Software Reliability as an Application of Martingale & Filtering Theory

G. Koch
University of Rome, Rome
P.J.C. Spreij
Mathematisch Centrum, Amsterdam

Key Words—**Software reliability, Martingale, Filtering, Conditional Distribution.**

Reader Aids—
Purpose: Widen state of the art
Special math needed for explanations: Probability, Statistics, Bayesian Inference, Filtering
Special math needed to use results: Probability, Statistics
Results useful to: Software reliability theoreticians

Abstract—**We show how martingale theory provides a comprehensive framework to obtain models for software reliability. Conditional distributions are obtained by solving filtering problems. By means of these distributions several reliability measures can be computed.**

1. INTRODUCTION

It is the purpose of the present paper to show how we can use martingale theory as a tool to describe software reliability measures. Thus we obtain an elegant and comprehensive framework within which we can not only derive existing results but also handle more complex situations in a natural way.

Software reliability deals with the following situation: A computer program which presumably contains a certain number of errors is tested over a given period of time, in order to infer some conclusions about its behaviour when used by future customers.

The basic problem then is to give an appropriate reliability measure and to exhibit a procedure to actually compute this quantity, given the history of the program during the testing period. Several authors have discussed this problem and have proposed certain models [2, 3, 5, 7]. For a survey we refer to [5].

We present a rather general model in which we distinguish between different error classes, and use a Bayes approach. Our model contains existing models. In the computation of reliability measures or the estimation of parameters, filtering problems automatically arise. Their solutions use martingale theory as a natural tool [1, 6].

In section 2 we give a general description of the model. In section 3 we study the situation of a single error class and section 4 treats the case of multiple error classes.

Proofs are not given. For these the reader is referred to [4].

2. GENERAL DESCRIPTION OF THE MODEL

We consider the situation in which a given program, with some possible errors in it is available to a population of customers. When a customer uses it, a path through the program is selected according to the specific customer's input. Along this path some errors might be encountered, which then cause failures in the output. The randomness in selecting the path can be taken into account by assuming that to each error a certain rate constant ν is attached, so that—

$$\Pr\{\text{this error is met in the time interval } (t, t + \Delta t)\}$$

$$= \nu\Delta t + o(\Delta t) \tag{2.1}$$

and by assuming that any two errors are met s-independently one of the other in a given time interval.

The rate constant ν describes at a time the randomness of the path population and of the location of the error in the structure of the program. For example, an error in an output procedure will have a high ν, while an error in a part of a program which is less likely to be used will be given a lower ν.

All possible errors can then be distributed among K classes ($K \leq \infty$) such that the same value of ν_j is shared by all errors in the same class j. Because of the assumption of s-independence among errors, we only have to consider the number X_{jt} of remaining errors in class j at time t. From (2.1) we then have—

$$\Pr\{\text{an error of the } j\text{-th class is met in } (t, t + \Delta t)\}$$

$$= \nu_j X_{jt}\Delta t + o(\Delta t). \tag{2.2}$$

It follows that we can model for each j the failure process n_{jt}, generated by the errors of class j as a counting process with semimartingale representation—

$$dn_{jt} = \lambda_{jt}dt + dm_{jt},\ n_{j0} = 0 \tag{2.3}$$

where $\lambda_{jt} = \nu_j X_{jt}$ and m_{jt} are independent martingales.

For a more compact formulation we define the vector processes—

$$N_t = [n_{1t}, \ldots, n_{Kt}]^T,\ X_t = [X_{1t}, \ldots, X_{Kt}]^T,$$

$$M_t = [m_{1t}, \ldots, m_{Kt}]^T,$$

and the matrix, $A = \operatorname{diag}(\nu_1, \ldots, \nu_K)$. Then (2.3) is summarized by the observation equation:

$$dN_t = AX_t dt + dM_t,\ N_0 = 0. \tag{2.4}$$

It might happen that we are able to observe only the sum, $n_t = \Sigma_1^K n_{jt}$ in which case (2.4) becomes—

