Title	Sensitivity-Analysis and Estimating Number-of-Faults in Removal Debugging
Author(s)	Yip, PSF; Zhejiang, LX; Fong, DYT; Hayakawa, Y
Citation	Ieee Transactions On Reliability, 1999, v. 48 n. 3, p. 300-305
Issued Date	1999
URL	http://hdl.handle.net/10722/43617
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Sensitivity-Analysis and Estimating Number-of-Faults in Removal Debugging

Paul S.F. Yip
The University of Hong Kong, Hong Kong

Liqun Xi
Zhejiang Institute of Silk and Textile, Hangzhou

Daniel Y.T. Fong
The University of Hong Kong, Hong Kong

Y. Hayakawa
Victoria University of Wellington, Wellington

Key Words — Martingale estimating equation, Removal debugging, Seeded fault, Sensitivity analysis.

Summary & Conclusions — Estimating the number of faults in a computer program is important in software debugging. A martingale equation is used to estimate the number of faults in removal-debugging by assuming a known proportionality constant between the failure rate of a 'newly detected fault' and a 'seeded fault'. The sensitivity of the assumption is examined, and the results are generalized to allow an unknown proportionality. The information of the proportionality is shown to be crucial in the precision & availability of the estimates. It is advisable to obtain the information about the proportionality constant from external sources in order to improve the efficiency of the method in this paper.

1. INTRODUCTION

Acronyms and Abbreviations

StdDev standard deviation
ZMM zero-mean martingale

Notation[1]

$\mathcal{F}_t$ history of the process during $[0, t]$
ν number of real faults in the system (parameter of interest)
D number of seeded faults (known)
λ_t, β_t failure intensity for [real, seeded] faults
U_{u-}, M_{u-} number of [real, seeded] faults detected/removed in $[0, u)$
U_t, M_t number of [real, seeded] faults detected/removed in $[0, t]$
$\mathcal{U}_t$, $\mathcal{M}_t$ zero-mean martingale
$\langle \mathcal{U}, \mathcal{M} \rangle_t$ variation process of $\mathcal{U}_t$, $\mathcal{M}_t$
$d\mathcal{R}_u$ see (2)
$\mathcal{R}_t^*$ $\int_0^t W_{u-}\, d\mathcal{R}_u$
θ λ_t/β_t: proportionality
Av($\hat{\nu}$) average of the 2000 simulated values
SD($\hat{\nu}$) StdDev of the 2000 simulated values
τ termination time of the experiment

[1]Other, standard notation is given in "Information for Readers and Authors" at the rear of each issue.

Software debugging usually focuses on estimating the number of indigenous faults. A computer software with ν unknown distinct faults is made to work in a given period of time τ. During τ, the software is examined with a controlled input of data, and any fault in the output data is recorded. This is the data-domain approach in [2], where it is assumed that,

- all indigenous faults are detectable.
- the number of such faults is constant throughout the period of examination.
- the time framework is continuous,
- no two faults can be detected at the same time.

All faults are removed upon detection and thus the reliability of the software increases with time. A traditional model [6] assumed all faults have a constant failure rate with respect to time. Likelihood estimates are available in [7, 8].

Ref [10], however, approaches the problem differently. A known number D of 'seeded' faults are inserted randomly into a computer program before the debugging process begins; see also [11]. The program then consists of both 'real' and 'seeded' faults with failure rates possibly dependent on time. By assuming a known constant proportionality, θ, between the failure rates of the two types of faults, [12] derived a martingale estimator for ν. The θ measures how alike the indigenous and seeded faults are. The more alike they are, the closer is θ to 1; $\theta > 1$ ($\theta < 1$) suggests that failure rate of the seeded fault is smaller (larger) than the real fault. A misspecification of θ would

0018-9529/99/$10.00

lead to serious bias in the resulting estimates of ν. A careful examination and clear understanding of how the estimates of ν behave when θ is misspecified, is therefore highly desirable. The estimation procedure for ν when θ is unknown, has not been studied in the literature. Thus, this paper:

- assesses the sensitivity of the ν estimates to misspecification of θ,
- proposes a new method to estimate ν when no information of θ is available, by using martingale estimating equations.

