Austral. J. Statist., **30(3)**, 1988, 342–358

BOOTSTRAPPED MULTI-DIMENSIONAL PRODUCT LIMIT PROCESS

BRIAN S. YANDELL

Departments of Statistics and of Horticulture,
University of Wisconsin-Madison, Madison, Wisconsin, USA 53706

AND

LAJOS HORVÁTH

Department of Statistics, University of Wisconsin-Madison

Summary

We approximate the limit process for a multivariate censored survival distribution using the bootstrap. The empirical process has a complicated covariance structure depending on the survival and censoring. The bootstrapped process provides a means to develop distribution-free procedures including simultaneous confidence bands. Results extend to comparison of multivariate survival distributions. A gallstone study is examined in some detail.

Key words: Survival; random censorship; strong approximation; cholecystectomy.

1. Introduction

We show that the bootstrap can be used to approximate the limit process for a multivariate survival distribution based on censored data. Since the limit process has a complicated covariance function which depends on the true survival and censoring distributions, the bootstrapped process provides a means to develop distribution-free procedures leading to simultaneous confidence bands or to comparisons of two multivariate distributions. Examples of multivariate survival distributions include multiple nonfatal pathologies on the same individual, breakdown of parts on a machine, and stages of a progressive disease or crop damage. All endpoints may be observed or censored, making this distinct from the competing risks problem. We apply our results to part of the National Cooperative Gallstone Study (Schoenfield *et al.*, 1981) to examine the efficacy of the drug chenodiol. We compare the bivariate distributions of times to bilary pain and to cholecystectomy for the placebo group and the high dose group.

Self-consistent estimators of multivariate survival distributions were introduced by Hanley & Parnes (1983), Campbell (1981) and Campbell & Földes (1982) based on writing the joint bivariate survival distribution as the product of a univariate survival distribution and a conditional survival distribution. Campbell & Földes (1982) pointed out that an estimator based on this is path dependent and may not even be a distribution function. However, it is asymptotically equivalent to other forms, and converges to a Gaussian process (Burke, 1988; Campbell, 1982). Lo & Wang (1985) showed weak convergence of the self-consistent survival estimator, and suggested a modification to this estimator which guarantees that it is a distribution function. Tsai, Leurgans & Crowley (1984) proposed a kernel estimator which is not path dependent. Burke (1987) suggested another bivariate estimator for randomly censored data, and showed its strong uniform consistency. Wei & Lachin (1984) proposed asymptotically distribution-free tests for equality of multivariate survival distributions based on analogs to linear rank tests. Related references can be found in these works.

We present our results in terms of bivariate distributions; the multivariate results follow in a completely analogous fashion. Section 2 presents results on the weak convergence for the bivariate bootstrap. The two sample problem is addressed in Section 3. Section 4 contains simulations and an application to disease progression. Although our method has been developed for censored data, it works equally well for comparing distributions of two populations when there is no censoring.

2. Weak Convergence of Bivariate Bootstrap

For convenience denote $\min(x,y)$ by $x \wedge y$ and $\max(x,y)$ by $x \vee y$. Also $X_n = o_p(1)$ means $P\{|X_n| > c\} \to 0$ for every positive constant c. Let $\{X_i^0, Y_i^0\}_{i=1}^{\infty}$ be a sequence of independent 2-dimensional random vectors with common continuous survival function $F^0(x,y) = P\{X_i^0 \geq x, Y_i^0 \geq y\}$, which we wish to estimate. In practice, we usually observe the censored vector, $\{X_i, Y_i, \delta_i, \mu_i\}_{i=1}^{n}$. Here, $X_i = X_i^0 \wedge C_i$ and $Y_i = Y_i^0 \wedge D_i$, in which $\{C_i, D_i\}_{i=1}^{\infty}$ is a sequence of independent random vectors, independent of $\{X_i^0, Y_i^0\}_{i=1}^{\infty}$, with common censoring distribution $H(c,d) = P\{C_i \geq c, D_i \geq d\}$. The indicators $\delta_i = I[X_i = X_i^0]$ and $\mu_i = I[Y_i = Y_i^0]$ specify whether the entire lifetime was observed or not.

Denote the survival function of $\{X_i, Y_i\}$ by $F(x,y) = P\{X_i \geq x, Y_i \geq$

$y\}$ and define the empirical survival function at the nth stage as $F_n(x,y) = (1/n)\#\{1 \le i \le n : X_i \ge x, Y_i \ge y\}$. Using the relation

$$F^0(x,y) = P\{X_i^0 \ge x\} P\{Y_i^0 \ge y \mid X_i^0 \ge x\}\,, \tag{2.1}$$

we can define the bivariate product-limit estimator proposed by Campbell (1981) and Hanley & Parnes (1983),

$$\hat{F}_n^0(x,y) = \prod_{\substack{1 \le i \le n,\\ X_i < x}} \left(\frac{nF_n(X_i, -\infty) - 1}{nF_n(X_i, -\infty)} \right)^{\delta_i} \prod_{\substack{1 \le i \le n,\\ Y_i < y}} \left(\frac{nF_n(x, Y_i) - 1}{nF_n(x, Y_i)} \right)^{\mu_i} \tag{2.2}$$

if $F_n(x,y) > 0$, with $\hat{F}_n^0(x,y) = 0$ if $F_n(x,y) = 0$. Interchanging X_i and Y_i in (2.1) and (2.2) yields another, asymptotically equivalent, estimator of F^0 (cf. Lemma 4.4 of Campbell & Földes (1982)). Self-consistency of $\hat{F}_n^0$ was proven by Campbell (1981), Campbell & Földes (1982) and Hanley & Parnes (1983). Assuming regularity conditions on F^0 and H, Campbell & Földes (1982) obtained a rate of uniform almost sure convergence of $\hat{F}_n^0$ to F^0 on the interval

$$T = \{(x,y) : -\infty < x \le x_0 < \infty\,, -\infty < y \le y_0 < \infty \text{ and } F(x_0, y_0) > 0\}\,. \tag{2.3}$$

