Theoretical Mathematics & Applications, vol.3, no.2, 2013, 91-97
ISSN: 1792- 9687 (print), 1792-9709 (online)
Scienpress Ltd, 2013

# Probability Distribution of Testing Time of the Software Based on Markov Chains

**Shushuang Hao[1], Changwei Xu[2] and Jianshe Zhao[3]**

## Abstract

This Probability distribution of testing time of the software package is a problem waiting to be solved quickly. We transfer it into settling a series of problems which related to continuous-time Markov chains. Firstly, we construct a Q-matrix corresponding with the Markov chain $X(t)$. Secondly, we solve the Kolmogorov forward equation corresponding with the Q-matrix and obtain the transition function $P_{ij}(t)$ of the Markov chain $X(t)$. Finally, we get the probability distribution, expectation and variance of testing time of the Software.

**Mathematics Subject Classification:** 60J10

**Keywords:** Q-matrix, Kolmogorov forward equation, extinction time

---

[1] Institute of Information Engineering, Huanghe Science and Technology College, Zhengzhou, China.
[2] College of Science, Zhongyuan University of Technology, Zhengzhou, China.
[3] Peking University Founder Group Co.,Ltd, Beijing, China.

Article Info: *Received* : February 5, 2013. *Revised* : May 24, 2013
*Published online* : June 25, 2013

# 1 Introduction

A testing procedure is often used to eliminate the faults or bugs in package when a new software package is developed. The procedure is to test whether there are any errors result or not with trying the package on a set of well-known problems. This goes on for some fixed time, with all resulting errors being noted. Then the testing stops and the software engineers check carefully to determine the specific bugs that were responsible for the observed errors. Then they alter the package to remove these bugs [1]. Because we are not sure what time is that all the bugs in the package have been eliminated, a problem of great importance is probability distribution of testing time of the software package.

# 2 Preliminary Notes

Let us denote by $X(t)$ the number of undiscovered bugs after that the package was been run for t time units. Suppose that a bug is responsible for an error for the sake of simplicity, and when we discover an error, the package is to be altered to remove the corresponding bug immediately. So the number of undiscovered bugs after t time units is no more than that at t time units. This problem meets Markov property, so there is a probability space $(\Omega, F, P)$, on it $\{X(t), t \in [0, +\infty)\}$ is a continuous-time Markov chain with state space $Z^+ \cup \{0\}$. The transition function corredponding with the state space is $P_{ij}(t) = P(X(t) = j \mid X(0) = i)$. It satisfies temporally homogeneous property

$$P(X(t) = j \mid X(s) = i) = P(X(t-s) = j \mid X(0) = i), \forall t > s. \tag{1}$$

There is a one-to-one correspondence between the Markov chain $X(t)$ and the transition function $P_{ij}(t)$ [2].

In this problem, the state 0 is absorbed, so we can say that extinction occurs when the Markov chain $X(t)$ reach state 0. Now there is no bug in this package. Let us define

$$\tau_0 = \begin{cases} \inf\{t > 0 \mid X(t) = 0\}, & \exists t > 0, s.t. X(t) = 0; \\ +\infty, & otherwise. \end{cases} \tag{2}$$

to be the extinction time of the Markov chain $X(t)$.

## 3 The Distribution of the Extinction Time

First we construct the Q-matrix corresponding with the Markov chain $X(t)$. Let us suppose that there are $N$ parts in the software package, initially it has an unknown number, $n$, of parts which contain one bug respectively ($n \gg N$). So the probability of which no error occurs after the package was used for $m$ times is $\left(\frac{N-n}{N}\right)^m$. Let us suppose that the package is used for $k$ times in 1 times unit, and $k$ is directly proportional to $N$, we note that $\mu = \frac{k}{N}$. So the probability of which no error occurs at t time unit.

And because that $\ln\left(\frac{N-n}{N}\right)^m = k\ln\left(1\text{-}\frac{n}{N}\right) \approx -k\frac{n}{N} = -n\mu,$ suppose that $\mu_n = n\mu, \quad n \ge 1, \mu_0 = 0,$ we get that the probability of which no error occurs at $t$ time unit is $e^{\mu_n t}$.

Thus the Q-matrix corresponding with the transition function $P_{ij}(t)$ is

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ \mu & -\mu & 0 & 0 & \cdots \\ 0 & 2\mu & -2\mu & 0 & \cdots \\ 0 & 0 & 3\mu & -3\mu & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3}$$

The relationship between the Q-matrix and the transition function is that: there is only one Q-matrix corresponding with a transition function $P_{ij}(t)$, but there may be many different Q-functions $P_{ij}(t)$ [2]. the following lemma show that there is a one-to-one correspondence between the Q-matrix and the transition function $P_{ij}(t)$.

**Lemma 3.1** Suppose that the Q-matrix is (3), there exist only one Markov chain $X(t)$ corresponding with it, the transition function of $X(t)$ is noted as $P_{ij}(t)$.

