The Annals of Statistics
1995, Vol. 23, No. 2, 422–439

THE CENTRAL LIMIT THEOREM UNDER RANDOM CENSORSHIP[1]

BY WINFRIED STUTE

Universität Giessen

Let $\hat{F}_n$ be the Kaplan–Meier estimator of a distribution function F computed from randomly censored data. We show that under optimal integrability assumptions on a function φ, the Kaplan–Meier integral $\int \varphi \, d\hat{F}_n$, when properly standardized, is asymptotically normal.

1. Introduction and main results. Assume that $X_1, \dots, X_n, \dots$ are independent and identically distributed (i.i.d.) random variables on the real line, defined over some probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Let F denote their common distribution function (d.f.). Under $\int x^2 F(dx) < \infty$, put

$$\mu = \int x F(dx) \quad \text{and} \quad \sigma^2 = \operatorname{Var} X_1 = \int x^2 F(dx) - \mu^2.$$

The CLT then states that for each real a,

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^{n} (X_i - \mu) \le a \right) = \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} \exp\left(-\frac{z^2}{2} \right) dz. \tag{1.1}$$

In applications (1.1) is often needed for proper transformations of the X's rather than the X's themselves. So, let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be any measurable function such that $\int \varphi^2 \, dF < \infty$. Since along with $X_1, \dots, X_n, \dots$ also $\varphi(X_1), \dots, \varphi(X_n), \dots$ are i.i.d., (1.1) may be applied to the transformed random variables as well. Introducing the empirical d.f.'s

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \le x\}}, \qquad x \in \mathbb{R},$$

Equation (1.1) becomes

$$n^{1/2} \int \varphi \, d(F_n - F) \to \mathscr{N}(0, \sigma^2) \quad \text{in distribution,} \tag{1.2}$$

where now

$$\sigma^2 = \int \varphi^2 \, dF - \left(\int \varphi \, dF \right)^2.$$

Received April 1993; revised August 1994.

[1]Supported by the Deutsche Forschungsgemeinschaft. This work was done while the author was on leave at the University of California, Davis.

1991 *subject classifications*. Primary 60F15; secondary 60G42, 62G30.

Key words and phrases. Censored data, CLT, Kaplan–Meier integral.

The objective of the present paper is to extend the CLT in full generality, that is, for a general transformation φ, to the random censorship model. Censorship typically comes up in the analysis of lifetime data. Rather than $X_1, \dots, X_n$, the variables of interest, one observes

$$Z_i = \min(X_i, Y_i) \quad \text{and} \quad \delta_i = 1_{\{X_i \le Y_i\}}, \qquad 1 \le i \le n,$$

where $Y_1, \dots, Y_n$ is another i.i.d. sequence from some (censoring) d.f. G being also independent of the X's; δ_i indicates whether X_i has been observed or not. Clearly, when the X's are not all available, the CLT (1.1) is of no help, for example, to establish confidence intervals for the "mean lifetime" $\mu = \int xF(dx)$. In view of (1.2), however, an alternative approach would be to replace F_n by any efficiently chosen substitute for F_n which is computable from (Z_i, δ_i), $1 \le i \le n$. Now, it is well known that the nonparametric maximum likelihood estimator of F is given by the time-honoured Kaplan–Meier (1958) product-limit estimator (PLE) defined by

$$(1.3) \qquad 1 - \hat{F}_n(x) = \prod_{i=1}^{n} \left[1 - \frac{\delta_{[i;n]}}{n - i + 1}\right]^{1_{\{Z_{i:n} \le x\}}}.$$

Here, $Z_{1:n} \le \cdots \le Z_{n:n}$ are the ordered Z-values, where ties within lifetimes or within censoring times are ordered arbitrarily and ties among lifetimes and censoring times are treated as if the former precedes the latter. $\delta_{[i:n]}$ is the concomitant of the ith order statistic, that is, $\delta_{[i:n]} = \delta_j$ if $Z_{i:n} = Z_j$. As before, let φ be any F square integrable function. Put

$$S_n = \int \varphi \, d\hat{F}_n.$$

It is easily seen from (1.3) that

$$S_n = \sum_{i=1}^{n} W_{in} \varphi(Z_{i:n}),$$

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

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

is the mass attached to the ith order statistic $Z_{i:n}$ under $\hat{F}_n$. With no censoring present, all δ's equal 1 so that each W collapses to $1/n$. Since also $Z_i = X_i$ we are back to the sample mean. Under censoring, however, S_n becomes a function of the Z order statistics properly weighted by the random W's.

As a main consequence of our Theorem 1.1 we shall obtain (1.2) with $\hat{F}_n$ in place of F_n. This is the CLT under random censoring. Of course, due to censoring effects, some changes for σ^2 will be needed. Also it may happen that the limit of the Kaplan–Meier integrals $\int \varphi \, d\hat{F}_n$ is no longer $\int \varphi \, dF$. Actually, we shall prove more than asymptotic normality. Theorem 1.1 provides a representation of a Kaplan–Meier integral as a sum of i.i.d. random variables plus a remainder. Such representations are extremely useful in applications and give more insight than just distributional convergence if one wants to study the joint distribution of (several) Kaplan–Meier integrals and other statistics of interest. The results are stated under minimal integrability assumptions on φ. If there is no censoring, they will reduce to the familiar condition $\int \varphi^2 \, dF < \infty$. The desired asymptotic normality is stated in Corollary 1.2. Also no continuity assumptions on F or G will be necessary. To formulate our result, some additional notation will be required. Let H denote the d.f. of the observed Z's. By independence of X_1 and Y_1,

$$(1 - H) = (1 - F)(1 - G).$$

Write $F\{a\} = F(a) - F(a-)$. Let A be the set of all atoms of H, possibly empty. Under $\int |\varphi| \, dF < \infty$ it was shown in Stute and Wang (1993), that with probability 1 and in the mean,

$$\lim_{n\to\infty} \int \varphi \, d\hat{F}_n = \int_{\{x < \tau_H\}} \varphi(x) F(dx) + 1_{\{\tau_H \in A\}} \varphi(\tau_H) F\{\tau_H\}. \tag{1.4}$$

Here

$$\tau_H = \inf\{x \colon H(x) = 1\} \le \infty$$

is the least upper bound for the support of H. That paper also contained a detailed discussion of various situations in which the right-hand side of (1.4) equals $\int \varphi \, dF$. For the general case, introduce the (sub-) d.f.