- Section 2 derives the estimators for ν for a known θ.
- Section 3 provides the estimators for ν and θ.
- Section 4 investigates the sensitivity of the estimators from section 2 to misspecification of θ and gives simulation results for estimating ν and θ simultaneously.
- Section 5 discusses the limitations of the method of the estimating equation in estimating the number of faults in removal debugging.

Assumptions

1. A known number of faults, D, is seeded in the system at the beginning of the experiment.
2. λ_t and β_t can be time-dependent.
3. θ need not be known.
4. The same failure intensity is applied to each type of fault in the system.

5a. Faults are removed (without introducing new faults or affecting the existing faults) immediately after detection.

5b. Faults are detected/removed one at a time. (This allows a continuous time formulation).

6. U_{u-}, M_{u-} are measurable with respect to $\mathcal{F}_{u-}$.

2. MARTINGALE ESTIMATING EQUATION

Assumptions

7. U_t and M_t are right-continuous so that both $dU_t = U_t - U_{t-}$ and $dM_t = M_t - M_{t-}$ take only the values 0 or 1.
8. The model uses multiplicative intensity with $\lambda_t = \theta \cdot \beta_t$ for all t. ◀

This section uses notation similar to that in [12]. Consider a bivariate counting process $\{U_t,\ M_t;\ t \in [0, \tau]\}$. By Doob-Meyer decomposition, each equation in (1) is ZMM [1, 14].

$$\mathcal{U}_t = U_t - \int_0^t \lambda_u \cdot (\nu - U_{u-})\, du \tag{1}$$

$$\mathcal{M}_t = M_t - \int_0^t \beta_u \cdot (D - M_{u-})\, du$$

The $\mathcal{U}_t$ and $\mathcal{M}_t$ are orthogonal:

$\langle \mathcal{U},\ \mathcal{M} \rangle_t = 0.$

Consider the martingale difference:

$$d\mathcal{R}_u = (D - M_{u-})\, dU_u - \theta \cdot (\nu - U_{u-})\, dM_u$$

with $\mathrm{E}\{d\mathcal{R}_u | \mathcal{F}_{u-}\} = 0$ [12]. Consequently the stochastic integral:

$$\mathcal{R}_t^* = \int_0^t W_{u-}\, d\mathcal{R}_u \tag{2}$$

$$= \int_0^t W_{u-} \cdot [(D - M_{u-})\, dU_u - \theta \cdot (\nu - U_{u-}) dM_u];$$

W_{u-} is any locally bounded and predictable process with respect to $\mathcal{F}_{u-}$, and is a ZMM. Equate (2) to zero and evaluate it at time τ; then a class of estimators for ν is obtained:

$$\hat{\nu}_w = \frac{\int_0^\tau W_{u-} [(D - M_{u-})\, dU_u + \theta \cdot U_{u-}\, dM_u]}{\theta \cdot \int_0^\tau W_{u-}\, dM_u} \tag{3}$$

which depends on the choice of W_{u-}. For the simple weight function, $W_{u-} = 1$:

$$\hat{\nu}_1 = \frac{\int_0^\tau [(D - M_{u-})\, dU_u + \theta \cdot U_{u-}\, dM_u]}{\theta \cdot M_\tau}$$

$$= \hat{\nu}_\tau + \frac{\left(\frac{1}{\theta} - 1\right) \cdot \int_0^\tau (D - M_{u-})\, dU_u}{M_\tau} \tag{4}$$

$$\hat{\nu}_\tau \equiv \frac{\int_0^\tau [(D - M_{u-})\, dU_u + U_{u-}\, dM_u]}{M_\tau}.$$

$\hat{\nu}_\tau$ is the estimator of ν when $\theta = 1$ with the condition $\lambda_t = \beta_t$ for all t. Hence, by (4), when $\hat{\nu}_\tau$ is used to estimate ν, but the intensity of seeded faults is smaller or larger than real faults; *ie*,

$\theta = \dfrac{\lambda_t}{\beta_t} > 1$ or < 1,

it leads to over-estimation or under-estimation of ν. This effect from misspecification of θ using (4) is examined further by a Monte Carlo study in section 4.

