypothesis] or 13 [composite null hypothesis]. No preference can be given for one-sided tests in the absence of censoring. The same is true for two-sided tests, since higher sensitivity to the left or right is no criterium for two-sided tests, where the user is supposed to be ignorant of direction.

Appendix A

On Wieand's theorem

This appendix contains the first three sections of Kallenberg and Koning (1993), which materialized in an effort to generalize Theorem 11, stated in section 4.3 on page 65.

Wieand's theorem on equivalence of limiting approximate Bahadur efficiency and limiting Pitman efficiency is extended in several ways. Conditions on monotonicity and continuity are obviated, composite null hypotheses are incorporated, and the implications of a weaker form of Wieand's Condition III* are investigated.

A.1 Introduction

To compare the performance of two sequences of test statistics, many efficiency concepts have been proposed. Probably the most widely used is asymptotic relative Pitman efficiency.

In Bahadur (1960) the concept of exact Bahadur efficiency was proposed. This concept requires large deviation results, the derivation of which often becomes the stumbling-block in the application. As a "quick and dirty" variant Bahadur simultaneously proposed approximate Bahadur efficiency, valid for comparison of so-called standard sequences. Since it is rather easy to verify whether a sequence of test statistics actually is standard, approximate Bahadur efficiency in spite of its apparent shortcomings [see Bahadur (1960)] has become quite popular.

In favour of approximate Bahadur efficiency, Bahadur argued that for many well-known test statistics the limiting [as the alternative approaches the null hypothesis] approximate Bahadur efficiency is equal to the asymptotic relative Pitman efficiency. Working with an extended version of asymptotic relative Pitman efficiency, Wieand (1976) elaborated this point by presenting a condition under which the limiting approximate Bahadur efficiency, coincided with the limiting [as the size of the test tends to zero] asymptotic relative Pitman efficiency.

Unfortunately, Wieand only proved the theorem which stated this coincidence for the simple null hypothesis. Moreover, he requires continuity and strict mono-

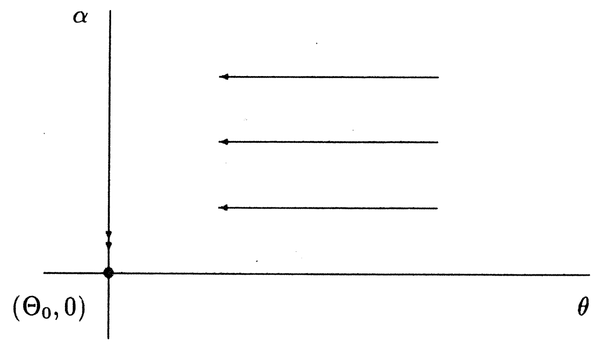

Figure A.1: Limiting Pitman efficiency.

tonicity of the tail of the asymptotic null distribution of the sequence of test statistics. In this note we extend Wieand's theorem to composite null hypotheses, while discarding continuity and strict monotonicity conditions.

Evaluating limiting Pitman efficiency means that the alternative is sent to the null hypothesis and afterwards the size α to zero. This is outlined in Figure A.1.

The order of the operations is reversed for the limiting approximate Bahadur efficiency, as is sketched in Figure A.2.

Under Wieand's Condition III* it holds for standard sequences of test statistics that both ways of approaching $(\Theta_0, 0)$ yield the same result. In fact it can be shown that for *any* way in which (θ, α) tend to $(\Theta_0, 0)$ the result is the same, provided that Wieand's Condition III* holds.

Often, the knowledge of the behavior of the test statistics is available in great

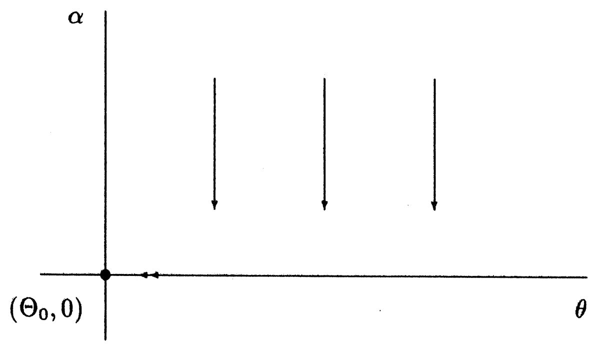

Figure A.2: Limiting Bahadur efficiency.

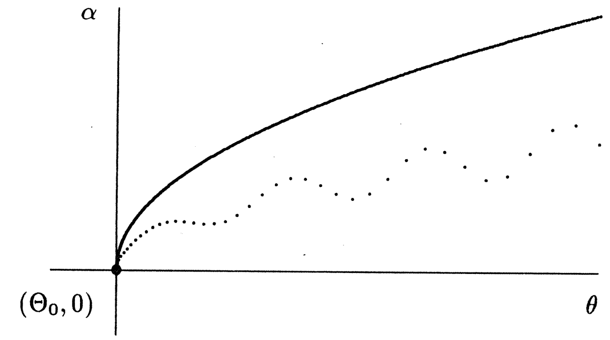

Figure A.3: The restricted area of admitted trajectories.

detail under the null hypothesis, but scarse under the alternative hypothesis. This may lead to problems with the verification of Wieand's Condition III*, since it deals explicitly with the behavior under the alternative hypothesis. However, if a weaker form of the condition holds [cf. Definition 13 in section A.2] it still can be shown that we come close to the answer of the Bahadur approach for any trajectory in a restricted region, for instance trajectories as sketched in Figure A.3.

The closer we want to be to the answer of the Bahadur approach, the smaller the area of admitted trajectories. This is reflected by a lowering of the curve depicted in Figure A.3, and implies that we should restrict ourselves to tests of even smaller size.

Thus, the quality of the approximation of the finite sample relative efficiency by the approximate Bahadur efficiency for alternatives close to the null hypothesis is only guaranteed in case of (very) small levels.

The organization is as follows. In section A.2 notation is introduced and definitions are given, while section A.3 contains the results. Proofs can be found in Kallenberg and Koning (1993).

A.2 Preliminaries

Consider the situation in which we have two infinite sequences of test statistics $\{T_{1n}\}_{n=1}^{\infty}$ and $\{T_{2n}\}_{n=1}^{\infty}$, rejecting the null hypothesis $H_0 : \theta \in \Theta_0$ for large values of the test statistics. The number n here refers to the number of available observations. The whole parameter space is denoted by Θ, which is a metric space with metric d. For each $\alpha \in (0,1)$ the critical value is denoted by $t_{in}(\alpha)$. So, H_0 is rejected if $T_{in} > t_{in}(\alpha)$, accepted if $T_{in} < t_{in}(\alpha)$, and

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_{in} > t_{in}(\alpha)) \leq \alpha \leq \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_{in} \geq t_{in}(\alpha)). \qquad \text{(A.1)}$$

Definition 11 *For each $\alpha, \beta \in (0,1)$ and $\theta \in \Theta - \Theta_0$, let $N_i(\alpha, \beta, \theta)$ be the largest sample size such that the power at θ of the size α test based on $\{T_{in}\}_{n=1}^{\infty}$ is less than β.*

The finite sample relative efficiency of $\{T_{1n}\}_{n=1}^{\infty}$ with respect to $\{T_{2n}\}_{n=1}^{\infty}$ is denoted by $N_2(\alpha,\beta,\theta)/N_1(\alpha,\beta,\theta)$. A value larger than 1 indicates that $\{T_{1n}\}_{n=1}^{\infty}$ is preferred to $\{T_{2n}\}_{n=1}^{\infty}$, since $\{T_{1n}\}_{n=1}^{\infty}$ needs less observations for the same performance.

Definition 12 *$\{T_{in}\}_{n=1}^{\infty}$ is said to be a standard sequence if the following three conditions are satisfied.*

a *There exists a nondecreasing function G_i such that $\inf_{\theta_0 \in \Theta_0} P_{\theta_0}(T_{in} \le t)$ tends to $G_i(t)$ for all continuity points of G_i.*

b *There exists a constant $a_i > 0$ such that*

$$\lim_{t \to \infty} t^{-2} \log(1 - G_i(t)) = -a_i/2. \tag{A.2}$$

c *There exists a positive function $b_i(\theta)$ such that $|n^{-1/2}T_{in} - b_i(\theta)|$ converges to zero in P_θ-probability for every $\theta \in \Theta - \Theta_0$.*

The approximate Bahadur slope of a standard sequence $\{T_{in}\}_{n=1}^{\infty}$ is defined as $a_i(b_i(\theta))^2$. The approximate Bahadur efficiency of a standard sequence $\{T_{1n}\}_{n=1}^{\infty}$ with respect to another standard sequence $\{T_{2n}\}_{n=1}^{\infty}$ is defined as the ratio of their respective Bahadur slopes $a_1(b_1(\theta))^2/[a_2(b_2(\theta))^2]$.

This is a slightly weaker form of the definition of standard sequences given in Bahadur (1960), not requiring continuity of G_i.

