RIMS

A note on sensitivity analysis for PH approximation

Jiahao Zhang*, Junjun Zheng **, Hiroyuki Okamura *** and Tadashi Dohi †

Abstract

This paper presents the moment-based approximation for model depandability when the uncertainty of model parameters is considered. The propagation of uncertainty of model parameters can be estimated by regarding the model parameters as random variables. However, statistical model often involves the non-exponential distribution such as Weibull distribution, which leads to the high computation cost of uncertainty analysis. In this paper, we focus on the Phase-type(PH) distribution to overcome the difficulty of computation for model contains Weibull distribution.

§ 1. Introduction

Phase-type (PH) distribution is defined as the time to absorption in a finite Markov chain with one absorbing state. Because PH distribution can approximate any other distribution with high accuracy, it has been widely used for approximating non-exponential distribution which is hard to compute when considering the first or higher-order derivative. The application of replacing original distribution with PH distribution is called phase (PH) expansion [1], whose main idea is to determine the PH parameters to approximate the original distribution with the fitted PH distribution. In general, the PH expansion is commonly applied to solve the stationary and transient solutions of non-Markovian models, such as the semi-Markov process and Markov regenerative process that involve the non-exponential distributions like the Weibull distribution. To the best of our knowledge, the PH parameters are usually estimated from: (i) observed

2020 Mathematics Subject Classification(s):
*Jiahao Zhang, Hiroshima University, Higashihiroshima 739-8527, Japan.
e-mail: jiahaozhang@hiroshima-u.ac.jp
**Junjun Zheng, Ritsumeikan University, Nojihigashi 525-8577, Japan.
e-mail: jzheng@asl.cs.ritsumei.ac.jp
***Hiroyuki Okamura, Hiroshima University, Higashihiroshima 739-8527, Japan.
e-mail: okamu@hiroshima-u.ac.jp
†Tadashi Dohi, Hiroshima University, Higashihiroshima 739-8527, Japan.
e-mail: dohi@hiroshima-u.ac.jp

data; and (ii) theoretical probability density function (p.d.f.) based on the knowledge of experts. In this paper, we focus on the PH fitting from the information of p.d.f.. However, the parameters of the original distribution usually contain estimation errors, which should not be ignored. Since the fitted PH parameters are estimated from the original density function, it is highly reasonable to say that the estimation errors in the original parameters may bring effect to the corresponding PH fitted parameters. So it is necessary and important to consider the effects of the variation in the parameters from the original distribution on the PH parameters. In such a case, the effect is regarded as local sensitivity. The local sensitivity, in general, evaluates the local impact of the changes in the input factors on the output measure.

On the other hand, when estimating PH parameters from the the information of p.d.f., the approximation is essentially a maximization problem. In general, the maximization problem should satisfy some optimality conditions or restrictions. But the optimality conditions can be easily ignored so that we cannot ensure whether the optimal solution exists or not. To overcome this issue, the Karush-Kuhn-Tucker (KKT) conditions [5] for restriction are considered to guarantee that the obtained optimal solutions are reliable.

In this paper, we focus on the local sensitivity of the fitted PH parameters with respect to the parameters of the original distribution. In particular, we apply the KKT conditions to represent the actual restriction in the PH parameters during solving the maximization problem.

§ 2. PH Expansion

The purpose of PH expansion is to determine the PH parameters to approximate the original distribution with the fitted PH distribution. In general, the PH distribution is classified into continuous and discrete PH distributions. This paper focuses on continuous PH distribution. Without loss of generality, the infinitesimal generator $\boldsymbol{Q}$ of continuous-time Markov chain (CTMC) is assumed as follows:

$$\boldsymbol{Q} = \left(\begin{array}{c|c} \boldsymbol{T} & \boldsymbol{\tau} \\ \hline \boldsymbol{0} & 0 \end{array}\right), \tag{2.1}$$