On the other hand, W_{u-} can be chosen optimally to minimize the size of s-confidence bounds for the estimator for ν. An optimal estimating equation for ν for known θ is [12]:

$$\int_0^\tau \frac{(D - M_{u-})\, dU_u - \theta \cdot (\nu - U_{u-})\, dM_u}{(\nu - U_{u-}) \cdot [(D - M_{u-}) + (\nu - U_{u-}) \cdot \theta]} = 0 \tag{5}$$

The optimal estimator $\hat{\nu}$ is then the solution of (5) with an approximate standard deviation:

$$\mathrm{StdDev}(\hat{\nu}) = \Psi_1(t)/\Psi_2(\tau);$$

$$\Psi_1(t) \equiv \int_0^t \frac{(D - M_{u-})^2\, dU_u + \theta^2 \cdot (\hat{\nu} - U_{u-})^2\, dM_u}{(\hat{\nu} - U_{u-})^2 \cdot [(D - M_{u-}) + \theta \cdot (\hat{\nu} - U_{u-})]^2},$$

$$\Psi_2(\tau) \equiv \int_0^\tau \frac{\theta\, dM_u}{(\hat{\nu} - U_{u-}) \cdot [(D - M_{u-}) + \theta \cdot (\hat{\nu} - U_{u-})]}.$$

For θ known, both the simple weight, $\hat{\nu}_1$, and the optimal weight, $\hat{\nu}$, perform satisfactorily [12].

3. UNKNOWN θ

When θ is unknown, two weight functions can be chosen in (2) such that the resulting estimators of ν and θ are the most s-efficient (optimal) among the estimators derived by other weight functions. These two 'optimal' weight functions are [3]:

$$\frac{1}{(\nu - U_{u-}) \cdot [(D - M_{u-}) + \theta \cdot (\nu - U_{u-})]}$$

for estimating ν;

$$\frac{1}{\theta \cdot [(D - M_{u-}) + \theta \cdot (\nu - U_{u-})]}$$

for estimating θ.

Both of these weight functions are non-decreasing and put more weight on the later part of the experiment. Details of the theory of choosing the optimal weights for a given martingale difference of dimension k are in the appendix; also see [5]. Substitute these optimal weights into (2), and evaluate them at $t = \tau$:

$$\mathcal{R}_1^*(\tau) = \int_0^\tau \frac{(D - M_{u-})\, dU_u - \theta \cdot (\nu - U_{u-})\, dM_u}{(\nu - U_{u-}) \cdot [(D - M_{u-}) + \theta \cdot (\nu - U_{u-})]}; \quad (6)$$

$$\mathcal{R}_2^*(\tau) = \int_0^\tau \frac{(D - M_{u-})\, dU_u - \theta \cdot (\nu - U_{u-})\, dM_u}{\theta \cdot [(D - M_{u-}) + \theta \cdot (\nu - U_{u-})]}. \quad (7)$$

$\hat{\nu}$ and $\hat{\theta}$ can then be obtained by solving (6) & (7) to their mean, zero, simultaneously.

To measure the precision of the estimates, proceed as follows. Let

$\mathbf{v} = (\nu, \theta)'$

$\mathcal{R}(\mathbf{v}) = (\mathcal{R}_1^*(\tau), \mathcal{R}_2^*(\tau))'$.