Horváth (1983), while extending a univariate result of Csörgő & Horváth (1983), generalized this result in two ways by obtaining the same rate on T only assuming that F^0 is continuous, and showing almost sure convergence on the plane (with a different rate). Campbell (1982) proved weak convergence of

$$Z_n(x,y) = n^{1/2}(\hat{F}_n^0(x,y) - F^0(x,y))$$

to a Gaussian process, which was extended by Burke (1988) to a strong approximation summarized in the following:

Theorem 1. *(Campbell, 1982; Burke, 1988). Assume F^0 is continuous and T is defined as in (2.3). Then we can define a sequence of Gaussian processes $\{\Lambda(x,y), \Lambda_n(x,y)\}_{n=1}^{\infty}$ such that*

$$\sup_{(x,y)\in T} |Z_n(x,y) - \Lambda_n(x,y)| = o_p(1)\,,$$

as $n \to \infty$, and for each n, Λ_n has the same distribution as Λ on T.

The covariance function of Λ was calculated by Campbell (1982) and Burke (1988). However, it has a very complicated form which depends on the unknown F^0 and H. Therefore, one cannot directly derive a distribution-free statistical procedure from Theorem 1.

Efron (1981) introduced a bootstrap method in the case of univariate censored data which can be extended to the multivariate case. Given data $\{X_i, Y_i, \delta_i, \mu_i\}_{i=1}^n$, let

$$G_n(\mathbf{z}) = (1/n)\#\{1 \le i \le n : X_i \le z_1, Y_i \le z_2, \delta_i \le z_3, \mu_i \le z_4\}$$

with $\mathbf{z} = (z_1, z_2, z_3, z_4)$. Draw a bootstrap sample $\{X_i^*, Y_i^*, \delta_i^*, \mu_i^*\}_{i=1}^m$, which are conditionally independent with common distribution function G_n. Define the bootstrapped empirical survival function $F_{m,n}(x,y) = (1/m)\#\{1 \le i \le m : X_i^* \ge x, Y_i^* \ge y\}$ and the bootstrapped bivariate product-limit estimator $\hat{F}_{m,n}^0(x,y)$ analogously to (2.2). Our first result is the weak convergence of the bootstrapped bivariate product-limit process,

$$Z_{m,n}(x,y) = m^{1/2}\left(\hat{F}_{m,n}^0(x,y) - \hat{F}_n^0(x,y)\right) .$$

Theorem 2. *Assume the conditions of Theorem 1. Then we can define a sequence of Gaussian processes* $\{\Lambda_{m,n}(x,y)\}_{m,n=1}^\infty$ *such that*

$$\sup_{(x,y)\in T} |Z_{m,n}(x,y) - \Lambda_{m,n}(x,y)| = o_p(1) ,$$

as $m \wedge n \to \infty$ *and* $mn^{-2} \to 0$, *and for each* m, n, $\Lambda_{m,n}$ *has the same distribution as* Λ *on* T.

Horváth & Yandell (1987) proved a strong approximation of $Z_{m,n}$ with a convergence rate for the one-dimensional case.

Our next result leads to Cramèr-von Mises and Kolmogorov-Smirnov type procedures. Let $\mathbf{D}(T)$ be the space of all bounded real-valued functions on G equipped with the supremum norm (cf. Gaenssler, 1983) and let ϕ be a continuous functional defined on $\mathbf{D}(T)$. Theorems 1 and 2 imply that as $m \wedge n \to \infty$ and $mn^{-2} \to 0$, $\phi(Z_n)$ and $\phi(Z_{m,n})$ both converge in distribution to $\phi(\Lambda)$. In other words, the corresponding distributions, $U_n(t) = P\{\phi(Z_n) \le t\}$ and $U_{m,n}(t) = P\{\phi(Z_{m,n}) \le t\}$, respectively, converge to $U(t) = P\{\phi(\Lambda) \le t\}$ at every continuity point of U.

The distribution U can be approximated by the empirical distribution based on a large number of bootstrap samples. Let $\{Z_{m,n}^{(j)}\}_{j=1}^N$ be N bootstrap processes calculated from independent samples of size m drawn from G_n. Introduce the empirical distribution

$$U_{N,m,n}(t) = (1/N)\#\{1 \le j \le N : \phi(Z_{m,n}^{(j)}) \le t\} .$$

By the Glivenko-Cantelli theorem, $|U_{N,m,n}(t) - U_{m,n}(t)| \to 0$ almost surely as $N \to \infty$ for each m and n. Hence, $U_{N,m,n}(t) \to U(t)$ almost surely and consequently, $|U_{N,m,n}(t) - U_n(t)| \to 0$ almost surely as $N \to \infty$ and $m \wedge n \to \infty, mn^{-2} \to 0$, at every continuity point of U. Define the critical value $c(1-\alpha) = \inf\{t \geq 0 : U(t) \geq 1-\alpha\}$, $0 < \alpha < 1$, and define $c_n(1-\alpha)$ and $c_{N,m,n}(1-\alpha)$ similarly for U_n and $U_{N,m,n}$, respectively. In a similar fashion to the proof of Corollary 17.3 of Csörgő *et al.*, (1986),

$$|c_n(1-\alpha) - c_{N,m,n}(1-\alpha)| \to 0 \tag{2.4}$$

almost surely and

$$c_{N,m,n}(1-\alpha) \to c(1-\alpha) \tag{2.5}$$

almost surely as $N \to \infty$ and $m \wedge n \to \infty, mn^{-2} \to 0$, if $U(t)$ is continuous on its open support. The following corollary shows that the Cramèr-von Mises and Kolmogorov-Smirnov statistics have continuous distributions, leading to global statistical procedures based on the bootstrap.