where $\boldsymbol{T}$ and $\boldsymbol{\tau}$ correspond to transition rates between transient states and the exit rates from transient states to the absorbing state. Let $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_N)$ be the initial probability vector over the transient states, where $\sum_{i=1}^{N} \alpha_i = 1$. The c.d.f. and p.d.f. of PH distribution can be obtained as

$$F(t) = 1 - \boldsymbol{\alpha} \exp(\boldsymbol{T}t)\boldsymbol{1}, \quad f(t) = \boldsymbol{\alpha} \exp(\boldsymbol{T}t)\boldsymbol{\tau}, \tag{2.2}$$

where $\boldsymbol{1}$ is a column vector whose elements are all 1. The exit rate vector has the relationship can be obtained by $\boldsymbol{\tau} = -\boldsymbol{T}\boldsymbol{1}$. In particular, the transient states here are defined as phases.

Phase expansion needs to estimate PH parameters from the information of p.d.f. of original distribution, then the fitted PH distribution can approximate original distribution accurately. It is well-known that the parameters of the original distribution are usually estimated from empirical data and typically involve estimation errors, so it is necessary and important to investigate how the variations in original parameters affect the fitted PH parameters. Such effect is regarded as local sensitivity, which will be discussed in detail in the following section.

§ 3. Local sensitivity of PH Expansion

As mentioned in the introduction, the local sensitivity can be computed by the first or higher-order derivative of functions. In this section, we present the derivatives of PH parameters in the context of PH approximation provided the the information of p.d.f.. Let $f(t;\boldsymbol{\nu})$ be a p.d.f. to be approximated, where $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_L)$ is a parameter vector. Also $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_M)$ is defined as a vector of PH parameters. Then the PH approximation provides the following PH parameters:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}} \int_0^\infty f(t;\boldsymbol{\nu}) \log f_{PH}(t;\boldsymbol{\beta}) dt. \tag{3.1}$$

Suppose that there exist the estimates of PH parameters which maximize the above expected log-likelihood function (LLF). To represent the actual restriction in parameters, the Karush-Kuhn-Tucker (KKT) conditions are considered. The KKT conditions are first derivative tests/first-order necessary conditions for a solution in nonlinear programming to be optimal, provided that some regularity conditions are satisfied [5]. The assumption is shown as follows.

$$\begin{aligned} &\max \int_0^\infty f(t;\boldsymbol{\nu}) \log f_{PH}(t;\boldsymbol{\beta}) dt \\ &s.t. \quad -\alpha_i \le 0, \quad i = 1, 2, \ldots, N, \\ &\qquad -\gamma_j \le 0, \quad j = 1, 2, \cdots, M-N, \\ &\qquad \sum_{i=1}^{N} \alpha_i = 1, \end{aligned} \tag{3.2}$$

where α_i and γ_j are PH parameters in $\boldsymbol{\beta}$. The corresponding Lagrangian function satisfies

$$\begin{aligned} L(\boldsymbol{\beta}, \mu, \lambda) = &\int_0^\infty f(t;\boldsymbol{\nu}) \log f_{PH}(t;\boldsymbol{\beta}) dt + \mu_0 (1 - \sum_{i=1}^{N} \alpha_i) \\ &+ \sum_{i=1}^{n} \mu_i \alpha_i + \sum_{j=1}^{n} \lambda_j \gamma_j, \end{aligned} \tag{3.3}$$

where the constants μ_i and λ_j are KKT multipliers. Then we can get the following KKT conditions:

$$U(\hat{\boldsymbol{\beta}};\boldsymbol{\nu}) + \mu_0 - \mu_i = \mathbf{0}, \quad U(\hat{\boldsymbol{\beta}};\boldsymbol{\nu}) + \lambda_j = \mathbf{0}, \tag{3.4}$$

$$\mu_i \alpha_i = 0, \quad \lambda_j \gamma_j = 0, \tag{3.5}$$

in which $U(\boldsymbol{\beta};\boldsymbol{\nu})$ is a score function for the weighted LLF:

