NONPARAMETRIC INFERENCE UNDER MINIMAL REPAIR

Myles Hollander,* Brett Presnell,*† and Jayaram Sethuraman†

Department of Statistics
Florida State University
Tallahassee, Florida 32306-3033

FSU Technical Report No. M-818
AFOSR Technical Report No. 90-245
USARO Technical Report No. D-111

February 1990

* Research supported by Air Force Office of Scientific Research Grant AFOSR 88-0040.

† Research supported by Army Research Office Grant DAAL03-86-K-0094.

AMS (1980) subject classifcations. Primary 62N05, 90B25; secondary 62E20, 62G05, 62G10, 62G15.

Key words and phrases: Imperfect repair, life distribution, product integral.

Nonparametric Inference Under Minimal Repair

Myles Hollander
Brett Presnell
Jayaram Sethuraman
Department of Statistics
Florida State University
Tallahassee, FL 32306-3033

Abstract

This paper summarizes the results presented at the Army Research Workshop held at Monterey, CA in October, 1989. A more detailed version will appear elsewhere.

In the age-dependent minimal repair model of Block, Borges, and Savits (1985), a system failing at age t undergoes one of two types of repair. With probability $p(t)$, a perfect repair is performed, and the system is returned to the "good-as-new" state, while with probability $1 - p(t)$, a minimal repair is performed, and the system is repaired, but is only as good as a working system of age t. Whitaker and Samaniego (1989) propose an estimator for the system life distribution F when data are collected under this model.

Using the product integral representation of the survival function, a basic result of Block, Borges, and Savits concerning the waiting time until the first perfect repair is extended to allow for discontinuous distributions. Then using counting process techniques, the large sample theorems of Whitaker and Samaniego are extended to the whole line. These results are used to derive confidence bands for F, and to determine a sufficient condition for their applicability on the whole line. Simulation results for the bands are provided. An extension of the Wilcoxon two-sample test to the minimal repair model is also examined.

1 The Minimal Repair Model

To fix notation, let F be a life distribution, let τ_F be the upper endpoint of the support of F (possibly infinite), and let $\Lambda(t) = \int_{(0,t]}(\bar{F}(s-))^{-1}dF(s)$ be the cumulative hazard function of F, where $\bar{F} = 1 - F$.

Now, for $j = 1, \ldots, n$, let $\{X_{j,0} \equiv 0, X_{j,1}, X_{j,2}, \ldots\}$ be independent record value processes from F. These are Markov processes with $P(X_{j,k} > t \mid X_{j,0}, \ldots, X_{j,k-1}) = \bar{F}(t)/\bar{F}(X_{j,k-1})$, for $t > X_{j,k-1}$, $k \geq 1$. If $\Delta F(\tau_F) > 0$, define $X_{j,l} = \infty$ for all l larger than the first k for which $X_{j,k} = \tau_F$. In all cases we take $p(\tau_F) = 1$. These processes represent the failure ages of n systems under a "forever minimal repair" scheme.

Perfect repair is introduced into this model by the use of independent uniform random variables. This facilitates the construction of the σ-field structure (filtrations) necessary to our analysis of the model through martingale methods. Thus we let $\{ U_{j,k} : 1 \leq j \leq n, k \geq 1 \}$, be i.i.d. uniform r.v.'s, and define

$$\delta_{j,k} = I(U_{j,k} \leq p(X_{j,k})),$$
$$\nu_j = \inf\{ k : \delta_{j,k} = 1 \}.$$

Thus observing $\{ (X_{j,1}, \ldots, X_{j,\nu_j}); j = 1, \ldots, n \}$, is equivalent to observing n independent copies of the age-dependent minmal repair process of Block, Borges, and Savits (BBS)(1985), each until the time of its first perfect repair.

This structure provides us with a concrete starting point for a statistical analysis of the BBS model. However, we need conditions which are sufficient to assure the finiteness of X_{j,ν_j}. Such conditions are given by the following result, which generalizes a result of BBS to the case of possibly discontinuous F. Though this generalization may not be important for modeling system failures, it will be useful to us in proving large sample results. Also, the proof of this result, which we sketch below, is more straightforward than the original proof of BBS. The reader is referred to Hollander, Proschan, and Sethuraman (1989) (HPS), for detailed proofs of this and other results in this paper.