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

Then the right-hand side of (1.4) becomes $\int \varphi \, d\tilde{F}$. So the proper extension of (1.2) would incorporate

$$\int \varphi \, d\big(\hat{F}_n - \tilde{F}\big).$$

An important role in our analysis of $\int \varphi d(\hat{F}_n - \tilde{F})$ will be played by the subdistribution functions

$$\tilde{H}^0(z) = \mathbb{P}(Z \le z, \delta = 0) = \int_{-\infty}^{z} (1 - F(y)) G(dy)$$

and

$$\tilde{H}^1(z) = \mathbb{P}(Z \le z, \delta = 1) = \int_{-\infty}^{z} (1 - G(y-)) F(dy), \qquad z \in \mathbb{R}.$$

Define

$$\gamma_0(x) = \exp\left\{\int_{-\infty}^{x-} \frac{\tilde{H}^0(dz)}{1 - H(z)}\right\},$$

$$\gamma_1(x) = \frac{1}{1 - H(x)} \int 1_{\{x < w\}} \varphi(w)\gamma_0(w)\tilde{H}^1(dw)$$

and

$$\gamma_2(x) = \int\int \frac{1_{\{v < x,\, v < w\}}\varphi(w)\gamma_0(w)}{[1 - H(v)]^2} \tilde{H}^0(dv)\tilde{H}^1(dw).$$

In the definition of γ_0 (and also in the proof section), $\int_{-\infty}^{x-}$ denotes integration on $(-\infty, x)$. The following assumptions will be needed in Theorem 1.1:

$$\int \varphi^2(x)\gamma_0^2(x)\tilde{H}^1(dx) = \int [\varphi(Z)\gamma_0(Z)\delta]^2 \, d\mathbb{P} < \infty \tag{1.5}$$

and

$$\int |\varphi(x)| C^{1/2}(x)\tilde{F}(dx) < \infty. \tag{1.6}$$

Here

$$C(x) = \int_{-\infty}^{x-} \frac{G(dy)}{[1 - H(y)][1 - G(y)]}.$$

Condition (1.5) is the properly modified "second moment" (or variance) assumption on φ, while (1.6) only incorporates the "first φ-moment." It is mainly to control the bias of $\int \varphi \, d\hat{F}_n$, which is a function of φ rather than φ^2. Stute (1994a) gives a detailed account of this issue. Among other things, it was shown there that though the bias tends to zero, the rate of convergence may be worse than $n^{-1/2}$. Hence for the general situation considered in this paper, (1.6) cannot be dispensed with. This does not mean that for a particular φ one might have in mind, (1.6) is implied by (1.5), or vice versa. A detailed discussion of $\gamma_0 - \gamma_2$ as well as of (1.5) and (1.6) will be postponed until the end of this section. The function $C(x)$ comes from the variance of a process (evaluated at x) related to but not identical to the cumulative hazard function of the censored data. In particular, this process will be a stochastic integral such that its integrand is not predictable.

THEOREM 1.1. *Under* (1.5) *and* (1.6), *we have*

$$\begin{aligned} \int \varphi \, d\hat{F}_n = n^{-1} \sum_{i=1}^{n} \varphi(Z_i)\gamma_0(Z_i)\delta_i + n^{-1} \sum_{i=1}^{n} \gamma_1(Z_i)(1 - \delta_i) \\ - n^{-1} \sum_{i=1}^{n} \gamma_2(Z_i) + R_n, \end{aligned} \tag{1.7}$$

where

$$|R_n| = o_{\mathbb{P}}(n^{-1/2}).$$

REMARK. It is easily seen that in (1.7) the summands of the first sum have expectation $\int \varphi \, d\tilde{F}$, while the summands of the second and third sum have identical expectations. In other words, (1.7) yields a representation

$$\int \varphi \, d\big(\hat{F}_n - \tilde{F}\big) = n^{-1} \sum_{i=1}^{n} U_i + R_n, \tag{1.8}$$

where the U's are i.i.d. with mean zero. Put $\sigma^2 = \operatorname{Var} U_1$.

COROLLARY 1.2. *Under* (1.5) *and* (1.6),

$$n^{1/2} \int \varphi \, d\big(\hat{F}_n - \tilde{F}\big) \to \mathscr{N}(0, \sigma^2) \quad \textit{in distribution}.$$

Note that

$$\sigma^2 = \operatorname{Var}\{\varphi(Z)\gamma_0(Z)\delta + \gamma_1(Z)(1-\delta) - \gamma_2(Z)\}.$$

So far asymptotic normality has been mainly established for indicators $\varphi = 1_{(-\infty, x]}$, $x < \tau_H$. See, for example, Breslow and Crowley (1974), Lo and Singh (1986) and Major and Rejtő (1988). Upon integrating by parts, their result may be readily extended to φ's which are of bounded variation and which vanish to the right of some $T < \tau_H$. See Gijbels and Veraverbeke (1991) for an extension of this technique to multiple Kaplan–Meier integrals. Susarla and Van Ryzin (1980) apparently were the first to treat a φ, namely, $\varphi(x) = x$, which is of unbounded variation. For technical reasons they proposed to truncate the Kaplan–Meier integral at some finite $M = M_n$, but such that $M_n \to \infty$. Note, however, that $\int \varphi \, d\hat{F}_n$ for $\varphi \geq 0$ is biased downward [see Mauro (1985), Stute and Wang (1993) and Stute (1994a)], so that every truncation only leads to a further increase of the bias.