For a standard sequence $\{T_{in}\}_{n=1}^{\infty}$ define for each $0 < \alpha < 1$

$$\underline{q}_i(\alpha) = \inf\{t : 1 - G_i(t) \le \alpha\}, \qquad \overline{q}_i(\alpha) = \sup\{t : 1 - G_i(t) \ge \alpha\}. \tag{A.3}$$

By this definition and the monotonicity of G_i we obtain

$$\begin{aligned} 1 - G_i(t) &> \alpha \quad \text{for all } t < \underline{q}_i(\alpha), & 1 - G_i(t) &\ge \alpha \quad \text{for all } t < \overline{q}_i(\alpha), \\ 1 - G_i(t) &\le \alpha \quad \text{for all } t > \underline{q}_i(\alpha), & 1 - G_i(t) &> \alpha \quad \text{for all } t > \overline{q}_i(\alpha). \end{aligned}$$

It is easily seen that

$$\underline{q}_i(\alpha) \le \overline{q}_i(\alpha). \tag{A.4}$$

In view of Definition 12a, denote for each $\epsilon > 0$ and $0 < \alpha < 1$ by $\overline{n}_i = \overline{n}_i(\alpha, \epsilon)$ the smallest number such that for all $n \ge \overline{n}_i(\alpha, \epsilon)$

$$1 - \inf_{\theta_0 \in \Theta_0} P_{\theta_0}(T_{in} \le \overline{q}_i(\alpha)(1+\epsilon)^{1/2}) < \alpha. \tag{A.5}$$

Similarly denote for each $\epsilon > 0$ and $0 < \alpha < 1$ by $\underline{n}_i = \underline{n}_i(\alpha, \epsilon)$ the smallest number such that for all $n \geq \underline{n}_i(\alpha, \epsilon)$

$$1 - \inf_{\theta_0 \in \Theta_0} P_{\theta_0}(T_{in} < \underline{q}_i(\alpha)(1-\epsilon)^{1/2}) > \alpha. \tag{A.6}$$

Define $n_i(\alpha, \epsilon)$ by

$$n_i(\alpha, \epsilon) = \max(\overline{n}_i(\alpha, \epsilon), \underline{n}_i(\alpha, \epsilon)). \tag{A.7}$$

Finally we present the weaker form of Wieand's Condition III*.

Definition 13 *A standard sequence $\{T_{in}\}_{n=1}^{\infty}$ is said to be a $\{s_n\}_{n=1}^{\infty}$-Wieand sequence if there is a constant $\epsilon_i^* > 0$ such that for every $\epsilon > 0$ and $\delta \in (0,1)$ there exists a constant $C_i(\epsilon, \delta)$ such that*

$$P_\theta(|n^{-1/2}T_{in} - b_i(\theta)| \geq \epsilon b_i(\theta)) < \delta \tag{A.8}$$

for every $\theta \in \Theta - \Theta_0$ satisfying $\inf_{\theta_0 \in \Theta_0} d(\theta, \theta_0) < \epsilon_i^*$ and $n^{1/2} b_i(\theta) > C_i(\epsilon, \delta) s_n$.

It can be shown that a standard sequence $\{T_{in}\}_{n=1}^{\infty}$ is a $\{s_n\}_{n=1}^{\infty}$-Wieand sequence if there exists a nondefective cumulative distribution function Q such that

$$P_\theta(|T_{in} - n^{1/2} b_i(\theta)| > x s_n) \leq 1 - Q(x) \tag{A.9}$$

for every $x > 0$ and $\theta \in \Theta - \Theta_0$ satisfying $\inf_{\theta_0 \in \Theta_0} d(\theta, \theta_0) < \epsilon_i^*$ [choose $C_i(\epsilon, \delta)$ so as to satisfy $Q(\epsilon C_i(\epsilon, \delta)) > 1 - \delta$]. Remark that if s_n remains bounded, Definition 13 reduces to Wieand's Condition III* [in case of a simple null hypothesis], and the observation above is a consequence of the lemma given in section 4 in Wieand (1976). In the sequel we shall use the phrase "under Wieand's Condition III*" to indicate that s_n remains bounded, thereby tacitly extending Wieand's original definition to general null hypotheses.

We close this section by introducing a set which plays a role in section A.3. For each $\alpha, \epsilon, \delta \in (0,1)$ let $A(\alpha, \epsilon, \delta)$ be the set of points $\theta \in \Theta - \Theta_0$ for which

$$n \geq \frac{-\log \alpha}{a_i\{b_i(\theta)\}^2} \quad \text{implies} \quad n^{1/2} b_i(\theta) > C_i(\epsilon, \delta) s_n \tag{A.10}$$

for $i = 1, 2$. Lemma 5 sheds light on the relation between the set $A(\alpha, \epsilon, \delta)$ and Wieand's Condition III*. Observe that in case we have

$$\inf_{\theta \in \Theta - \Theta_0} b_i(\theta) = 0,$$

this lemma implies that $A(\alpha, \epsilon, \delta)$ equals $\Theta - \Theta_0$ for α sufficiently small if and only if s_n is bounded.

Lemma 5 *If s_n remains bounded then for every $\epsilon, \delta > 0$ there exists $\alpha(\epsilon, \delta) > 0$ such that $A(\alpha, \epsilon, \delta)$ coincides with $\Theta - \Theta_0$ for every $0 < \alpha < \alpha(\epsilon, \delta)$. If $\inf_{\theta \in A(\alpha, \epsilon, \delta)} b_i(\theta) = 0$ for some $\alpha, \epsilon, \delta > 0$, then s_n remains bounded.*

A.3 Results

Our main result is the following theorem, which shows where the finite sample relative efficiency $N_2(\alpha,\beta,\theta)/N_1(\alpha,\beta,\theta)$ may be approximated by the approximate Bahadur efficiency $a_1\{b_1(\theta)\}^2/[a_2\{b_2(\theta)\}^2]$. (If limits are taken with θ tending to some subset of Θ_0, then of course θ runs through $\Theta-\Theta_0$; further the notation is explained in section A.2.)

Theorem 14 *Let $\{T_{1n}\}_{n=1}^{\infty}$ and $\{T_{2n}\}_{n=1}^{\infty}$ be two $\{s_n\}_{n=1}^{\infty}$-Wieand sequences, and let Θ_0^* be a subset of Θ_0. Suppose that the limit*

$$\lim_{\theta\to\Theta_0^*} a_1\{b_1(\theta)\}^2/[a_2\{b_2(\theta)\}^2] =: e(\Theta_0^*) \tag{A.11}$$

exists. For each $\alpha,\epsilon,\delta\in(0,1)$, let $A^*(\alpha,\epsilon,\delta)$ be a subset of $A(\alpha,\epsilon,\delta)$ satisfying

$$\lim_{\epsilon\to 0}\limsup_{\alpha\to 0}\sup_{\theta\in A^*(\alpha,\epsilon,\delta)}\inf_{\theta_0\in\Theta_0^*} d(\theta,\theta_0)=0. \tag{A.12}$$

Assume that

$$\limsup_{\epsilon\to 0}\limsup_{\alpha\to 0}\sup_{\theta\in A^*(\alpha,\epsilon,\delta)}\frac{n_i(\alpha,\epsilon)}{\left[\frac{-2\log\alpha}{a_i\{b_i(\theta)\}^2}\right]}<1. \tag{A.13}$$

Then for $\beta\in[\delta,1-\delta]$ we have

$$\lim_{\epsilon\to 0}\liminf_{\alpha\to 0}\inf_{\theta\in A^*(\alpha,\epsilon,\delta)}\frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)}=\lim_{\epsilon\to 0}\limsup_{\alpha\to 0}\sup_{\theta\in A^*(\alpha,\epsilon,\delta)}\frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)}=e(\Theta_0^*). \tag{A.14}$$

When considering a specified trajectory $\theta(\alpha,\epsilon,\delta)$, along which Θ_0^* is approached as $\alpha\to 0$, it seems obvious to choose $A^*(\alpha,\epsilon,\delta)$ equal to the singleton containing $\theta(\alpha,\epsilon,\delta)$.

If Wieand's Condition III* does not hold, the requirement that $\theta(\alpha,\epsilon,\delta)$ should be an element of $A(\alpha,\epsilon,\delta)$ limits the trajectories admitted for a fixed $\epsilon>0$. The lower ϵ, the smaller the area through which the trajectories are allowed to run. As already observed in section A.2, this is reflected in Figure A.3 by a lowering of the curve. For instance, if $n^{1/2}/s_n$ is a monotone increasing function g of n, condition (A.10) holds if and only if

$$g\left(\frac{-\log\alpha}{a_i\{b_i(\theta)\}^2}\right)b_i(\theta)>C_i(\epsilon,\delta)$$

for $i=1,2$.

If Wieand's Condition III* does hold, the set $A(\alpha,\epsilon,\delta)$ equals $\Theta-\Theta_0$ if $0<\alpha<\alpha(\epsilon,\delta)$, where $\alpha(\epsilon,\delta)$ is given by Lemma 5. It now easily follows that for every sequence $\{(\theta_j,\alpha_j)\}_{j=1}^{\infty}$ tending to $(\Theta_0^*,0)$ we obtain

$$\frac{N_2(\alpha_j,\beta,\theta_j)}{N_1(\alpha_j,\beta,\theta_j)}\to e(\Theta_0^*), \tag{A.15}$$

provided that

$$\limsup_{\epsilon\to 0} \lim_{j\to\infty} \frac{n_i(\alpha_j,\epsilon)}{\left[\frac{-2\log\alpha_j}{a_i\{b_i(\theta_j)\}^2}\right]} < 1,$$

as is formally seen by taking $A^*(\alpha,\epsilon,\delta)$ equal to $\{\theta_j\}$ if $\alpha_j \leq \alpha < \alpha_{j-1}$, $j = 1, 2, \ldots$ $[\alpha_0 = 1]$. This is closely related to intermediate efficiency [cf. Kallenberg (1983)].

