The Annals of Statistics
1996, Vol. 24, No. 6, 2679–2704

THE JACKKNIFE ESTIMATE OF VARIANCE OF A KAPLAN–MEIER INTEGRAL[1]

BY WINFRIED STUTE

University of Giessen

Let $\hat{F}_n$ be the Kaplan–Meier estimator of a distribution function F computed from randomly censored data. It is known that, under certain integrability assumptions on a function φ, the Kaplan–Meier integral $\int \varphi \, d\hat{F}_n$, when properly standardized, is asymptotically normal. In this paper it is shown that, with probability 1, the jackknife estimate of variance consistently estimates the (limit) variance.

1. Introduction and main results. Let $X_1, \ldots, X_n$ be a sample of independent random variables from some unknown distribution function (d.f.) F. In survival analysis the Xs usually may be interpreted as the times from the entry into a followup study until failure. Due to other causes of failure, it may then happen that not all Xs are observable. Rather one obtains variables $Z_1, \ldots, Z_n$ and indicators $\delta_1, \ldots, \delta_n$ such that $Z_i = X_i$ only if $\delta_i = 1$. The nonparametric estimator of F that adapts the (unknown) empirical d.f. F_n of the Xs to the present situation is the time-honoured Kaplan–Meier product limit estimator $\hat{F}_n$ defined as

$$1 - \hat{F}_n(x) = \prod_{i=1}^{n} \left[1 - \frac{\delta_{[i:n]}}{n-i+1}\right]^{1_{\{Z_{i:n} \le x\}}}.$$

Here $Z_{1:n} \le \cdots \le Z_{n:n}$ are the order statistics of the Z-sample and $\delta_{[i:n]}$ is the δ associated with $Z_{i:n}$, that is, $\delta_{[i:n]} = \delta_j$ if $Z_{i:n} = Z_j$. When ties are present, lifetimes within censored ($\delta_i = 0$) and uncensored ($\delta_i = 1$) data are ordered arbitrarily, while ties among lifetimes and censored times are treated as if the former precedes the latter.

In statistics the investigation of estimators or test statistics may often be reduced, via a linearization and Cramér–Slutsky, to empirical integrals

$$\int \varphi \, dF_n = n^{-1} \sum_{i=1}^{n} \varphi(X_i).$$

Here φ is a proper score function depending on the quantity of interest. Consistency and asymptotic normality may then be obtained, of course, from the classical SLLN and CLT applied to the independent and identically distributed (i.i.d.) random variables $\varphi(X_1), \ldots, \varphi(X_n)$.

Naturally, under random censorship, the role of $\int \varphi \, dF_n$ is played by the Kaplan–Meier integral $\int \varphi \, d\hat{F}_n$. Unfortunately, unlike the sample mean,

---

Received December 1994; revised November 1995.
[1]Work supported by the "Deutsche Forschungsgemeinschaft."
AMS 1991 *subject classifications*. Primary 62G05, 62G09; secondary 62G30, 60G42.
Key words and phrases. Censored data, Kaplan–Meier integral, variance, jackknife.

Kaplan–Meier integrals are no longer (weighted) sums of i.i.d. random variables. For this, let W_{in} denote the mass attached to $Z_{i:n}$ under $\hat{F}_n$. It is readily seen that

$$W_{in} = \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1} \right]^{\delta_{[j:n]}}.$$

Hence

$$S_n \equiv \int \varphi \, d\hat{F}_n = \sum_{i=1}^{n} W_{in} \varphi(Z_{i:n})$$

is a sum of randomly weighted functions of the ordered Z-sample. Interestingly enough, the assertions of the SLLN and the CLT could be extended to the more complicated setup. To properly describe the results, introduce the distributional characteristics of the observed (Z, δ), the d.f. of Z and the conditional expectation of δ given Z:

$$H(x) = \mathbb{P}(Z \le x), \qquad m(x) = \mathbb{P}(\delta = 1 | Z = x).$$

H and m together determine the joint distribution of (Z, δ). Stute and Wang (1993) then showed that, with probability 1 and in the mean,

$$\text{(1.1)} \quad \lim_{n\to\infty} \int \varphi \, d\hat{F}_n \equiv S = \int \varphi(t) m(t) \exp\left\{ \int_{-\infty}^{t-} \frac{1-m(x)}{1-H(x)} H(dx) \right\} H(dt),$$

provided that $\int |\varphi| \, dF < \infty$. Traditionally, it has always been assumed in the literature that $Z_i = \min(X_i, Y_i)$, where $Y_1, \ldots, Y_n$ is another sequence of independent random variables from some unknown censoring d.f. G, being also independent of the Xs. In this situation the limit in (1.1) collapses to $\int \varphi \, d\tilde{F}$, where

$$\text{(1.2)} \quad \tilde{F}(x) = \begin{cases} F(x), & \text{if } x < \tau_H, \\ F(x-) + 1_{\{\tau_H \in A\}} [F(\tau_H) - F(\tau_H -)], & \text{if } x \ge \tau_H, \end{cases}$$

$\tau_H = \inf\{x\colon\ H(x) = 1\}$ $(\le \infty)$ is the least upper bound for the support of H and A is the set of H-atoms (possibly empty).

Tsiatis (1975), among others, showed the fully nonparametric character of this model in that independence of X and Y cannot be checked from the data. Hence it might be preferable to obtain results like (1.1), which do not rely on the assumption of independent competing risks.

The CLT for censored data is due to Stute (1995). Again, the methodology works for general censorship and not just for $Z_i = \min(X_i, Y_i)$:

$$\text{(1.3)} \qquad n^{1/2} \left[\int \varphi \, d\hat{F}_n - S \right] \longrightarrow \mathcal{N}(0, \sigma_1^2) \quad \text{in distribution.}$$

Of course, the limit variance σ_1^2 equals the familiar expression

$$\sigma_0^2 = \int \varphi^2 \, dF - \left[\int \varphi \, dF \right]^2$$

if there is no censorship ($m \equiv 1$), but becomes much more complicated in the general case. It is an important quantity if one wants to compute confidence intervals for the target S, respectively, $\int \varphi\, dF$. For example, if $\varphi(x) = x$, $\int \varphi\, d\hat{F}_n$ is an estimator of the mean lifetime. For $\varphi = 1_{(-\infty,\, x_0]}$, $x_0 < \tau_H$, the resulting integral becomes $\hat{F}_n(x_0)$. Other examples are mentioned in Stute and Wang (1993). Unlike the indicator, most other φs are not of bounded variation (on the whole line) nor do they vanish right of some $T < \tau_H$. When these two assumptions are satisfied, Kaplan–Meier integrals can be readily handled using integration by parts and the Breslow and Crowley (1974) invariance principle.

Now, (1.3) suggests that

$$\operatorname{Var}\left[\int \varphi\, d\hat{F}_n\right] \sim \frac{\sigma_1^2}{n} \quad \text{as } n \to \infty.$$

It is the purpose of the present paper to show that the jackknife, under optimal conditions on φ, provides a second order consistent estimate of σ_1^2:

$$n\, \widehat{\operatorname{Var}}(\text{Jack}) \to \sigma_1^2 \quad \text{with probability 1.} \tag{1.4}$$

Before we give details, we need to (very) briefly discuss several aspects of the jackknife in the "complete data" situation. For more details, see Gray and Schucany (1972) and Efron and Tibshirani (1993). The jackknife has been proposed to serve two purposes [cf. Quenouille (1956) and Tukey (1958)], namely, to provide a methodology to reduce a possible bias of an estimator $S(F_n)$ and, second, to yield an approximation for its variance. The jackknife incorporates the so-called pseudovalues which result from applying the statistic of interest to the n subsamples of $X_1, \dots, X_n$ with X_k, $1 \le k \le n$, deleted one after another.

The crucial thing about this approach is that the statistic of interest is a function of F_n and therefore attaches mass $1/n$ to each of the data. Consequently deletion of one point just results in a change of the mass from $1/n$ to $1/(n-1)$. For the Kaplan–Meier integral, the situation is completely different, since now the statistic is a sum of (functions of) order statistics weighted by the complicated (random) W_{in}. So, if we denote by $\hat{F}_n^{(k)}$ the Kaplan–Meier estimator from the entire sample except $(Z_{k:n}, \delta_{[k:n]})$, then $S_n^{(k)} \equiv S(\hat{F}_n^{(k)})$ does not only involve changes of the standard weights, but also incorporates replacement of the weights W_{in} by new ones depending on the labels $\delta_{[i:n]}$, $1 \le i \le n$. This may be one reason why the jackknife under random censorship has been dealt with only in a few papers. Gaver and Miller (1983) showed that the jackknife-corrected Kaplan–Meier estimator at a fixed point $x < \tau_H$ has the same limit distribution as $\hat{F}_n(x)$. Stute and Wang (1994) investigated an arbitrary φ and derived an explicit formula for the jackknife modification of $\int \varphi\, d\hat{F}_n$. In particular, the mean of the pseudovalues equals

$$\bar{S}_n = S_n - \varphi(Z_{n:n}) \frac{\delta_{[n:n]}(1 - \delta_{[n-1:n]})}{n} \prod_{i=1}^{n-2} \left[\frac{n-i-1}{n-i}\right]^{\delta_{[i:n]}}.$$

Hence $\bar{S}_n$ does not necessarily coincide with S_n. This will have some impact on the statistical properties of $\widehat{\mathrm{Var}}(\mathrm{Jack})$ to the effect that, for (1.4), some slight modification of $\widehat{\mathrm{Var}}(\mathrm{Jack})$ will be required. Moreover, the bias-corrected jackknife substitute for S_n equals

$$\tilde{S}_n = S_n + \frac{n-1}{n}\varphi(Z_{n:n})\delta_{[n:n]}(1-\delta_{[n-1:n]})\prod_{i=1}^{n-2}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i:n]}},$$

that is, $\tilde{S}_n$ results from S_n by replacing W_{nn} with

$$W_{nn}^J = W_{nn} + \frac{n-1}{n}\delta_{[n:n]}(1-\delta_{[n-1:n]})\prod_{i=1}^{n-2}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i:n]}}.$$

The second summand of $\bar{S}_n$ appeared for the first time in Lemma 2.2 of Stute and Wang (1993). It also played an important role in analyzing the bias of a Kaplan–Meier integral [see Stute (1994)] and, as seen above, in the expressions for $\tilde{S}_n$ and $\bar{S}_n$.