Gill (1983) considered the Kaplan–Meier integral for the identity function over the whole real line and for nonnegative continuous nonincreasing φ's satisfying further integrability assumptions. Schick, Susarla and Koul (1988) obtained, for this class of φ's, a weak representation of $\int \varphi \, d\hat{F}_n$ in terms of a sum of i.i.d. random variables plus a remainder. In all these papers, integration by parts was essential. To the best of our knowledge, no CLT for a reasonably larger class of φ's is available at present. What is more, application of counting process techniques needed further truncation, which required additional assumptions in the analysis of $\int \varphi \, d(\hat{F}_n - \tilde{F})$. See, for example, Corollary 3.2 in Gill (1983). Finally, often an $\hat{F}_n$ was considered properly modified so as to become a true d.f. irrespective of whether the largest datum is censored or not. As pointed out by Wellner (1985), this artificial handling has some serious drawbacks on the bias of $\int \varphi \, d\hat{F}_n$. Stute and Wang (1994) showed that the jackknife-corrected Kaplan–Meier integral is much more cautious about attributing weights to $Z_{n:n}$ when $\delta_{[n:n]} = 0$.

In this paper, a general φ will be considered. As a consequence, no integration by parts can be applied nor do we have to stop the Kaplan–Meier process at the largest Z.

We now make some comments on our proof. Lemma 2.1 presents a derivation of $\int \varphi\, d\hat{F}_n$ which does not follow the traditional route of expressing $\hat{F}_n$ in terms of the Aalen–Nelson estimator of the cumulative hazard function of F. In particular, for arbitrary φ rather than indicators of intervals, the aforementioned relation seems unsuitable because of its nonlinearity. Further, we shall make heavy use of U-statistics theory. Projection techniques will be applied on several occasions.

We close this section with a discussion of the technical assumptions (1.5) and (1.6).

First, when there is no censoring at all, we may formally set $G = \delta_\infty$ = Dirac at infinity. In this case, all δ's equal 1. Furthermore,

$$\tilde{H}^0 \equiv 0, \qquad \tilde{H}^1 = H = F, \qquad \gamma_0 \equiv 1 \quad \text{and} \quad \gamma_2 \equiv 0$$

on the real line. Consequently, the right-hand side of (1.7) collapses to

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

Under censorship, the assumptions (1.5) and (1.6) are clearly satisfied whenever

$$\int \varphi^2\, dF < \infty \quad \text{and} \quad \varphi(x) = 0 \quad \text{on some } T < x \le \tau_H.$$

The second condition, however, rules out many examples so that an analysis must be carried through which also allows for handling those φ's which have noncompact support. In the following we shall discuss censoring only when F and G are continuous. This is just for convenience since under continuity, identification of the involved quantities is a little simpler. Identification of $\int \varphi\, d\tilde{F}$, for example, in the general case may be found in Stute and Wang (1993). Now, for each $x < \tau_H$, we have

$$\gamma_0(x) = \frac{1}{1 - G(x)}.$$

Conclude that

$$\gamma_1(x) = \frac{1}{1 - H(x)} \int_x^{\tau_H} \varphi(w) F(dw)$$

and

$$\begin{aligned}\gamma_2(x) &= \int_{-\infty}^{\tau_H} \int_{-\infty}^{\tau_H} \frac{\varphi(w) 1_{\{v < x \wedge w\}}}{[1 - H(v)][1 - G(v)]} G(dv) F(dw) \\ &= \int_{-\infty}^{\tau_H} \varphi(w) C(x \wedge w) F(dw).\end{aligned}$$

Furthermore,

$$\begin{aligned}\mathbb{E}[\gamma_1(Z)(1-\delta)] &= \mathbb{E}[\gamma_2(Z)]\\ &= \int\int \frac{1_{\{v<w\}}\varphi(w)\gamma_0(w)}{1-H(v)}\tilde{H}^0(dv)\tilde{H}^1(dw)\\ &= \int_{-\infty}^{\tau_H}\int_{-\infty}^{\tau_H}\frac{1_{\{v<w\}}\varphi(w)}{1-G(v)}G(dv)F(dw).\end{aligned}$$

Since also

$$\mathbb{E}[\varphi(Z)\gamma_0(Z)\delta] = \int_{-\infty}^{\tau_H}\varphi(x)F(dx) = \int \varphi(x)\tilde{F}(dx),$$

the U's appearing in (1.8) are indeed centered.

Condition (1.5) just states that $\varphi(Z)\gamma_0(Z)\delta$ has a finite second moment. Equivalently,

$$\int \varphi^2(1-G)^{-1}\, d\tilde{F} < \infty. \tag{1.9}$$

As to the function C, note that, for example,

$$C(x) \le \frac{1}{1-F(x)}\int_{-\infty}^{x}\frac{G(dy)}{[1-G(y)]^2} \le \frac{1}{1-H(x)}.$$

Hence (1.6) is implied by

$$\int \frac{|\varphi|\, d\tilde{F}}{(1-H)^{1/2}} < \infty. \tag{1.10}$$

To further illustrate (1.6), assume that (apart from continuity)

$$1-F \sim c(1-G)^{\beta} \quad \text{in a neighborhood of } \tau_H, \tag{1.11}$$

for some $c > 0$ and $\beta > 0$; large values of β indicate heavier tails in the censoring distribution. Condition (1.10) then is implied by

$$\int \frac{|\varphi|\, d\tilde{F}}{(1-G)^{\alpha}} < \infty \quad \text{with } \alpha = \frac{1+\beta}{2}. \tag{1.12}$$

Clearly, (1.9) and (1.12) may be achieved for a large class of φ's, F's and G's. Only for further illustration, assume that φ is bounded, but not necessarily of bounded support. Then (1.9) and (1.12) hold true if $\beta > 1$, that is, if, as mentioned earlier, there is enough information on F in the tails. For $\beta = 1$, we may include logarithmic factors in (1.11) to make (1.9) and (1.10) still hold true. Without such a modification, that is, for (1.11) with $\beta = 1$, the bias may be of the order $an^{-1/2}$ with some nonvanishing factor a; see Stute (1994a). This indicates that Theorem 1.1 is now no longer valid.