0018-9529/83/1000-0342$01.00

$$dn_t = \lambda_t dt + dm_t, n_0 = 0 \tag{2.5}$$

$$\lambda_t \equiv \sum_1^K \lambda_{jt}, m_t \equiv \sum_1^K m_{jt}.$$

Finally we need a dynamic equation for the error process X_t. The error process model has to account both for the mechanism by which errors are corrected after detection of a failure and for a random or deterministic character of involved parameters. Thus leaving aside possibly unknown parameters in a priori distributions, the relevant quantities are the initial numbers of errors in each class and the rate constants ν_j. A discussion on this point appeared in [5]. In our opinion there is enough motivation to assume a random character for all of them, at least unless one deals with just one program and/or one well-defined input sequence. As far as the correction mechanism is concerned, throughout this paper we assume immediate full correction of an error which has caused an observed failure. This amounts to saying that $dX_t = -dN_t$ thus leading to the following model of the error process:

$$dX_t = -AX_t dt - dM_t. \tag{2.6}$$

Indeed this error/failure model (2.4), (2.6) is more general than other existing models that assume the same error correction mechanism, such as Goel & Okumoto [2], Jelinski & Moranda [3], Littlewood [5].

The computation of reliability measures at time t, such as the number of errors left or mean time to next failure, involves (apart from knowledge of parameters in a priori distributions) estimation of the random quantities X_t, λ_t, ν_j, given the knowledge of N_s, $0 \le s \le t$ or n_s, $0 \le s \le t$. Stated otherwise, we have to solve a filtering problem. For several cases this problem is precisely stated and solved in the next sections.

3. RESULTS FOR ONE-ERROR CLASS

In this case ($K = 1$) our model (2.4), (2.6) becomes

$$dn_t = \nu X_t dt + dm_t, n_0 = 0, \tag{3.1}$$

$$dX_t = -\nu X_t dt - dm_t, X_0. \tag{3.2}$$

Assume (X_0, ν) to be a random quantity which has a priori pdf p_0 with respect to the measure μ on $\mathbb{N} \times \mathbb{R}_+$, μ being the product of counting and Lebesgue measure. The filtering problem then is to give an expression for the a posteriori pdf p_t of (X_t, ν), conditioned upon $\{n_s, 0 \le s \le t\}$. The solution of this problem involves deterministic parameters of the a priori pdf p_0. If these are not known they can be estimated according to the maximum likelihood principle.

The likelihood functional Λ_t in this situation is known to take the form [1]:

$$\Lambda_t = \exp[-\int_0^t (\hat{\lambda}_s - 1)ds + \int_0^t \log \hat{\lambda}_{s-} dn_s], \tag{3.3}$$

where $\hat{\lambda}_s$ is the conditional s-expectation of λ_s given n_σ, $0 \le \sigma \le s$: $\hat{\lambda}_s = E\{\lambda_s | n_\sigma, 0 \le \sigma \le s\}$, which can be computed as soon as we know p_s: $\hat{\lambda}_s = \int x\nu p_s(x, \nu)d\mu(x, \nu)$.

The following theorem is the main result.

THEOREM. Suppose that n_s, $0 \le s \le t$ is observed and that at the times $t_j (j = 1, \ldots, n_t)$ a failure occurred. Given these observations we have for the conditional pdf p_t of (X_t, ν)

$$p_t(x, \nu) = cp_0(x + n_t, \nu) \frac{(x + n_t)!}{x!} \nu^{n_t} \exp[-\nu(xt + \sum_j t_j)] \tag{3.4}$$

where c is such that $\int p_t(x, \nu)d\mu(x, \nu) = 1$.

PROOF. This is an application of appropriate martingale techniques [4].

3.1 Specific cases.

Specific cases can be derived from (3.4) according to different choices of p_0. These are discussed below.

i. Assume that X_0, ν are s-independent random variables, $X_0 \sim \text{bin}(N, p)$, $\nu \sim \Gamma(\alpha, \beta)$:

$$p_0(x, \nu) = \binom{N}{x} p^x(1 - p)^{N-x} \beta^\alpha \nu^{\alpha-1} e^{-\beta\nu}/\Gamma(\alpha),$$

then one obtains

$$p_t(x, \nu) = c_1 \binom{N - n_t}{x} p^x(1 - p)^{N-n_t-x} \nu^{\alpha+n_t-1} e^{-\nu(\beta+xt+\Sigma_j t_j)}. \tag{3.5}$$

From (3.5) it is clear that (X_t, ν) are not s-independent anymore for $t > 0$; however the marginal distribution for ν is a binomial mixture of gamma distributions. Conditional mean values for X_t and λ_t can be easily computed from (3.5).