Needless to say, as $n \to \infty$, $\bar{S}_n$ eventually coincides with S_n almost surely provided that φ vanishes right of some $x_0 < \tau_H$. For finite sample size n, $\bar{S}_n$ may differ from S_n, even if this support condition is satisfied. As we will see from Theorem 1.1, the extra term in $\bar{S}_n$ may destroy the consistency of $\widehat{\mathrm{Var}}(\mathrm{Jack})$ for a general φ not satisfying the support condition or may lead to an unstable behavior for a fixed sample size. Now, whatever φ may be, $\bar{S}_n$ also coincides with S_n unless

$$(*) \qquad \delta_{[n-1:n]} = 0 \quad \text{and} \quad \delta_{[n:n]} = 1.$$

So if, under $(*)$, we artificially set $\delta^*_{[n:n]} = 0$, we arrive at a modified estimate of variance which will be denoted $\widehat{\mathrm{Var}}^*(\mathrm{Jack})$. By construction the corresponding S_n^* and $\bar{S}_n^*$ coincide. This redefinition seems a little surprising in view of the fact that in the past many authors have proposed to reduce the bias of the Kaplan–Meier estimator by artificially putting $\delta^*_{[n:n]} = 1$ if $\delta_{[n:n]}$ happens to be zero, which is just the other way around. In this paper, however, our principal goal is to estimate the variance and not to reduce the bias of S_n. Also our choice of $\delta^*_{[n:n]}$ may be well motivated by the finite sample formula in Theorem 1.1, while for the bias reduction, our formula for $\tilde{S}_n$ points out that the jackknife is much more cautious about attributing mass to the largest Z.

Now, compared with the bias, a rigorous treatment of the variance estimator is much more complicated. Since

$$\begin{aligned} n\,\widehat{\mathrm{Var}}(\mathrm{Jack}) &= (n-1)\sum_{k=1}^{n}[S_n^{(k)} - \bar{S}_n]^2 \\ &= (n-1)\sum_{k=1}^{n}[S_n^{(k)}]^2 - n(n-1)\bar{S}_n^2, \end{aligned}$$

we need a careful investigation of the sum of squares. This will be done in Theorem 1.1. In order to compare the resulting expression for $n\,\widehat{\mathrm{Var}}(\mathrm{Jack})$

with σ_1^2, we finally have to discuss the structure of σ_1^2 and the assumptions needed for (1.3). For this, introduce the subdistribution functions

$$\tilde{H}^0(x) = \mathbb{P}(Z \le x,\ \delta = 0) = \int_{-\infty}^{x} (1-m)\,dH$$

and

$$\tilde{H}^1(x) = \mathbb{P}(Z \le x,\ \delta = 1) = \int_{-\infty}^{x} m\,dH.$$

Set

$$\begin{aligned}
\gamma_0(x) &= \exp\left\{\int_{-\infty}^{x-} \frac{\tilde{H}^0(dz)}{1-H(z)}\right\} = \exp\left\{\int_{-\infty}^{x-} \frac{1-m}{1-H}\,dH\right\},\\
\gamma_1(x) &= \frac{1}{1-H(x)} \int 1_{\{x<w\}}\varphi(w)\gamma_0(w)\tilde{H}^1(dw)\\
&= \frac{1}{1-H(x)} \int 1_{\{x<w\}}\varphi(w)\gamma_0(w)m(w)H(dw)
\end{aligned}$$

and

$$\begin{aligned}
\gamma_2(x) &= \int\int \frac{1_{\{v<x,\, v<w\}}\varphi(w)\gamma_0(w)}{[1-H(v)]^2}\tilde{H}^0(dv)\tilde{H}^1(dw)\\
&= \int_{-\infty}^{x-} \frac{\gamma_1(v)}{1-H(v)}[1-m(v)]H(dv).
\end{aligned}$$

In terms of these quantities, Corollary 1.2 in Stute (1995) then implies that (1.3) holds with

$$\begin{aligned}
\sigma_1^2 &= \int \varphi^2(x)\gamma_0^2(x)m(x)H(dx) - \int \gamma_1^2(x)(1-m(x))H(dx) - S^2\\
&\quad + \int \frac{\gamma_1^2(v)(1-m(v))^2}{1-H(v)} H\{v\}H(dv),
\end{aligned} \tag{1.5}$$

where $H\{v\} = H(v) - H(v-)$. Clearly, the last integral vanishes for a continuous H.

As to the assumptions needed for (1.3), recall that for the classical CLT one requires $\int \varphi^2\,dF < \infty$. Under censorship, two conditions are needed. The first is

$$\int \varphi^2(x)\gamma_0^2(x)m(x)H(dx) < \infty. \tag{1.6}$$

This integral is part of σ_1^2. Hence (1.6) guarantees that σ_1^2 is finite. When there is no censorship, $m \equiv 1$ and, therefore, $\gamma_0 \equiv 1$. In other words, (1.6) becomes the familiar $\int \varphi^2\,dF < \infty$. The other condition needed for (1.3) is a first moment assumption. It is mainly to control the bias of a Kaplan–Meier integral, which is not an issue when the Xs are completely observable, but is of some concern when the data are at risk of being censored. See Stute (1994) for a detailed discussion of this topic. Since the present paper only deals with the variance, it is likely that the variant of (1.4) holds only under the condition (1.6). As will be made clear by Theorem 1.2, this will, in fact, be the case.

It is instructive to compute σ_1^2 when $Z = \min(X, Y)$, where $Y \sim G$ is independent of X. To simplify things, assume continuity throughout. Then we have [cf. Stute and Wang (1993)] $\gamma_0(x) = 1/(1 - G(x))$ and, therefore,

$$\begin{aligned} \sigma_1^2 &= \int \frac{\varphi^2(x)}{1 - G(x)} \tilde{F}(dx) - \left[\int \varphi(x) \tilde{F}(dx) \right]^2 \\ &\quad - \int \left[\int_x^\infty \varphi(w) \tilde{F}(dw) \right]^2 \frac{1 - F(x)}{[1 - H(x)]^2} G(dx), \end{aligned} \tag{1.7}$$

where $\tilde{F}$ has been defined in (1.2). Condition (1.6), which now becomes

$$\int \frac{\varphi^2(x)}{1 - G(x)} \tilde{F}(dx) < \infty, \tag{1.8}$$

ensures (among other things) that, in the right tails, censoring does not have enough of a dominating influence that the variability of the estimator would get out of control.

From (1.7) the asymptotic variance of $\hat{F}_n(x_0)$, $x_0 < \tau_H$, is obtained if we set $\varphi = 1_{(-\infty, x_0]}$. The plug-in estimator of σ_1^2 together with a correction for discontinuities then immediately leads to the time-honoured Greenwood formula [cf. Kaplan and Meier (1958) and Klein (1991)].

As a conclusion of our discussion, we see that the jackknife, when applied to indicators, yields a new estimate of variance which (i) is strongly consistent under no conditions on F and G and (ii) compared with the already existing estimators, is not a plug-in estimator based on asymptotic expressions, but, in view of the definition of $\widehat{\mathrm{Var}}^*$(Jack), automatically adapts for the variability of the data. Moreover, and most importantly, the jackknife (iii) may be applied for much more general φs and (iv) is valid without the assumption of independent competing risks.

We are now in the position to formulate Theorem 1.1. As mentioned earlier, Theorem 1.1 provides a proper representation of the sum of squares, $\sum_{k=1}^n [S_n^{(k)}]^2$. Put, for $1 \le j \le n - 2$,

$$b_j = b_{jn} = \frac{1}{(n - j - 1)^2} - \frac{1}{(n - j)^2} + \frac{1}{(n - j - 1)} - \frac{1}{n - j}$$

and recall W_{in}.