Another consequence of Lemma 5 is that under Wieand's Condition III* we may take $A^*(\alpha,\epsilon,\delta)$ to be the intersection of $\Theta - \Theta_0$ with an arbitrary environment of Θ_0^*, shrinking to Θ_0^* as $\alpha \to 0$, and [afterwards] $\epsilon \to 0$. More precisely, let

$$A^*(\alpha,\epsilon,\delta) = \{\theta \in A(\alpha,\epsilon,\delta) \ : \ \inf_{\theta_0\in\Theta_0^*} d(\theta,\theta_0) < \epsilon, \ \frac{n_i(\alpha,\epsilon)}{1+\epsilon} < \frac{-2\log\alpha}{a_i\{b_i(\theta)\}^2}\}.$$

Then (A.12) and (A.13) are satisfied. Now consider a sequence $\{\theta_j\}_{j=1}^{\infty}$ which tends to Θ_0^* as $j \to \infty$. Assume that

$$\lim_{j\to\infty} b_i(\theta_j) = 0.$$

Let $\alpha, \epsilon, \delta > 0$ satisfy $\alpha < \alpha(\epsilon,\delta)$, with $\alpha(\epsilon,\delta)$ as in Lemma 5. Then there exists $j(\alpha,\epsilon,\delta)$ such that for all $j \geq j(\alpha,\epsilon,\delta)$ we have $\theta_j \in A^*(\alpha,\epsilon,\delta)$, and hence it follows that

$$\begin{aligned} \inf_{\theta\in A^*(\alpha,\epsilon,\delta)} \frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)} &\leq \liminf_{j\to\infty} \frac{N_2(\alpha,\beta,\theta_j)}{N_1(\alpha,\beta,\theta_j)} \\ &\leq \limsup_{j\to\infty} \frac{N_2(\alpha,\beta,\theta_j)}{N_1(\alpha,\beta,\theta_j)} \leq \sup_{\theta\in A^*(\alpha,\epsilon,\delta)} \frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)}. \end{aligned} \tag{A.16}$$

Applying Theorem 14 we get

$$\begin{aligned} e(\Theta_0^*) &= \lim_{\epsilon\to 0} \liminf_{\alpha\to 0} \inf_{\theta\in A^*(\alpha,\epsilon,\delta)} \frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)} \\ &\leq \liminf_{\alpha\to 0} \liminf_{j\to\infty} \frac{N_2(\alpha,\beta,\theta_j)}{N_1(\alpha,\beta,\theta_j)} \leq \limsup_{\alpha\to 0} \limsup_{j\to\infty} \frac{N_2(\alpha,\beta,\theta_j)}{N_1(\alpha,\beta,\theta_j)} \\ &\leq \lim_{\epsilon\to 0} \limsup_{\alpha\to 0} \sup_{\theta\in A^*(\alpha,\epsilon,\delta)} \frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)} = e(\Theta_0^*) \end{aligned}$$

and therefore

$$\lim_{\alpha\to 0} \lim_{j\to\infty} \frac{N_2(\alpha,\beta,\theta_j)}{N_1(\alpha,\beta,\theta_j)} = e(\Theta_0^*).$$

Since the sequence $\{\theta_j\}_{j=1}^{\infty}$ was arbitrarily chosen, it follows that

$$\lim_{\alpha\to 0} \lim_{\theta\to\Theta_0^*} \frac{N_2(\alpha,\beta,\theta)}{N_1(\alpha,\beta,\theta)} = e(\Theta_0^*).$$

This extends Wieand's theorem to testing problems with composite null hypotheses. Moreover, note that Wieand's conditions on continuity and monotonicity are not needed to obtain this result.

There exist alternative definitions of limiting Pitman efficiency which involve $\underline{N}_i(\alpha, \beta, \theta)$, the smallest sample size such that the power at θ of the size α test based on $\{T_{in}\}_{n=1}^{\infty}$ is greater than or equal to β. Observe that there is a disproportionate effect of the behavior of T_{i1} on $\underline{N}_i(\alpha, \beta, \theta)$. Definition 13 only has meaning for sample sizes which are sufficiently large, and is therefore not capable of providing lower bounds for $\underline{N}_i(\alpha, \beta, \theta)$. It follows that the fact that $\{T_{in}\}_{n=1}^{\infty}$ is a $\{s_n\}_{n=1}^{\infty}$-Wieand sequence does not shed light on asymptotic relative Pitman efficiency as defined in for instance Rothe (1981).

In absence of Wieand's Condition III* the verification of (A.13) requires some effort. In case the test statistics are functionals of either the empirical process or the partial sum process, the approximation theorems available for these processes [see Komlós, Major and Tusnády (1975)] may lead to probability inequalities of the type considered in the following lemma.

Lemma 6 *Suppose that $\{T_{in}\}_{n=1}^{\infty}$ is a standard sequence for which there exist a random variable T_i with distribution function G_i, positive constants τ, γ, λ and a sequence $k_n = o(n^{\gamma/\tau})$ such that*

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(|T_{in} - T_i| > n^{-\gamma}(k_n + x)^{\tau}) \leq e^{-\lambda x}. \tag{A.17}$$

Then for every $\epsilon > 0$ there exist a constant $c(\epsilon)$ such that (A.13) holds for

$$A^*(\alpha, \epsilon, \delta) = \{\theta \in A(\alpha, \epsilon, \delta) \ : \ a_i\{b_i(\theta)\}^2 < |\log \alpha|^{1-\tau/\gamma}/c(\epsilon)\}.$$

Typically, one finds $\gamma \leq 1/2$, $\tau \geq 1$ and k_n proportional to $\log n$ [see Chapters 2–4, Koning (1993), and Csörgő and Horváth (1993)].

The magnitude of k_n allowed in Lemma 6 strongly suggests that it is essentially the tail rather than the center of the distribution of $n^{\gamma}|T_{in} - T_i|$ which determines the shape of $A^*(\alpha, \epsilon, \delta)$.

Bibliography

Aalen, O.O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. Statist.* **3**, 15–27.

Adler, R.J. and Samorodnitsky, G. (1987). Tail behaviour for the suprema of Gaussian processes with applications to empirical processes. *Ann. Probab.* **4**, 1339–1351.

Aki, S. (1986). Some test statistics based on the martingale term of the empirical distribution function. *Ann. Inst. Statist. Math.* **38A**, 1–21.

Aki, S. and Kashiwagi, N. (1989). Asymptotic properties of some goodness-of-fit tests based on the L_1 norm. *Ann. Inst. Statist. Math.* **41**, 753–764.

Akritas, M.G. (1988). Pearson-type goodness-of-fit tests: the univariate case. *J. Amer. Statist. Assoc.* **83**, 222–230.

Aronszajn, N. (1950). The theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68**, 337–404.

Bahadur, R.R. (1960). Stochastic comparisons of tests. *Ann. Math. Statist.* **31**, 276–295.

Barlow, R.E. and Proschan, F. (1969). A note on tests for monotone failure rate based on incomplete data. *Ann. Math. Statist.* **40**, 595–600.

Barlow, R.E. and Campo, R. (1975). Total time on test processes and applications to failure data analysis. *Reliability and Fault Tree Analysis*, 451–481.

Birnbaum, Z.W. and Tingey, F.H. (1951). One sided confidence contours for probability distribution functions. *Ann. Math. Statist.* **22**, 592–596.

Birnbaum, Z.W. and McCarty, R.C. (1958). A distribution-free upper confidence bound for $P(Y < X)$ based on independent samples of X and Y. *Ann. Math. Statist.* **29**, 558–562.

Borell, C. (1975). The Brunn-Minkowsky inequality in Gauss space. *Invent. Math.* **30**, 207–216.

Borgan, O. (1984). Maximum likelihood estimation in counting process models, with applications to censored failure time data. *Scand. J. Statist.* **11**, 1–16.

Burke, M.D., Csörgő, S., Horváth, L. (1981). Strong approximations of some biometric estimates under random censorship. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **56**, 87–112.

Breslow, N.E. (1975). Analysis of survival data under proportional hazards model. *Int. Stat. Rev.* **43**, 45–58.

Bretagnolle, J., and Massart, P. (1989). Hungarian constructions from the nonasymptotic viewpoint. *Ann. Probab.* **17**, 239–256.

Castelle, N. (1991). *Principes d'Invariance et Application à la Statistique de Modèles Censurés.* Ph. D. Thesis, Université Paris-Sud.

Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics.* Wiley, New York.

Devroye, L.P. and Wise, G.L. (1979). On the recovery of discrete probability densities from imperfect measurements. *J. Franklin Inst.* **307**, 1–20.

Donsker, M.D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **23**, 277–281.

Doob, J.L. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **20**, 393–403.

Dudley, R.M. (1967). The sizes of compact sets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal. 1*, 290–330.

Durbin, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function.* SIAM, Philadelphia.