A Taylor's series expansion of $\mathcal{R}(\mathbf{v})$ at $\mathbf{v} = \hat{\mathbf{v}}$, gives:

$$\mathcal{R}(\mathbf{v}) \approx \mathcal{R}(\hat{\mathbf{v}}) + \mathcal{R}'(\hat{\mathbf{v}}) \cdot (\mathbf{v} - \hat{\mathbf{v}})$$

$\mathcal{R}'(\hat{\mathbf{v}}) \equiv$ first derivative of $\mathcal{R}(\mathbf{v})$ evaluated at $\hat{\mathbf{v}}$.

The error for this approximation is $O\left((\nu - \hat{\nu})^2\right)$, which is negligible when $\hat{\nu}$ is sufficiently close to $\mathbf{v}$. Then, equivalently,

$$(\mathbf{v} - \hat{\mathbf{v}}) \approx \frac{\mathcal{R}(\mathbf{v})}{\mathcal{R}'(\hat{\mathbf{v}})}.$$

Thus a measure of the dispersion of $(\hat{\mathbf{v}} - \mathbf{v})$ is:

$$[\mathcal{R}'(\hat{\mathbf{v}})]^{-1} V(\hat{\mathbf{v}}) [\mathcal{R}'(\hat{\mathbf{v}})]^{-T} \quad (8)$$

$A^{-T} \equiv$ transpose of the inverse of matrix A

$V(\hat{\mathbf{v}}) \equiv$ dispersion matrix of $\mathcal{R}(\hat{\mathbf{v}})$.

Before computing (8), note that:

$$V(\mathbf{v}) = \begin{bmatrix} \langle \mathcal{R}_1^*, \mathcal{R}_1^* \rangle(\tau) & \langle \mathcal{R}_1^*, \mathcal{R}_2^* \rangle(\tau) \\ \langle \mathcal{R}_2^*, \mathcal{R}_1^* \rangle(\tau) & \langle \mathcal{R}_2^*, \mathcal{R}_2^* \rangle(\tau) \end{bmatrix}.$$

$$\langle \mathcal{R}_1^*, \mathcal{R}_1^* \rangle(\tau) = \int_0^\tau d\langle \mathcal{R}_1^*, \mathcal{R}_1^* \rangle(u) = \int_0^\tau \frac{(D - M_{u-})^2 dU_u}{(\nu - U_{u-})^2 \cdot \phi_1} + \int_0^\tau \frac{\theta^2\, dM_u}{\phi_1},$$

$$\langle \mathcal{R}_1^*, \mathcal{R}_2^* \rangle(\tau) = \int_0^\tau d\langle \mathcal{R}_1^*, \mathcal{R}_2^* \rangle(u) = \int_0^\tau \frac{(D - M_{u-})^2\, dU_u}{\theta \cdot (\nu - U_{u-}) \phi_1} + \int_0^\tau \frac{\theta \cdot (\nu - U_{u-})\, dM_u}{\phi_1},$$

$$\langle \mathcal{R}_2^*, \mathcal{R}_2^* \rangle(\tau) = \int_0^\tau d\langle \mathcal{R}_2^*, \mathcal{R}_2^* \rangle(u) = \int_0^\tau \frac{(D - M_{u-})^2\, dU_u}{\theta^2 \cdot \phi_1} + \int_0^\tau \frac{(\nu - U_{u-})^2\, dM_u}{\phi_1}.$$

$$\phi_1 \equiv [(D - M_{u-}) + \theta \cdot (\nu - U_{u-})]^2$$

$$\mathcal{R}'(\mathbf{v}) = \begin{bmatrix} \dfrac{\partial \mathcal{R}_1^*(\tau)}{\partial \nu} & \dfrac{\partial \mathcal{R}_1^*(\tau)}{\partial \theta} \\ \dfrac{\partial \mathcal{R}_2^*(\tau)}{\partial \nu} & \dfrac{\partial \mathcal{R}_2^*(\tau)}{\partial \theta} \end{bmatrix}.$$

$$\frac{\partial \mathcal{R}_1^*(\tau)}{\partial \nu} = -\int_0^\tau \frac{(D - M_{u-}) \cdot \phi_2\, dU_u}{(\nu - U_{u-})^2 \cdot \phi_1} + \int_0^\tau \frac{\theta^2 dM_u}{\phi_1},$$