Corollary 1. *Let*

$$\phi(\Lambda) = \sup_{(x,y)\in T} \Lambda(x,y)\,, \quad \text{or} \quad \phi(\Lambda) = \sup_{(x,y)\in T} |\Lambda(x,y)|\,,$$

$$\text{or} \quad \phi(\Lambda) = \iint_{(x,y)\in T} \Lambda^2(x,y) w^2(x,y)\,dx\,dy$$

$$\text{with} \quad \iint_{(x,y)\in T} w^2(x,y)\,dx\,dy < \infty\,.$$

Then $U(t) = P\{\phi(\Lambda) \leq t\}$ is continuous on its open support.

Consider the Kolmogorov-Smirnov functional $\phi(Z) = \sup|Z|$ and a specified null distribution $\tilde{F}^0$. Let $\tilde{Z}_n = n^{1/2}(\tilde{F}_n^0 - \tilde{F}^0)$ and $\tilde{U}_n = P\{\phi(\tilde{Z}_n) \leq t\}$. Note that (2.5) holds as $N \to \infty$ and $m \wedge n \to \infty, mn^{-2} \to 0$ regardless of whether or not $\tilde{F}^0$ is the true distribution. If it is not, then $\phi(\tilde{Z}_n)$ diverges as $n \to \infty$ and $\tilde{U}_n(c_{N,m,n}(1-\alpha)) \to 0$. If $\tilde{F}^0$ is true, then $\tilde{U}_n(c_{N,m,n}(1-\alpha)) \to 1-\alpha$. Thus our bootstrap procedure is asymptotically consistent. The asymptotic consistency of statistical procedures based on other functionals ϕ can be discussed in a similar way.

3. Two Sample Bivariate Bootstrap

Consider samples from two bivariate distributions which we wish to compare. Let $\{{}^{(1)}X_i^0, {}^{(1)}Y_i^0\}_{i=1}^{\infty}$ denote a bivariate sequence with common continuous survival function ${}^{(1)}F^0(x,y) = P\{{}^{(1)}X_i^0 \geq x\,, {}^{(1)}Y_i^0 \geq y\}$. We actually observe the censored data $\{{}^{(1)}X_i, {}^{(1)}Y_i, {}^{(1)}\delta_i, {}^{(1)}\mu_i\}_{i=1}^{n_1}$, defined in an analogous way to Section 2. We define a second sample, independent from the first, in a similar fashion with common continuous survival function ${}^{(2)}F^0$.

We wish to test whether ${}^{(1)}F^0 = {}^{(2)}F^0$, or to estimate the difference ${}^{(1)}F^0 - {}^{(2)}F^0$. We proceed in an analogous manner to Section 2 in defining the empirical survival functions ${}^{(1)}\hat{F}^0_{n_1}$ and ${}^{(2)}\hat{F}^0_{n_2}$. Introduce

$$ {}^{(j)}Z_{n_j}(x,y) = n_j^{1/2}\left({}^{(j)}\hat{F}^0_{n_j}(x,y) - {}^{(j)}F^0(x,y)\right)\,, \quad j = 1,2\,. $$

Because our samples are independent, we have an immediate corollary to Theorem 1.

Corollary 2. *Assume* ${}^{(j)}F^0$, $j = 1,2$, *are continuous,*

$$ \begin{aligned} T^* = \{(x,y) : -\infty < x \leq x_0^*\,, -\infty < y \leq y_0^* \\ \text{and} \quad {}^{(j)}F^0(x_0^*, y_0^*) > 0\,, \quad j = 1,2\}\,. \end{aligned} \tag{3.1} $$

Then we can define two independent sequences of Gaussian processes $\{{}^{(j)}\Lambda(x,y), {}^{(j)}\Lambda_n(x,y)\}_{n=1}^{\infty}$, $j = 1,2$, *such that*

$$ \sup_{(x,y)\in T^*} |{}^{(j)}Z_n(x,y) - {}^{(j)}\Lambda_n(x,y)| = o_p(1)\,, \quad j = 1,2\,, $$

as $n \to \infty$, *and for each* n, j, ${}^{(j)}\Lambda_n$ *and* ${}^{(j)}\Lambda$ *have the same distribution.*

The same problem arises here as in the one-sample case, that the asymptotic covariance functions depend on the survival functions ${}^{(j)}F^0$, $j = 1,2$, and on the censoring distributions, in a complicated way. Therefore it is also natural to use the bootstrap method in the two-sample case. Using similar notation to Section 2, we can draw independent bootstrap samples of sizes m_1 and m_2 from the given data samples. We define the bootstrapped bivariate product-limit estimators ${}^{(j)}\hat{F}^0_{m_j,n_j}$ $j = 1,2$, and the bootstrapped bivariate product-limit processes

$$ {}^{(j)}Z_{m_j,n_j}(x,y) = m_j^{1/2}\left({}^{(j)}\hat{F}^0_{m_j,n_j}(x,y) - {}^{(j)}\hat{F}^0_{n_j}\right)\,, \quad j = 1,2\,. $$

Again, because the samples are independent, and because the bootstrap samples for the two distributions are drawn independently, the following corollary to Theorem 2 is immediate:

Corollary 3. *Assume the conditions of Corollary 2. Then we can define two independent sequences of Gaussian processes* $\{{}^{(j)}\Lambda_{m,n}(x,y)\}_{m,n=1}^{\infty}$, $j=1,2$, *such that*

$$\sup_{(x,y)\in T^*} |{}^{(j)}Z_{m,n}(x,y) - {}^{(j)}\Lambda_{m,n}(x,y)| = o_p(1)\,, \quad j=1,2\,,$$

as $m \wedge n \to \infty$ *with* $mn^{-2} \to 0$, *and for each* m, n *and* j, ${}^{(j)}\Lambda_{m,n}$ *and* ${}^{(j)}\Lambda$ *have the same distribution.*

For simplicity we focus on Kolmogorov-Smirnov type procedures. Suppose that $n_2/(n_1+n_2) \to \lambda$, $0 < \lambda < 1$, as $n_1, n_2 \to \infty$. The maximal deviation

$$\begin{aligned} D_{n_1,n_2} &= \sup_{(x,y)\in T^*} \Bigg| \left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2} [({}^{(1)}\hat{F}^0_{n_1}(x,y) - {}^{(2)}\hat{F}^0_{n_2}(x,y)) \\ &\qquad - ({}^{(1)}F^0(x,y) - {}^{(2)}F^0(x,y))] \Bigg| \\ &= \sup_{(x,y)\in T^*} \Bigg| \left(\frac{n_2}{n_1+n_2}\right)^{1/2} {}^{(1)}Z_{n_1}(x,y) \\ &\qquad - \left(\frac{n_1}{n_1+n_2}\right)^{1/2} {}^{(2)}Z_{n_2}(x,y)) \Bigg| \end{aligned}$$

converges in distribution to

$$D = \sup_{(x,y)\in T^*} |\lambda^{1/2(1)}\Lambda(x,y) - (1-\lambda)^{1/2(2)}\Lambda(x,y)|\,.$$

The distribution of D can be approximated by the empirical distribution based on N pairs of bootstrap samples. Let $\{{}^{(j)}Z^{(k)}_{m_j,n_j}\}_{k=1}^{N}$, $j=1,2$, be $2N$ bootstrap processes calculated from independent samples of size m from the given data. The distribution of the bootstrap maximal deviations,

$$\begin{aligned} D^{(k)}_{m_1,k_2,n_1,n_2} &= \sup_{(x,y)\in T^*} \Bigg| \left(\frac{m_2}{m_1+m_2}\right)^{1/2} {}^{(1)}Z^{(k)}_{m_1,n_1}(x,y) \\ &\qquad - \left(\frac{m_1}{m_1+m_2}\right)^{1/2} {}^{(2)}Z^{(k)}_{m_2,n_2}(x,y)) \Bigg|\,, \end{aligned}$$

converges to that of D provided $m_2/(m_1+m_2) \to \lambda$ as $m_j \wedge n_j \to \infty$ with $m_j n_j^{-2} \to 0$ for $j=1,2$. Note that if ${}^{(1)}F^0 = {}^{(2)}F^0 = F^0$, then D has the same distribution as $\sup_{(x,y)\in T^*} |\Lambda(x,y)|$.

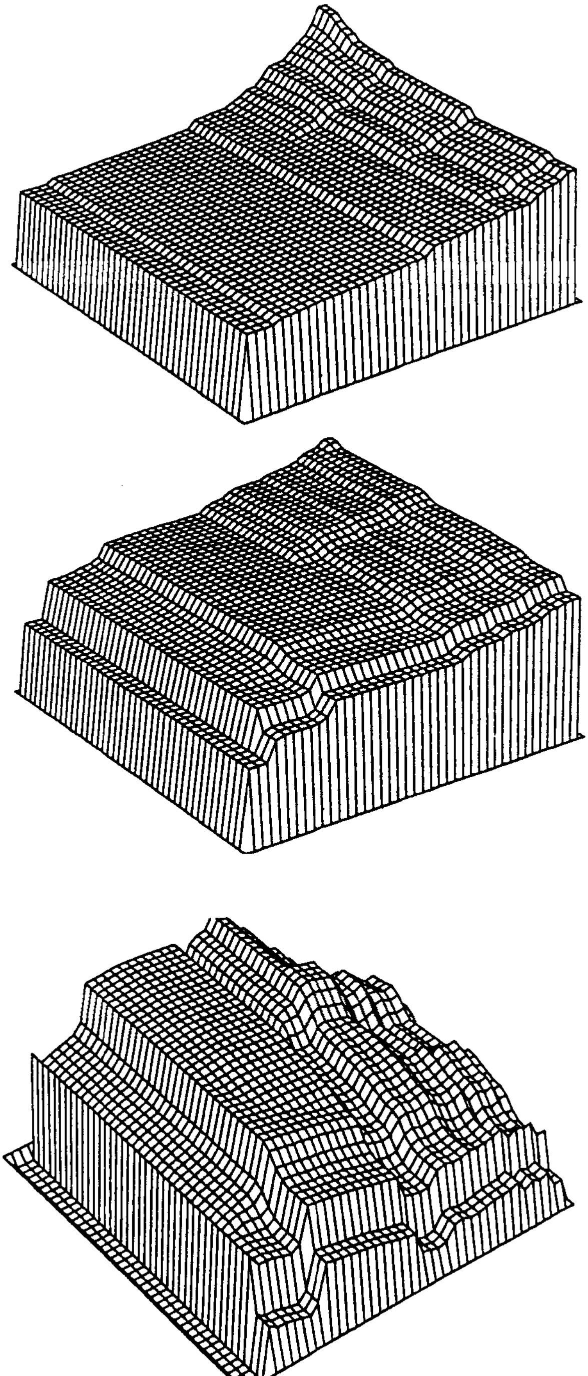
Arguing as in Section 2, we can show that our two sample bootstrap approach based on the empirical distribution of $\{D^{(k)}_{m_1,m_2,n_1,n_2}\}_{k=1}^{N}$ is asymptotically consistent for the Kolmogorov-Smirnov procedure. Procedures based on other functionals with continuous distributions could be developed in a similar fashion.