Durbin, J. (1985). The first-passage density of a continuous Gaussian process to a general boundary. *J. Appl. Prob.* **22**, 99–122.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27**, 642–669.

Einmahl, J.H.J. and Koning, A.J. (1992). Limit theorems for a general weighted process under random censoring. *Canad. J. Statist.* **20**, 77–89.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications, Vol. 2.* Wiley, New York.

Fernique, X. (1970). Intégrabilité des vecteurs gaussiens. *C. R. Acad. Sci. Paris Ser. A* **270**, 1698–1699.

Fernique, X. (1971). Régularité des processus gaussiens. *Invent. Math.* **12**, 304–320.

Fleming, T.R. and Harrington, D.P. (1981). A class of hypothesis tests for one and two samples of censored data. *Comm. Stat. Theor. Meth.* **13**, 2469–2486.

Fleming, T.R., Harrington, D.P. and O'Sullivan, M. (1987). Supremum versions of the log-rank and generalized Wilcoxon statistics. *J. Amer. Statist. Assoc.* **82**, 312–320.

Fleming, T.R., O'Fallon, J.R., O'Brien, P.C. and Harrington, D.P. (1980). Modified Kolmogorov Smirnov test procedures with applications to arbitrarily right-censored data. *Biometrics* **36**, 607–625.

Gill, R.D. (1980). *Censoring and Stochastic Integrals.* Math. Centre Tracts 124. Mathematisch Centrum, Amsterdam.

Gill, R.D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* **11**,49–58.

Gill, R.D. (1986). The total time on test plot and the cumulative time on test statistic for a counting process. *Ann. Stat.* **14**, 1234–1239.

Gnedenko, B.V., Korolyuk, V.S. and Skorohod, A.V. (1961). Asymptotic expansions in probability theory. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **2**, 153–170. Univ. California Press, Berkeley, California.

Harter, H.L. (1980). Modified asymptotic formulas for critical values of the Kolmogorov test statistic. *Am. Statist.* **34**, 110–111.

Harrington, D.P. and Fleming, T.R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69**, 553–566.

Hjort, N.L. (1984). Weak convergence of cumulative intensity processes when parameters are estimated, with applications to goodness-of-fit tests in models with censoring. Report no. 764. Norwegian Computing Center.

Hjort, N.L. (1985). Comment on "Counting process models for life history data: a review" by P.K Andersen and O. Borgan. *Scand. J. Statist.* **12**, 141–150.

Hjort, N.L. (1990). Goodness of fit tests in models for life history dat based on cumulative hazard rates. *Ann. Statist.* **18**, 1221–1258.

Hu, I. (1985). A uniform bound for the tail probability of Kolmogorov-Smirnov statistics. *Ann. Statist.* **13**, 821–826.

Inglot, T. and Ledwina, T. (1990). On probabilities of excessive deviations for Kolmogorov-Smirnov, Cramér-von Mises and chi-square statistics. *Ann. Statist.* **18**, 1491–1495.

Inglot, T. and Ledwina, T. (1989). Large and moderate deviations for some functionals of weighted empirical process. Technical Report.

Kallenberg, W.C.M. (1983). Intermediate efficiency, theory and examples. *Ann. Statist.* **11**, 170–182.

Kallenberg, W.C.M., Koning, A.J. (1993). On Wieand's theorem. Research paper.

Khamaladze, E.V. (1981). Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26**, 246–265.

Khamaladze, E.V. (1982). Some applications of the theory of martingales to statistics. *Uspekhi Mat. Nauk* **37**:6, 193–212. Translation: *Russian Math. Surveys* **37**:6, 215–237.

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent R.V.'s and the sample F. I. *Z. Wahrsch. verw. Gebiete* **34**, 33–58.

Koning, A.J. (1991). *Stochastic Integrals and Goodness-of-fit tests.* Ph. D. Thesis, University of Twente.

Koning, A.J. (1992). Approximation of stochastic integrals with applications to goodness–of–fit tests. *Ann. Statist.* **20**, 428–454.

Koning, A.J. (1993). Approximation of the basic martingale. *Ann. Statist.*, accepted for publication.

Landau, H.J. and Shepp, L.A. (1971). On the supremum of a Gaussian process. *Sankhyā Ser. A* **32**, 369–378.

Lauwerier, H.A. (1963). The asymptotic expansion of the statistical distribution of N.V. Smirnov. *Z. Wahrsch. verw. Geb.* **2**, 61–68.

Marcus, M.B. and Shepp, L.A. (1971). Sample behaviour of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2**, 423–441. Univ. California Press, Berkeley, California.

Marcus, M.B. and Zinn, J. (1984). The bounded law of the iterated logarithm for the weighted empirical distribution process in the non-i.i.d. case. *Ann. Probab.* **12**, 335–360.

Mason, D.M. and van Zwet, W.R. (1987). A refinement of the KMT-inequality for the uniform empirical process. *Ann. Probab.* **15**, 871–884.

Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer Wolfowitz inequality. *Ann. Probab.* **18**, 1269–1283.

Penkov, B.I. (1976). Asymptotic distribution of Pyke's statistic. *Theor. Prob. Appl.* **21**, 370–374.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer Series in Statistics. Springer-Verlag, New York.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics* **5**, 375–383.

Rao, C.R. (1963). Criteria of estimation in large samples. *Sankhyā Ser. A* **25**, 189–206.

Rothe, G. (1981). Some properties of the asymptotic relative Pitman efficiency. *Ann. Statist.* **9**, 663–669.

Rubin, H. and Sethuraman, J. (1965). Bayes risk efficiency. *Sankhyā Ser. A* **27**, 347–356.

Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

Skorohod, A.V. (1956). Limit theorems for stochastic processes. *Theor. Prob. Appl.* **1**, 261–290.

Smirnov, N.V. (1944). Approximate laws of distribution of random variables from empirical data. *Usp. Mat. Nauk* **10**, 179–206 (in Russian).

Stephens, M.A. (1970). Use of the Kolmogorov-Smirnov, Crámer-von Mises and related statistics without extensive tables. *J. Roy. Statist. Soc. B* **32**, 115–122.

Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **3**, 211–226.

Talagrand, M. (1988). Small tails for the supremum of a Gaussian process. *Ann. Inst. Henri Poincaré* **24**, 307–315.

Wieand, H.S. (1976). A condition under which the Pitman and Bahadur approaches to efficiency coincide. *Ann. Statist.* **4** , 1003–1011.

Woolson, R.F. (1981). Rank tests and a one-sample log rank test for comparing observed survival data to a standard population. *Biometrics* **37**, 687–696.

Author Index

Subject Index

List of Symbols

a constant describing [asymptotic] tail behavior of $T(Q_n(\cdot)/q(\cdot)))$, 28

a_R constant describing tail behavior of $T_R(Q_n)$, 34

$a_R(v_0)$ constant describing tail behavior of $T_R(Q_n(\cdot; v_0))$, 66

a_S constant describing tail behavior of $T_S(Q_n)$, 34

$a_S(v_0)$ constant describing tail behavior of $T_S(Q_n(\cdot; v_0))$, 66

$\breve{a}(v_0)$ constant reflecting [asymptotic] tail behavior of $Q_n(t; v^{(n)})$, 60

$B_n^1(t)$ Gaussian process used to construct $W_n(t)$, 37

$B_n(t)$ Gaussian process used to construct $W_n(t)$, 37

$\tilde{B}_n(t)$ Brownian bridge approximating $\tilde{U}_n(t)$, 19

$D(t; \theta_0, \theta)$ function reflecting distance between $F(t; \theta)$ and $F(t; \theta_0)$, 24

$D(t; v', \theta_0, v, \theta)$ function reflecting distance between $\Lambda(t; v, \theta)$ and $\Lambda(s; v', \theta_0)$, 54

$D(t; v, \theta)$ abbreviation for $D(t; \pi(v, \theta), \theta_0, v, \theta)$, 54

$D[0, \infty)$ function space, 20

$e(h)$ limiting approximate Bahadur, limiting asymptotic relative Pitman and weak intermediate efficiency of test based on $T(Q_n(\cdot)/q(\cdot))$, 33

$e(h; v_0)$ limiting approximate Bahadur, limiting asymptotic relative Pitman and weak intermediate efficiency of $T(Q_n(\cdot; v^{(n)}))$, 65

$e_R(h)$ asymptotic relative Pitman, limiting approximate Bahadur and weak intermediate efficiency of test based on $T_R(Q_n)$, 34

$e_R(h, v_0)$ asymptotic relative Pitman, limiting approximate Bahadur and weak intermediate efficiency of $T_R(Q_n(\cdot; v_0))$, 66

$e_S(h)$ limiting approximate Bahadur, limiting asymptotic relative Pitman and weak intermediate efficiency of test based on $T_S(Q_n)$, 34

$e_S(h, v_0)$ limiting approximate Bahadur, limiting asymptotic relative Pitman, and weak intermediate efficiency of $T_S(Q_n(\cdot; v_0))$, 66

$e_{12}^{\tilde{\beta}}(h)$ asymptotic relative Pitman efficiency in the direction h, 31