ii. If we choose in i instead of a bionomial distribution a Poisson (μ) distribution then we get instead of (3.5) its limiting expression for $N \to \infty$, $Np \to \mu$:

$$p_t(x, \nu) = c \frac{\mu^x}{x!} e^{-\mu} \nu^{\alpha+n_t} e^{-\nu(\beta+xt+\Sigma_j t_j)}. \tag{3.6}$$

iii. Suppose ν is taken to be deterministic, say ν_0 and X_0 has an a priori Poisson (μ) distribution. Then it results from (3.4) that the a posteriori distribution of X_t is Poisson (μ_t), where $\mu_t = \mu e^{-\nu_0 t}$. Hence $\hat{\lambda}_t = \nu_0 \hat{X}_t = \nu_0 \mu e^{-\nu_0 t}$, independent of n_t. We conclude that this situation actually describes the Goel & Okumoto model [2].

iv. Assume that one is dealing with a population of just one program. Then X_0 may well be taken as deterministic. This amounts to $p = 1$ in case i. Now we see from (3.5) that the distribution of X_t is degenerate in the value $N - n_t$, and we are left with a pdf for ν which is gamma Γ (α

$+ n_t, \beta + (N - n_t)t + \Sigma_j t_j)$. Then $\hat{\lambda}_t = (N - n_t)(\alpha + n_t)/(\beta + (N - n_t)t + \Sigma_j t_j)$. Using this result we compute the reliability measure consisting of the pdf of the time θ_t to the next failure after t. Since given a certain value of ν, the density for θ_t takes the form $\nu X_t e^{-\nu X_t \theta}$, $\theta > 0$, we have the following expression for the posterior pdf of θ_t, given the history of the failure process:

$$\int_0^\infty \nu X_t e^{-\nu X_t \theta} p_t(\nu) d\nu = \frac{(\alpha + n_t)(N - n_t)(\beta + (N - n_t)t + \Sigma_j t_j)^{n_t + \alpha}}{(\beta + (N - n_t)(\theta + t) + \Sigma_j t_j)^{n_t + \alpha + 1}}. \quad (3.7)$$

From (3.7) follows that, if the conditional mean of θ_t exists, it happens to be—

$$\hat{\theta}_t = \frac{\beta + (N - n_t)t + \Sigma_j t_j}{(n_t + \alpha - 1)(N - n_t)}. \quad (3.8)$$

In all these cases maximum likelihood estimates of the parameters in the a priori distributions (if they are unknown) can be obtained by substituting the relevant expression for $\hat{\lambda}_t$ in (3.3).

4. RESULTS FOR MULTIPLE ERROR CLASSES

The procedure in the previous section, where we treated the situation with only one error class ($K = 1$), also applies if we have to deal with multiple error classes ($K > 1$, even $K = \infty$). Explicit formulas for the solution p_t of the filtering problem for $K > 1$ can be found in [4]. In section 4.1 we list some examples.

The remainder of this section is devoted to the restricted situation where each error class contains at most one element. We then have the following important properties [4], which we state as a—

THEOREM

a. Let $K < \infty$. Assume that the pairs (X_{j0}, ν_j), $j = 1, \ldots, K$ are s-independent under the initial distribution p_0. Then they keep being s-independent for each $t > 0$ under the posterior distribution p_t. Moreover all pairs (X_{jt}, ν_j) corresponding to a class where no failure has occurred are identically distributed (if they were such under p_0). Thus it is sufficient for the joint posterior pdf of $(X_t, \nu)(\nu = [\nu_1, \ldots, \nu_K])$ to specify the posterior pdf of the (X_{jt}, ν_j)'s separately.

b. For reliability purposes and maximum likelihood estimation it is sufficient to know n_s, $0 \le s \le t$ only instead of the complete failure process N_s, $0 \le s \le t$.