$$\frac{\partial \mathcal{R}_1^*(\tau)}{\partial \theta} = \frac{\partial \mathcal{R}_2^*(\tau)}{\partial \nu} = -\int_0^\tau \frac{D - M_{u-}}{\phi_1} [dM_u + dU_u]$$

$$\frac{\partial \mathcal{R}_2^*(\tau)}{\partial \theta} = -\int_0^\tau \frac{(D - M_{u-}) \cdot \phi_2\, dU_u}{\theta^2 \cdot \phi_1} + \int_0^\tau \frac{(\nu - U_{u-})^2\, dM_u}{\phi_1}.$$

$$\phi_2 \equiv (D - M_{u-}) + 2\theta \cdot (\nu - U_{u-})$$

Thus the estimated variance-covariance matrix for $(\hat{\mathbf{v}} - \mathbf{v})'$ can be computed from (8) by replacing all the unknown parameters with the corresponding estimates.

4. SENSITIVITY-ANALYSIS AND SIMULATION

Notation

Av(SD($\hat{\nu}$))	average StdDev($\hat{\nu}$) of the 2000 simulated trials
Coverage	proportion of the estimate in each simulated trial lies between the 95% s-confidence limits
P	fraction of the seeded faults removed: the stopping criterion
Prop	fraction of failures in solving for ν and θ, based on (6) – (7).

Assumption

9. The failure intensities for real and seeded faults follow a homogeneous Poisson process. ◁

Monte Carlo methods are used to examine

- (when θ is known) the sensitivity of $\hat{\nu}_1$ in (4) to misspecification of θ,

· (when θ is unknown) the performance of the optimal estimators for ν and θ obtained from (6) – (7).

Table 1 gives the simulation results for the effect of a misspecified θ with $\nu = 400$ and $D = 100$. The simulation size is 2000. The true value is in column #1. The four computed statistics computed are:

$\text{Av}(\hat{\nu})$, $\text{SD}(\hat{\nu})$, Av(SD), Coverage.

When θ is larger than the true value, ν is underestimated. Though it gives a smaller standard error, the coverage was very unsatisfactory.

On the other hand, when θ is smaller than the true value, ν is over-estimated. The coverage was also not satisfactory.

Performance of ν for a correct value of θ improved as P increased from 0.5 to 0.9. The coverage is satisfactory with 95% level.

Table 2 gives the simulation results of estimating ν and θ simultaneously. When the removal proportion is not large (*ie*, $P = 0.5$) the failure proportion in solving for ν and θ, based on (6) – (7) was about 0.44. With a large p, the performance of $\hat{\nu}$ was acceptable. Performance improved with a large value of D. We have included the estimate of ν for a known θ in table 2. The standard deviations of ν were much smaller than the ones derived from an unknown value of θ. Knowing θ is crucial in solving for ν using the martingale equations.

Table 1: Simulation Results for ν with θ Known and the Effect of Misspecification

($\nu = 400$, $D = 400$)
($\tilde{\theta} \Rightarrow$ trial value of θ, Cov $\Rightarrow$ Coverage)

$p = 0.5$

$\tilde{\theta}$	$\text{Av}(\hat{\nu})$	$\text{SD}(\hat{\nu})$	$\text{Av}(\text{SD}(\hat{\nu}))$	Cov
		True $\theta = 0.5$		
.3	640.0	101.0	99.2	.23
.4	494.7	76.3	74.2	.85
.5	408.0	61.5	59.2	.94
1.0	238.0	31.8	29.5	.02
1.5	184.9	22.0	19.8	.00
		True $\theta = 1.0$		
.5	700.2	95.7	94.5	.01
.8	478.0	60.0	58.2	.84
1.0	406.0	48.0	46.2	.94
1.5	314.5	31.9	30.5	.25
2.0	272.8	23.9	22.9	.02
		True $\theta = 1.5$		
.5	911.7	119.1	119.4	.00
1.0	525.0	59.1	57.3	.38
1.5	405.0	38.5	37.2	.94
1.7	378.7	33.6	32.6	.82
2.0	360.7	28.1	27.5	.51