$e_{12}^{\tilde{\beta}}(h; v_0)$ asymptotic relative Pitman efficiency, 64

$\mathcal{E}_0$ expectation under P_0, 23

$\mathcal{E}_n$ expectation under P_n, 23, 52

$\mathcal{E}_{v_0}$ expectation under P_{v_0}, 51

$F(t)$ cumulative distribution function of X_i, 2

$F(t; \theta_n)$ cumulative distribution function, simple null hypothesis, 5, 23

$F(t; v_n, \theta_n)$ cumulative distribution function of X_i, composite null hypothesis, 51

$F_n(t)$ empirical distribution function, 2

$G(t)$ cumulative distribution function of Y_i, 23, 51

$H(t;\theta_n)$ cumulative distribution function of Z_1 under P_n, simple null hypothesis, 24

$H(t;\upsilon_n,\theta_n)$ cumulative distribution function of Z_1 under P_n, 53

$H^0(t;\theta_n)$ cumulative distribution function of censored failure times, 36

$H^0(t;\upsilon_n,\theta_n)$ cumulative distribution function of censored failure times, 72

$H^1(t;\theta_n)$ cumulative distribution function of observed failure times under P_n, simple null hypothesis, 5, 24

$H^1(t;\upsilon_n,\theta_n)$ cumulative distribution function of observed failure times under P_n, composite null hypothesis, 53

$H_n^1(t)$ empirical distribution function of observed failure times, 4

$H_n(t)$ empirical distribution function of all observations, 4

$I(\theta_0)$ Fisher information matrix, simple null hypothesis, 35

$I(\upsilon_0,\theta_0)$ Fisher information matrix, composite null hypothesis, 67

I_{00} $r\times r$ submatrix of $I(\upsilon_0,\theta_0)$, 67

I_{0a} $r\times p$ submatrix of $I(\upsilon_0,\theta_0)$, 67

I_{a0} $p\times r$ submatrix of $I(\upsilon_0,\theta_0)$, 67

I_{aa} $p\times p$ submatrix of $I(\upsilon_0,\theta_0)$, 67

$J_n(t)$ jump process approximating $W_n(t)$, 38

$K_0(t;\upsilon,\theta)$ r-dimensional vector function describing effect of $\upsilon^{(n)}$ on behavior $Q_n(t;\upsilon^{(n)})$, 57

$K_a(t)$ p-dimensional vector function, 30

$K_{a|0}(t;\upsilon_0)$ p-dimensional vector function, 62

$\tilde{K}(t,n)$ Kiefer process approximating $\tilde{U}_n(t)$, 19

$L(t;\theta_n)$ deterministic function approximating $L_n(t)$ under P_n, 26

$L(t;\upsilon,\theta)$ limiting function of $L_n(t;\upsilon)$, 57

$L_i^{(1)}(t;\theta)$ first order partial derivative of $L(t;\theta)$ with respect to the i^{th} component of θ, 30

$L_i^{(1)}(t;\upsilon,\theta)$ first order partial derivative of $L(t;\upsilon,\theta)$ with respect to the i^{th} component of (υ,θ), 62

$L_{ni}^{(1)}(t;\upsilon)$ first order partial derivative of $L_n(t;\upsilon)$ with respect to the i^{th} component of υ, 57

$L_{ij}^{(2)}(t;\theta)$ second order partial derivative of $L(t;\theta)$ with respect to the i^{th} and j^{th} components of θ, 30

$L_{ij}^{(2)}(t;\upsilon,\theta)$ second order partial derivative of $L(t;\upsilon,\theta)$ with respect to the i^{th} and j^{th} components of (υ,θ), 62

$L_{nij}^{(2)}(t;\upsilon)$ second order partial derivative of $L_n(t;\upsilon)$ with respect to the i^{th} and j^{th} components of υ, 75

$L_i^{[1]}(t;\upsilon,\theta)$ limiting function of $L_{ni}^{(1)}(t;\upsilon)$, 57

$L_0(s,t;\upsilon_0)$ version of $L(s;\upsilon_0,\theta_0,\upsilon_0,\theta_0)$, in which the influence of the M-estimator at the stochastic integral evaluated at point t is accounted for, 58

$L_n(t)$ weight process, integrand stochastic integral $Q_n(t)$, 24

$L_n(t;\upsilon)$ weight process, integrand stochastic integral $Q_n(t;\upsilon)$, 52

$M_n(t;\theta_0)$ basic martingale, simple null hypothesis, 5, 23

$M_n(t;\upsilon_0,\theta_0)$ basic martingale, composite null hypothesis, 52

$\underline{\mathrm{N}}_i^{\tilde{\beta}}(\theta)$ smallest sample size such that the power of the test is at least as desired under P_θ, 31

$\underline{\mathrm{N}}_i^{\tilde{\beta}}(\upsilon,\theta)$ smallest sample size such that the power of the test is at least as desired under $P_{(\upsilon,\theta)}$, 64

$\overline{N}_i^{\tilde{\beta}}(\theta)$ largest sample size such that the power of the test is less than desired under P_θ, 31

$\overline{N}_i^{\tilde{\beta}}(v,\theta)$ largest sample size such that the power of the test is less than desired under $P_{(v,\theta)}$, 63

P_0 probability measure under the null hypothesis, simple null hypothesis, 23

P_n actual probability measure, 23, 51

$P_{(v,\theta)}$ probability measure under fixed alternative $v_n = v$ and $\theta_n = \theta$, 61

P_θ probability measure under fixed alternative $\theta_n = \theta$, 29

P_{v_0} probability measure under the null hypothesis, composite null hypothesis, 51

$Q_n^{(1)}(t;v)$ r-dimensional vector with elements $Q_{ni}^{(1)}(t;v)$, 76

$Q_{ni}^{(1)}(t;v)$ first order partial derivative of $Q_n(t;v)$ with respect to the i^{th} component of v, 75

$Q_n^{(2)}(t;v)$ $r \times r$ matrix with elements $Q_{nij}^{(2)}(t;v)$, 76

$Q_{nij}^{(2)}(t;v)$ second order partial derivative of $Q_n(t;v)$ with respect to the i^{th} and j^{th} components of v, 75

$Q_n(t)$ stochastic integral with respect to $M_n(t;\theta_0)$, 24

$Q_n(t;v)$ stochastic integral with respect to $M_n(t;v,\theta_0)$, 52

$\tilde{Q}_{ni}^{(1)}(t;v)$ approximation of $Q_{ni}^{(1)}(t;v)$, 75

$\tilde{Q}_{nij}^{(2)}(t;v)$ approximation of $Q_{nij}^{(2)}(t;v)$, 75

$R_n(v)$ r-dimensional function derived from M-equations, satisfying $v^{(n)} = R_n(v^{(n)})$, 71

S set of Strassen functions, 21

S_n random variable describing behavior of M-estimator, 55, 70

T_R generalized rank functional, 8

T_S supremum functional, 8

$U_n^1(t;\theta_n)$ empirical process based on $H_n^1(t)$, 24

$U_n^1(t;v_n,\theta_n)$ empirical process, 70

$U_n(t)$ empirical process, 2

$U_{n-}(t;\theta_n)$ empirical process based on $H_{n-}(t)$, 24

$U_{n-}(t;v_n,\theta_n)$ empirical process, 70

$\tilde{U}_n(t)$ uniform empirical process, 17, 72

V_n closed ball in Υ with centre v_{n0} and radius c_{59}, 70

$W(t)$ standard Wiener process, 20

$W_n(t)$ Gaussian process approximating [a centered version of] the basic martingale, 26, 37

$W_n(t;v_{n0})$ mean zero Gaussian processes involved in approximation of $Q_n(t;v^{(n)})$, 58

$X(t;\theta_0)$ limit in P_0-distribution of $Q_n(t)$ in q-metric, 27

X_i failure time, 4

$\check{X}(t;v_n,\theta_n)$ mean zero Gaussian process approximating centered version of $Q_n(t;v^{(n)})$, 59, 78

Y_i censoring time, 4

Z_i censored failure time, 4

$\tilde{Z}_i$ transformation of (Z_i,δ_i) to standard uniform random variable, 36, 72

α constant, 25

$\tilde{\beta}(\theta)$ desired power under P_θ, 31

$\tilde{\beta}(v,\theta)$ desired power under $P_{(v,\theta)}$, 63

δ_i censoring indicator, 4

θ_0 value of parameter of interest under the null hypothesis, 23, 51

θ_n actual value of parameter of interest, 23, 51

θ parameter of interest, 23, 51

$\lambda(t;\theta)$ hazard function, simple null hypothesis, 8, 29

$\lambda(t;v,\theta)$ hazard function, composite null hypothesis, 53

$\Lambda(t;\theta)$ cumulative hazard function belonging to X_i, simple null hypothesis, 5, 23

$\Lambda(t;\upsilon,\theta)$ cumulative hazard function belonging to $F(t;\upsilon,\theta)$, composite null hypothesis, 52

$\Lambda_i^{(1)}(t;\upsilon,\theta)$ first order partial derivative of $\Lambda(t;\upsilon,\theta)$ with respect to the i^{th} component of (υ,θ), 62

$\mu_\beta(\upsilon)$ integral, 88

$\hat{\mu}_\beta(\upsilon)$ estimator of $\mu_\beta(\upsilon)$, 93

$\Xi_0(\upsilon,\theta)$ $r \times r$ matrix describing the behavior of M-estimator, 54