THEOREM 1.1. *We have*

$$
\begin{aligned}
\sum_{k=1}^{n}[S_n^{(k)}]^2 \qquad&\\
= \Bigg\{ n\left[\frac{n}{n-1}\right]^{2\delta_{[1:n]}}\left[\frac{n-2}{n}\right]^{\delta_{[1:n]}} &+ (\delta_{[1:n]}-1)\frac{n}{(n-1)^2}\Bigg\} S_n^2\\
+ \sum_{j=1}^{n-2}(\delta_{[j:n]}-1)b_j \prod_{k=1}^{j-1}&\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k:n]}}\left[\sum_{i=j+1}^{n}\varphi(Z_{i:n})W_{in}\right]^2\\
+ \sum_{i=1}^{n-1}\varphi^2(Z_{i:n})\delta_{[i:n]}\frac{1}{(n-i)^2}&\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n]}} + R_n\varphi^2(Z_{n:n})W_{nn}^2,
\end{aligned}
\tag{1.9}
$$

where

$$
\begin{aligned}
R_n = \frac{n}{(n-1)^2} &- \sum_{j=1}^{n-2}(\delta_{[j:n]}-1)b_j\prod_{k=1}^{j-1}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k:n]}}\\
&+ \sum_{j=1}^{n-1}\delta_{[j:n]}\frac{n-j+1}{(n-j)^3}\prod_{k=1}^{j-1}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k:n]}}\\
&- \prod_{k=1}^{n-1}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k:n]}}.
\end{aligned}
$$

Before we proceed it is worthwhile to compare the right-hand side of (1.9) with (1.5), assuming continuity of H for a moment. The first term of (1.9), when inserted into the formula for $\widehat{\mathrm{Var}}(\mathrm{Jack})$, will contribute to $-S^2$. The second sum will be in charge of $-\int \gamma_1^2(1-m)\,dH$, while the last sum corresponds to $\int \varphi^2\gamma_0^2 m\,dH$. Finally the term containing R_n will be asymptotically negligible. As a recommendation, if one has no time to go through the proof, it would be instructive to check (1.9) for $n = 2$ and $n = 3$ though.

To formulate the main result of this paper, recall $\widehat{\mathrm{Var}}^*(\mathrm{Jack})$, the jackknife estimate of variance with $\delta^*_{[n:n]} = 0$ if $(*)$ holds.

THEOREM 1.2. *Under* (1.6) [*respectively*, (1.8)], *we have*, *with probability* 1,

$$
\lim_{n\to\infty} n\,\widehat{\mathrm{Var}}^*(\mathrm{Jack}) = \sigma_1^2.
$$

The proof of Theorem 1.2 proceeds by showing that each of the three leading terms of (1.9) corresponds to a slightly disturbed (sub-, super-) martingale in reverse time, to which a proper convergence theorem may be applied (Lemma A from the Appendix), and that the remainder term is negligible.

It should be noted that, for the plug-in estimator, no general result corresponding to Theorem 1.2 is known so far. Simulation results for the mean lifetime estimator indicate that it might be very unstable.

REMARK. A reader familiar with the so-called delta method [see Andersen, Borgan, Gill and Keiding (1993), page 109] may wonder about the necessity to investigate linear functionals of the Kaplan–Meier estimator. Actually, the discussion of the (truncated) mean-lifetime estimator on page 275 of Andersen, Borgan, Gill and Keiding (1993) may be misleading in that one gets the impression that it is only the nonlinearity of a statistic which causes troubles and that a clever linearization could do the job. This is not so, just because the concept of compact differentiability as elaborated there only involves smooth functionals of the Kaplan–Meier estimator when restricted to a given compact set. The mean-lifetime estimator is the simplest of many other examples of a (linear) functional of the Kaplan–Meier estimator, which is not even continuous (in the weak topology). To this author, a general theory of quadratic functionals of $\hat{F}_n$, for example, is completely missing. A detailed study of such basic statistics would require deep probabilistic investigations, however, which could not be replaced by purely topological arguments.

2. Proofs. Since n will be kept fixed throughout, we may omit it notationally and write

$$W_i \equiv W_{in} = \frac{\delta_{[i]}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j]}}$$

for short. Similarly, with $Z_{(i)} \equiv Z_{i:n}$,

$$S_n = \sum_{i=1}^{n} W_i \varphi(Z_{(i)})$$

and, for $1 \le k \le n$,

$$\begin{aligned} S_n^{(k)} = & \sum_{i=1}^{k-1} \frac{\varphi(Z_{(i)})\delta_{[i]}}{n-i} \prod_{j=1}^{i-1} \left[\frac{n-j-1}{n-j}\right]^{\delta_{[j]}} \\ & + \sum_{i=k+1}^{n} \frac{\varphi(Z_{(i)})\delta_{[i]}}{n-i+1} \prod_{j=1}^{k-1} \left[\frac{n-j-1}{n-j}\right]^{\delta_{[j]}} \prod_{j=k+1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j]}}. \end{aligned}$$

After some simple algebraic manipulations, we obtain

$$\begin{aligned} \sum_{k=1}^{n} [S_n^{(k)}]^2 = & \, 2 \sum_{1 \le i < r \le n} \varphi(Z_{(i)})\varphi(Z_{(r)})\delta_{[i]}\delta_{[r]}\{A_{ir} + B_{ir} + C_{ir}\} \\ & + \sum_{i=2}^{n-1} \varphi^2(Z_{(i)})\delta_{[i]}\{D_i + E_i\} + \varphi^2(Z_{(1)})\delta_{[1]}D_1 + \varphi^2(Z_{(n)})\delta_{[n]}E_n \\ \equiv & \, \mathrm{I} + \mathrm{II} + \varphi^2(Z_{(1)})\delta_{[1]}D_1 + \varphi^2(Z_{(n)})\delta_{[n]}E_n. \end{aligned}$$

Here, for $1 \le i < r \le n$,

$$A_{ir} = \frac{1}{n-i} \prod_{j=1}^{i-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}} \prod_{j=i}^{r-1} \left[\frac{n-j-1}{n-j}\right]^{\delta_{[j]}},$$

$$B_{ir} = \sum_{k=i+1}^{r-1} \frac{1}{(n-i)(n-r+1)} \prod_{j=1}^{i-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}} \times \prod_{j=i}^{k-1} \left[\frac{n-j-1}{n-j}\right]^{\delta_{[j]}} \prod_{j=k+1}^{r-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j]}},$$

$$C_{ir} = \sum_{k=1}^{i-1} \frac{1}{(n-i+1)(n-r+1)} \prod_{j=1}^{k-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}} \times \prod_{j=k+1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j]}} \prod_{j=i}^{r-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j]}},$$

$$D_i \equiv D_{in} = \frac{1}{n-i} \prod_{j=1}^{i-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}, \qquad 1 \le i \le n-1,$$

and

$$E_i \equiv E_{in} = \sum_{k=1}^{i-1} \frac{1}{(n-i+1)^2} \prod_{j=1}^{k-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}} \times \prod_{j=k+1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j]}}, \qquad 2 \le i \le n.$$

We shall first simplify the sum I. Lemma 2.1 provides a recursion formula in r for the factor $\{\cdots\}$, with i fixed.

LEMMA 2.1. *We have, for* $1 \le i < r < n$,

$$A_{i,r+1} + B_{i,r+1} + C_{i,r+1} = \left(\frac{n-r+1}{n-r}\right)\left(\frac{n-r}{n-r+1}\right)^{\delta_{[r]}} (A_{ir} + B_{ir} + C_{ir}).$$

PROOF. Check that

$$B_{i,r+1} = \frac{n-r+1}{n-r}\left(\frac{n-r}{n-r+1}\right)^{\delta_{[r]}} B_{ir} + \frac{1}{n-r} A_{ir},$$

$$C_{i,r+1} = \frac{n-r+1}{n-r}\left(\frac{n-r}{n-r+1}\right)^{\delta_{[r]}} C_{ir},$$

and

$$A_{i,r+1} = \left(\frac{n-r-1}{n-r}\right)^{\delta_{[r]}} A_{ir}.$$

Since

$$\frac{1}{n-r} + \left[\frac{n-r-1}{n-r}\right]^{\delta_{[r]}} = \frac{n-r+1}{n-r}\left[\frac{n-r}{n-r+1}\right]^{\delta_{[r]}},$$

the conclusion of the lemma follows. □

COROLLARY 2.2. *For $1 \le i < r \le n$, we have*

$$A_{ir} + B_{ir} + C_{ir} = \frac{n-i}{n-r+1} \prod_{j=i+1}^{r-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j]}} \{A_{i,\,i+1} + C_{i,\,i+1}\}.$$

PROOF. Apply Lemma 2.1 and use induction on r. Also note that $B_{i,\,i+1} = 0$. □

For the following text, put

$$A_i := A_{i,\,i+1} = \frac{1}{n-i} \prod_{j=1}^{i-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}} \left[\frac{n-i-1}{n-i}\right]^{\delta_{[i]}}$$

and

$$C_i := C_{i,\,i+1} = \sum_{k=1}^{i-1} \frac{1}{(n-i+1)(n-i)} \prod_{j=1}^{k-1} \left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}} \times \prod_{j=k+1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j]}} \left[\frac{n-i}{n-i+1}\right]^{\delta_{[i]}}.$$

The next lemma provides a useful representation of $A_i + C_i$.