4. Gallstone Efficacy

We examine data on floating gallstones from the National Cooperative Gallstone Study (Schoenfield *et al.*, 1981) on the effect of chenodiol on the times to bilary pain and to cholecystectomy. We limit attention to a subgroup of 113 patients who had a high expected incidence of efficacy (disappearance of gallstones). The control group of 48 patients only received a placebo, while the treatment group of 65 received a high dose of chenodiol, which was hoped to obviate the need for surgery. We show estimates of the bivariate distributions for both groups and the difference between the distributions. Our test of the difference in distributions is consistent with the earlier findings of Wei & Lachin (1984).

Figure 1(a–b) shows the bivariate survival distributions for the placebo and the high dose groups. The placebo group has a rapid drop in bivariate survival time early, and levels out at 50% survival at about 500 days to bilary pain. The high dose group shows a less marked drop in bivariate survival, with the 50% time occurring at around 750 days to bilary pain. Note that survival never goes very low for either group, as many of the patients had not developed bilary pain or had a cholecystectomy by the end of the study (20/48 censored time to bilary pain in placebo group; 39/65 censored time to bilary pain in the high dose group). The improvement in bivariate survival for the high dose group can be seen in the standardized difference of the survival curves (Figure 1(c)), which is positive over most of the domain. The largest differences occur around 400 days for bilary pain, and up to 600 days for cholestoctomy.

Goodness-of-fit tests or tests for equality of distribution can be performed using the bootstrapped empirical distributions of the maximal deviation statistics. The one-sample and two-sample distributions were based on 1000 independent bootstrap trials with bootstrap samples of size $m_j = kn_j$ for various multiples k (Figure 2(a–c)). Note that the choice of m_j is not too critical, even for small sample sizes. The empirical 90% critical values for the placebo bivariate product limit process were 1.15, 1.12 or 1.15, depending on the choice of $k = 1, 2$ or 4. The empirical 90% critical values for the placebo process were 1.78, 1.85 or 1.79 and for the



(a)

(b)

(c)

Fig. 1.— Placebo group (a) and high dose group (b) survival distributions; (c) normalized differences between high dose and placebo survival distributions. Maximum height for (c) is 1.45. All perspectives run from 0 to 800 days and have a trim of height 0. Bilary pain axis runs back to left; cholestoctomy axis runs back to right.

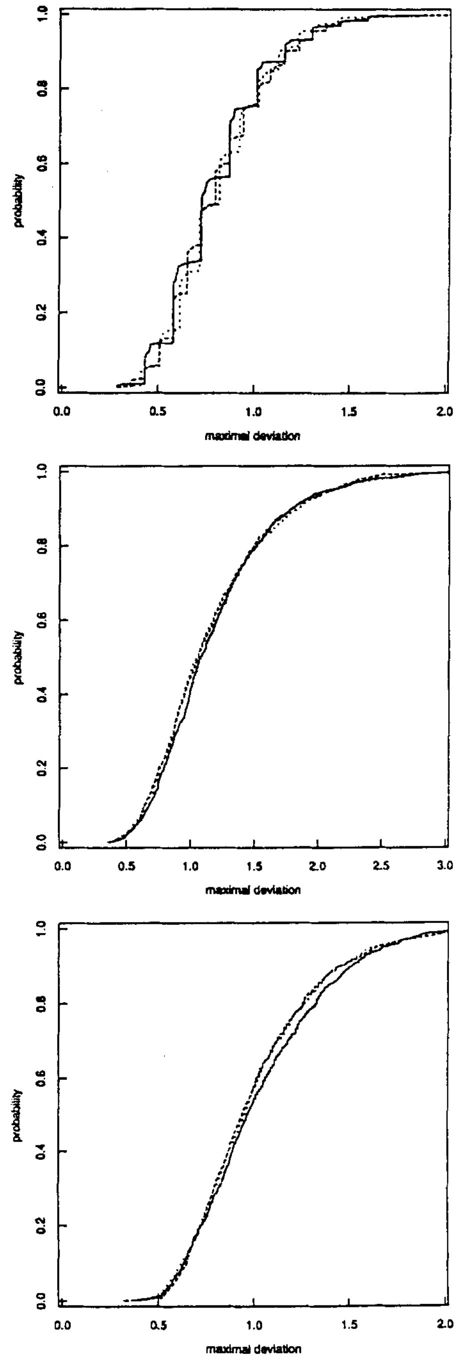


Fig. 2.—Bootstrapped empirical distribution of maximal deviation statistics with $N=1000$, $n_1=48$, $n_2=65$ and $m_j=kn_j$, $k=1$ (solid), 2 (dashed), 4 (dotted): (a) placebo group; (b) high dose group; and (c) two-sample difference.

two sample comparison, the 90% critical values were 1.52, 1.46 or 1.46 for $k = 1, 2$ or 4. A test for difference in distribution between the 2 groups using the normed maximal deviaion of 1.45 yielded a two-sided p-value 0.128, 0.100 or 0.100, for $k = 1, 2$ or 4. Thus we have slim evidence for a difference in bivariate product limit curves. As noted by Wei & Lachin (1984), a test streamlined to detect early differences in bilary pain (as evidenced by the peak difference at 400 days) would have higher power for this situation.