**Proof.** We know that the coefficients of a birth, death, immigration (BDI) process [3,P103]are $\lambda_n = a + n\lambda,\ \mu_n = n\mu,\ n \ge 0,\ a, \lambda, \mu \ge 0.$ Let $\lambda = 0, a = 0,$ the Q-matrix of the BDI process is (3). we see that a BDI process without immigration(i.e. with $a = 0$) is regular. In this case there exist only one Markov chain $X(t)$, the corresponding minimal Q-function is honest and is the only Q-function [3, p. 81], noted as $P_{ij}(t)$. □

**Theorem 3.1** Suppose that $X(t)$ is the only one Markov chain corresponding with the Q-matrix (3),and $X(0) = n$, the probability distribution of the extinction time $\tau_0$ is

$$F_1(t) = (1 - e^{-\mu t})^n. \tag{4}$$

**Proof.** At first we propose to actually find the minimal Q-function $P_{ij}(t)$, which we know to be honest and unique. We can do this by solving the Kolmogorov forward equations. They are

$$P_{ij}^{'}(t) = \begin{cases} -\mu j P_{ij}(t) + \mu(j+1)P_{i,j+1}(t), & j \ge 1; \\ \mu P_{ij}(t), & j = 1. \end{cases} \tag{5}$$

We now transform the system of equations in (5) to an equivalent partial differential equation. To this effect, define the probability generating function $P_i(t,s) = \sum_{j=0}^{+\infty} P_{ij}(t) s^j$. Multiply the first equation in (5) by $s^j$, and sum the result over $j \geq 1$, then add the second equation to the result. After some rearrangement, we have

$$\frac{\partial P_i(t,s)}{\partial t} = \mu(1-s)\frac{\partial P_i(t,s)}{\partial s} \tag{6}$$

with initial condition $P_i(0,s) = s^j$.

Now the equation in (6) is a standard first-order linear partial differential equation. we have $P_i(t,s) = (1-e^{-\mu t} + se^{-\mu t})^i$. The method of solution of this equation can be found in John (1977). Inversion gives

$$P_{n1}(t) = ne^{-\mu t}(1-e^{-\mu t})^{n-1}$$

and

$$P_{n0}(t) = (1-e^{-\mu t})^n.$$

So the probability distribution of the extinction time $\tau_0$ is

$$F_1(t) = P\{\tau_0 \leq t \mid X(0) = n\} = P_{n0}(t) = (1-e^{-\mu t})^n. \qquad \square$$

**Note 3.1** In the case of $X(0) = n$ in theorem 1, the expectation and variance of the extinction time $\tau_0$ is

$$E(\tau_0) = \int_0^{+\infty} tdF_1(t) = \int_0^{+\infty}(1-F_1(t))dt = \sum_{i=1}^{+\infty} C_n^i(-1)^{i+1}\frac{1}{i\mu} \tag{7}$$

and

$$D(\tau_0) = C_n^i(-1)^{i+1}\frac{2}{i^2\mu^2} - \left(\sum_{i=1}^{+\infty} C_n^i(-1)^{i+1}\frac{1}{i\mu}\right)^2 \tag{8}$$

respectively.

**Corollary 3.1** Suppose that $X(t)$ is the only one Markov chain corresponding with the Q-matrix (3), and $X(0)$ is distributed with mean *n*, the probability distribution of the extinction time $\tau_0$ is $F(t) = e^{-\lambda e^{-\mu t}}$.

**Proof.** From the conclusion of the theorem 2, we have

$$
\begin{aligned}
F(t) = P(\tau_0 \le t) &= \sum_{n=0}^{+\infty} P\{\tau_0 \le t, X(0) = n\} \\
&= \sum_{n=0}^{+\infty} P\{\tau_0 \le t \mid X(0) = n\} P\{X(0) = n\} \\
&= \sum_{n=0}^{+\infty} (1 - e^{-\mu t})^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda e^{-\mu t}}.
\end{aligned}
$$

**Note 3.2** In the case of that $X(0)$ is distributed with mean $n$ in corollary 1, the expectation and variance of the extinction time $\tau_0$ is

$$E(\tau_0) = \frac{1}{\mu} \sum_{i=1}^{+\infty} (-1)^{i+1} \frac{\lambda^i}{i \cdot i!} \tag{9}$$

and

$$D(\tau_0) = \sum_{i=1}^{+\infty} (-1)^{i+1} \lambda^i \frac{2}{i^2 \mu^2 \cdot i!} - \left( \frac{1}{\mu} \sum_{i=1}^{+\infty} (-1)^{i+1} \frac{\lambda^i}{i \cdot i!} \right)^2 \tag{10}$$

respectively.

## 4 Conclusion

This Probability distribution of testing time of the software package is a very important problem. Under the assumptions of we construct a Q-matrix corresponding with the Markov chain $X(t)$ and get the probability distribution, expectation and variance of the Software testing time.

**ACKNOWLEDGEMENTS.** This work is supported by Scientific and Technological Project of Henan Province, China (No.082102210041) and Natural Science Function of the Education Department of Henan Province, China (No. 2012B110016).

## References

[1] Sheldon M. Ross, *Introduction to Probability Models*, Ninth edition, Elsevier Pte led, Singapore, 2007.

[2] Zhenting Hou, *The unique rule of Q-processes*, Hunan Science and Technology Press, Changsha, China, 1982.

[3] William J. Anderson, *Continuous-Time Markov Chains*, Springer-Verlag, NewYork, 1991.

[4] P.J. Brockwell, The Extinction Time of a General Birth and Death Process with Cata strophes, *J. Appl. Prob.*, **14**(1), (1986), 709-731.

[5] Zikun Wang, *The General Theory of Stochastic Processes*, Peking Normal University Press, Peking, China, 1996.

[6] Y. Baba, Maximum Likelihood Estimation of Parameters in Birth and Death Process by Poisson Sampling, *J. Oper. Res. Soc.*, **25**(1), (1982), 99-112.