Proposition 1 *Let $H(t) = P(X_\nu \leq t, v < \infty)$. Then*

$$\begin{aligned} \bar{H}(t) &= \prod\nolimits_{(0,t]}(1 - d\Lambda_H) \\ &= \exp\left(-\int_{(0,t]} p(s)\frac{dF^c(s)}{\bar{F}(s-)}\right) \prod_{s \leq t}\left(1 - p(s)\frac{\Delta F(s)}{\bar{F}(s-)}\right) \end{aligned}$$

Moreover, if either

(i) $\Delta F(\tau_F) > 0$ (and $p(\tau_F) = 1$),

or

(ii) $F(\tau_F -) = 1$ and $\int_0^{\tau_F} p(s)\frac{dF(s)}{\bar{F}(s-)} = +\infty$,

then H is a proper distribution function and ν is almost surely finite. Conversely, if H is a proper distribution function, then either (i) or (ii) must hold.

Proof. (Sketch) Note that

$$\begin{aligned} \bar{H}(t) &= 1 - P(X_\nu \leq t, \nu < \infty) \\ &= 1 - \sum_{j=1}^{\infty} P(X_j \leq t, \nu = j). \end{aligned}$$

A conditioning argument shows that

$$\frac{\bar{H}(t)}{\bar{F}(t)} = 1 + \sum_{j=1}^{\infty} \int_{0<t_1<\cdots<t_j\leq t} \cdots \int d\alpha(t_1)\cdots d\alpha(t_j),$$

where

$$\alpha(t) = \int_{(0,t]} (1 - p(s)) \frac{dF(s)}{\bar{F}(s)}.$$

This is equivalent to

$$\frac{\bar{H}(t)}{\bar{F}(t)} = \prod\nolimits_{(0,t]}(1 + d\alpha) = \exp\left(\alpha^c(t)\right) \prod_{s \le t}(1 + \Delta\alpha(s)),$$

where α^c is the continuous part of α and $\Delta\alpha(t)$ is the jump in α at t. Here, $\prod_{(0,t]}(1 + d\alpha)$ represents a product integral. The theory of product integration with applications in statistics is reviewed in Gill and Johansen (1987). The result follows from the last equation after some algebra. □

We will say that a pair satisfying either (i) or (ii) describes a *regular* repair scheme.

2 The Whitaker-Samaniego Estimator

In this section, we derive a martingale representation for the Whitaker-Samaniego (1989) estimator (WSE). This representation is then used in conjunction with Rebolledo's Martingale Central Limit Theorem and the techniques of Gill (1983) to derive limit theorems for the WSE.

The Basic Martingale

Define

$$N_j^*(t) = \#\{\, k : X_{j,k} \le t \,\},$$

and

$$\begin{aligned} \mathcal{F}_t &= \sigma\left(\{\, N_j^*(s) : s \le t, 1 \le j \le n \,\}\right) \\ &\quad \vee \quad \sigma\left(\{\, U_{j,k} : k \ge 1, 1 \le j \le n \,\}\right). \end{aligned}$$

For the rest of this paper, $(\mathcal{F}_t)_{t \ge 0}$ will serve as the underlying filtration for all martingales.

Now let

$$N(t) = \#\{\, (j,k) : X_{j,k} \le t, k \le \nu_j, 1 \le j \le n \,\},$$
$$Y(t) = \#\{\, j : X_{j,\nu_j} \ge t, 1 \le j \le n \,\},$$

and

$$M(t) = N(t) - \int_{(0,t]} Y(s)\, d\Lambda(s).$$

In HPS, it is shown that M is a locally square-integrable martingale with predictable quadratic variation given by

$$\langle M \rangle(t) = \int_{(0,t]} Y(s)(1 - \Delta\Lambda(s))\, d\Lambda(s). \tag{1}$$

This provides the basic martingale structure for further analysis of the minimal repair model.

A Martingale Representation for the WSE

Assume that F is continuous and that the pair (F, p) describes a regular repair scheme. Let $X_{(k)}$ be the k^{th} ordered value of the set $\{ X_{j,k} : k \leq \nu_j, 1 \leq j \leq n \}$, let

$$T = \min\{ X_{(k)} : Y(X_{(k)}) = 1 \},$$

and let $J(s) = I(s \leq T)$. Then the Whitaker-Samaniego estimator (WSE) can be written as

$$\hat{\bar{F}}(t) = \prod_{(0,t]}(1 - d\hat{\Lambda}) = \prod_{s \leq t} \left(1 - \Delta\hat{\Lambda}(s)\right),$$

where

$$\hat{\Lambda}(t) = \int_{(0,t]} \frac{J(s)}{Y(s)}\, dN(s).$$

Using Duhammel's equation (Gill and Johansen, 1989), $(\hat{F} - F)/\bar{F}$ can be expressed as an integral with respect to the martingale M:

$$\frac{\hat{F}(t) - F(t)}{\bar{F}(t)} = \int_0^t \frac{\hat{\bar{F}}(s-)}{\bar{F}(s)Y(s)}\, dM(s).$$