We only mention that bootstrap versions of Theorem 1.1 and Corollary 1.2 are readily available by imitating the arguments in Section 2 for the bootstrap sample.

2. Proofs. Some further notation will be needed. Denote by

$$H_n(z) = n^{-1} \sum_{i=1}^{n} 1_{\{Z_i \le z\}}$$

and

$$\tilde{H}_n^j(z) = n^{-1} \sum_{i=1}^{n} 1_{\{Z_i \le z,\, \delta_i = j\}}, \qquad j = 0, 1,$$

the empirical (sub-) distribution function estimators of H, $\tilde{H}^0$ and $\tilde{H}^1$, respectively. Recall

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

Our first lemma provides a useful expression of

$$\int \varphi \, d\hat{F}_n = \sum W_{in} \varphi(Z_{i:n})$$

in terms of the (unordered) (Z, δ)'s and the functions H_n, $\tilde{H}_n^0$ and $\tilde{H}_n^1$. To understand its background, observe that (under continuity)

$$\begin{aligned} \int \varphi(x) \tilde{F}(dx) &= \int \varphi(X) \frac{1_{\{X \le Y\}}}{1 - G(X)} \, d\mathbb{P} \\ &= \int \varphi(w) \exp\left\{ \int_{-\infty}^{w} \frac{\tilde{H}^0(dz)}{1 - H(z)} \right\} \tilde{H}^1(dw). \end{aligned}$$

This is a proper representation of $\int \varphi \, d\tilde{F}$ in terms of estimable quantities. Lemma 2.1 below constitutes its empirical analog. It will be stated for a continuous H. This is only for notational convenience since then there will be no ties among the Z's so that their ranks may be easily expressed in terms of H_n. In the final proof of Theorem 1.1 we will see how the general case may be traced back to the present one.

LEMMA 2.1. *For a continuous H, we have*

$$\sum_{i=1}^{n} W_{in} \varphi(Z_{i:n}) = \int \varphi(w) \exp\left\{ n \int_{-\infty}^{w-} \ln\left[1 + \frac{1}{n(1 - H_n(z))} \right] \tilde{H}_n^0(dz) \right\} \tilde{H}_n^1(dw).$$

PROOF. Check that

$$\begin{aligned}
\sum_{i=1}^{n} W_{in}\varphi(Z_{i:n}) &= \sum_{i=1}^{n} \frac{\varphi(Z_{i:n})\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{[j:n]}} \\
&= \sum_{i=1}^{n} \frac{\varphi(Z_{i:n})\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left(1 - \frac{\delta_{[j:n]}}{n-j+1} \right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi(Z_{i:n})\delta_{[i:n]} \prod_{j=1}^{i-1} \left(1 + \frac{1-\delta_{[j:n]}}{n-j} \right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi(Z_i)\delta_i \prod_{j=1}^{n} \left(1 + \frac{1-\delta_j}{n - \operatorname{Rank} Z_j} \right)^{1_{\{Z_j < Z_i\}}} \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi(Z_i)\delta_i \exp\left\{ \sum_{j=1}^{n} 1_{\{Z_j < Z_i\}} \ln\left[1 + \frac{1-\delta_j}{n(1-H_n(Z_j))} \right] \right\} \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi(Z_i)\delta_i \\
&\quad \times \exp\left\{ \sum_{j=1}^{n} 1_{\{Z_j < Z_i\}}(1-\delta_j) \ln\left[1 + \frac{1}{n(1-H_n(Z_j))} \right] \right\},
\end{aligned}$$

whence the result comes. □

To give an outline of the proof, replace the $\ln(1+x)$ term by x, which is legitimate at least for all small $x > 0$. The exponential term then becomes

$$\exp\left\{ \int_{-\infty}^{w-} \frac{\tilde{H}_n^0(dz)}{1-H_n(z)} \right\}.$$

Integration w.r.t. $\tilde{H}_n^1$ and further expansion finally leads to a U-statistic of degree 3. Its Hájek projection is the desired leading term in (1.7) and therefore determines the limit distribution of a Kaplan–Meier integral. Further clarifying remarks will be made whenever it seems appropriate.

We now make things rigorous. Expand the exponential term in Lemma 2.1 as follows: with $w = Z_i$,

$$\begin{aligned}
\exp\{\cdots\} = \exp\left\{ \int_{-\infty}^{Z_i-} \frac{\tilde{H}^0(dz)}{1-H(z)} \right\} \Bigg[1 &+ n\int_{-\infty}^{Z_i-} \ln\left[1 + \frac{1}{n(1-H_n(z))} \right] \tilde{H}_n^0(dz) \\
&- \int_{-\infty}^{Z_i-} \frac{\tilde{H}^0(dz)}{1-H(z)} \Bigg] \\
&+ \frac{1}{2} e^{\Delta_i} \left\{ n\int_{-\infty}^{Z_i-} \ln\left[1 + \frac{1}{n(1-H_n(z))} \right] \tilde{H}_n^0(dz) - \int_{-\infty}^{Z_i-} \frac{\tilde{H}^0(dz)}{1-H(z)} \right\}^2,
\end{aligned}$$