4.1 Specific Cases

In all of the following cases we assume to have observed n_s, $0 \le s \le t$ and we denote by $t_j \le t$ $(j = 1, \ldots, n_t)$ the time at which a failure occurred in the ℓ_j-th class.

i. Let $K < \infty$. Assume that (X_{j0}, ν_j) are independent for $j = 1, \ldots, K$ and identically distributed: $X_{j0} \sim \text{Bin}(1, p)$, $\nu_j \sim \Gamma(\alpha, \beta)$.

$$p_0(X_j, \nu_j) = p^{x_j}(1 - p)^{1 - x_j} \beta^\alpha \nu_j^{\alpha - 1} e^{-\beta \nu_j} / \Gamma(\alpha). \quad (4.1)$$

The posterior density $p_t^{\ell_j}$ of each of the pairs $(X_{\ell_j t}, \nu_{\ell_j})$ is degenarate for $X_{\ell_j t}$ at $X_{\ell_j t} = 0$ and for ν_{ℓ_j} we have a gamma $\Gamma(\alpha + 1, \beta + t_j)$ distribution. Hence $\hat{\nu}_{\ell_j} = (\alpha + 1)/(\beta + t_j)$. This reflects the easily understood circumstance that errors which are debugged first are expected to have a larger rate constant. For the other pairs $(X_{jt}, \nu_j)(j \notin \{\ell_1, \ldots, \ell_{n_t}\})$ we have for each of them the posterior density p_t^j:

$$p_t^j(x_j, \nu_j) = \frac{\nu_j^{\alpha - 1} e^{-\nu_j(\beta + t x_j)}}{\Gamma(\alpha)} \beta^{\alpha(1 - x_j)} (\beta + t)^{\alpha x_j} p(t)^{x_j} (1 - p(t))^{1 - x_j} \quad (4.2)$$

where

$$p(t) = \frac{p\beta^\alpha}{p\beta^\alpha + (1 - p)(\beta + t)^\alpha}.$$

The interpretation is $p(t) = \text{Prob}\{X_{jt} = 1\}$. As a consequence we have $\Sigma_j X_{jt}$ is conditionally $\text{Bin}(K - n_t, p(t))$. Hence we estimate it by $\widehat{\Sigma_j X_{jt}} = (K - n_t)p(t)$. Furthermore $\hat{\lambda}_t = \alpha(K - n_t)p(t)/(\beta + t)$. As in case (iv) of section 3.1 we can compute the a posteriori distribution function of θ_t:

$$P(\theta_t \le \theta) = 1 - \left(p(t)\left(\frac{\beta + t}{\beta + t + \theta}\right)^\alpha + 1 - p(t)\right)^{K - n_t} \quad (4.3)$$

which is defective unless $p = 1$.

ii. Similar to case *iv* in 3.2, when dealing with one single program, we can derive from the previous case by setting $p = 1$ the following conditional distributions $\nu_{\ell_j} \sim \Gamma(\alpha + 1, \beta + t_j)$, $\nu_j \sim \Gamma(\alpha, \beta + t)(j \notin \{\ell_1, \ldots, \ell_{n_t}\})$, $\Sigma_j X_{jt}$ degenerated at $K - n_t$, and $\lambda_t \sim \Gamma(\alpha(K - n_t), \beta + t)$. This case is indeed the model discussed by Littlewood [5].

iii. This case allows K to be infinite. Here we consider of course only conditional pdf's for λ_t. A natural choice for p_0, the initial pdf for λ_0, follows from (i) by assuming that $\Sigma_j X_{j0} \sim \text{Poisson}(\mu)$ and that λ_0 conditioned upon $\{\Sigma_j X_{j0} = h\}$ is $\Gamma(\alpha h, \beta)$:

$$p_0(\ell) = \sum_h \frac{\mu^h}{h!} e^{-\mu} \Gamma(\alpha h, \beta). \quad (4.4)$$

Then we have as a result that $\Sigma_j X_{jt} \sim$ Poisson $(\mu(t))$ and

$$p_t(\ell) = \sum_h \frac{(\mu(t))^h}{h} e^{-\mu(t)} \Gamma(\alpha h, \beta + t) \quad (4.5)$$

where $\mu(t) \equiv \mu\beta^\alpha/(\beta + t)^\alpha$. Hence $\hat{\lambda}_t = \alpha\mu(t)/(\beta + t)$.

As in section (3.2) maximum likelihood estimates of the parameters in the a priori distributions can be obtained by substituting the relevant expression for $\hat{\lambda}_t$ in (3.3).

5. CONCLUSIONS

In this paper we showed how a martingale approach can be usefully adopted in formulating models for the error and failure processes under different circumstances and in solving estimation problems and computation of reliability measures. This approach provides a unifying theoretical framework which not only contains known results, but also yields new results.