$\Xi_b(\upsilon,\theta)$ $r \times r$ matrix, 62

Ξ_{0n} abbreviation for $\Xi_0(\upsilon_n,\theta_n)$, 55

$\pi(\upsilon,\theta)$ projection related to M-estimator, mapping $\Upsilon \times \Theta$ into Υ, 54

Σ_{e0} abbreviation for $\Xi_0(\upsilon_0,\theta_0)$, 55

Σ_{ee} covariance matrix of $\tilde{\Phi}_n$, 56

$\upsilon^{(n)}$ M-estimator, 52

υ_0 value of nuisance parameter under P_{υ_0}, 51

υ_n actual value of nuisance parameter, 51

υ_{n0} abbreviation for $\pi(\upsilon_n,\theta_n)$, 55

υ nuisance parameter, 51

$\phi(t;\upsilon)$ r-dimensional vector with elements $\phi_i(t;\upsilon)$, 55

$\phi_{ijk}^{(2)}(t;\upsilon)$ second order partial derivative of $\phi_i(t;\upsilon)$ with respect to the j^{th} and the k^{th} components of υ, 69

$\phi_i(t;\upsilon)$ function used for constructing an M-estimator, 53

$\Phi_n^{(1)}(\upsilon)$ $r \times r$ matrix with elements $\Phi_{nij}^{(1)}(\upsilon)$, 69

$\Phi_{nij}^{(1)}(\upsilon)$ first order partial derivative of $\Phi_{ni}(\upsilon)$ with respect to the j^{th} component of υ, 69

$\Phi_{ni}^{(1)}(\upsilon_{n0})$ i^{th} row of $\Phi_n^{(1)}(\upsilon_{n0})$, 70

$\Phi_{nijk}^{(2)}(\upsilon)$ second order partial derivative of $\Phi_{ni}(\upsilon)$ with respect to the j^{th} and k^{th} components of υ, 69

$\Phi_{ni}^{(2)}(\upsilon)$ $r \times r$ matrix with elements $\Phi_{nijk}^{(2)}(\upsilon)$, 69

$\Phi_{ni}(\upsilon)$ shorthand notation for $\Phi_{ni}(\infty;\upsilon)$, 69

$\Phi_{ni}(t;\upsilon)$ stochastic integral with integrand $\phi_i(t;\upsilon)$, involved in equations determining M-estimator, 53

$\tilde{\Phi}_n$ r-dimensional Gaussian random vector approximating $n^{1/2}\Sigma_{e0}(\upsilon^{(n)} - \upsilon_{n0})$, 56

$\psi(t;\upsilon,\theta)$ r-dimensional vector with components $\psi_i(t;\upsilon,\theta)$, 58

$\psi_{ij}^{(1)}(t;\theta)$ second order partial derivative of $\lambda(t;\theta)$ with respect to the i^{th} and j^{th} components of θ, 29

$\psi_{jk}^{(1)}(t;\upsilon)$ first order partial derivative of $\psi_j(t;\upsilon)$ with respect to the k^{th} component of υ, 69

$\psi_a(t;\theta_0)$ score function, simple null hypothesis, 8

$\psi_a(t;\upsilon_0,\theta_0)$ p-dimensional vector function with elements $\psi_{r+i}(t;\upsilon_0,\theta_0)$, 62

$\psi_i(t;\theta)$ first order partial derivative of $\lambda(t;\theta)$ with respect to the i^{th} component of θ, 29

$\psi_i(t;\upsilon,\theta)$ first order partial derivative of $\lambda(t;\upsilon,\theta)$ with respect to the i^{th} component of (υ,θ), 53

$\psi_{a|0}(t;\upsilon_0,\theta_0)$ effective score function, 67

Ω_n subset of Ω, 76

$(\Omega,\mathcal{A},\mathcal{P})$ underlying probability space, 17, 23, 51

MATHEMATICAL CENTRE TRACTS

1 T. van der Walt. *Fixed and almost fixed points.* 1963.

2 A.R. Bloemena. *Sampling from a graph.* 1964.

3 G. de Leve. *Generalized Markovian decision processes, part I: model and method.* 1964.

4 G. de Leve. *Generalized Markovian decision processes, part II: probabilistic background.* 1964.

5 G. de Leve, H.C. Tijms, P.J. Weeda. *Generalized Markovian decision processes, applications.* 1970.

6 M.A. Maurice. *Compact ordered spaces.* 1964.

7 W.R. van Zwet. *Convex transformations of random variables.* 1964.

8 J.A. Zonneveld. *Automatic numerical integration.* 1964.

9 P.C. Baayen. *Universal morphisms.* 1964.

10 E.M. de Jager. *Applications of distributions in mathematical physics.* 1964.

11 A.B. Paalman-de Miranda. *Topological semigroups.* 1964.

12 J.A.Th.M. van Berckel, H. Brandt Corstius, R.J. Mokken, A. van Wijngaarden. *Formal properties of newspaper Dutch.* 1965.

13 H.A. Lauwerier. *Asymptotic expansions.* 1966, out of print; replaced by MCT 54.

14 H.A. Lauwerier. *Calculus of variations in mathematical physics.* 1966.

15 R. Doornbos. *Slippage tests.* 1966.

16 J.W. de Bakker. *Formal definition of programming languages with an application to the definition of ALGOL 60.* 1967.

17 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 1.* 1968.

18 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 2.* 1968.

19 J. van der Slot. *Some properties related to compactness.* 1968.

20 P.J. van der Houwen. *Finite difference methods for solving partial differential equations.* 1968.

21 E. Wattel. *The compactness operator in set theory and topology.* 1968.

22 T.J. Dekker. *ALGOL 60 procedures in numerical algebra, part 1.* 1968.

23 T.J. Dekker, W. Hoffmann. *ALGOL 60 procedures in numerical algebra, part 2.* 1968.

24 J.W. de Bakker. *Recursive procedures.* 1971.

25 E.R. Paërl. *Representations of the Lorentz group and projective geometry.* 1969.

26 European Meeting 1968. *Selected statistical papers, part I.* 1968.

27 European Meeting 1968. *Selected statistical papers, part II.* 1968.

28 J. Oosterhoff. *Combination of one-sided statistical tests.* 1969.

29 J. Verhoeff. *Error detecting decimal codes.* 1969.

30 H. Brandt Corstius. *Exercises in computational linguistics.* 1970.

31 W. Molenaar. *Approximations to the Poisson, binomial and hypergeometric distribution functions.* 1970.

32 L. de Haan. *On regular variation and its application to the weak convergence of sample extremes.* 1970.

33 F.W. Steutel. *Preservation of infinite divisibility under mixing and related topics.* 1970.

34 I. Juhász, A. Verbeek, N.S. Kroonenberg. *Cardinal functions in topology.* 1971.

35 M.H. van Emden. *An analysis of complexity.* 1971.

36 J. Grasman. *On the birth of boundary layers.* 1971.

37 J.W. de Bakker, G.A. Blaauw, A.J.W. Duijvestijn, E.W. Dijkstra, P.J. van der Houwen, G.A.M. Kamsteeg-Kemper, F.E.J. Kruseman Aretz, W.L. van der Poel, J.P. Schaap-Kruseman, M.V. Wilkes, G. Zoutendijk. *MC-25 Informatica Symposium.* 1971.

38 W.A. Verloren van Themaat. *Automatic analysis of Dutch compound words.* 1972.

39 H. Bavinck. *Jacobi series and approximation.* 1972.

40 H.C. Tijms. *Analysis of (s,S) inventory models.* 1972.

41 A. Verbeek. *Superextensions of topological spaces.* 1972.

42 W. Vervaat. *Success epochs in Bernoulli trials (with applications in number theory).* 1972.

43 F.H. Ruymgaart. *Asymptotic theory of rank tests for independence.* 1973.

44 H. Bart. *Meromorphic operator valued functions.* 1973.

45 A.A. Balkema. *Monotone transformations and limit laws.* 1973.

46 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 1: the language.* 1973.

47 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 2: the compiler.* 1973.

48 F.E.J. Kruseman Aretz, P.J.W. ten Hagen, H.L. Oudshoorn. *An ALGOL 60 compiler in ALGOL 60, text of the MC-compiler for the EL-X8.* 1973.

49 H. Kok. *Connected orderable spaces.* 1974.

50 A. van Wijngaarden, B.J. Mailloux, J.E.L. Peck, C.H.A. Koster, M. Sintzoff, C.H. Lindsey, L.G.L.T. Meertens, R.G. Fisker (eds.). *Revised report on the algorithmic language ALGOL 68.* 1976.

51 A. Hordijk. *Dynamic programming and Markov potential theory.* 1974.

52 P.C. Baayen (ed.). *Topological structures.* 1974.

53 M.J. Faber. *Metrizability in generalized ordered spaces.* 1974.

54 H.A. Lauwerier. *Asymptotic analysis, part 1.* 1974.

55 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 1: theory of designs, finite geometry and coding theory.* 1974.

56 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 2: graph theory, foundations, partitions and combinatorial geometry.* 1974.

57 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 3: combinatorial group theory.* 1974.