where Δ_i is between the two terms in brackets. Write

$$A_{in} \equiv n\int_{-\infty}^{Z_i-}\ln\left[1+\frac{1}{n(1-H_n(z))}\right]\tilde{H}_n^0(dz)-\int_{-\infty}^{Z_i-}\frac{\tilde{H}^0(dz)}{1-H(z)} \equiv B_{in}+C_{in},$$

with

$$B_{in} := n\int_{-\infty}^{Z_i-}\ln\left[1+\frac{1}{n(1-H_n(z))}\right]\tilde{H}_n^0(dz)-\int_{-\infty}^{Z_i-}\frac{\tilde{H}_n^0(dz)}{1-H_n(z)}$$

and

$$C_{in} := \int_{-\infty}^{Z_i-}\frac{\tilde{H}_n^0(dz)}{1-H_n(z)}-\int_{-\infty}^{Z_i-}\frac{\tilde{H}^0(dz)}{1-H(z)}.$$

As in the first section, put

$$\gamma_0(x) \equiv \gamma(x) = \exp\left\{\int_{-\infty}^{x-}\frac{\tilde{H}^0(dz)}{1-H(z)}\right\}.$$

In terms of these quantities we have thus obtained

$$\begin{aligned}\sum_{i=1}^{n} W_{in}\varphi(Z_{i:n}) &= n^{-1}\sum_{i=1}^{n}\varphi(Z_i)\gamma(Z_i)\delta_i[1+B_{in}+C_{in}]\\ &\quad+n^{-1}\sum_{i=1}^{n}\tfrac{1}{2}\varphi(Z_i)\delta_i e^{\Delta_i}\{B_{in}+C_{in}\}^2.\end{aligned}\tag{2.1}$$

We shall first study C_{in} in greater detail. This quantity is closely related to the cumulative hazard function process for the censored data evaluated at Z_i. We need a special representation of C_{in}, which will lead us to a proper decomposition of

$$n^{-1}\sum_{i=1}^{n}\varphi(Z_i)\gamma(Z_i)\delta_i C_{in}.$$

For this, note that for $z < Z_{n:n}$,

$$\frac{1}{1-H_n(z)} = -\frac{1-H_n(z)}{[1-H(z)]^2}+\frac{2}{1-H(z)}+\frac{[H_n(z)-H(z)]^2}{[1-H(z)]^2[1-H_n(z)]}.$$

Hence

$$\begin{aligned}C_{in} &= -\int_{-\infty}^{Z_i-}\frac{1-H_n(z)}{[1-H(z)]^2}\tilde{H}_n^0(dz)+\int_{-\infty}^{Z_i-}\frac{2}{1-H(z)}\tilde{H}_n^0(dz)\\ &\quad-\int_{-\infty}^{Z_i-}\frac{1}{1-H(z)}\tilde{H}^0(dz)\\ &\quad+\int_{-\infty}^{Z_i-}\frac{[H_n(z)-H(z)]^2}{[1-H(z)]^2[1-H_n(z)]}\tilde{H}_n^0(dz).\end{aligned}\tag{2.2}$$

Conclude that

$$n^{-1}\sum_{i=1}^{n}\varphi(Z_i)\gamma(Z_i)\delta_i C_{in}$$

$$= -\iiint \frac{1_{\{v<u,\, v<w\}}\varphi(w)\gamma(w)}{[1-H(v)]^2} H_n(du)\tilde{H}_n^0(dv)\tilde{H}_n^1(dw)$$

$$+ 2\iint \frac{1_{\{v<w\}}\varphi(w)\gamma(w)}{1-H(v)}\tilde{H}_n^0(dv)\tilde{H}_n^1(dw)$$

$$-\iint \frac{1_{\{v<w\}}\varphi(w)\gamma(w)}{1-H(v)}\tilde{H}^0(dv)\tilde{H}_n^1(dw) + R_{n1},$$

where

$$R_{n1} = \iint \varphi(w)\gamma(w)1_{\{z<w\}}\frac{[H_n(z)-H(z)]^2}{[1-H(z)]^2[1-H_n(z)]}\tilde{H}_n^0(dz)\tilde{H}_n^1(dw).$$

As we shall see, the first three integrals will contribute to the expansion of $\int \varphi d\hat{F}_n$, while the remaining terms will become negligible. The next two lemmas will provide a representation of the first two integrals as sums of i.i.d. random variables (plus remainder). Of course, the third one is a sum already.

NOTE. For Lemmas 2.2–2.7 below we temporarily make the assumption

$$\varphi(x) = 0 \quad \text{for all } T < x \text{ and some } T < \tau_H. \tag{2.3}$$

As a consequence of (2.3), all denominators appearing in proofs are bounded away from below. Thus they will not cause any troubles and may be handled along the same lines as in, for example, Földes and Rejtő (1981) or Csörgő and Horváth (1983). The integrability condition reduces to $\int \varphi^2\, dF < \infty$.

LEMMA 2.2. *Under* $\int \varphi^2\, dF < \infty$ *and* (2.3),

$$\iiint \frac{1_{\{v<u,\, v<w\}}\varphi(w)\gamma(w)}{[1-H(v)]^2} H_n(du)\tilde{H}_n^0(dv)\tilde{H}_n^1(dw)$$

$$= \iiint \frac{1_{\{v<u,\, v<w\}}\varphi(w)\gamma(w)}{[1-H(v)]^2}$$

$$\times\big[H_n(du)\tilde{H}^0(dv)\tilde{H}^1(dw) + H(du)\tilde{H}_n^0(dv)\tilde{H}^1(dw)$$

$$+H(du)\tilde{H}^0(dv)\tilde{H}_n^1(dw) - 2H(du)\tilde{H}^0(dv)\tilde{H}^1(dw)\big] + R_{n2},$$

where R_{n2} *satisfies*

$$|R_{n2}| = \begin{cases} O(n^{-1}), & \textit{in probability},\\ O(n^{-1}\ln n), & \textit{with probability } 1.\end{cases}$$