ACKNOWLEDGMENTS

We are pleased to express our thanks to H. Kwakernaak, G.J. Olsder, and R.C.W. Strijbos for several enlightening discussions, as well as to G. Moek and J.T. v.d. Hoven of NLR for providing us some insight in the daily life practice of program testing. Also we gratefully acknowledge the referees and J.H. van Schuppen for helpful comments.

This work was carried out when both authors were with the Dept. of Applied Math., Twente Univ. of Technology, The Netherlands.

A related paper was presented at the 11-th Conf. on Stochastic Processes and Their Applications in 1982 at Clermont Ferrand, France. A related report appeared internally as Research Memorandum #384 in the Dept. of Mathematics at the Twente Univ. of Technology.

REFERENCES

[1] P. Bremaud, *Point Processes and Queues*, New York: Springer Verlag, 1981.

[2] A.L. Goel, K. Okumoto, "Time-Dependent Error-Detection Rate Model for Software Reliability and other Performance Measures," *IEEE Trans. Reliability*, vol R-28, 1979 Aug, pp 206-211.

[3] Z. Jelinski, P.B. Moranda, "Software Reliability Research," in *Statistical Computer Performance Evaluation*, W. Freiberger, Ed. Academic Press, 1972 pp 465-484.

[4] G. Koch, and P.J.C. Spreij, *A Martingale Approach to Software Reliability*, Twente University of Technology, Technical Report no. 384, 1982 April.

[5] B. Littlewood, "Theories of software reliability: How good are they and how can they be improved?", *IEEE Trans. Software Engineering*, vol SE-5, 1980 Sep, pp 489-500.

[6] R.S. Liptser, A.N. Shiryayev, *Statistics of Random Processes*, vol II: *Applications*, Springer Verlag, 1978.

[7] J.D. Musa, "A theory of software reliability and its application," *IEEE Trans. Software Engineering*, vol SE-1, 1975 Sep, pp 312-327.

AUTHORS

G. Koch; Instituto Matematico "G. Castelnuovo"; Universita di Roma; Citta Universitaria; 00100 Roma, ITALY.

G. Koch was born in 1942. In 1966 he received a degree in Electronic Engineering from the University of Rome, Italy and in 1971 the MS degree in System Optimization from the University of California, Los Angeles. Presently he is a professor in Probability at the University of Rome. Recently he spent some months at the Twente University of Technology, Enschede, The Netherlands. His interests are in Reliability, Stochastic Processes, Filtering and Biological Systems. He is a member of IEEE.

P.J.C. Spreij; Mathematisch Centrum; POBox 4079; 1009 AB Amsterdam, The Netherlands.

P.J.C. Spreij was born in 1956. In 1979 he received a degree in Mathematics from the Free University, Amsterdam, The Netherlands. He was with the Free University, Amsterdam, The Twente University of Technology, Enschede, The Netherlands, and presently he is with the Mathematisch Centrum, Amsterdam, The Netherlands. His interests are in Reliability, Stochastic Processes, Filtering and Estimation.

Manuscript TR82-53 received 1982 May 25; revised 1983 October 5.

★★★

(continued from page 341)

Pragmatic Testing Protocols

AUTHORS

Dr. Peter Kubat; Dewey Hall; Graduate School of Management; The University of Rochester; Rochester, New York 14627 USA.

Peter Kubat is an Assistant Professor of Operations Research and Operations Management at the Graduate School of Management, University of Rochester. He received a D.Sc. in Statistics from the Technion, Israel Institute of Technology. His research interests include order statistics, the theory of extreme values and its applications, life testing, quality control, reliability and applied stochastic processes.

Dr. Harvey S. Koch; General Electric Corp., Valley Forge Space Center, P.O. Box 8555, Philadelphia, PA 19101, USA.

Harvey S. Koch is a System Engineer at the Valley Forge Space Center, General Electric Corp., Philadelphia. He received a Ph.D. in Computer Science from the Pennsylvania State University. His current research interests are how managers can use software quality measures for effective decision-making and development of new computer auditing capabilities for time-sharing systems. He is an affiliate member of IEEE and a member of ACM and IIA.

Manuscript TR82-55 received 1982 June 1; revised 1983 June 16. ★★★