58 W. Albers. *Asymptotic expansions and the deficiency concept in statistics.* 1975.

59 J.L. Mijnheer. *Sample path properties of stable processes.* 1975.

60 F. Göbel. *Queueing models involving buffers.* 1975.

63 J.W. de Bakker (ed.). *Foundations of computer science.* 1975.

64 W.J. de Schipper. *Symmetric closed categories.* 1975.

65 J. de Vries. *Topological transformation groups, 1: a categorical approach.* 1975.

66 H.G.J. Pijls. *Logically convex algebras in spectral theory and eigenfunction expansions.* 1976.

68 P.P.N. de Groen. *Singularly perturbed differential operators of second order.* 1976.

69 J.K. Lenstra. *Sequencing by enumerative methods.* 1977.

70 W.P. de Roever, Jr. *Recursive program schemes: semantics and proof theory.* 1976.

71 J.A.E.E. van Nunen. *Contracting Markov decision processes.* 1976.

72 J.K.M. Jansen. *Simple periodic and non-periodic Lamé functions and their applications in the theory of conical waveguides.* 1977.

73 D.M.R. Leivant. *Absoluteness of intuitionistic logic.* 1979.

74 H.J.J. te Riele. *A theoretical and computational study of generalized aliquot sequences.* 1976.

75 A.E. Brouwer. *Treelike spaces and related connected topological spaces.* 1977.

76 M. Rem. *Associons and the closure statement.* 1976.

77 W.C.M. Kallenberg. *Asymptotic optimality of likelihood ratio tests in exponential families.* 1978.

78 E. de Jonge, A.C.M. van Rooij. *Introduction to Riesz spaces.* 1977.

79 M.C.A. van Zuijlen. *Emperical distributions and rank statistics.* 1977.

80 P.W. Hemker. *A numerical study of stiff two-point boundary problems.* 1977.

81 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 1.* 1976.

82 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 2.* 1976.

83 L.S. van Benthem Jutting. *Checking Landau's "Grundlagen" in the AUTOMATH system.* 1979.

84 H.L.L. Busard. *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia (?), books vii-xii.* 1977.

85 J. van Mill. *Supercompactness and Wallman spaces.* 1977.

86 S.G. van der Meulen, M. Veldhorst. *Torrix I, a programming system for operations on vectors and matrices over arbitrary fields and of variable size.* 1978.

88 A. Schrijver. *Matroids and linking systems.* 1977.

89 J.W. de Roever. *Complex Fourier transformation and analytic functionals with unbounded carriers.* 1978.

90 L.P.J. Groenewegen. *Characterization of optimal strategies in dynamic games.* 1981.

91 J.M. Geysel. *Transcendence in fields of positive characteristic.* 1979.

92 P.J. Weeda. *Finite generalized Markov programming.* 1979.

93 H.C. Tijms, J. Wessels (eds.). *Markov decision theory.* 1977.

94 A. Bijlsma. *Simultaneous approximations in transcendental number theory.* 1978.

95 K.M. van Hee. *Bayesian control of Markov chains.* 1978.

96 P.M.B. Vitányi. *Lindenmayer systems: structure, languages, and growth functions.* 1980.

97 A. Federgruen. *Markovian control problems; functional equations and algorithms.* 1984.

98 R. Geel. *Singular perturbations of hyperbolic type.* 1978.

99 J.K. Lenstra, A.H.G. Rinnooy Kan, P. van Emde Boas (eds.). *Interfaces between computer science and operations research.* 1978.

100 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1.* 1979.

101 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2.* 1979.

102 D. van Dulst. *Reflexive and superreflexive Banach spaces.* 1978.

103 K. van Harn. *Classifying infinitely divisible distributions by functional equations.* 1978.

104 J.M. van Wouwe. *Go-spaces and generalizations of metrizability.* 1979.

105 R. Helmers. *Edgeworth expansions for linear combinations of order statistics.* 1982.

106 A. Schrijver (ed.). *Packing and covering in combinatorics.* 1979.

107 C. den Heijer. *The numerical solution of nonlinear operator equations by imbedding methods.* 1979.

108 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 1.* 1979.

109 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 2.* 1979.

110 J.C. van Vliet. *ALGOL 68 transput, part I: historical review and discussion of the implementation model.* 1979.

111 J.C. van Vliet. *ALGOL 68 transput, part II: an implementation model.* 1979.

112 H.C.P. Berbee. *Random walks with stationary increments and renewal theory.* 1979.

113 T.A.B. Snijders. *Asymptotic optimality theory for testing problems with restricted alternatives.* 1979.

114 A.J.E.M. Janssen. *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes.* 1979.

115 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 1.* 1979.

116 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 2.* 1979.

117 P.J.M. Kallenberg. *Branching processes with continuous state space.* 1979.

118 P. Groeneboom. *Large deviations and asymptotic efficiencies.* 1980.

119 F.J. Peters. *Sparse matrices and substructures, with a novel implementation of finite element algorithms.* 1980.

120 W.P.M. de Ruyter. *On the asymptotic analysis of large-scale ocean circulation.* 1980.

121 W.H. Haemers. *Eigenvalue techniques in design and graph theory.* 1980.

122 J.C.P. Bus. *Numerical solution of systems of nonlinear equations.* 1980.

123 I. Yuhász. *Cardinal functions in topology - ten years later.* 1980.

124 R.D. Gill. *Censoring and stochastic integrals.* 1980.

125 R. Eising. *2-D systems, an algebraic approach.* 1980.

126 G. van der Hoek. *Reduction methods in nonlinear programming.* 1980.

127 J.W. Klop. *Combinatory reduction systems.* 1980.

128 A.J.J. Talman. *Variable dimension fixed point algorithms and triangulations.* 1980.

129 G. van der Laan. *Simplicial fixed point algorithms.* 1980.

130 P.J.W. ten Hagen, T. Hagen, P. Klint, H. Noot, H.J. Sint, A.H. Veen. *ILP: intermediate language for pictures.* 1980.

131 R.J.R. Back. *Correctness preserving program refinements: proof theory and applications.* 1980.

132 H.M. Mulder. *The interval function of a graph.* 1980.

133 C.A.J. Klaassen. *Statistical performance of location estimators.* 1981.

134 J.C. van Vliet, H. Wupper (eds.). *Proceedings international conference on ALGOL 68.* 1981.

135 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part I.* 1981.

136 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part II.* 1981.

137 J. Telgen. *Redundancy and linear programs.* 1981.

138 H.A. Lauwerier. *Mathematical models of epidemics.* 1981.

139 J. van der Wal. *Stochastic dynamic programming, successive approximations and nearly optimal strategies for Markov decision processes and Markov games.* 1981.

140 J.H. van Geldrop. *A mathematical theory of pure exchange economies without the no-critical-point hypothesis.* 1981.

141 G.E. Welters. *Abel-Jacobi isogenies for certain types of Fano threefolds.* 1981.

142 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 1.* 1981.

143 J.M. Schumacher. *Dynamic feedback in finite- and infinite-dimensional linear systems.* 1981.

144 P. Eijgenraam. *The solution of initial value problems using interval arithmetic; formulation and analysis of an algorithm.* 1981.

145 A.J. Brentjes. *Multi-dimensional continued fraction algorithms.* 1981.

146 C.V.M. van der Mee. *Semigroup and factorization methods in transport theory.* 1981.

147 H.H. Tigelaar. *Identification and informative sample size.* 1982.

148 L.C.M. Kallenberg. *Linear programming and finite Markovian control problems.* 1983.

149 C.B. Huijsmans, M.A. Kaashoek, W.A.J. Luxemburg, W.K. Vietsch (eds.). *From A to Z, proceedings of a symposium in honour of A.C. Zaanen.* 1982.

150 M. Veldhorst. *An analysis of sparse matrix storage schemes.* 1982.

151 R.J.M.M. Does. *Higher order asymptotics for simple linear rank statistics.* 1982.

152 G.F. van der Hoeven. *Projections of lawless sequences.* 1982.

153 J.P.C. Blanc. *Application of the theory of boundary value problems in the analysis of a queueing model with paired services.* 1982.

154 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part I.* 1982.

155 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part II.* 1982.

156 P.M.G. Apers. *Query processing and data allocation in distributed database systems.* 1983.

157 H.A.W.M. Kneppers. *The covariant classification of two-dimensional smooth commutative formal groups over an algebraically closed field of positive characteristic.* 1983.

158 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 1.* 1983.

159 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 2.* 1983.

160 A. Rezus. *Abstract AUTOMATH.* 1983.

161 G.F. Helminck. *Eisenstein series on the metaplectic group, an algebraic approach.* 1983.

162 J.J. Dik. *Tests for preference.* 1983.

163 H. Schippers. *Multiple grid methods for equations of the second kind with applications in fluid mechanics.* 1983.

164 F.A. van der Duyn Schouten. *Markov decision processes with continuous time parameter.* 1983.

165 P.C.T. van der Hoeven. *On point processes.* 1983.

166 H.B.M. Jonkers. *Abstraction, specification and implementation techniques, with an application to garbage collection.* 1983.

167 W.H.M. Zijm. *Nonnegative matrices in dynamic programming.* 1983.

168 J.H. Evertse. *Upper bounds for the numbers of solutions of diophantine equations.* 1983.

169 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 2.* 1983.