$p = 0.9$

$\tilde{\theta}$	$\text{Av}(\hat{\nu})$	$\text{SD}(\hat{\nu})$	$\text{Av}(\text{SD}(\hat{\nu}))$	Cov
		True $\theta = 0.5$		
.3	570.1	50.5	53.0	.03
.4	462.2	38.0	38.6	.69
.5	401.3	30.3	30.1	.94
1.0	304.4	14.4	14.6	.00
1.5	286.5	10.9	10.7	.00
		True $\theta = 1.0$		
.5	555.2	36.6	38.5	.00
.8	430.5	20.8	20.3	.75
1.0	400.3	15.0	15.0	.92
1.5	374.3	8.8	9.3	.28
2.0	367.1	7.1	7.3	.04
		True $\theta = 1.5$		
.5	626.9	37.5	41.7	.00
1.0	431.0	14.6	13.3	.33
1.5	400.2	6.7	6.9	.91
1.7	396.2	5.5	5.8	.80
2.0	392.9	4.6	4.8	.60

5. DISCUSSION

The sensitivity of $\hat{\nu}_1$ is analyzed for a misspecified θ. Any misspecified value leads to either overestimation or underestimation of ν. The information of θ is therefore important in,

· searching for the solution of ν from the estimating equations,

· determining the precision of the estimates.

The use of estimating equations to estimate $\hat{\nu}$ and $\hat{\theta}$ simultaneously performs satisfactorily only if the removal proportion is large.

During removal debugging of a computer software, it is advisable to obtain more information about θ from some external sources or prior knowledge such that (5) can be used; (5) provides reliable information for ν. Another option is to insert seeded faults as alike as possible to the real faults so that θ can be taken as 1. Otherwise, a sufficient removal proportion of faults is necessary before the number of real faults can be estimated together with θ by using (6) – (7). To decide whether this sufficiency is reached, one could keep debugging until the estimates can be obtained which still provide reasonably good coverage of the corresponding true values. This topic is discussed further in [4, 9, 13].

ACKNOWLEDGMENT

We gratefully acknowledge the useful comments from the three referees. The work was supported by a RGC grant.

APPENDIX

Martingale Estimating Functions

Consider a stochastic process that develops in time.

Table 2: Simulation Results for ν with an Unknown θ

($\nu = 400$, $D = 100$) (Underlined numbers represent the result for ν when θ is known)
(Cov ⇒ Coverage)
(Prop ⇒ proportion of the simulations which cannot provide estimates for ν & θ based on (6) & (7))

$p = 0.7$

θ	Av($\hat{\nu}$)	SD($\hat{\nu}$)	Av(SD($\hat{\nu}$))	Cov(ν)	Av($\hat{\theta}$)	SD($\hat{\theta}$)	Av(SD($\hat{\theta}$))	Cov	Prop
0.5	349.7	114.2	205.8	.74	.7	.31	.42	.96	.28
	<u>401.9</u>	<u>44.1</u>	<u>42.5</u>	<u>.93</u>					
1.0	418.4	105.6	124.5	.85	1.1	.38	.41	.97	.06
	<u>401.5</u>	<u>30.0</u>	<u>29.4</u>	<u>.93</u>					
1.5	414.1	65.7	58.4	.87	1.5	.44	.43	.95	.00
	<u>401.4</u>	<u>20.7</u>	<u>20.2</u>	<u>.93</u>					