$$U(\boldsymbol{\beta};\boldsymbol{\nu}) = \frac{\partial}{\partial \boldsymbol{\beta}} \int_0^\infty f(t;\boldsymbol{\nu}) \log f_{PH}(t;\boldsymbol{\beta}) dt \tag{3.6}$$

and $\mathbf{0}$ is a column vector whose entries are 0. By applying least squares method, we obtain the KKT multipliers μ_i and λ_j satisfying

$$U(\hat{\boldsymbol{\beta}};\boldsymbol{\nu}) = \mathbf{0}. \tag{3.7}$$

In Eq. (3.7), we apply the fist derivative in terms of ν_i. Since $\hat{\boldsymbol{\beta}}$ is a function of $\boldsymbol{\nu}$, then we have

$$\frac{\partial U(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})}{\partial \nu_i} + U^{(2)}(\hat{\boldsymbol{\beta}};\boldsymbol{\nu}) \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \nu_i} = \mathbf{0}, \tag{3.8}$$

where $U^{(2)}(\boldsymbol{\beta};\boldsymbol{\nu})$ is a matrix for the second derivative of the weighted LLF:

$$U^{(2)}(\boldsymbol{\beta};\boldsymbol{\nu}) = \frac{\partial^2}{\partial \boldsymbol{\beta}^2} \int_0^\infty f(t;\boldsymbol{\nu}) \log f_{PH}(t;\boldsymbol{\beta}) dt. \tag{3.9}$$

Therefore, the local sensitivity (first derivative) of $\hat{\boldsymbol{\beta}}$ with respect to ν_i is given by

$$\frac{\partial \hat{\boldsymbol{\beta}}}{\partial \nu_i} = -U^{(2)}(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})^{-1} \frac{\partial U(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})}{\partial \nu_i}. \tag{3.10}$$

Similarly, the second derivative of $\hat{\boldsymbol{\beta}}$ is given by

$$\begin{aligned} \frac{\partial^2 \hat{\boldsymbol{\beta}}}{\partial \nu_i \partial \nu_j} = -U^{(2)}(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})^{-1} \Bigg(& \frac{\partial^2 U(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})}{\partial \nu_i \partial \nu_j} \\ & + \frac{\partial U^{(2)}(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})}{\partial \nu_i} \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \nu_j} + \frac{\partial U^{(2)}(\hat{\boldsymbol{\beta}};\boldsymbol{\nu})}{\partial \nu_j} \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \nu_i} \\ & + \left(\frac{\partial \hat{\boldsymbol{\beta}}}{\partial \nu_i} \right)^T \sum_l \frac{\partial U^{(2)}(\hat{\boldsymbol{\beta}})}{\partial \beta_P} \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \nu_j} \Bigg). \end{aligned} \tag{3.11}$$

§ 4. Numerical Experiment

This section we illustrate the experiment about effects of original parameters on PH parameters when the model is given by a two-state CTMC. Meanwhile, the failure time follows the Weibull distribution as follows.

$$f(t;\beta,\eta) = \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1}e^{-(\frac{x}{\eta})^{\beta}}, \tag{4.1}$$

where $\beta > 0$ and $\eta > 0$ are the shape parameter and scale parameter, respectively. The CTMC only contains the 'UP' state and 'DOWN' state which means system is operational and failed. In particular, we set $\beta = 1.5$, $\eta = 1.0$. The repair rate of CTMC is 2.0.

In the experiment, the number of phases = 2 and 4. The PH parameters e.g., when phases = 2 are shown below:

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2), \quad \boldsymbol{T} = \begin{pmatrix} -\gamma_1 & \gamma_1 \\ 0 & -\gamma_2 \end{pmatrix}, \quad \boldsymbol{\tau} = \begin{pmatrix} 0 \\ \gamma_2 \end{pmatrix}.$$

In particular, to obtain the accuray obviously, we consider a nuerical derivative method for comparison. First, set $\delta = 0.00001$ as the variation, that is, we change the shape parameter and scale parameter such as $\beta - \delta$ and $\eta + \delta$. Second, re-estimate model parameters to see how much the fitted PH parameters changed. The results are shown as follows.

