A STOPPING RULE FOR TIME-DOMAIN SOFTWARE TESTING

Mehmet Sahinoglu, IEEE Senior Member; Abdülsattar Al-Khalidi,
Troy State University Montgomery (TSUM), Montgomery, AL 36103, mesa@tsum.edu
Department of Statistics, Dokuz Eylül Univ., Izmir, Turkey; salim@gediz.sci.deu.edu.tr

ABSTRACT

An efficient and economical stopping rule using empirical-Bayesian principles for the Poisson counting process compounded with logarithmic-series distribution(LSD) for clump-size is derived and satisfactorily applied to time-domain sequential software testing. The resulting compound Poison distribution for X (number of failures) is also known as a Negative Binomial distribution(NBD) provided a certain underlying assumption[2]. For each checkpoint in time, either the software satisfies a desired reliability attached to an economic criterion, or else the software testing is allowed to continue. The proposed stopping rule is applied to five different software failure calendar-data in terms of weekly intervals in the time-domain problem. Sensitivity studies are conducted by varying the cost-criteria and Bayesian prior parameters of a generalized beta.

SEQUENTIAL STOPPING RULE AND NUMERICAL APPLICATIONS

If the expected incremental difference between sequential steps i=1,2,.. where i denotes testing interval in terms of days, weeks in time-domain or test-cases in effort-domain etc. exceeds a given economic criterion “d”, then continue testing. Otherwise stop testing . Observe below in one-step-ahead formula, where α, β are prior parameters, k is the NBD parameter and x_i is the cumulative of failures until step i, X is the for total number of failures.

$$e(x)=E(X_{i+1})-E(X_i)\leq d \qquad (1)$$

Equation(1) can be rearranged into Equation(2) by utilizing Equation(3), [3]

$$e(x)=k_{i+1}\frac{(\alpha+x_{i+1})}{(\beta+k_{i+1})}-k_i\frac{(\alpha+x_i)}{(\beta+k_i)}\leq d \qquad (2)$$

$$E(X)=k\left(\frac{\alpha+x}{\beta+k}\right) \qquad (3)$$

whereas incorporating the Generalized Beta prior, as in Equation(2), we get (4) below:

$$e(x)=k_{i+1}\frac{(\theta_2-\theta_1)(\alpha+x_{i+1})}{(\alpha+\beta+x_{i+1}+k_{i+1})-(\theta_2-\theta_1)(\alpha+x_{i+1})}-$$

$$k_i\frac{(\theta_2-\theta_1)(\alpha+x_i)}{(\alpha+\beta+x_i+k_i)-(\theta_2-\theta_1)(\alpha+x_i)}\leq d$$

where, $d=\dfrac{c}{a-b}$ and α, β, k_i, x_i, θ_2 (upper correlation limit) and θ_1(lower correlation limit) are all input values at each step i where c= cost of each unit time testing, a, b= costs per failure correction after and before release. Note that Equation(4) defaults to Equation(2) for θ_1 =0 and θ_2=1 when no constraint is assumed on the correlation among failures. The narrowing of the correlation span will provide more information than none. The trend will be obvious in Table 2. The following Tables 1 and 2 will tabulate input and output for the 3 weekly Data Sets WD1,3,5 under varying parameters. Stopping occurs earlier if d is reduced. Other parameters are not as effective; no obvious trends. However one discovers The independent weekly simulated JPL data: WD1 (Galileo), WD3 (Alaska), WD5 (Alaska SAR) respectively have a total of 131 failures in 64 weeks, 340 failures in 41

when to stop.

weeks and 366 failures in 50 weeks [4]. The rule in Table 2. commands when to stop, or start next testing strategy, instead of indefinite testing until the end of mission.

Data	E(w)	σ(w)	Var(w)	E(w^2)	q	θ	α	β
WD1	2.047	2.675	7.155	11.345	5.54	0.82	8.2(5)	1.8(1)
WD3	8.293	7.16	51.26	102.53	14.32	0.92	9.2(15)	0.8(3)
WD5	7.32	4.254	18.09	7.167	9.79	0.90	9.0(4)	1.0(1)

Table 1. α, β in WD1, WD3, WD5 for Cases 1(idealized) & 2(Method of Moments)

			$\theta_2 - \theta_1$=1.0	$\theta_2 - \theta_1$=.8	$\theta_2 - \theta_1$=.5	$\theta_2 - \theta_1$=.2
α=8.2	β=1.8	d=2x10^{-3}	X(12)=25	X(30)=74	X(38)=85	X(21)=49
α=8.2	β=1.8	d=2x10^{-7}	X(13)=33	X(34)=82	X(49)=121	X(55)=128
α=5	β=1	d=2x10^{-3}	X (9) =22	X(30)=74	X(38)=85	X(21) =49
α=5	β=1	d=2x10^{-7}	X (9) =22	X(34)=82	X(50)=124	X(56)=129
α=9.2	β=0.8	d=2x10^{-3}	X(4) =40	X(40)=339	X(41)=340	X(41)=340
α=9.2	β=0.8	d=2x10^{-7}	X(4) =40	X(41)=340	X(41)=340	X(41)=340
α=15	β=3	d=2x10^{-3}	X(5) =68	X(16)=180	X(36)=338	X(25)=301
α=15	β=3	d=2x10^{-7}	X(5) =68	X(28)=320	X(41)=340	X(41)=340
α=9	β=1	d=2x10^{-3}	X(7)=52	X(42)=309	X(48)=352	X(35)=267
α=9	β=1	d=2x10^{-7}	X(7)=52	X(44)=323	X(50)=366	X(50)=366
α=4	β=1	d=2x10^{-3}	X(5)=23	X(27)=201	X(38)=283	X(29)=209
α=4	β=1	d=2x10^{-7}	X(5)=23	X(28)=268	X(45)=330	X(50)=366

Table 2. Stopping Rules X(.) of WD1(rows 1-4), WD3(rows 5-8), WD5(rows 9-12) for Cases 1, 2 for d=c/(a-b)=.002, c=.01$, a=6$, b=1$ and d=c/(a-b)= 2.E-7,c=1.E-6$

CONCLUSIONS

The contribution of this methodology lies in an empirically updated Bayesian approach to determine an economic stopping rule in a testing environment modeled by a Compound Poisson process to study the accumulation of correlated failures in clumps at each step. Correlation constraints are imposed. It is a useful and practical stopping rule algorithm given the penalties for correcting faults prior (b$) and posterior(a$) to the release of the software module, and based on the capital cost (c$) of testing per unit time. It also recognizes the educated correlation constraints within aggregates(weeks here). A version of this rule has been applied to coverage-testing[3].

REFERENCES

[1] Randolph P.and M. Sahinoglu, ”A Stopping Rule for a Compound Poisson Random Variable”, Applied Stoch. Models and Data Analysis., Vol. 11, 135-143, 1995

[2] M. Sahinoglu, “Compound Poisson Software Reliability Model”, IEEE Trans. Software Eng., Vol. 18, 624-630, 1992

[3] M. Sahinoglu, A. Von Mayrhauser, A. Hajjar, T. Chen, Ch. Anderson “How Much Testing is Enough? Applying Stopping Rules to Behavioral Model Testing”, 4th International Symposium on High Assurance System Engineering, Washington D.C., Nov.,17-19,1999

[4] Handbook of Software Reliability Engineering, Ed. by M. Lyu, IEEE Computer Society Press, McGraw Hill, 1995