From this and (1) it follows $(\hat{F} - F)/\bar{F}$ is itself a locally square-integrable martingale with predictable quadratic variation process given by

$$\left\langle \frac{\hat{F} - F}{\bar{F}} \right\rangle = \int_{(0,t]} \left(\frac{\hat{\bar{F}}(s-)}{\bar{F}(s-)}\right)^2 \frac{dF(s)}{\bar{F}(s)Y(s)}.$$

This quadratic variation process essentially serves to identify the covariance structure of the limiting Gaussian processes derived in the next section.

Large Sample Results

With the above representation, Rebolledo's martingale CLT and the methods of Gill(1983) yield the following result, which extends Theorem 3.3 of Whitaker and Samaniego (1989) to the whole line.

Theorem 1 *Let (F, p) describe a regular repair scheme, with F continuous. Then the following hold:*

(i) *As $n \to \infty$,*

$$\sqrt{n}\,(\hat{F} - F) \xrightarrow{\mathcal{D}} \bar{F} \cdot B(C) \quad \text{in } D[0, \infty],$$

where B is Brownian motion on $[0, \infty)$, and

$$C(t) = \int_0^t \frac{dF(s)}{\bar{H}(s-)\bar{F}(s)}.$$

(ii) *As* $n \to \infty$,

$$\sqrt{n}\,\frac{\bar{K}}{\bar{F}}(\hat{F} - F) \xrightarrow{\mathcal{D}} B^0(K) \quad in\ D[0,\infty],$$

where B^0 *is Brownian bridge on* $[0,1]$*, and* $K = C/(1+C)$.

Details of the proof of this theorem are given in HPS. We note here that the proof of (i) does not require any additional conditions beyond regularity of the repair scheme. This is in contrast with the analogous result of Gill (1983) for the Kaplan-Meier estimator in the usual censored survival data model, where some condition on the amount of censoring is needed. We will see below however, that an additional condition limiting the amount of imperfect repair is needed to assure convergence of the expression in (ii) when an estimate is substituted for $\bar{K}/\bar{F}$.

3 Applications

In this section, the asymptotic results of the last section are used to derive large sample confidence bands for F and to obtain the limiting distribution of an extension of the Mann-Whitney-Wilcoxon test statistic to the minimal repair model.

Confidence Bands

The result in part (ii) of Theorem 1 suggests confidence bands based on the distribution of the supremum of Brownian bridge. It is necessary however to estimate $\bar{K}/\bar{F}$ in order to construct the bands. Let $\widehat{H}$ be the empirical cdf of the X_{j,ν_j}, and let $\widehat{K} = \widehat{C}/(1+\widehat{C})$, where $\widehat{C}$ is defined by

$$\widehat{C}(t) = \int_{(0,t]} \frac{d\hat{F}(s)}{\bar{\widehat{H}}(s-)\bar{\hat{F}}(s)},$$

We would like to have

$$\sqrt{n}\,\frac{\bar{\widehat{K}}}{\bar{\hat{F}}}(\hat{F} - F) \xrightarrow{\mathcal{D}} B^0(K) \quad \text{in } D[0,\infty], \text{ as } n \to \infty, \tag{2}$$

in order to justify asymptotic $(1-\alpha) \times 100\%$ confidence bands for F of the form

$$\hat{F} \pm \sqrt{n}\,\lambda_\alpha \bar{\hat{F}}/\bar{\widehat{K}},$$

where λ_α is the upper α^{th} quantile of the distribution of $\sup|B^0(t)|$.

We can show that (2) holds on $[0,\tau]$ for any $\tau < \tau_F$, but for the complete result, some additional condition seems to be needed. Using the result of Prop.1, it is shown in HPS that $\bar{K}/\bar{F}$ and $\bar{\widehat{K}}/\bar{\hat{F}}$ are nondecreasing and that

$$1 \le \frac{\bar{K}}{\bar{F}} \le \frac{\bar{H}_-}{\bar{F}_-} \qquad \text{and} \qquad 1 \le \frac{\bar{\widehat{K}}}{\bar{\hat{F}}} \le \frac{\bar{\widehat{H}}_-}{\bar{\hat{F}}_-}.$$

Using this, it can be shown that a sufficient condition for (2) is that

$$\frac{\bar{H}(\tau_F-)}{\bar{F}(\tau_F-)} = \int_{(0,\tau_F)} (1 - p(s))\, d\Lambda(s) < \infty.$$

This condition requires that $p(t) \to 1$ as $t \uparrow \tau_F$ (at a rate sufficient for the convergence of the integral), and hence provides a limit on the amount of imperfect repair.

