In these notes, we write $\hat{S}$ for the estimated survival function and $\hat{\Lambda}$ for the estimated cumulative hazard function, with the understanding that the notation covers both the KM and NA estimators (which are asymptotically equivalent if the survival time has a continuous distribution).

Results for $\hat{\Lambda}$ can be translated into results for $\hat{S}$ using the Taylor approximation

$$\hat{S}(t) - S(t) \doteq -S(t)(\hat{\Lambda}(t) - \Lambda(t)), \tag{1}$$

which follows from the relation $S(t) = \exp(-\Lambda(t))$ and the Taylor approximation $e^b - e^a = e^a(e^{b-a} - 1) \doteq e^a(b-a)$.

The first result is

$$\sqrt{n}(\hat{\Lambda}(t) - \Lambda(t)) \rightsquigarrow B(V(t)), \tag{2}$$

where $B(\cdot)$ is the Brownian motion process, $\rightsquigarrow$ denotes stochastic process convergence (more precisely, weak convergence in the space D as described in Billingsley's *Convergence of Probability Measures*), and

$$V(t) = \lim_{n\to\infty} n\hat{V}(t) = \int_0^t \frac{\lambda(s)}{r(s)} ds,$$

with $\hat{V}(t)$ given by Greenwood's formula and $r(s) = \Pr(X \geq s)$. Note that $V(t)$ is a monotone increasing function with $V(0) = 0$.

Next, for any $\tau > 0$, the above result and the fact that $V(\tau)^{1/2} B(V(t)/V(\tau))$ has the same distribution (as a process) as $B(V(t))$ (see the notes on Brownian motion and Brownian bridge), we obtain

$$\sqrt{n}\, V(\tau)^{-1/2}(\hat{\Lambda}(t) - \Lambda(t)) \rightsquigarrow B(V(t)/V(\tau)). \tag{3}$$

A consequence of this is that

$$\sup_{t\in[0,\tau]} \sqrt{n}\, V(\tau)^{-1/2} |\hat{\Lambda}(t) - \Lambda(t)| \xrightarrow{d} \max_{s\in[0,1]} |B(s)|. \tag{4}$$

Hence, we can construct a simultaneous $100(1-\alpha)\%$ confidence band for $\Lambda(t)$ over the interval $t \in [0, \tau]$ as

$$\hat{\Lambda}(t) \pm c(\alpha, B)\sqrt{\frac{\hat{V}(\tau)}{n}}, \tag{5}$$

where $c(\alpha, B)$ is the value of c such that

$$\Pr(\max_{s\in[0,1]} |B(s)| \leq c) = 1 - \alpha.$$

The notes on Brownian motion and Brownian bridge contain a table of critical values. This confidence band can be converted into a confidence band for $S(t)$ using the relation $S(t) = \exp(-\Lambda(t))$.

The results (2) and (3) can be combined with the approximation (1) to yield the following results for $\hat{S}(t)$.

$$\frac{\sqrt{n}(\hat{S}(t) - S(t))}{S(t)} \rightsquigarrow B(V(t)), \tag{6}$$

$$\frac{\sqrt{n}(\hat{S}(t) - S(t))}{S(t)V(\tau)^{1/2}} \rightsquigarrow B(V(t)/V(\tau)). \tag{7}$$

The second of these two results yields an alternate simultaneous $100(1-\alpha)\%$ confidence band for $S(t)$ over the interval $t \in [0, \tau]$, in the form

$$\hat{S}(t) \pm c(\alpha, B)\hat{S}(t)\sqrt{\frac{\hat{V}(\tau)}{n}}. \tag{8}$$

This confidence band was proposed in Gill's 1980 book *Censoring and Stochastic Integrals*.

Define $K(t) = V(t)/(1 + V(t))$. We then have the following further result.

$$\frac{\sqrt{n}(1 - K(t))(\hat{S}(t) - S(t))}{S(t)} \rightsquigarrow B^{\circ}(K(t)), \tag{9}$$

where $B^{\circ}(\cdot)$ is the Brownian bridge process. This result follows from the result (6) and the fact that $(1 + V(t))B^{\circ}(V(t)/(1 + V(t))$ has the same distribution (as a process) as $B(V(t))$ (see the notes on Brownian motion and Brownian bridge). A consequence of this is that

$$\sup_{t\in[0,\tau]} \frac{\sqrt{n}(1 - K(t))|\hat{S}(t) - S(t)|}{S(t)} \xrightarrow{d} \max_{u\in[0,K(t)]} |B^{\circ}(u)|. \tag{10}$$

Hence, we can construct an alternate simultaneous $100(1-\alpha)\%$ confidence band for $S(t)$ over the interval $t \in [0, \tau]$ as

$$\hat{S}(t) \pm \frac{c(\alpha, \hat{K}(t), B^{\circ})\hat{S}(t)}{(1 - \hat{K}(t))\sqrt{n}}, \tag{11}$$

where $c(\alpha, \kappa, B^\circ)$ is the value of c such that

$$\Pr\left(\max_{u \in [0,\kappa]} |B^\circ(u)| \leq c\right) = 1 - \alpha.$$

The notes on Brownian motion and Brownian bridge contain a table of critical values. This confidence band was proposed by Hall and Wellner (1980, *Biometrika*). It has two advantages relative to the other bands. First, in the case where there is no censoring and $\tau = \infty$, it reduces to the classical Kolmogorov band. Second, in the presence of censoring, the value of $\hat{V}(\tau)$ can be very large for large τ, making the Hall-Wellner interval more practicable than the intervals that involve $\hat{V}(\tau)$.