TABLE 1

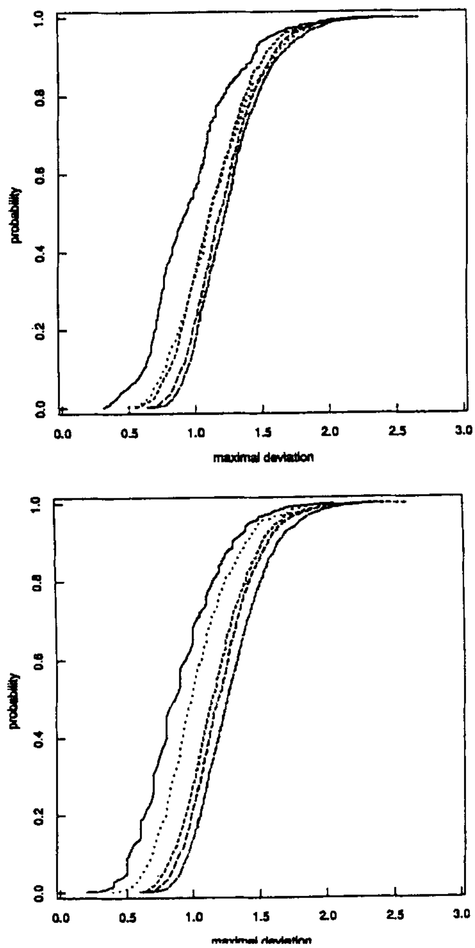
Simulated 90% bootstrap critical values

Maximal deviation for bivariate exponential						
n	m	censoring				
			uniform		exponential	
		0%	20%	50%	20%	50%
20	20	1.34	1.43	1.70	1.46	1.82
		(.03)	(.09)	(.17)	(.08)	(.18)
20	100	1.32	1.43	1.76	1.45	1.80
		(.04)	(.06)	(.19)	(.06)	(.17)
50	50	1.41	1.53	1.97	1.54	2.08
		(.02)	(.05)	(.19)	(.05)	(.20)
50	100	1.41	1.53	1.98	1.54	2.08
		(.03)	(.05)	(.18)	(.05)	(.22)

We were concerned that our estimators were path-dependent and were not strictly distribution functions. We computed the estimates with bilary pain and cholestoctomy times interchanged, and found no appreciable differences in the estimates over the domain.

5. Simulations

Simulations were performed to assess the importance of sample size. Exponential survival data were simulated with 0%, 20% and 50% censoring (uniform and exponential) for sample sizes n = 10, 20, 50, 100 and 200 using N = 1000 Monte Carlo trials with $m = n$ and m = 100. Table 1 shows the 90% critical value for the empirical distribution of the maximal deviation statistic $U_{N,m,n}$ for n = 20, 50; the size of m does not appear



(a)

(b)

Fig. 3.—Simulations of bootstrapped maximal deviation statistics $U_{1000,m,n}$, for single samples from an independent bivariate exponential distribution with 20% exponential censoring, n=10, 20, 50, 100, 200 and (a) m=n or (b) m=100.

to affect the statistic. Figure 3(a) shows realizations of $U_{N,m,n}$ for independent bivariate exponential data with 20% censoring over the range of n with $m = n$, while Figure 3(b) shows similar realizations with $m = 100$. These simulations show that the distribution seems to settle down above $n = 50$, and support the idea of using $m = 100$ in practice for the range of n considered. Similar findings were obtained for 20% and 50% uniform censoring and for 50% exponential censoring, though some realizations with 50% censoring had uniformly large deviations.

6. Appendix

We prove Theorem 2 and Corollary 1 after some preliminary results. For each n we have data $\{X_i, Y_i, \delta_i, \mu_i\}_{i=1}^n$, and for each m and n a bootstrap sample $\{X_i^*, Y_i^*, \delta_i^*, \mu_i^*\}_{i=1}^m$. Define the distribution $G(\mathbf{z}) = P\{X_1 \le z_1, Y_1 \le z_2, \delta_1 \le z_3, \mu_1 \le z_4\}$ and, recalling the definition of G_n, introduce the bootstrap distribution

$$G_{m,n}(\mathbf{z}) = (1/m)\#\{1 \le i \le m : X_i^* \le z_1, Y_i^* \le z_2, \delta_i^* \le z_3, \mu_i^* \le z_4\}$$

and the bootstrap empirical process

$$\alpha_{m,n}(\mathbf{z}) = m^{1/2}\left(G_{m,n}(\mathbf{z}) - G_n(\mathbf{z})\right) .$$

We use the following special case of Theorem 4.3 of Gaenssler (1986).

Theorem A1. *(Gaenssler, 1986). We can define a sequence of Gaussian processes* $\{B(\mathbf{z}), B_{m,n}(\mathbf{z}), \mathbf{z} \in \mathbf{R}^4\}$ *such that*

$$\sup_{\mathbf{z}\in\mathbf{R}^4} |\alpha_{m,n}(\mathbf{z}) - B_{m,n}(\mathbf{z})| = o_p(1)$$

as $m \wedge n \to \infty$, *and for each* m, n, $B_{m,n}(\mathbf{z})$ *and* $B(\mathbf{z})$ *have the same distribution over* $\mathbf{R}^4$. *Further, for* $\mathbf{z}, \mathbf{z}^* \in \mathbf{R}^4$, $EB(\mathbf{z}) = 0$ *and*

$$EB(\mathbf{z})B(\mathbf{z}^*) = G(\mathbf{z} \wedge \mathbf{z}^*) - G(\mathbf{z})G(\mathbf{z}^*) .$$

Recall that $F(x,y) = P\{X_i \ge x\,,\, Y_i \ge y\}$ and define the sub-distribution functions

$$K(x,y) = P\{X_1 < x, Y_1 \ge y, \delta_1 = 1\}$$
$$\text{and} \quad L(x,y) = P\{X_1 \ge x, Y_1 < y, \mu_1 = 1\} .$$