Simulation results for the bands computed over finite intervals (in the case of constant p) indicate that coverage probabilities are quite good for sample sizes of 50 or more. This will of course vary with the parameters of the model. Simulations were carried out with both Gamma and Weibull F, with varying shape parameters, and with various values of p, various interval lengths, and various nominal confidence levels. As an example, the following table gives the simulated coverage probabilities for nominal 95% confidence bands over the interval $[0, 4.744]$ when the underlying F is Gamma with shape parameter 2. (Note that 4.774 is the ninety-fifth percentile of Gamma(2).) More extensive tables are provided in HPS.

n	$p = .50$	$p = .25$	$p = .10$
10	.9025	.8660	.8710
20	.9270	.9125	.9187
30	.9460	.9287	.9327
50	.9515	.9398	.9395
100	.9528	.9540	.9452
200	.9515	.9517	.9495

An Extension of the Mann-Whitney-Wilcoxon Test

Using part (i) of Theorem 1, it is also possible to obtain the limiting distribution for an adaptation of the Mann-Whitney-Wilcoxon two-sample statistic to the minimal repair model. Here we assume that for $i = 1, 2$, we observe n_i BBS processes from (F_i, p_i), each until its first perfect repair. In general we wish to test the null hypothesis $\mathrm{H}_0 : F_1 = F_2$, with typical one-sided alternatives specifying $\int F_1\, dF_2 > 1/2$, and two-sided alternatives specifying $\int F_1\, dF_2 \neq 1/2$.

A statistic analogous to the Mann-Whitney form of the Wilcoxon two-sample test statistic is W, as given by

$$\begin{aligned} W &= \int \widehat{F}_1\, d\widehat{F}_2 \\ &= \sum_{\Delta N_2(s) > 0} \widehat{F}_1(s) \widehat{F}_2(s-) \frac{\Delta N_2(s)}{Y_2(s)}, \end{aligned}$$

where $\widehat{F}_i$ is the WSE, $\Delta N_i(s)$ is the number of failures at age s, and $Y(s)$ is the number of items at risk at age s in the i^{th} sample. This statistic is a natural estimator of $\int F_1\, dF_2 = P(X_1 \leq X_2)$, where X_1 and X_2 are independent random variables, with $X_i \sim F_i$. Assuming continuous distributions, $P(X_1 \leq X_2) = 1/2$ under H_0, and in the one-sided case, significantly large values of W provide evidence against H_0 in the direction of $\int F_1\, dF_2 > 1/2$. For large sample sizes, we have the following result, which is proven in HPS:

Theorem 2 *If F_1 and F_2 are continuous, and the pairs (F_1, p_1) and (F_2, p_2) describe regular repair schemes, and if $n_1, n_2 \to \infty$ in such a way that $\frac{n_1}{n_1+n_2} \to \lambda$, $0 < \lambda < 1$, then*

$$\sqrt{n_1+n_2}\left[W - \int F_1\, dF_2\right] \xrightarrow{\mathcal{D}} N\left(0\,,\, \frac{1}{\lambda}\sigma_1^2 + \frac{1}{1-\lambda}\sigma_2^2\right), \tag{3}$$

where

$$\begin{aligned}\sigma_1^2 &= 2\int_0^\infty \int_t^\infty \bar{F}_1(s)\bar{F}_1(t)C_1(t)\, dF_2(s)dF_2(t),\\ \sigma_2^2 &= 2\int_0^\infty \int_t^\infty \bar{F}_2(s)\bar{F}_2(t)C_2(t)\, dF_1(s)dF_1(t).\end{aligned}$$

Under the null hypothesis, $\mathrm{H}_0 : F_1 = F = F_2$,

$$\sigma_1^2 = 2\int_0^\infty \bar{F}(t)C_1(t)\left(\int_t^\infty \bar{F}(s)\, dF(s)\right)\, dF(t) = \frac{1}{4}\int_0^\infty \frac{\bar{F}^3(s)}{\bar{H}_1(s-)}\, dF(s)\,.$$

For purposes of testing the null hypothesis in the large sample case, we thus propose referring the test statistic

$$Z = \left(W - \frac{1}{2}\right) \Big/ \left(\frac{\widehat{\sigma}_1^2}{n_1} + \frac{\widehat{\sigma}_2^2}{n_2}\right)^{\frac{1}{2}}$$

to a standard normal distribution, where

$$\widehat{\sigma}_i^2 = \frac{1}{4}\int_0^\infty \frac{\widehat{\bar{F}}_i^3(s)}{\widehat{\bar{H}}_i(s-)}\, d\widehat{F}_i(s) = x\frac{1}{4}\sum_{\Delta N_i(s)>0} \frac{n\widehat{\bar{F}}_i^3(s)\widehat{\bar{F}}_i(s-)\Delta N_i(s)}{Y_i^2(s)},$$

and H_i is the empirical distribution of the perfect repair ages in the i^{th} sample.