$p = 0.9$

θ	Av($\hat{\nu}$)	SD($\hat{\nu}$)	Av(SD($\hat{\nu}$))	Cov(ν)	Av($\hat{\theta}$)	SD($\hat{\theta}$)	Av(SD($\hat{\theta}$))	Cov	Prop
0.5	410.2	97.1	120.2	.84	.5	.17	.20	.96	.08
	<u>400.4</u>	<u>30.2</u>	<u>29.8</u>	<u>.94</u>					
1.0	405.2	37.1	30.7	.88	1.0	.21	.21	.95	.00
	<u>400.2</u>	<u>15.0</u>	<u>15.0</u>	<u>.92</u>					
1.5	401.1	10.9	9.8	.88	1.5	.25	.24	.94	.00
	<u>400.1</u>	<u>6.7</u>	<u>6.8</u>	<u>.91</u>					

Notation

- $\mathcal{F}_t$ history of the stochastic process up to time t
- E_t s-expectation, conditional on the history $\mathcal{F}_t$
- G_t any r.v. depending only on data up to time t
- dG_t the change in G_t over $[t, t+dt)$
- H_t a p-dimensional ZMM
- I_t $p \times p$ random matrix: the information in H_t
- D_t $k \times p$ random matrix; see entries in (A-5)
- V_t $k \times k$ random matrix; see entries in (A-5)

The G_t is ZMM if

$$\mathrm{E}_t\{dG_t\} = 0, \quad \mathrm{E}_t\{dG_t^2\} < \infty \tag{A-1}$$

The G_t is an accumulation of mean-zero finite-variance r.v. More generally, let
$\mathbf{dG}_t = (dG_1, ..., dG_k)^T$ be a martingale difference of dimension k. For each t let
W_t be a $p \times k$ 'weight' matrix which depends only on data up to time t. Then, conditional on $\mathcal{F}_t$, these matrices are non-random; thus

$$H_t = \int_0^t W_{s-}\ \mathbf{dG}_s \tag{A-2}$$

When $\mathbf{dG}_t$ depends on an unknown parameter β of dimension p then the solution $\hat{\beta}_t$ of $H_t = 0$ is a generalization of method of moments estimation. Details of both finite-sample and asymptotic-theory of such estimators are in [3]. They establish mild conditions under which
$I_t^{1/2}(\hat{\beta}_t - \beta)$ converges to multivariate standard Gaussian variable. In the partial ordering of matrices, I_t is dominated by

$$I_t^* = \int_0^t D_s^T \times V_s^{-1} \times D_s\ . \tag{A-3}$$

The weight matrix which achieves this bound is

$$W_{t-}^* = D_t^T \times V_t^{-1}. \tag{A-4}$$

$$D_{i,j} \equiv \mathrm{E}_t\left\{\frac{\partial dG_i}{\partial \beta_j}\right\}, \quad V_{i,j} \equiv \mathrm{E}_t\{dG_i \cdot dG_j\}. \tag{A-5}$$

The optimal set of equations is

$$H_t^* = \int_0^t D_s^T \times V_s^{-1}\ dG_s = 0.$$

Amongst the family of estimating functions (A-2) for given differences dG_t, the optimized function H_t^* has maximal correlation with the true likelihood score function. The asymptotic variance matrix of the optimal estimator $\hat{\beta}_t^*$ is the inverse of the optimal information matrix (A-3). The D_t and V_t can be replaced by their s-expectations in (A-3) whenever:

$$\frac{D_t}{\mathrm{E}_t\{D_t\}} \to 1, \qquad \frac{V_t}{\mathrm{E}_t\{V_t\}} \to 1.$$

REFERENCES

[1] P.K. Andersen, O. Borgan, R.D. Gill, N. Keiding, *Statistical Models Based On Counting Processes*, 1993; Springer-Verlag.

[2] J.S. George, R.R. Ray, "An analysis of competing software reliability models", *IEEE Trans. Software Eng'g*, vol SE-4, 1978, pp 104 – 120.

[3] V.P. Godambe, C.C. Heyde, "Quasi-likelihood and optimal estimation", *Int'l Statistical Review*, vol 55, 1987, pp 231 – 244.