Let K_n and L_n denote the empirical sub-distributions based on the data and $K_{m,n}$ and $L_{m,n}$ denote the empirical sub-distributions based on the bootstrap sample. Theorem A1 implies the joint weak convergence of the following processes:

$$t_{m,n}(x,y) = m^{1/2}\left(F_{m,n}(x,y) - F_n(x,y)\right) ,$$
$$k_{m,n}(x,y) = m^{1/2}\left(K_{m,n}(x,y) - K_n(x,y)\right) ,$$
$$l_{m,n}(x,y) = m^{1/2}\left(L_{m,n}(x,y) - L_n(x,y)\right) .$$

Theorem A2. *We can define three sequences of bivariate Gaussian processes*

$$\left\{(\Gamma, \Gamma^{(1)}, \Gamma^{(2)}), (\Gamma_{m,n}, \Gamma^{(1)}_{m,n}, \Gamma^{(2)}_{m,n})\right\}_{m,n=1}^{\infty},$$

such that

$$\sup_{-\infty<x,y<\infty} |t_{m,n}(x,y) - \Gamma_{m,n}(x,y)| = o_p(1),$$
$$\sup_{-\infty<x,y<\infty} |k_{m,n}(x,y) - \Gamma^{(1)}_{m,n}(x,y)| = o_p(1),$$
$$\sup_{-\infty<x,y<\infty} |l_{m,n}(x,y) - \Gamma^{(2)}_{m,n}(x,y)| = o_p(1),$$

as $m \wedge n \to \infty$, *and for each* m, n, $(\Gamma_{m,n}, \Gamma^{(1)}_{m,n}, \Gamma^{(2)}_{m,n})$ *and* $(\Gamma, \Gamma^{(1)}, \Gamma^{(2)})$ *have the same joint distribution. In addition,*

$$\begin{aligned} E\Gamma &= E\Gamma^{(1)} = E\Gamma^{(2)} = 0 \\ E\Gamma(x,y)\Gamma(s,t) &= F(x \vee s, y \vee t) - F(x,y)F(s,t) \\ E\Gamma^{(1)}(x,y)\Gamma^{(1)}(s,t) &= K(x \vee s, y \wedge t) - K(x,y)K(s,t) \\ E\Gamma^{(2)}(x,y)\Gamma^{(2)}(s,t) &= L(x \wedge s, y \vee t) - L(x,y)L(s,t) \end{aligned}$$

and

$$E\Gamma^{(1)}(x,y)\Gamma^{(2)}(s,t) = P\{s \le X_1 < x, y \le Y_1 < t, \delta_1 = 1, \mu_1 = 1\} - K(x,y)L(s,t).$$

Similar formulae hold for $E\Gamma\Gamma^{(i)}$, $i = 1, 2$.

Proof of Theorem 2. We first consider the bivariate empirical cumulative hazard functions

$$H_n^{(1)}(x,y) = \int_{-\infty}^{x} F_n^{-1}(u,y)\, d_u K_n(u,y),$$
$$H_n^{(2)}(x,y) = \int_{-\infty}^{y} F_n^{-1}(x,v)\, d_v L_n(x,v),$$
$$H_n(x,y) = H_n^{(1)}(x,-\infty) + H_n^{(2)}(x,y),$$

and the bootstrapped cumulative hazard functions $H^{(1)}_{m,n}$, $H^{(2)}_{m,n}$ and $H_{m,n}$, defined in an analogous way using the bootstrap distributions. Let

$$\begin{aligned}\Lambda^{(1)}_{m,n}(x,y) &= \int_{-\infty}^{x} F^{-1}(u,y)\, d_u\, \Gamma^{(1)}_{m,n}(u,y) \\ &\quad - \int_{-\infty}^{x} \Gamma_{m,n}(u,y) F^{-2}(u,y)\, d_u\, K(u,y) \\ \Lambda^{(2)}_{m,n}(x,y) &= \int_{-\infty}^{y} F^{-1}(x,v)\, d_v\, \Gamma^{(2)}_{m,n}(x,v) \\ &\quad - \int_{-\infty}^{y} \Gamma_{m,n}(x,v) F^{-2}(x,v)\, d_v\, L(x,v)\,.\end{aligned}$$

We can prove, in a fashion similar to the one-dimensional case, that

$$\sup_{(x,y)\in T} \left| m^{1/2}\left(H^{(i)}_{m,n}(x,y) - H^{(i)}_{n}(x,y)\right) - \Lambda^{(i)}_{m,n}(x,y)\right| = o_p(1) \qquad \text{(A.1)}$$

for $i = 1, 2$. Combining these two results, we can approximate the cumulative hazard process

$$\left\{ m^{1/2}\left(H_{m,n}(x,y) - H_n(x,y)\right)\,,\ (x,y) \in T \right\}\,.$$

Using an argument similar to Breslow & Crowley (1974), we can show that

$$\begin{aligned}0 &\le -\log \hat{F}^0_{m,n}(x,y) - H_{m,n}(x,y) \\ &\le \tfrac{1}{m} \int_{-\infty}^{x} \left(F_{m,n}(u,-\infty) - 1/m\right)^{-2} d_u\, K_{m,n}(u,-\infty) \\ &\quad + \tfrac{1}{m} \int_{-\infty}^{y} \left(F_{m,n}(x,v) - 1/m\right)^{-2} d_v\, L_{m,n}(x,v)\end{aligned}$$

for all m and n, and $(x,y) \in T$ (cf. Horváth (1983), p.207). Immediately we have

$$\sup_{(x,y)\in T} \left| \hat{F}^0_{m,n}(x,y) - \exp(-H_{m,n}(x,y))\right| = O_p(1/m)\,.$$

By the original Breslow & Crowley (1974) lemma,

$$\sup_{(x,y)\in T} \left| \hat{F}^0_{n}(x,y) - \exp(-H_{n}(x,y))\right| = O_p(1/n)\,.$$