PROOF. The integral on the left-hand side is a V-statistic of the bivariate data (Z_i, δ_i), $1 \le i \le n$; the one on the right-hand side is related to its Hájek projection. The result follows from Theorems 5.3.2 and 5.3.3 in Serfling (1980) and Berk's (1966) SLLN for U-statistics (needed to control the "diagonal" $u = w$). Note that an application of these results neither requires symmetry of the U-kernel nor real-valued observations. Actually, symmetry is often assumed only to make proofs smoother [see the comment in Serfling (1980), page 172], and U-statistics of multivariate random vectors were already considered in Hoeffding (1948). □

LEMMA 2.3. *Under* $\int \varphi^2\, dF < \infty$ *and* (2.3),

$$\iint \frac{1_{\{v<w\}}\varphi(w)\gamma(w)}{1-H(v)} \tilde{H}_n^0(dv)\tilde{H}_n^1(dw)$$
$$= \iint \frac{1_{\{v<w\}}\varphi(w)\gamma(w)}{1-H(v)} \big[\tilde{H}^0(dv)\tilde{H}_n^1(dw) + \tilde{H}_n^0(dv)\tilde{H}^1(dw) - \tilde{H}^0(dv)\tilde{H}^1(dw)\big] + R_{n3},$$

where the remainder satisfies

$$|R_{n3}| = \begin{cases} O(n^{-1}), & \textit{in probability}, \\ O(n^{-1}\ln n), & \textit{with probability } 1. \end{cases}$$

PROOF. Similar to before. □

Since some of the triple integrals in Lemma 2.2 cancel with some of the double integrals in Lemma 2.3, we finally obtain the following result.

COROLLARY 2.4. *Under* $\int \varphi^2\, dF < \infty$ *and* (2.3),

$$n^{-1}\sum_{i=1}^{n} \varphi(Z_i)\gamma(Z_i)\delta_i C_{in}$$
$$= -\iiint \frac{1_{\{v<u,\, v<w\}}\varphi(w)\gamma(w)}{[1-H(v)]^2} H_n(du)\tilde{H}^0(dv)\tilde{H}^1(dw)$$
$$+ \iint \frac{1_{\{v<w\}}\varphi(w)\gamma(w)}{1-H(v)} \tilde{H}_n^0(dv)\tilde{H}^1(dw)$$
$$+ R_{n1} - R_{n2} + 2R_{n3}.$$

In view of (2.1) we end up with

$$\int \varphi\, d\hat{F}_n = \int \varphi(w)\gamma(w)\tilde{H}_n^1(dw) + \int\int \frac{1_{\{v<w\}}\varphi(w)\gamma(w)}{1-H(v)}\tilde{H}^1(dw)\tilde{H}_n^0(dv)$$
$$-\int\int\int \frac{1_{\{v<u,\, v<w\}}\varphi(w)\gamma(w)}{[1-H(v)]^2}\tilde{H}^0(dv)\tilde{H}^1(dw)H_n(du)$$
$$+ n^{-1}\sum_{i=1}^n \varphi(Z_i)\gamma(Z_i)\delta_i B_{in} + n^{-1}\sum_{i=1}^n \frac{1}{2}\varphi(Z_i)\delta_i e^{\Delta_i}\{B_{in}+C_{in}\}^2$$
$$+ R_{n1} - R_{n2} + 2R_{n3}.$$

To prove the assertion of Theorem 1.1 under continuity of H and the additional assumption (2.3), it suffices to bound R_{n1} and the last two sums.

LEMMA 2.5. *Under* $\int \varphi^2\, dF < \infty$ *and* (2.3),

$$|R_{n1}| = O\left(\frac{\ln n}{n}\right) \quad \textit{with probability } 1.$$

PROOF. Immediate consequence of the LIL for empirical measures, the ordinary SLLN and assumption (2.3). □

To obtain a proper bound for

$$S_{n1} \equiv n^{-1}\sum_{i=1}^n \varphi(Z_i)\gamma(Z_i)\delta_i B_{in},$$

note that

$$x - \frac{x^2}{2} \le \ln(1+x) \le x \quad \text{for } x \ge 0.$$

As a consequence

$$-\frac{1}{2n}\int_{-\infty}^{Z_i-} \frac{\tilde{H}_n^0(dz)}{[1-H_n(z)]^2} \le B_{in} \le 0. \tag{2.4}$$

Apply (2.3), (2.4), Glivenko–Cantelli and the SLLN to get the following lemma.

LEMMA 2.6. *Under* $\int \varphi^2\, dF < \infty$ *and* (2.3),

$$|S_{n1}| = O\left(\frac{1}{n}\right) \quad \textit{with probability } 1.$$

Finally, set

$$S_{n2} = n^{-1}\sum_{i=1}^n \tfrac{1}{2}|\varphi(Z_i)|\delta_i e^{\Delta_i}\{B_{in}+C_{in}\}^2.$$

Under (2.3), inequality (2.4), Glivenko–Cantelli and the LIL for cumulative hazard functions on compacta immediately yield the following lemma.

LEMMA 2.7. *Under* $\int \varphi^2 \, dF < \infty$ *and* (2.3),

$$|S_{n2}| = O\left(\frac{\ln n}{n}\right) \quad \textit{with probability } 1.$$

To summarize the results obtained so far, under $\int \varphi^2 \, dF < \infty$, continuity of H and (2.3), we have

$$\begin{aligned} \int \varphi \, d\hat{F}_n &= n^{-1} \sum_{i=1}^{n} \varphi(Z_i) \gamma_0(Z_i) \delta_i \\ &\quad + n^{-1} \sum_{i=1}^{n} \gamma_1(Z_i)(1 - \delta_i) - n^{-1} \sum_{i=1}^{n} \gamma_2(Z_i) + R_n, \end{aligned} \tag{2.5}$$

where

$$|R_n| = O(n^{-1} \ln n) \qquad \mathbb{P}\text{-a.s.}$$

Conclude that Theorem 1.1 follows under (2.3).