[4] I.B.J. Goudie, "A plant-recapture approach for achieving complete coverage of a population", *Communication in Statistics, A*, vol 24, 1995, pp 1293 – 1305.

[5] C.C. Heyde, *Quasi-Likelihood And Its Application*, 1997; Springer Verlag.

[6] Z. Jelinski, P.B. Moranda, *Software Reliability Research In Statistical Computer Performance Evaluation* (W. Freiberger, *Ed*), 1972, pp 465 – 484; Academic Press.

[7] B. Littlewood, "Stochastic reliability-growth: A model for fault-removal in computer-programs and hardware-designs", *IEEE Trans. Reliability*, vol R-30, 1981 Oct, pp 313 – 320.

[8] B. Littlewood, J.L. Verall, "Likelihood function of a debugging model for computer software reliability", *IEEE Trans. Reliability*, vol R-30, 1981 Jun, pp 145 – 148.

[9] C. Lloyd, P.S.F. Yip, K.S. Chan, "A comparison of recapture, removal and resightly design for population size estimation", *J. Statistics Planning and Inference*, vol 71, 1998, pp 363 – 375.

[10] H.D. Mills, "On the statistical validation of computer programs", *IBM FSD Rep.*, 1970 Jul.

[11] G. McGraw, *Software Fault Injection*, 1998; John Wiley & Sons.

[12] P.S.F. Yip, "Estimating number of errors in a computer program", *IEEE Trans. Reliability*, vol 44, 1995 Jun, pp 322 – 326.

[13] P.S.F. Yip, "Effect of plant capture in a capture-recapture experiment", *Communications in Statistics*, vol 25, 1996, pp 2025 – 2038.

[14] P.S.F. Yip, D.Y.T. Fong, K. Wilson, "Estimating population size by recapturing sampling via estimating function", *Stochastic Models*, vol 9, 1993, pp 179 – 193.

AUTHORS

Dr. Paul S.F. Yip; Dep't of Statistics and Actuarial Science; The Univ. of Hong Kong, HONG KONG.
Internet (e-mail): sfpyip@hku.hk

Paul Yip is a Senior Lecturer in the Department of Statistics and Actuarial Science of the University of Hong Kong. His research interests are in reliability studies, and statistical modeling of biological & ecological systems.

Liqun Xi; Mathematical Section; Zhejiang Inst. of Silk and Textile; Hangzhou, P.R. CHINA.
Internet (e-mail): liqunxi@263.net

Liqun Xi is a lecturer in the Mathematical Section of Zhejiang Institute of Silk and Textile of China. He has a MSc in Mathematics from Zhejiang University. His interests are in applying martingale theory to make inferences on stochastic processes.

Dr. Daniel Y.T. Fong; Clinical Trials Centre, Faculty of Medicine; The Univ. of Hong Kong; HONG KONG.
Internet (e-mail): dytfong@hku.hk

Daniel Y.T. Fong is a Research Assistant Professor at the Clinical Trials Centre of the Faculty of Medicine, The University of Hong Kong. He received a BSc (1st Class Honours) in Mathematics and an MPhil in Statistics from The University of Hong Kong. He received a PhD (1997) in Statistics from the University of Waterloo, Canada in 1997. His research interests are baseline adjustment in clinical trials, analysis of repeated measurements, recurrent-event studies, and lifetime data analysis.

Dr. Y. Hayakawa; Inst. of Statistics and Operational Research; Victoria Univ. of Wellington; Wellington, NEW ZEALAND.
Internet (e-mail): yuhayakawa@vuw.ac.nz

Yu Hayakawa is a lecturer in the School of Mathematical and Computing Sciences at Victoria University of Wellington, New Zealand. She received the an MS (1988) and PhD (1992) in Operations Research from the University of California at Berkeley. Her research interests include reliability theory and Bayes statistics.

Manuscript TR1998-123 received: 1998 August 14;
revised: 1999 May 9, May 20.
Responsible editor: M.R. Lyu
Publisher Item Identifier S 0018-9529(99)07230-9