LEMMA 2.3. *We have, for $1 \le i < n$,*

$$A_i + C_i = A_1 \prod_{j=2}^{i} a_j + \sum_{j=1}^{i-1} A_j \frac{\delta_{[j]} - \delta_{[j+1]}}{(n-j-1)^2} \prod_{k=j+2}^{i} a_k,$$

where, for $2 \le k < n$,

$$a_k = \frac{n-k+2}{n-k}\left[\frac{n-k+1}{n-k+2}\right]^{\delta_{[k-1]}} \left[\frac{n-k}{n-k+1}\right]^{\delta_{[k]}}.$$

PROOF. The assertion is trivially true for $i = 1$. The general case follows by induction on i. Actually, we have

$$A_{i+1} = \frac{n-i}{n-i-1}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i]}} \left[\frac{n-i-2}{n-i-1}\right]^{\delta_{[i+1]}} A_i$$

and

$$C_{i+1} = \frac{n-i+1}{n-i-1}\left[\frac{n-i}{n-i+1}\right]^{\delta_{[i]}}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i+1]}} C_i$$
$$+ \frac{1}{(n-i)(n-i-1)}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i+1]}}.$$

We conclude that

$$A_{i+1} + C_{i+1} = a_{i+1}C_i + \left\{\frac{n-i}{n-i-1}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i]}}\left[\frac{n-i-2}{n-i-1}\right]^{\delta_{[i+1]}}\right.$$
$$\left. + \frac{1}{n-i-1}\left[\frac{n-i}{n-i-1}\right]^{\delta_{[i]}}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i+1]}}\right\} A_i.$$

Now, for any choice of $\delta_{[i]}$ and $\delta_{[i+1]}$, the term in brackets equals

$$\frac{n-i+1}{n-i-1}\left[\frac{n-i}{n-i+1}\right]^{\delta_{[i]}}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i+1]}} + \frac{\delta_{[i]}-\delta_{[i+1]}}{(n-i-1)^2}.$$

From this, we get

$$A_{i+1} + C_{i+1} = a_{i+1}(A_i + C_i) + (\delta_{[i]} - \delta_{[i+1]})\frac{A_i}{(n-i-1)^2}.$$

Apply induction on i to complete the proof. □

For further analysis, note that

$$A_1 \prod_{j=2}^{i} a_j = \frac{n}{(n-i)(n-i+1)}\left[\frac{n-2}{n}\right]^{\delta_{[1]}}\prod_{j=2}^{i-1}\left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j]}}\left[\frac{n-i}{n-i+1}\right]^{\delta_{[i]}}$$

and, under $\delta_{[j]} - \delta_{[j+1]} \neq 0$, that is, $\delta_{[j]} + \delta_{[j+1]} = 1$,

$$\frac{A_j}{(n-j-1)^2}\prod_{k=j+2}^{i} a_k = \frac{n-j}{(n-j-1)^2(n-i)(n-i+1)}\left[\frac{n-i}{n-i+1}\right]^{\delta_{[i]}}$$
$$\times \prod_{k=1}^{j}\left[\frac{n-k-1}{n-k}\right]^{2\delta_{[k]}}\prod_{k=j+1}^{i-1}\left[\frac{n-k}{n-k+1}\right]^{2\delta_{[k]}}.$$

Thus, from Corollary 2.2 and Lemma 2.3,

$$\mathrm{I} = 2\sum_{1\le i<r\le n}\varphi(Z_{(i)})\varphi(Z_{(r)})W_i W_r\{\cdots\}_i,$$

with

$$\{\cdots\}_i = n\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}\left[\frac{n-2}{n}\right]^{\delta_{[1]}}$$
$$+ \sum_{j=1}^{i-1}(\delta_{[j]} - \delta_{[j+1]})\frac{n-j}{(n-j-1)^2}\prod_{k=1}^{j}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k]}}.$$

Change the order of summation to get

$$\mathrm{I} = 2n\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}\left[\frac{n-2}{n}\right]^{\delta_{[1]}}\sum_{1\le i<r\le n}\varphi(Z_{(i)})\varphi(Z_{(r)})W_iW_r$$
$$+2\sum_{j=1}^{n-2}(\delta_{[j]}-\delta_{[j+1]})\frac{n-j}{(n-j-1)^2}\prod_{k=1}^{j}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k]}}I_j,$$

where

$$I_j=\sum_{j<i<r}\varphi(Z_{(i)})\varphi(Z_{(r)})W_iW_r.$$

We now rewrite the second sum in the representation of I in such a way that it will be tractable in an asymptotic analysis.

LEMMA 2.4. *We have*

$$\sum_{j=1}^{n-2}(\delta_{[j]}-\delta_{[j+1]})\frac{n-j}{(n-j-1)^2}\prod_{k=1}^{j}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k]}}I_j$$
$$=(\delta_{[1]}-1)I_1\frac{n}{(n-1)^2} \tag{2.1}$$
$$+\sum_{j=1}^{n-2}(\delta_{[j]}-1)b_j\prod_{k=1}^{j-1}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k]}}I_j,$$

where, *as in Section* 1,

$$b_j\equiv b_{jn}:=\frac{1}{(n-j-1)^2}-\frac{1}{(n-j)^2}+\frac{1}{(n-j-1)}-\frac{1}{n-j}.$$

PROOF. Summation by parts readily shows that (2.1) equals

$$\delta_{[1]}\frac{n-1}{(n-2)^2}\left[\frac{(n-2)n}{(n-1)^2}\right]^{2\delta_{[1]}}I_1-\delta_{[n-1]}\frac{2}{1^2}\prod_{k=1}^{n-2}[\cdots]^{2\delta_{[k]}}I_{n-2}$$
$$+\sum_{j=2}^{n-2}\delta_{[j]}\frac{n-j+1}{(n-j)^3}\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}I_j \tag{2.2}$$
$$+\sum_{j=2}^{n-2}\delta_{[j]}\frac{n-j+1}{(n-j)^2}\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}(I_j-I_{j-1}).$$

Note that

$$I_j-I_{j-1}=-\sum_{r=j+1}^{n}\varphi(Z_{j:n})\varphi(Z_{r:n})W_jW_r$$

and W_j vanishes unless $\delta_{[j]} = 1$. Hence the sum (2.2) remains the same if in each summand the factor $\delta_{[j]}$ is replaced by 1. Another summation by parts yields, for (2.2), the expression

$$-\sum_{j=2}^{n-2}\delta_{[j]}\frac{n-j+1}{(n-j)^3}\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}I_j - \sum_{j=2}^{n-2}(1-\delta_{[j]})b_j\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}I_j$$
$$+\frac{2}{1^2}\prod_{k=1}^{n-2}[\cdots]^{2\delta_{[k]}}I_{n-2} - \frac{n-1}{(n-2)^2}[\cdots]^{2\delta_{[1]}}I_1.$$

Since the second term in the expansion of (2.1) also remains unchanged if we replace $\delta_{[n-1]}$ by 1, collecting terms completes the proof of the lemma. □

Obviously

$$(\delta_{[1]}-1)\frac{n}{(n-1)^2}I_1 = (\delta_{[1]}-1)\frac{n}{(n-1)^2}\sum_{1\le i<r\le n}\varphi(Z_{(i)})\varphi(Z_{(r)})W_iW_r,$$

whence, from Lemma 2.4,

$$\mathrm{I} = \left\{n\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}\left[\frac{n-2}{n}\right]^{\delta_{[1]}} + (\delta_{[1]}-1)\frac{n}{(n-1)^2}\right\}\sum_{i\neq r}\varphi(Z_{(i)})\varphi(Z_{(r)})W_iW_r$$
$$+2\sum_{j=1}^{n-2}(\delta_{[j]}-1)b_j\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}I_j. \tag{2.3}$$

We now derive the corresponding representation for II. To this end note that, for $2 \le i \le n-1$,

$$\delta_{[i]}E_i = \delta_{[i]}C_i$$

and

$$\delta_{[i]}D_i = \delta_{[i]}A_i + \delta_{[i]}\frac{1}{(n-i)^2}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}.$$

Hence

$$\mathrm{II} = \sum_{i=2}^{n-1}\varphi^2(Z_{(i)})\delta_{[i]}\{A_i + C_i\}$$
$$+\sum_{i=2}^{n-1}\varphi^2(Z_{(i)})\delta_{[i]}\frac{1}{(n-i)^2}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}.$$

Apply Lemmas 2.3 and 2.4 to show, in complete analogy to the arguments leading to (2.3), that the first sum equals

$$\left\{n\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}\left[\frac{n-2}{n}\right]^{\delta_{[1]}}+(\delta_{[1]}-1)\frac{n}{(n-1)^2}\right\}\sum_{i=2}^{n-1}\varphi^2(Z_{(i)})W_i^2$$
$$+\sum_{j=1}^{n-2}(\delta_{[j]}-1)b_j\prod_{k=1}^{j-1}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k]}}J_j,$$

with

$$J_j=\sum_{i=j+1}^{n-1}\varphi^2(Z_{(i)})W_i^2.$$

Collecting terms we obtain

$$\begin{aligned}\sum_{k=1}^{n}[S_n^{(k)}]^2 &= \left\{n\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}\left[\frac{n-2}{n}\right]^{\delta_{[1]}}+(\delta_{[1]}-1)\frac{n}{(n-1)^2}\right\}S_n^2\\ &\quad-\{\cdots\}\varphi^2(Z_{(1)})W_1^2-\{\cdots\}\varphi^2(Z_{(n)})W_n^2\\ &\quad+\sum_{j=1}^{n-2}(\delta_{[j]}-1)b_j\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}\left[\sum_{i=j+1}^{n}\varphi(Z_{(i)})W_i\right]^2\\ &\quad-\sum_{j=1}^{n-2}(\delta_{[j]}-1)b_j\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}\varphi^2(Z_{(n)})W_n^2\\ &\quad+\varphi^2(Z_{(1)})\delta_{[1]}D_1+\varphi^2(Z_{(n)})\delta_{[n]}E_n\\ &\quad+\sum_{i=2}^{n-1}\varphi^2(Z_{(i)})\delta_{[i]}\frac{1}{(n-i)^2}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}.\end{aligned}\tag{2.4}$$

In the next lemma we obtain a useful representation of E_n to the effect that some of the terms in the last expansion cancel out.