REMARK. Note that (2.5) also yields the LIL for $\int \varphi \, d\hat{F}_n$, under (2.3), for an arbitrary square integrable φ. The LIL is also valid under much weaker conditions, but its proof requires considerably more effort.

We are now in the position to give the proof of Theorem 1.1.

PROOF OF THEOREM 1.1. Assume first that H is continuous. In particular, $H\{\tau_H\} = 0$. For a given $\epsilon > 0$ choose an F square integrable $\tilde{\varphi}$ vanishing outside of $(-\infty, T]$, for some $T < \tau_H$, such that

$$\int (\varphi - \tilde{\varphi})^2 \gamma^2 \, d\tilde{H}^1 \le \varepsilon \tag{2.6}$$

and

$$\int |\varphi - \tilde{\varphi}| C^{1/2} \, d\tilde{F} \le \varepsilon, \tag{2.7}$$

which is possible under (1.5) and (1.6). Set $\varphi_1 = \varphi - \tilde{\varphi}$. We shall prove that, as $n \to \infty$,

$$n^{1/2}\left[\int \varphi_1 \, d\hat{F}_n - \int \varphi_1 \, d\tilde{F}\right] = O_{\mathbb{P}}(\varepsilon^{1/2}).$$

In view of (2.5), a Cramér–Slutzky type argument will then complete the proof of Theorem 1.1. Now, using our previous notation, Lemma 2.1 yields

$$\int \varphi_1 d\hat{F}_n = n^{-1} \sum_{i=1}^{n} \varphi_1(Z_i) \delta_i \gamma(Z_i) \exp[B_{in} + C_{in}].$$

Hence

$$n^{1/2}\int\varphi_1\,d\big(\hat{F}_n-\tilde{F}\big)=n^{-1/2}\sum_{i=1}^{n}\left[\varphi_1(Z_i)\,\delta_i\gamma(Z_i)-\int\varphi_1\,d\tilde{F}\right]$$

$$+\,n^{-1/2}\sum_{i=1}^{n}\varphi_1(Z_i)\,\delta_i\gamma(Z_i)\big[\exp(B_{in}+C_{in})-1\big].$$

By (2.6), the variance of the first sum is less than or equal to ε. The second sum is bounded in absolute value by

$$n^{-1/2}\sum_{i=1}^{n}|\varphi_1(Z_i)|\delta_i\gamma(Z_i)|A_{in}|\exp\big[|B_{in}|+|C_{in}|\big].$$

From (2.4),

$$|B_{in}|\le\frac{1}{n(1-H_n(Z_i-))}\le 1. \tag{2.8}$$

Since continuity of H implies continuity of F and G on $x<\tau_H$ [cf. the discussion in Stute and Wang (1993), page 1604], Theorem 2.1 in Zhou (1991) may be applied to get

$$\sup_{1\le i\le n}|C_{in}|=O_{\mathbb{P}}(1).$$

Actually, Zhou's result was formulated for the cumulative hazard function process, but it may be easily seen that the difference between his process evaluated at Z_i and our C_{in} is $O_{\mathbb{P}}(1)$ uniformly in $1\le i\le n$. See Stute (1994b) for a further discussion of the cumulative hazard function process on large sets, which may be equally well applied to bound the C's.

Altogether we see that the exponential factor is $O_{\mathbb{P}}(1)$ uniformly in $1\le i\le n$. It remains to bound

$$n^{-1/2}\sum_{i=1}^{n}|\varphi_1(Z_i)|\delta_i\gamma(Z_i)|A_{in}|.$$

Since $A_{in}=B_{in}+C_{in}$, it suffices to bound the sums with B_{in} and C_{in} in place of A_{in}. By (2.4) and (2.8),

$$n^{-1/2}\sum_{i=1}^{n}|\varphi_1(Z_i)|\delta_i\gamma(Z_i)|B_{in}|$$

$$\le n^{-1}\sum_{i=1}^{n}|\varphi_1(Z_i)|\delta_i\gamma(Z_i)\left[\int_{-\infty}^{Z_i-}\frac{\tilde{H}_n^0(dz)}{[1-H_n(z)]^2}\right]^{1/2}.$$

Since $(1-H)/(1-H_n)$ is stochastically bounded from above on $z<Z_{n:n}$, as $n\to\infty$ [cf. Shorack and Wellner (1986), page 415], it remains to bound

$$n^{-1}\sum_{i=1}^{n}|\varphi_1(Z_i)|\delta_i\gamma(Z_i)\left[\int_{-\infty}^{Z_i-}\frac{\tilde{H}_n^0(dz)}{[1-H(z)]^2}\right]^{1/2} \tag{2.9}$$

By a symmetry argument, the expectations of the summands in (2.9) are all equal. For $i = 1$, say, condition on Z_1 and apply Hölder's inequality to $\mathbb{E}[\cdots]^{1/2}$ to finally bound the expectation of (2.9) from above by

$$\int |\varphi_1(z)| C^{1/2}(z) \tilde{F}(dz) \le \varepsilon. \tag{2.10}$$

We now bound the sum corresponding to C_{in}. For this, recall the expansion (2.2). As a consequence it is required to bound each of the sums

$$n^{-1/2} \sum_{i=1}^{n} |\varphi_1(Z_i)| \delta_i \gamma(Z_i) \int_{-\infty}^{Z_i-} \frac{[H_n(z) - H(z)]^2}{[1 - H(z)]^2 [1 - H_n(z)]} \tilde{H}_n^0(dz), \tag{2.11}$$

$$n^{-1/2} \sum_{i=1}^{n} |\varphi_1(Z_i)| \delta_i \gamma(Z_i) \int_{-\infty}^{Z_i-} \frac{\tilde{H}_n^0(dz) - \tilde{H}^0(dz)}{1 - H(z)}, \tag{2.12}$$

$$n^{-1/2} \sum_{i=1}^{n} |\varphi_1(Z_i)| \delta_i \gamma(Z_i) \int_{-\infty}^{Z_i-} \frac{H_n(z) - H(z)}{[1 - H(z)]^2} \tilde{H}_n^0(dz). \tag{2.13}$$