CWI TRACTS

1 D.H.J. Epema. *Surfaces with canonical hyperplane sections.* 1984.

2 J.J. Dijkstra. *Fake topological Hilbert spaces and characterizations of dimension in terms of negligibility.* 1984.

3 A.J. van der Schaft. *System theoretic descriptions of physical systems.* 1984.

4 J. Koene. *Minimal cost flow in processing networks, a primal approach.* 1984.

5 B. Hoogenboom. *Intertwining functions on compact Lie groups.* 1984.

6 A.P.W. Böhm. *Dataflow computation.* 1984.

7 A. Blokhuis. *Few-distance sets.* 1984.

8 M.H. van Hoorn. *Algorithms and approximations for queueing systems.* 1984.

9 C.P.J. Koymans. *Models of the lambda calculus.* 1984.

10 C.G. van der Laan, N.M. Temme. *Calculation of special functions: the gamma function, the exponential integrals and error-like functions.* 1984.

11 N.M. van Dijk. *Controlled Markov processes; time-discretization.* 1984.

12 W.H. Hundsdorfer. *The numerical solution of nonlinear stiff initial value problems: an analysis of one step methods.* 1985.

13 D. Grune. *On the design of ALEPH.* 1985.

14 J.G.F. Thiemann. *Analytic spaces and dynamic programming: a measure theoretic approach.* 1985.

15 F.J. van der Linden. *Euclidean rings with two infinite primes.* 1985.

16 R.J.P. Groothuizen. *Mixed elliptic-hyperbolic partial differential operators: a case-study in Fourier integral operators.* 1985.

17 H.M.M. ten Eikelder. *Symmetries for dynamical and Hamiltonian systems.* 1985.

18 A.D.M. Kester. *Some large deviation results in statistics.* 1985.

19 T.M.V. Janssen. *Foundations and applications of Montague grammar, part 1: Philosophy, framework, computer science.* 1986.

20 B.F. Schriever. *Order dependence.* 1986.

21 D.P. van der Vecht. *Inequalities for stopped Brownian motion.* 1986.

22 J.C.S.P. van der Woude. *Topological dynamix.* 1986.

23 A.F. Monna. *Methods, concepts and ideas in mathematics: aspects of an evolution.* 1986.

24 J.C.M. Baeten. *Filters and ultrafilters over definable subsets of admissible ordinals.* 1986.

25 A.W.J. Kolen. *Tree network and planar rectilinear location theory.* 1986.

26 A.H. Veen. *The misconstrued semicolon: Reconciling imperative languages and dataflow machines.* 1986.

27 A.J.M. van Engelen. *Homogeneous zero-dimensional absolute Borel sets.* 1986.

28 T.M.V. Janssen. *Foundations and applications of Montague grammar, part 2: Applications to natural language.* 1986.

29 H.L. Trentelman. *Almost invariant subspaces and high gain feedback.* 1986.

30 A.G. de Kok. *Production-inventory control models: approximations and algorithms.* 1987.

31 E.E.M. van Berkum. *Optimal paired comparison designs for factorial experiments.* 1987.

32 J.H.J. Einmahl. *Multivariate empirical processes.* 1987.

33 O.J. Vrieze. *Stochastic games with finite state and action spaces.* 1987.

34 P.H.M. Kersten. *Infinitesimal symmetries: a computational approach.* 1987.

35 M.L. Eaton. *Lectures on topics in probability inequalities.* 1987.

36 A.H.P. van der Burgh, R.M.M. Mattheij (eds.). *Proceedings of the first international conference on industrial and applied mathematics (ICIAM 87).* 1987.

37 L. Stougie. *Design and analysis of algorithms for stochastic integer programming.* 1987.

38 J.B.G. Frenk. *On Banach algebras, renewal measures and regenerative processes.* 1987.

39 H.J.M. Peters, O.J. Vrieze (eds.). *Surveys in game theory and related topics.* 1987.

40 J.L. Geluk, L. de Haan. *Regular variation, extensions and Tauberian theorems.* 1987.

41 Sape J. Mullender (ed.). *The Amoeba distributed operating system: Selected papers 1984-1987.* 1987.

42 P.R.J. Asveld, A. Nijholt (eds.). *Essays on concepts, formalisms, and tools.* 1987.

43 H.L. Bodlaender. *Distributed computing: structure and complexity.* 1987.

44 A.W. van der Vaart. *Statistical estimation in large parameter spaces.* 1988.

45 S.A. van de Geer. *Regression analysis and empirical processes.* 1988.

46 S.P. Spekreijse. *Multigrid solution of the steady Euler equations.* 1988.

47 J.B. Dijkstra. *Analysis of means in some non-standard situations.* 1988.

48 F.C. Drost. *Asymptotics for generalized chi-square goodness-of-fit tests.* 1988.

49 F.W. Wubs. *Numerical solution of the shallow-water equations.* 1988.

50 F. de Kerf. *Asymptotic analysis of a class of perturbed Korteweg-de Vries initial value problems.* 1988.

51 P.J.M. van Laarhoven. *Theoretical and computational aspects of simulated annealing.* 1988.

52 P.M. van Loon. *Continuous decoupling transformations for linear boundary value problems.* 1988.

53 K.C.P. Machielsen. *Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space.* 1988.

54 L.C.R.J. Willenborg. *Computational aspects of survey data processing.* 1988.

55 G.J. van der Steen. *A program generator for recognition, parsing and transduction with syntactic patterns.* 1988.

56 J.C. Ebergen. *Translating programs into delay-insensitive circuits.* 1989.

57 S.M. Verduyn Lunel. *Exponential type calculus for linear delay equations.* 1989.

58 M.C.M. de Gunst. *A random model for plant cell population growth.* 1989.

59 D. van Dulst. *Characterizations of Banach spaces not containing* l^1. 1989.

60 H.E. de Swart. *Vacillation and predictability properties of low-order atmospheric spectral models.* 1989.

61 P. de Jong. *Central limit theorems for generalized multilinear forms.* 1989.

62 V.J. de Jong. *A specification system for statistical software.* 1989.

63 B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems, part I.* 1989.

64 B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems, part II.* 1989.

65 B.M.M. de Weger. *Algorithms for diophantine equations.* 1989.

66 A. Jung. *Cartesian closed categories of domains.* 1989.

67 J.W. Polderman. *Adaptive control & identification: Conflict or conflux?.* 1989.

68 H.J. Woerdeman. *Matrix and operator extensions.* 1989.

69 B.G. Hansen. *Monotonicity properties of infinitely divisible distributions.* 1989.

70 J.K. Lenstra, H.C. Tijms, A. Volgenant (eds.). *Twenty-five years of operations research in the Netherlands: Papers dedicated to Gijs de Leve.* 1990.

71 P.J.C. Spreij. *Counting process systems. Identification and stochastic realization.* 1990.

72 J.F. Kaashoek. *Modeling one dimensional pattern formation by anti-diffusion.* 1990.

73 A.M.H. Gerards. *Graphs and polyhedra. Binary spaces and cutting planes.* 1990.

74 B. Koren. *Multigrid and defect correction for the steady Navier-Stokes equations. Application to aerodynamics.* 1991.

75 M.W.P. Savelsbergh. *Computer aided routing.* 1992.

76 O.E. Flippo. *Stability, duality and decomposition in general mathematical programming.* 1991.

77 A.J. van Es. *Aspects of nonparametric density estimation.* 1991.

78 G.A.P. Kindervater. *Exercises in parallel combinatorial computing.* 1992.

79 J.J. Lodder. *Towards a symmetrical theory of generalized functions.* 1991.

80 S.A. Smulders. *Control of freeway traffic flow.* 1993.

81 P.H.M. America, J.J.M.M. Rutten. *A parallel object-oriented language: design and semantic foundations.* 1992.

82 F. Thuijsman. *Optimality and equilibria in stochastic games.* 1992.

83 R.J. Kooman. *Convergence properties of recurrence sequences.* 1992.

84 A.M. Cohen (ed.). *Computational aspects of Lie group representations and related topics. Proceedings of the 1990 Computational Algebra Seminar at CWI, Amsterdam.* 1991.

85 V. de Valk. *One-dependent processes.* 1993.

86 J.A. Baars, J.A.M. de Groot. *On topological and linear equivalence of certain function spaces.* 1992.

87 A.F. Monna. *The way of mathematics and mathematicians.* 1992.

88 E.D. de Goede. *Numerical methods for the three-dimensional shallow water equations.* 1993.

89 M. Zwaan. *Moment problems in Hilbert space with applications to magnetic resonance imaging.* 1993.

90 C. Vuik. *The solution of a one-dimensional Stefan problem.* 1993.

91. E.R. Verheul. *Multimedians in metric and normed spaces.* 1993.

92. J.L.M. Maubach. *Iterative methods for non-linear partial differential equations.* 1993.

93. A.W. Ambergen. *Statistical uncertainties in posterior probabilities.* 1993.

94. P.A. Zegeling. *Moving-grid methods for time-dependent partial differential equations.* 1993.

95. M.J.C. van Pul. *Statistical analysis of software reliability models.* 1993.

96. J.K. Scholma. *A Lie algebraic study of some integrable systems associated with root systems.* 1993.

97. J.L. van den Berg. *Sojourn times in feedback and processor sharing queues.* 1993.

98. A.J. Koning. *Stochastic integrals and goodness-of-fit tests.* 1993.