LEMMA 2.5. *We have*

$$\begin{aligned}E_n &= \left\{n\left[\frac{n-2}{n}\right]^{\delta_{[1]}}\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}+\delta_{[1]}\frac{n^2}{(n-1)^3}\right\}\prod_{j=1}^{n-1}\left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j]}}\\ &\quad\times\left\{\frac{1}{2}\delta_{[n-1]}-(1-\delta_{[n-1]})\right\}\prod_{j=1}^{n-2}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}\\ &\quad+\sum_{j=2}^{n-2}\delta_{[j]}\frac{n-j+1}{(n-j)^3}\prod_{k=1}^{j-1}\left[\frac{n-k-1}{n-k}\right]^{2\delta_{[k]}}\prod_{k=j}^{n-1}\left[\frac{n-k}{n-k+1}\right]^{2\delta_{[k]}}.\end{aligned}$$

PROOF. We will rewrite E_n in terms of A_{n-1} and C_{n-1} and then apply Lemma 2.3. To this end observe that

$$E_n = 2\left(\frac{1}{2}\right)^{\delta_{[n-1]}}(A_{n-1} + C_{n-1}) - \prod_{j=1}^{n-2}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}(-1)^{\delta_{[n-1]}}.$$

By Lemma 2.3, the right-hand side becomes, after some by now obvious manipulations,

$$\begin{aligned}
&n\left[\frac{n-2}{n}\right]^{\delta_{[1]}}\left[\frac{n}{n-1}\right]^{2\delta_{[1]}}\prod_{j=1}^{n-1}\left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j]}} \\
&\quad + \sum_{j=1}^{n-2}(\delta_{[j]} - \delta_{[j+1]})\frac{n-j}{(n-j-1)^2}\prod_{k=1}^{j}\left[\frac{n-k-1}{n-k}\right]^{2\delta_{[k]}}\prod_{k=j+1}^{n-1}\left[\frac{n-k}{n-k+1}\right]^{2\delta_{[k]}} \\
&\quad - \prod_{j=1}^{n-2}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}(-1)^{\delta_{[n-1]}}.
\end{aligned}$$

The middle sum is readily seen to be equal to

$$\begin{aligned}
&\delta_{[1]}\frac{n-1}{(n-2)^2}\left(\frac{n-2}{n-1}\right)^{2\delta_{[1]}}\prod_{k=2}^{n-1}\left[\frac{n-k}{n-k+1}\right]^{2\delta_{[k]}} \\
&\quad - \delta_{[n-1]}2\prod_{k=1}^{n-2}\left[\frac{n-k-1}{n-k}\right]^{2\delta_{[k]}}\left[\frac{1}{2}\right]^{2\delta_{[n-1]}} \\
&\quad + \sum_{j=2}^{n-2}\delta_{[j]}\frac{n-j+1}{(n-j)^3}\prod_{k=1}^{j-1}[\cdots]^{2\delta_{[k]}}\prod_{k=j}^{n-1}[\cdots]^{2\delta_{[k]}},
\end{aligned}$$

whence the result. □

PROOF OF THEOREM 1.1. Theorem 1.1 follows from (2.4) and Lemma 2.5 upon collecting and rearranging terms. □

We now study the sequence

$$T_n = (n-1)\sum_{i=1}^{n-1}\delta_{[i:n]}\psi(Z_{i:n})(n-i)^{-2}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n]}},$$

with $\psi = \varphi^2 \geq 0$. Since we will investigate the stochastic behavior of T_n as n varies in time, it will sometimes be necessary to use the notation $Z_{i:n}$ and $\delta_{[i:n]}$ again instead of $Z_{(i)}$ and $\delta_{[i]}$. Put

$$\mathscr{F}_n = \sigma(Z_{i:n}, \delta_{[i:n]},\ 1 \leq i \leq n,\ Z_i, \delta_i,\ i > n).$$

Clearly, T_n is adapted to $\mathscr{F}_n$. Moreover, the $\mathscr{F}_n$s are nonincreasing. In the following text we study the conditional expectation of T_n w.r.t. $\mathscr{F}_{n+1}$. For this, note that (for a continuous H)

$$\begin{aligned}
&\mathbb{E}\left\{\delta_{[i:n]}\psi(Z_{i:n})(n-i)^{-2}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n]}}|\mathscr{F}_{n+1}\right\}\\
&\quad=\mathbb{E}\{\cdots 1_{\{Z_{i:n+1}<Z_{n+1}\}}|\mathscr{F}_{n+1}\}+\sum_{k=1}^{i}\mathbb{E}\{\cdots 1_{\{Z_{n+1}=Z_{k:n+1}\}}|\mathscr{F}_{n+1}\}\\
&\quad=\delta_{[i:n+1]}\psi(Z_{i:n+1})\frac{n-i+1}{(n-i)^2(n+1)}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n+1]}}\\
&\qquad+\delta_{[i+1:n+1]}\psi(Z_{i+1:n+1})\frac{1}{(n-i)^2(n+1)}\\
&\qquad\quad\times\sum_{k=1}^{i}\prod_{j=1}^{k-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n+1]}}\prod_{j=k+1}^{i}\left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j:n+1]}}.
\end{aligned}$$

Conclude that

$$\begin{aligned}
\mathbb{E}\{T_n|\mathscr{F}_{n+1}\}&=(n-1)\sum_{i=1}^{n-1}\delta_{[i:n+1]}\psi(Z_{i:n+1})V_{i,n+1}\\
&\quad+(n-1)\delta_{[n:n+1]}\psi(Z_{n:n+1})\frac{1}{n+1}\sum_{k=1}^{n-1}\prod_{j=1}^{k-1}[\cdots]^{2\delta_{[j:n+1]}}\prod_{j=k+1}^{n-1}[\cdots]^{2\delta_{[j:n+1]}},
\end{aligned}$$

where, for $1\le i\le n-1$,

$$\begin{aligned}
V_{i,n+1}&=\frac{n-i+1}{(n-i)^2(n+1)}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n+1]}}\\
&\quad+\frac{1}{(n-i+1)^2(n+1)}\sum_{k=1}^{i-1}\prod_{j=1}^{k-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n+1]}}\prod_{j=k+1}^{i-1}\left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j:n+1]}}.
\end{aligned}$$

The following lemma will be crucial to show that $(T_n,\mathscr{F}_n)_{n\ge 2}$ (almost) is a reverse-time submartingale.

LEMMA 2.6. *For $1\le i\le n-1$, we have*

$$(n-1)V_{i,n+1}\ge n(n+1-i)^{-2}\prod_{j=1}^{i-1}\left[\frac{n-j}{n-j+1}\right]^{2\delta_{[j:n+1]}}.$$

PROOF. For $i=1$, the inequality reduces to $n/(n-1)(n+1)\ge 1/n$, which is trivially true. Use induction on i for the general case. Assuming that the inequality holds true for i, observe that

$$\begin{aligned}
&(n-1)V_{i+1,\,n+1}\\
&\quad=\frac{(n-1)(n-i)}{(n-i-1)^2(n+1)}\prod_{j=1}^{i}[\cdots]^{2\delta_{[j]}}\\
&\qquad+\frac{n-1}{(n-i)^2(n+1)}\sum_{k=1}^{i}\prod_{j=1}^{k-1}[\cdots]^{2\delta_{[j]}}\prod_{j=k+1}^{i}[\cdots]^{2\delta_{[j]}}\\
&\quad=\frac{n-1}{n+1}\left\{\frac{n-i}{(n-i-1)^2}\left[\frac{n-i-1}{n-i}\right]^{2\delta_{[i]}}+\frac{1}{(n-i)^2}\right\}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j]}}\\
&\qquad+\frac{n-1}{n+1}\frac{1}{(n-i)^2}\left[\frac{n-i}{n-i+1}\right]^{2\delta_{[i]}}\sum_{k=1}^{i-1}\prod_{j=1}^{k-1}[\cdots]^{2\delta_{[j]}}\prod_{j=k+1}^{i-1}[\cdots]^{2\delta_{[j]}}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\left[\frac{n-i}{(n-i-1)^2}\left[\frac{n-i-1}{n-i}\right]^{2\delta_{[i]}}+\frac{1}{(n-i)^2}\right]\frac{(n-i)^2}{n-i+1}\\
&\qquad\geq\left[\frac{n+1-i}{n-i}\right]^2\left[\frac{n-i}{n-i+1}\right]^{2\delta_{[i]}}.
\end{aligned}$$