In (2.11), first replace $1 - H_n$ by $1 - H$ in the denominator of the integrand. Then take expectations of each of the three sums. For the first, one readily sees that the resulting bound is of the same order as that for (2.9), namely, (2.10). Similarly, by Hölder, for (2.12). For (2.13) a little more work is needed. For the ith summand, condition on $Z_i = z_0$, apply Hölder again and write

$$\left\{ \int_{-\infty}^{z_0-} \frac{H_n(z) - H(z)}{[1 - H(z)]^2} \tilde{H}_n^0(dz) \right\}^2$$
$$= \int_{-\infty}^{z_0-} \int_{-\infty}^{z_0-} \frac{H_n(z_1) - H(z_1)}{[1 - H(z_1)]^2} \frac{H_n(z_2) - H(z_2)}{[1 - H(z_2)]^2} \tilde{H}_n^0(dz_1) \tilde{H}_n^0(dz_2).$$

The expectation of this quantity is again $O[C(z_0)]$ uniformly in z_0 so that also the expectation of (2.13) admits the bound (2.10) (up to constants). This completes the proof of Theorem 1.1 for a continuous H.

Next we consider the situation when F and G have no common jumps but may have separate discontinuities otherwise. This case (which is sufficient for practical purposes) has been studied in detail in Stute and Wang (1993). In particular, it was shown on page 1605 in Stute and Wang (1993) that a quantile transformation may be applied so as to trace everything back to uniformly distributed Z's. The same procedure also applies here. Finally, to handle common jumps, we apply an idea already elaborated in Gill [(1980), page 74]. If $\{x_i\}$ are the common jumps of F and G, replace each x_i by a small interval $[x_i, x_i + \Delta_i)$, where $\sum_i \Delta_i < \infty$, and move the G-mass of x_i to $x_i + \Delta_i$. Extend F and φ to the new time scale by putting, for example, $F(x) = F(x_i)$ on $x_i \le x \le x_i + \Delta_i$. Due to our convention that in W_{in} tied uncensored and censored observations are treated as if the first precede the latter, $\int \varphi \, d(\hat{F}_n - \tilde{F})$ remains unchanged on the extended time scale. Obviously the "no common jumps" condition is satisfied now. Also, our integrability conditions (1.5) and (1.6) hold true in the new context, so that in summary we obtain the representation (1.7) on the extended time scale. Check that the right-hand sides of (1.7) agree on both scales. This completes the proof of Theorem 1.1. □

REMARK. Our technique may be readily extended so as to yield, under the assumptions of Theorem 1.1, a functional CLT for the process $n^{1/2}(\int \varphi \, d(\hat{F}_n - \tilde{F}))$.

Acknowledgment. Thanks to a referee who suggested the last argument in the proof of Theorem 1.1 for removing the "no common jumps" condition.

REFERENCES

BERK, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51–58.

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.

CSÖRGŐ, S. and HORVÁTH, L. (1983). The rate of strong uniform consistency for the product-limit estimator. *Z. Wahrsch. Verw. Gebiete* **62** 411–426.

FÖLDES, A. and REJTŐ, L. (1981). A LIL type result for the product limit estimator. *Z. Wahrsch. Verw. Gebiete* **56** 75–86.

GIJBELS, I. and VERAVERBEKE, N. (1991). Almost sure asymptotic representation for a class of functionals of the Kaplan–Meier estimator. *Ann. Statist.* **19** 1457–1470.

GILL, R. D. (1980). *Censoring and Stochastic Integrals. Math. Centre Tracts* **124**. Math. Centrum, Amsterdam.

GILL, R. D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* **11** 49–58.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

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

LO, S. H. and SINGH, K. (1986). The product-limit estimator and the bootstrap: Some asymptotic representations. *Probab. Theory Related Fields* **71** 455–465.

MAJOR, P. and REJTŐ, L. (1988). Strong embedding of the estimator of the distribution function under random censorship. *Ann. Statist.* **16** 1113–1132.

MAURO, D. (1985). A combinatoric approach to the Kaplan–Meier estimator. *Ann. Statist.* **13** 142–149.

SCHICK, A., SUSARLA, V. and KOUL, H. (1988). Efficient estimation of functionals with censored data. *Statist. Decisions* **6** 349–360.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

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

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

STUTE, W. (1994b). Strong and weak representations of cumulative hazard function and Kaplan–Meier estimators on increasing sets. *J. Statist. Plann. Inference.* **42** 315–329.

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.

SUSARLA, V. and VAN RYZIN, J. (1980). Large sample theory for an estimator of the mean survival time from censored samples. *Ann. Statist.* **8** 1002–1016.

WELLNER, J. A. (1985). A heavy censoring limit theorem for the product limit estimator. *Ann. Statist.* **13** 150–162.

ZHOU, M. (1991). Some properties of the Kaplan–Meier estimator for independent nonidentically distributed random variables. *Ann. Statist.* **19** 2266–2276.

MATHEMATICAL INSTITUTE
UNIVERSITY OF GIESSEN
ARNDTSTRASSE 2
D-35392 GIESSEN
GERMANY