It is shown in HPS that the σ_i are consistent, which justifies the use of this test. If the p_i are constants (see Brown-Proschan (1983)), the above expressions simplify greatly under H_0. If $F_1 = F_2 = F$, then $\bar{H}_i = \bar{F}^{p_i}$, and the asymptotic variance in (3) reduces to

$$\frac{1}{\lambda}\sigma_1^2 + \frac{1}{1-\lambda}\sigma_2^2 = \frac{1}{\lambda}\left(\frac{1}{4(4-p_1)}\right) + \frac{1}{1-\lambda}\left(\frac{1}{4(4-p_2)}\right).$$

The p_i's are of course consistently estimated by their MLE's, $\hat{p}_i$, the ratio of n_i to the total number of failures in the i^{th} sample, and for large samples, the statistic Z', given by

$$Z' = \left(W - \frac{1}{2}\right) \Big/ \left[\frac{1}{4n_1(4-\hat{p}_1)} + \frac{1}{4n_2(4-\hat{p}_2)}\right]^{1/2},$$

can be referred to a standard normal distribution in order to test the null hypothesis. Note also that if $p_1 = p_2 = 1$, then we are in the usual i.i.d. two-sample model, the WSE's reduce to the empirical c.d.f.'s, and W is just a multiple of the Mann-Whitney form of the Wilcoxon rank-sum statistic. In this case, the above results yield

$$\left(W - \frac{1}{2}\right) \Big/ \left[\frac{1}{12}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]^{1/2} \xrightarrow{\mathcal{D}} \quad N(0,1),$$

in agreement with the usual results for the Mann-Whitney-Wilcoxon test.

REFERENCES

BLOCK, H. W., BORGES, W. S., and SAVITS, T. H. (1985). Age-dependent minimal repair. *J. Appl. Prob.* **22** 370–385.

BROWN, M. and PROSCHAN, F. (1983). Imperfect repair. *J. Appl. Prob.* **20** 851–859.

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

GILL, R. D. and JOHANSEN, S. (1987). Product integrals and counting processes. CWI Report MS-R8707, Centrum voor Wiskunde en Informatica, Amsterdam. To appear in *Ann. Statist.*

HOLLANDER, M., PRESNELL, B., and SETHURAMAN, J. (1989). Nonparametric Methods for Imperfect Repair Models. FSU Technical Report No. M-817, Department of Statistics, Florida State University, Tallahassee, Florida.

WHITAKER, L. R. and SAMANIEGO, F. J. (1989). Estimating the reliability of systems subject to imperfect repair. *J. Amer. Statist. Assoc.* **84** 301–309.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO 23699.26-MA	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) Nonparametric Inference Under Minimal Repair		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER M-818
7. AUTHOR(s) Myles Hollander Brett Presnell Jayaram Sethuraman		8. CONTRACT OR GRANT NUMBER(s) DAAL03-86-K-0094
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306-3033		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE February, 1990
		13. NUMBER OF PAGES 8
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Imperfect repair, life distribution, product integral.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper summarizes the results presented at the Army Research Workshop held at Monterey, CA in October, 1989. A more detailed version will appear elsewhere.

In the age-dependent minimal repair model of Block, Borges, and Savits(1985), A system failing at age t undergoes one of two types of repair. With probability p(t), a perfect repair is performed, and the system is returned to the "good-as-new" state, while with probability 1-p(t), a minimal repair is performed, and the system is repaired, but is only as good as (ctd)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (ctd) a working system of age t. Whitaker and Samaniego (1989) propose an estimator for the system life distribution F when data are collected under this model.

Using the product integral representation of the s rvival function, a basic result of Block, Borges, and Savits concerning the waiting time until the first perfect repair is extended to allow for discontinuous distributions. Then, using counting process techniques, the large sample theorems of Whitaker and Samaniego are extended to the whole line. These results are used to derive confidence bands for F, and to determine a sufficient condition for their applicability on the whole line. Simulation results for the bands are provided. An extnesion of the Wilcoxon two-sample test to the minimal repair model is also examined.