Actually, both sides equal 1 if $\delta_{[i]}=1$, while for $\delta_{[i]}=0$ the resulting inequality is satisfied if and only if

$$(n-i)^5+(n-i)^2(n-i-1)^2\geq(n-i+1)^3(n-i-1)^2.$$

The difference between the left- and the right-hand side is, however, $3(n-i)^2-(n-i)-1\geq 0$. It follows that

$$(n-1)V_{i+1,\,n+1}\geq(n-1)V_{i,\,n+1}\left[\frac{n+1-i}{n-i}\right]^2\left[\frac{n-i}{n-i+1}\right]^{2\delta_{[i]}}.$$

Now apply the induction hypothesis to complete the proof. □

Neglecting the last sum in the representation of $\mathbb{E}\{T_n|\mathscr{F}_{n+1}\}$ for a moment, Lemma 2.6 would imply

$$\mathbb{E}\{T_n|\mathscr{F}_{n+1}\}\geq T_{n+1},$$

that is, $(T_n,\mathscr{F}_n)_n$ is a reverse-time submartingale. Unfortunately, this last sum may be strictly less than the target value, as is readily seen for the first few n. The following lemma will be needed to show that the submartingale property is only slightly violated.

LEMMA 2.7. *For each choice of labels we have*

$$E_{nn} = \sum_{k=1}^{n-1} \prod_{j=1}^{k-1} \left[\frac{n-j-1}{n-j} \right]^{2\delta_{[j]}} \prod_{j=k+1}^{n-1} \left[\frac{n-j}{n-j+1} \right]^{2\delta_{[j]}}$$
$$\geq (n-1) \prod_{j=1}^{n-1} \left[\frac{n-j}{n-j+1} \right]^{2\delta_{[j]}}.$$

PROOF. The assertion will be shown by induction on n. For $n = 1$ both sides equal zero. Assume now that the lemma holds true for sample size n. We then obtain, for $n + 1$,

$$\sum_{k=1}^{n} \prod_{j=1}^{k-1} \left[\frac{n-j}{n-j+1} \right]^{2\delta_{[j]}} \prod_{j=k+1}^{n} \left[\frac{n-j+1}{n-j+2} \right]^{2\delta_{[j]}}$$
$$= \prod_{j=2}^{n} \left[\frac{n-j+1}{n-j+2} \right]^{2\delta_{[j]}} + \left(\frac{n-1}{n} \right)^{2\delta_{[1]}} \sum_{k=2}^{n} \prod_{j=2}^{k-1} [\cdots]^{2\delta_{[j]}} \prod_{j=k+1}^{n} [\cdots]^{2\delta_{[j]}}.$$

After a proper index transformation, the last term becomes

$$\prod_{j=2}^{n} \left[\frac{n-j+1}{n-j+2} \right]^{2\delta_{[j]}}$$
$$+ \left(\frac{n-1}{n} \right)^{2\delta_{[1]}} \sum_{k=1}^{n-1} \prod_{j=1}^{k-1} \left[\frac{n-j-1}{n-j} \right]^{2\delta_{[j+1]}} \prod_{j=k+1}^{n-1} \left[\frac{n-j}{n-j+1} \right]^{2\delta_{[j+1]}}.$$

Application of the induction hypothesis yields the lower bound

$$\prod_{j=2}^{n} \left[\frac{n-j+1}{n-j+2} \right]^{2\delta_{[j]}} + \left(\frac{n-1}{n} \right)^{2\delta_{[1]}} (n-1) \prod_{j=1}^{n-1} \left[\frac{n-j}{n-j+1} \right]^{2\delta_{[j+1]}}$$
$$= \left\{ 1 + (n-1) \left(\frac{n-1}{n} \right)^{2\delta_{[1]}} \right\} \prod_{j=2}^{n} \left[\frac{n-j+1}{n-j+2} \right]^{2\delta_{[j]}}.$$

Check that the term in brackets always exceeds

$$n \left(\frac{n}{n+1} \right)^{2\delta_{[1]}}.$$

The proof is complete. □

COROLLARY 2.8. *We have*

$$\mathbb{E}\{T_n | \mathscr{F}_{n+1}\} \geq T_{n+1} - 3\delta_{[n:n+1]} \psi(Z_{n:n+1}) \prod_{j=1}^{n-1} \left[\frac{n-j}{n-j+1} \right]^{2\delta_{[j:n+1]}}.$$

PROOF. Corollary 2.8 follows from Lemmas 2.6 and 2.7 upon noting that

$$\frac{(n-1)^2}{n+1} \ge n-3. \qquad \square$$

Next we study the sequence $R_n\varphi^2(Z_{n:n})W_{nn}^2$. It is readily seen that $R_n = O(1)$, so that

$$nR_n\varphi^2(Z_{n:n})W_{nn}^2 = O[n\varphi^2(Z_{n:n})W_{nn}^2]. \tag{2.5}$$

LEMMA 2.9. *With the same filtration as before*,

$$\begin{aligned}\mathbb{E}\{n\varphi^2(Z_{n:n})W_{nn}^2|\mathscr{F}_{n+1}\} \ge{}& (n+1)\varphi^2(Z_{n+1:n+1})W_{n+1,\,n+1}^2\\ &- 2\varphi^2(Z_{n+1:n+1})W_{n+1,\,n+1}^2.\end{aligned}$$

PROOF. As in the discussion of T_n, we obtain

$$\begin{aligned}&\mathbb{E}\{\cdots|\mathscr{F}_{n+1}\}\\ &\ge \frac{n}{n+1}\varphi^2(Z_{n+1:n+1})\delta_{[n+1:n+1]}\sum_{k=1}^{n}\prod_{i=1}^{k-1}\left[\frac{n-i}{n-i+1}\right]^{2\delta_{[i:n+1]}}\prod_{i=k+1}^{n}\left[\frac{n-i+1}{n-i+2}\right]^{2\delta_{[i:n+1]}}\\ &= \frac{n}{n+1}\varphi^2(Z_{n+1:n+1})\delta_{[n+1:n+1]}E_{n+1,\,n+1}\end{aligned}$$

and, therefore, by Lemma 2.7,

$$\begin{aligned}\mathbb{E}\{\cdots|\mathscr{F}_{n+1}\} &\ge \frac{n^2}{n+1}\varphi^2(Z_{n+1:n+1})W_{n+1,\,n+1}^2\\ &\ge (n+1)\varphi^2(Z_{n+1:n+1})W_{n+1,\,n+1}^2 - 2\varphi^2(Z_{n+1:n+1})W_{n+1,\,n+1}^2. \qquad \square\end{aligned}$$

By Corollary 2.8 and Lemma 2.9, T_n and $n\varphi^2(Z_{n:n})W_{nn}^2$, $n \ge 2$, are (slightly disturbed) reverse-time submartingales. A variant of the martingale convergence theorem (cf. Lemma A in the Appendix) will guarantee that both sequences converge with probability 1. Lemma 2.10 will be needed in order to justify the applicability of that lemma.

LEMMA 2.10. *Under* (1.6), *we have*:

(i) $n\mathbb{E}\{\varphi^2(Z_{n:n})W_{nn}^2\} \to 0$.
(ii) $\sum_{n\ge 2}\mathbb{E}\{\varphi^2(Z_{n:n})W_{nn}^2\} < \infty$.

PROOF. The method of proof is an adaptation of arguments elaborated, in a different context, in Stute (1994). First, by incorporating the usual quantile transformation, we may assume without loss of generality, that the Zs are uniformly distributed on $[0, 1]$. By the distributional theory of uniform order statistics and Lemma 2.1 in Stute and Wang (1993), we then obtain

$$\begin{aligned}
&\mathbb{E}\left\{\varphi^2(Z_{n:n})\delta_{[n:n]}\prod_{i=1}^{n-1}\left[\frac{n-i}{n-i+1}\right]^{2\delta_{[i:n]}}\right\}\\
&\quad = n\int_0^1 \varphi^2(u)m(u)\mathbb{E}\left\{1_{\{Z_{n-1:n-1}\le u\}}\prod_{i=1}^{n-1}\left[\frac{n-i}{n-i+1}\right]^{2\delta_{[i:n-1]}}\right\}du\\
&\quad = \frac{1}{n}\int_0^1 \varphi^2(u)m(u)\mathbb{E}\left\{1_{\{Z_{n-1:n-1}\le u\}}\prod_{i=1}^{n-1}\left[1+\frac{1-m(Z_{i:n-1})}{n-i}\right]^{21_{\{Z_{i:n-1}<u\}}}\right\}du.
\end{aligned}$$