Thus we can approximate the bootstrapped bivariate product-limit process

$$\sup_{(x,y)\in T} \left| Z_{m,n}(x,y) - m^{1/2}\left[\exp(-H_{m,n}(x,y)) - \exp(-H_n(x,y))\right]\right| = o_p(1) \qquad \text{(A.2)}$$

provided $mn^{-2} \to 0$ as $m \wedge n \to \infty$. Define the process

$$\Lambda_{m,n}(x,y) = -F^0(x,y)\big(\Lambda^{(1)}_{m,n}(x,-\infty) + \Lambda^{(2)}_{m,n}(x,y)\big)\,.$$

Using a one-term Taylor expansion we obtain

$$\sup_{(x,y)\in T} \big|Z_{m,n}(x,y) - \Lambda_{m,n}(x,y)\big| = o_p(1)\,.$$

By definition, the distribution of $\Lambda_{m,n}$ does not depend on m or n. Further the distribution of Λ is the same as that obtained by Campbell (1982) and Burke (1988).

Proof of Corollary 1. Arguing similarly to Proposition 17.5 in Csörgő, Csörgő & Horváth (1986), Theorem 1 of Tsirel'son (1975) implies the continuity of the distribution functions of the first two statistics. Consider the third, Cramèr-von Mises type, random variable. It follows from the L_2-decomposability of square-integrable processes (cf. Sato (1969)) that the characteristic function can be expressed as

$$\theta(t) = E\exp\Big(it \iint_{(x,y)\in T} w^2(x,y)\Lambda^2(x,y)\,dx\,dy\Big) = \prod_{k=1}^{\infty}\Big(\frac{\lambda_k}{\lambda_k - 2it}\Big)^{1/2}$$

in which $\lambda_1, \lambda_2, \ldots$ are the reciprocals of the eigenvalues of the covariance function of $w(x,y)\Lambda(x,y)$, with $\lambda_k > 0$, $k \geq 1$. It is easy to show that the function $|t|^q|\theta(t)|$ is bounded on the whole line for any $q > 0$. This implies (Corollary 11.6.1 in Kawata (1972)) continuity of the third distribution function of Corollary 1.

Acknowledgements

This work was supported in part by US Department of Agriculture CSRS grant 511-100, National Science Foundation grant DMS 8404970, and NSERC Canada grants. Computing was performed on the University of Wisconsin Statistics Research VAX 11/750 computer.

References

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

BURKE, M.D. (1987). Estimation of a bivariate distribution function under random censorship. Dept. of Mathematics and Statistics, U. of Calgary.

BURKE, M.D. (1988). An almost-sure approximation of a multivariate product-limit estimator under random censorship. *Statist. Decisions*, (to appear).

CAMPBELL, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* **68**, 417–422.

CAMPBELL, G. (1982). Asymptotic properties of several nonparametric multivariate distribution function estimators under random censorship. In *Survival Analysis*, eds. J. Crowley and R.A. Johnson, pp.243–256. IMS Lecture Notes Series 2.

CAMPBELL, G. & FÕLDES, A. (1982). Large-sample properties of nonparametric bivariate estimators with censored data. In *Nonparametric Statistical Inference, Colloquia Mathematica Societatis János Bolyai*, eds. B.V. Gnedenko, M.L. Puri and I. Vincze, p.32. Amsterdam: North Holland.

CSÕRGŐ, M., CSÕRGŐ, S. & HORVÁTH, L. (1986). *An Asymptotic Theory for Empirical Reliability and Concentration Processes.* Lecture Notes in Statistics, Vol.33. New York: Springer-Verlag.

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

EFRON, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.* **76**, 312–319.

GAENSSLER, P. (1983). *Empirical Processes.* IMS Lecture Notes Series 3.

GAENSSLER, P. (1986). Bootstrapping empirical measures indexed by Vapnik-Chervonenkis classes of sets. In *Prob. Theory and Math. Stat.* Vol.1, eds. Prohorov *et al.*, pp.467–481. VNU Science Press.

HANLEY, J.A. & PARNES, M.N. (1983). Estimation of multivariate distribution in the presence of censoring. *Biometrics* **39**, 129–139.

HORVÁTH, L. (1983). The rate of strong uniform consistency for the multivariate product-limit estimator. *J. Multivariate Anal.* **13**, 202–209.

HORVÁTH, L. & YANDELL, B.S. (1987). Convergence rates for the bootstrapped product limit process. *Ann. Statist.* **15**, 1155–1173.

KAWATA, T. (1972). *Fourier Analysis in Probability Theory.* New York: Academic Press.

LO, S.-H. & WANG, J.-L. (1985). I.I.D. representations of the bivariate product limit estimator. Technical Report, Rutgers U. and U. of California, Davis. (*IMS Bull.* **15**(1), 86t–3.)

SATO, H. (1969). Gaussian measure on a Banach space and abstract Wiener measure. *Nagoya Math. J.* **36**, 65–81.

SCHOENFIELD, L.J., LACHIN, J.M., the Steering Committee and the NCGS Group (1981). Chenodiol (chenadeoxycholic acid) for dissolution of gallstones: the National Cooperative Gallstone Study. *Annals of Internal Medicine* **95**, 257–282.

TSAI, W.-Y., LEURGANS, S. & CROWLEY, J. (1984). Nonparametric estimation of a bivariate survival function in the presence of censoring. *Ann. Statist.* **14**, 1351–1365.

TSIREL'SON, V.S. (1975). The density of the distribution of the maximum of a Gaussian process. *Theory Probab. Appl.* **20**, 847–856.

WEI, L.J. & LACHIN, J.M. (1984). Two-sample asymptotically distribution-free tests from incomplete multivariate observations. *J. Amer. Statist. Assoc.* **79**, 653–661.

Received November 1986; revised October 1987