Table 1. First derivative of β when phases = 2.

parameters	First Derivative	Numerical Derivative
α_1	0.503687	0.503940
α_2	-0.503687	-0.503940
γ_1	0.730342	0.736272
γ_2	0.730046	0.730194

Table 2. First derivative of η when phases = 2.

parameters	First Derivative	Numerical Derivative
α_1	-0.000380	$5.55111e-12$
α_2	-0.000380	$-2.42861e-12$
γ_1	1.441367	1.443436
γ_2	1.441451	1.443516

Table 3. First derivative of η when phases = 4.

parameters	First Derivative	Numerical Derivative
α_1	0.855005	0.777805
α_2	-0.438633	-0.699314
α_3	-0.260701	0.023301
α_4	-0.155671	-0.101792
γ_1	3.829656	4.569169
γ_2	3.829649	4.569036
γ_1	3.376594	6.022194
γ_2	-3.002208	-9.879515

Table 4. First derivative of η when phases = 4.

parameters	First Derivative	Numerical Derivative
α_1	0.000734	$1.343370e-09$
α_2	0.025226	$4.297951e-09$
α_3	-0.023501	$-5.073719e-09$
α_4	-0.002459	$-5.700301e-10$
γ_1	1.959289	2.039837
γ_2	1.959303	2.039850
γ_1	2.268672	2.291495
γ_2	2.894545	2.590062

In the table, 'First Serivative' means the results of local sensitivity of original parameters on PH parameters, the local sensitivity here is same to first derivative of model parmeters, whereas ' Numerical Derivative ' indicates the derivatives of numerical approach by estimating difference between the PH parameters estimated from original parameters changed with variation. From the table 1 and 2, no whether the derivative of β or η, the results based on these two approach are very close to each other. This indicates that our method has a high accuracy when the number of phases is 2. However, when the number of phases up to 4, shown from Table 3 and 4, the difference between two method begins to be large, especially when the derivative parameter is η. In fact, with the increase of the number of the phases, the accuracy of our method reduce gradually. In summary, our method can well apply when the number of phases is under 5, otherwise the results will change a lot.

§ 5. Conclusion

In this paper, we discussed the local sensitivity of original parameters on the PH parameters. To ensure the optimal value of original parameters is really exist, the KKT conditions were applied for constriction. The effect of original parameters on the PH parameters are obtained through the local sensitivity when the optimal value of original parameters were get. In the experiment, although the accuracy of local sensitivity is high when the number of phases is 2, there are still exist some errors when the number of phases increases. In the future, we will try to disscuss how to solve this problem and to obtain the effects of the original parameters on the final outputs.

References

[1] K. Trivedi and A. Bobbio.o "Reliability and Availability Engineering, Modeling, Analysis, and Applications," Cambridge University Press, 2017.

[2] H. Okamura, T. Dohi, and K. Trivedi, "Parametric Uncertainty Propagation through Dependability Models," *in Proc. 2018 Eighth Latin-American Sympo. on Dependable Computing*, IEEE CPS, pp. 10–18, 2018.

[3] H. Takahasi and M. Mori, "Double Exponential Formulas for Numerical Integration," *Publications of the Research Institute for Mathematical Sciences*, vol. 9, no. 3, pp.721–741, 1974.

[4] H. Okamura, Y. Kamahara, and T. Dohi, "Estimating Markov-modulated compound Poisson processes," *in Proc. the 2nd International Conference on Performance Evaluation Methodologies and Tools*, 2007, pp. 1-8.

[5] Z. Xiao "Generalized Karush-Kuhn-Tucker Conditions in Variational and Set-Valued Analysis," arXiv preprint arXiv:1912.05173, 2019.