For ease of notation we restrict ourselves to the case $Z = \min(X, Y)$ with $Y \sim G$. By Lemma 2.6 of Stute and Wang (1993), the inner product converges to $(1-G(u))^{-2}$ with probability 1 for each $0 < u < 1$. Apply Cauchy–Schwarz to bound the $\mathbb{E}$-term and use a dominated convergence argument to show that finally the above term is bounded from above, up to a constant factor, by

$$n^{-1}\int \varphi^2(u)m(u)(1-G(u))^{-2}u^{(n-1)/2}\,du = n^{-1}\int \frac{\varphi^2(x)}{1-G(x)}[H(x)]^{(n-1)/2}\tilde{F}(dx).$$

Details are omitted. Another application of dominated convergence yields (i). Item (ii) is also standard. □

COROLLARY 2.11. *Under* (1.6),

$$nR_n\varphi^2(Z_{n:n})W_{nn}^2 \to 0 \quad \textit{with probability } 1.$$

PROOF. The proof follows from an application of (2.5), Lemma A and Lemmas 2.9 and 2.10 upon noting that, by the Hewitt–Savage 0–1 law,

$$\lim_{n\to\infty} n\varphi^2(Z_{n:n})W_{nn}^2 = \lim_{n\to\infty} n\mathbb{E}\{\varphi^2(Z_{n:n})W_{nn}^2\} = 0. \qquad \square$$

REMARK. When $Z = \min(X, Y)$ and (1.8) holds, convergence in probability rather than convergence with probability 1 may also be obtained in the following way: Conclude from the Stute and Wang (1993) SLLN that, under (1.8),

$$\frac{\varphi^2(Z_{n:n})W_{nn}}{(1-G(Z_{n:n}))} \to 0 \quad \text{with probability 1.}$$

Then apply Gill's inequality [cf. Shorack and Wellner (1986), page 317] to the Kaplan–Meier estimator of G to get stochastic boundedness of

$$\frac{1-G(Z_{n:n})}{1-\hat{G}_n(Z_{n:n}-)} \quad \text{as } n\to\infty.$$

Finally

$$n(1-\hat{G}_n(Z_{n:n}-))W_{nn} = n(1-H_n(Z_{n:n}-)) = O_{\mathbb{P}}(1).$$

Note that this separation argument does not yield the assertion of Corollary 2.11, since, with probability 1, $(1-G)/(1-\hat{G}_n)$ is unbounded.

We now determine the limit T of T_n, as $n \to \infty$. Again, by the Hewitt–Savage 0–1 law, it will be constant. This constant will be part of the asymptotic variance σ_1^2. The following lemma provides a formula for the expectation of T_n. It corresponds to Lemma 2.4 in Stute and Wang (1993), where the first moment of a Kaplan–Meier integral was dealt with. Since we prefer to state our result for Zs without ties, continuity of H will be assumed. In the final proof of Theorem 1.2, this assumption may be dropped again.

LEMMA 2.12. *Assume that H is continuous. Then*

$$\mathbb{E}T_n = \frac{n}{n-1}\mathbb{E}[\varphi^2(Z)m(Z)g_{n-1}(Z)],$$

where

$$g_n(t) = \mathbb{E}\varphi_n(t)$$

and

$$\varphi_n(t) = \prod_{i=1}^{n-1}\left[1 + \frac{2(1-m(Z_{i:n}))}{n-i} + \frac{1-m(Z_{i:n})}{(n-i)^2}\right]^{1_{\{Z_{i:n}<t\}}} 1_{\{t<Z_{n:n}\}}.$$

PROOF. Let R_{jn} denote the rank of Z_j among $Z_1, \dots, Z_n$. By Lemma 2.1 of Stute and Wang (1993),

$$\begin{aligned}
\mathbb{E}\{T_n\} &= (n-1)\mathbb{E}\left\{\sum_{i=1}^{n-1}\delta_{[i:n]}\psi(Z_{i:n})(n-i)^{-2}\prod_{j=1}^{i-1}\left[1-\frac{\delta_{[j:n]}}{n-j}\right]^2\right\} \\
&= (n-1)\mathbb{E}\left\{\sum_{i=1}^{n-1}\delta_{[i:n]}\psi(Z_{i:n})(n-i)^{-2}\prod_{j=1}^{i-1}\left[1-\frac{2\delta_{[j:n]}}{n-j}+\frac{\delta_{[j:n]}}{(n-j)^2}\right]\right\} \\
&= (n-1)\mathbb{E}\left\{\sum_{i=1}^{n-1}m(Z_{i:n})\psi(Z_{i:n})(n-i)^{-2}\prod_{j=1}^{i-1}\left[1-\frac{2m(Z_{j:n})}{n-j}+\frac{m(Z_{j:n})}{(n-j)^2}\right]\right\}.
\end{aligned}$$

After some elementary algebraic manipulations, the last term becomes

$$\begin{aligned}
&\mathbb{E}\left\{\sum_{i=1}^{n-1}\frac{m(Z_{i:n})\psi(Z_{i:n})}{n-1}\prod_{j=1}^{i-1}\left[1+\frac{2(1-m(Z_{j:n}))}{n-j-1}+\frac{1-m(Z_{j:n})}{(n-j-1)^2}\right]\right\} \\
&\quad = \frac{n}{n-1}\mathbb{E}\Bigg[m(Z_1)\psi(Z_1)1_{\{Z_1<Z_{n:n}\}} \\
&\qquad\qquad \times \prod_{j=1}^{n}\left[1+\frac{2(1-m(Z_j))}{n-R_{jn}-1}+\frac{1-m(Z_j)}{(n-R_{jn}-1)^2}\right]^{1_{\{Z_j<Z_1\}}}\Bigg].
\end{aligned}$$

Since $Z_1 < Z_{n:n}$ if and only if $Z_1 < Z_{n-1:n-1}$, the largest order statistic in the sample $Z_2, \dots, Z_n$, and (in an obvious notation) $R_{jn} = R_{j,n-1}$ on $\{Z_j < Z_1\}$, $2 \le j \le n$, the assertion of the lemma follows by first conditioning on $Z_1 = t$ and then integrating out. □

For further analysis of $\varphi_n(t)$, fix t such that $H(t) < 1$. Then there exists a small $\varepsilon > 0$ and a subset Ω_0 of Ω with $\mathbb{P}(\Omega_0) = 1$ such that $Z_{[n(1-\varepsilon)]}(\omega) > t$ for all $\omega \in \Omega_0$ and $n \ge n_0(\omega)$. In such a situation, we may restrict the multiplication in the definition of φ_n from 1 to $[n(1-\varepsilon)]$. Conclude that

$$\varphi_n(t) - \varphi_n^0(t) \longrightarrow 0 \quad \text{as } n \to \infty \text{ with probability 1,}$$

where

$$\varphi_n^0(t) = \prod_{i=1}^{n-1}\left[1 + \frac{2(1 - m(Z_{i:n}))}{n-i}\right]^{1_{\{Z_{i:n} < t\}}}.$$

LEMMA 2.13. *With probability* 1, *for each* $t < \tau_H$,

$$\lim_{n\to\infty} \varphi_n^0(t) = \exp\left\{\int_{-\infty}^{t} \frac{2(1 - m(x))}{1 - H(x)} H(dx)\right\}.$$

The proof is similar to Lemma 2.6 in Stute and Wang (1993).

COROLLARY 2.14. *For a continuous* H *we have*, *under* (1.6), *with probability* 1,

$$\lim_{n\to\infty} T_n = \lim_{n\to\infty} \mathbb{E}T_n = \int \varphi^2(x)\gamma_0^2(x)m(x)H(dx).$$

PROOF. The proof is similar to that of Corollary 2.11. Just replace Lemma 2.9 by Corollary 2.8 and apply Lemmas 2.12 and 2.13. □

Note that the last integral is an essential part of σ_1^2.

PROOF OF THEOREM 1.2. We shall first give the proof for a continuous H when φ satisfies the condition

(c) φ vanishes right of x_0 for some $x_0 < \tau_H$.

Since $Z_{n:n} \to \tau_H$ with probability 1, $\varphi(Z_{n:n}) = 0$ eventually with probability 1 so that $\bar{S}_n = S_n$ and $\widehat{\mathrm{Var}}^*(\mathrm{Jack}) = \widehat{\mathrm{Var}}(\mathrm{Jack})$. Conclude that, by Theorem 1.1,

$$
\begin{aligned}
n\,\widehat{\mathrm{Var}}(\mathrm{Jack})& \\
&= (n-1)\sum_{k=1}^{n}[S_n^{(k)}]^2 - n(n-1)S_n^2 \\
&= \left\{(n-1)n\left[\frac{n}{n-1}\right]^{2\delta_{[1:n]}}\left[\frac{n-2}{n}\right]^{\delta_{[1:n]}} + (\delta_{[1:n]}-1)\frac{n}{n-1} - n(n-1)\right\}S_n^2 \\
&\quad + (n-1)\sum_{j=1}^{n-2}(\delta_{[j:n]}-1)b_j\prod_{k=1}^{j-1}\left[\frac{(n-k-1)(n-k+1)}{(n-k)(n-k)}\right]^{2\delta_{[k:n]}} \\
&\qquad \times\left[\sum_{i=j+1}^{n}\varphi(Z_{i:n})W_{in}\right]^2 \\
&\quad + (n-1)\sum_{i=1}^{n-1}\varphi^2(Z_{i:n})\delta_{[i:n]}\frac{1}{(n-i)^2}\prod_{j=1}^{i-1}\left[\frac{n-j-1}{n-j}\right]^{2\delta_{[j:n]}}.
\end{aligned}
\tag{2.6}
$$

Check that $\{\cdots\}$, for each choice of $\delta_{[1:n]}$, equals $-n/(n-1)$. Thus, by the SLLN for Kaplan–Meier integrals, the first term converges to $-S^2$ with probability 1 as $n \to \infty$. The last sum equals T_n, so that Corollary 2.14 applies. Finally, as to the middle term, we may restrict, under (c), summation to $1 \le j \le n(1-\varepsilon)$ for some appropriate $0 < \varepsilon < 1$, at least for all large n. It is then easy to see that the sum is, with probability 1, asymptotically equivalent to

$$
n\sum_{j=1}^{n}(\delta_{[j:n]}-1)(n-j-1)^{-2}\left[\sum_{i=j+1}^{n}\varphi(Z_{i:n})W_{in}\right]^2.
$$

This in turn may be rewritten as

$$
-\int\frac{1}{[1-H_n(x-)]^2}\left[\int_{\{y>x\}}\varphi(y)\hat{F}_n(dy)\right]^2\tilde{H}_n^0(dx),
$$

where

$$
H_n(x) = \frac{1}{n}\sum_{i=1}^{n}1_{\{Z_i\le x\}} \quad\text{and}\quad \tilde{H}_n^0(x) = \frac{1}{n}\sum_{i=1}^{n}1_{\{Z_i\le x,\,\delta_i=0\}}.
$$

For each fixed x, the SLLN for Kaplan–Meier integrals yields, with probability 1,

$$
\lim_{n\to\infty}\int_{\{y>x\}}\varphi(y)\hat{F}_n(dy) = \int_{\{y>x\}}\varphi(y)\gamma_0(y)m(y)H(dy).
\tag{2.7}
$$

A standard Glivenko–Cantelli argument may be applied to show that the convergence in (2.7) is uniform in x. By (c) we may restrict integration to x's which are bounded away from τ_H. In such a situation the factor $[1-H_n(x-)]^{-2}$ also creates no problems, so that in summary, with probability 1,

$$\lim_{n\to\infty}\int \frac{1}{[1-H_n(x-)]^2}\left[\int_{\{y>x\}}\varphi(y)\hat{F}_n(dy)\right]^2 \tilde{H}_n^0(dx)=\int \gamma_1^2(x)(1-m(x))H(dx).$$

This completes the proof of Theorem 1.2 when H is continuous and (c) holds. To deal with a general φ, note that, by Corollary 2.11,

$$nR_n\varphi^2(Z_{n:n})W_{nn}^2 \to 0 \quad \text{with probability 1.}$$

If $(*)$ applies, this remainder term is zero anyway. Therefore, it remains to study (2.6) with S_n^* and $\delta_{[n:n]}^*=0$ rather than S_n and $\delta_{[n:n]}$. Since along with S_n, S_n^* also converges to S and the last sum in (2.6) does not contain $\delta_{[n:n]}$ at all, only the sum containing the b_js needs some special care. This term, however, is always less than or equal to zero; it vanishes if there is no censorship. Since $\widehat{\mathrm{Var}}^*$(Jack) is nonnegative, we may conclude that $n\widehat{\mathrm{Var}}^*$(Jack) is, with probability 1 (uniformly in n), continuous in φ, whenever φ satisfies the integrability assumption (1.6). Because the functions satisfying the support condition (c) are dense with respect to the L^2-norm, this completes the proof of Theorem 1.2 when H is continuous. Finally, the general case is obtained by a quantile transformation. See, for example, Stute and Wang (1993) and Stute (1995). □

REMARK. We conclude this section by discussing a possible inconsistency of $\widehat{\mathrm{Var}}$(Jack). Whenever $\varphi(Z_{n:n})\neq 0$,

$$n\widehat{\mathrm{Var}}(\mathrm{Jack})=(n-1)\sum_{k=1}^{n}[S_n^{(k)}]^2-n(n-1)\bar{S}_n^2,$$

so that, in general, the expansion (2.6) contains an additional summand

$$+\,2(n-1)S_n\varphi(Z_{n:n})\delta_{[n:n]}(1-\delta_{[n-1:n]})\prod_{i=1}^{n-2}\left[\frac{n-i-1}{n-i}\right]^{\delta_{[i:n]}}.$$

Under $(*)$, this term is not negligible; on the contrary, it may dominate the other terms appearing in (2.6).

APPENDIX

The following lemma constitutes an adaptation of a result by Van Ryzin (1969) to the case of reverse-time submartingales. Let $\mathscr{F}_n$, $n\geq 1$, be a nonincreasing sequence of sub-σ-fields such that M_n is integrable and adapted to $\mathscr{F}_n$. Also, let M_n' be another integrable random variable adapted to $\mathscr{F}_n$.

LEMMA A. *Assume*:

(i) $\mathbb{E}(M_n|\mathscr{F}_{n+1})\geq M_{n+1}+M_{n+1}'$.

(ii) $\sum_{n=1}^{\infty} \mathbb{E}(|M'_n|) < \infty$.

(iii) $\inf_n \mathbb{E}(M_n) > -\infty$.

Then M_n converges almost surely and in the mean to a finite limit.

PROOF. Put

$$T_n = M_n + \mathbb{E}\biggl(\sum_{k=1}^{n} M'_k|\mathscr{F}_n\biggr).$$

From (i) and by monotonicity of $(\mathscr{F}_n)_n$,

$$\begin{aligned}\mathbb{E}(T_n|\mathscr{F}_{n+1}) &\geq M_{n+1} + M'_{n+1} + \mathbb{E}\biggl(\sum_{k=1}^{n} M'_k|\mathscr{F}_{n+1}\biggr)\\ &= T_{n+1},\end{aligned}$$

that is, $(T_n, \mathscr{F}_n)_n$ is a reverse submartingale. Moreover by (iii),

$$\begin{aligned}\mathbb{E}(T_n) &\geq \mathbb{E}(M_n) - \sum_{k=1}^{n} \mathbb{E}|M'_k|\\ &\geq \inf_k \mathbb{E}(M_k) - \sum_{k=1}^{\infty} \mathbb{E}|M'_k| > -\infty.\end{aligned}$$

Hence, $(T_n)_n$ converges almost surely and in the mean to a finite limit. However, $\sum_{k=1}^{n} M'_k \to M$ almost surely and in the mean. Hence, by the reverse martingale version of the corollary of Chow and Teicher [(1978), page 233],

$$\mathbb{E}\biggl(\sum_{k=1}^{n} M'_k|\mathscr{F}_n\biggr) \to \mathbb{E}(M|\mathscr{F}_\infty)$$

almost surely and in the mean, where $\mathscr{F}_\infty = \bigcap_{n\geq 1} \mathscr{F}_n$. This proves the lemma. □

Acknowledgments. Thanks to my students Silke Thies and Sandra Ziegler for their assistance in a simulation study. Without their help the special effects in the extreme right tails probably would not have been discovered.

REFERENCES

ANDERSEN, P. K., BORGAN, Ø., GILL, R. D. and KEIDING, N. (1993). *Statistical Models Based on Counting Processes.* Springer, New York.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life-table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory. Independence, Interchangeability, Martingales.* Springer, Berlin.

EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap.* Chapman and Hall, New York.

GAVER, D. P. and MILLER, R. G. (1983). Jackknifing the Kaplan–Meier survival estimator for censored data: simulation results and asymptotic analysis. *Comm. Statist. Theory Methods* **12** 1701–1718.

GRAY, H. L. and SCHUCANY, W. R. (1972). *The Generalized Jackknife Statistics.* Dekker, New York.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

KLEIN, J. P. (1991). Small sample moments of some estimators of the variance of the Kaplan–Meier and Nelson–Aalen estimators. *Scand. J. Statist.* **18** 333–340.

QUENOUILLE, M. H. (1956). Notes on bias in estimation. *Biometrika* **43** 353–360.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

STUTE, W. (1994). The bias of Kaplan–Meier integrals. *Scand. J. Statist.* **21** 475–484.

STUTE, W. (1995). The central limit theorem under random censorship. *Ann. Statist.* **23** 422–439.

STUTE, W. and WANG, J.-L. (1993). The strong law under random censorship. *Ann. Statist.* **21** 1591–1607.

STUTE, W. and WANG, J.-L. (1994). The jackknife estimate of a Kaplan–Meier integral. *Biometrika* **81** 602–606.

TSIATIS, A. A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci. U.S.A.* **72** 20–22.

TUKEY, J. W. (1958). Bias and confidence in not quite large samples (abstract). *Ann. Math. Statist.* **29** 614.

VAN RYZIN, J. (1969). On strong consistency of density estimates. *Ann. Math. Statist.* **40** 1765–1772.

MATHEMATISCHES INSTITUT
UNIVERSITÄT GIESSEN
ARNDTSTRASSE 2
D-35392 GIESSEN
GERMANY
E-MAIL: winfried.stute@math.uni-giessen.de