ESTIMATING MARKOVIAN TRANSITION PROBABILITIES FOR BRIDGE DETERIORATION FORECASTING

Yoshitane TSUDA[1], Kiyoyuki KAITO[2], Kazuya AOKI[3] and Kiyoshi KOBAYASHI[4]

[1]Member of JSCE, Graduate student, Dept. of Urban Management, Kyoto University
(Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan)
E-mail:y-tsuda@psa2.kuciv.kyoto-u.ac.jp
[2]Member of JSCE, Bridge Maintenance Consultant Corporation
(WBG Maribu West 25th Floor, 2-6, Nakase, Mihama-ku, Chiba 261-7125, Japan)
E-mail:kaito@psa.mbox.media.kyoto-u.ac.jp
[3]Member of JSCE, GIS Institute, PASCO Corporation
(2-8-10 Higashiyama, Meguro-ku, Tokyo 153-0043, Japan)
E-mail:kazuya_aoki@pasco.co.jp
[4]Member of JSCE, Professor, Graduate School of Management, Kyoto University
(Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan)
E-mail:kkoba@psa.mbox.media.kyoto-u.ac.jp

In this paper, a methodology to estimate the Markovian transition probability model is presented to forecast the deterioration process of bridge components. The deterioration states of the bridge components are categorized into several ranks, and their deterioration processes are characterized by hazard models. The Markovian transition probabilities between the deterioration states which are defined for the fixed intervals between the inspection points in time, are described by the exponential hazard models. The applicability of the estimation methodology presented in this paper is investigated by the empirical data set of steel bridges in New York city.

Key Words :*Markovian transition probability, bridge asset management, exponential hazard function, fixed intervals*

1. INTRODUCTION

A major problem in bridge asset management consists in finding an optimal repair strategy that minimizes life cycle cost. Predictions on repair demands of the bridge in future times also have to be made along with appropriate budget plans that cover maintenance and repair requirements. Estimations of both life cycle cost and repair demands heavily rely on deterioration forecasting models.

In order to forecast the deterioration progress of a bridge component two types of models have been proposed: 1) a statistical deterioration model based on past visual inspection results, and 2) a deterioration model based on dynamical mechanics. The selection of one model over the other should be determined according with the purposes of the problem. For example, the former model is preferred when maintenance strategies and budget management policies for the whole

infrastructure assets have to be defined. On the other hand, when estimating remaining life time or determining an optimum repair tactic for a concrete damaged member is the main purpose, a deterioration model based on dynamical mechanics is likely to be more effective. In some cases, however, this model requires too many indefinite factors so there is no choice but to use a statistical model.

A Markov chain model is a statistical model used to forecast the deterioration progress of a bridge component. In this model, a rank order as results of visual inspections represents the condition state of the component, and Markov transit on probabilities are estimated to characterize the deterioration progress between two consecutive states. The Markov chain model is advantageous to flexibility of modeling and high operability. It was firstly introduced in American standard applications such as PONTIS and it is now used in many other bridge management systems. In

many of these applications, the Markov transition probabilities are simply estimated using the relative transition frequency of the condition states, which is determined based on the results of visual inspections carried out at several times. However, there are many cases in which the information obtained from visual inspections does not make reference to differences in the structural characteristics of bridges, environmental conditions or even time intervals between inspections, compromising the accuracy of the Markov transition probability estimations.

In this research, a methodology for estimating the Markov transition probabilities for the deterioration of bridge components is proposed. It is based on the condition states ranks that result from visual inspections. The Markov transition probabilities are estimated after characterizing the deterioration process of each bridge component by hazard models. In Section **2.** the fundamental of this research is briefly explained. The Markov transition probability model based on hazard models is formulized in Section **3.** In Section **4.** a methodology for estimating Markov transition probabilities is described. Finally, an application example is explained in Section **5.**

2. FUNDAMENTALS

(1) Outline of the research

Among the vast literature related to deterioration forecasting, Kaito et al.[1] discuss a methodology to calculate the average deterioration curve of New York City bridges considering deterioration speeds. The deterioration speeds are obtained by using visual inspection data, and are then related with random variables so as to estimate a Markov process that reflects the past inspection data history. Other authors [2]−[7] have developed models based on dynamical mechanics to forecast the deterioration of bridge components. However, statistical deterioration models remain in early stages and there are still many problems that compromise the reliability of its results. Additionally, some researches have also been made in relation to Markov transition probability estimations. Among them, Lee et al. [8] estimate Markov transition probabilities between two condition states based on aggregate time series data. Although the methodology described in this research is simple to implement, it has the limitation that it cannot reflect the influence of different structural characteristics of each bridge component or its external conditions in the deterioration process. Concerned with this problem, this paper describes the deterioration process of each bridge component by means of a hazard model, which is based on visual inspection results. The hazard model was firstly developed in the field of reliability analysis to predict the life expectancy of facilities and machines, and since then, it has been applied and reported in many other fields. Detailed descriptions of hazard model estimation methods have been already compiled [9],[10]. In the field of asset management for example, Shin and Madanat [11] propose a Weibull deterioration hazard model to predict the time in which road pavement cracking initiates. However, in the traditional hazard model applied to facilities and machines, the deterioration state is expressed only by two values that define whether a failure exists or not. Therefore, it cannot be applied for bridge deterioration forecasting in which various condition states need to be handle at the same time. A multi destination type hazard model that considers various possible states following a transition has been proposed[9]. In a Markov transition probability that assumes a multi destination hazard model the transition from a given state to only one among mutually exclusive states occurs. However this hazard model can not be applied to deterioration of bridge components, since the condition state of them is irreversibly getting worse with deterioration progress. This research proposes a methodology to describe the transition process between two adjacent deterioration state ranks, related each other with a perpendicular transition, by using exponential hazard models. In addition, a methodology based on estimations of hazard models to estimate Markov transition probabilities that describe the deterioration process of bridge components is developed. That is, this research proposes a two steps methodology: the first step estimates, in a disaggregate way, the Markov transition probabilities that describe the condition states of each bridge component, and then using the previous results, the second step estimates the Markov transition probabilities describing the average deterioration process of the bridge components as a whole.

(2) Condition state and periodical inspection scheme

In order to estimate a model to forecast the deterioration of bridge components it is necessary to accumulate time series data on the condition states of the components. The historical deterioration process of a bridge component is described in **Fig.1**. This figure shows the deterioration progress of a component that has not been repaired yet. In reality, there exists uncertainty in
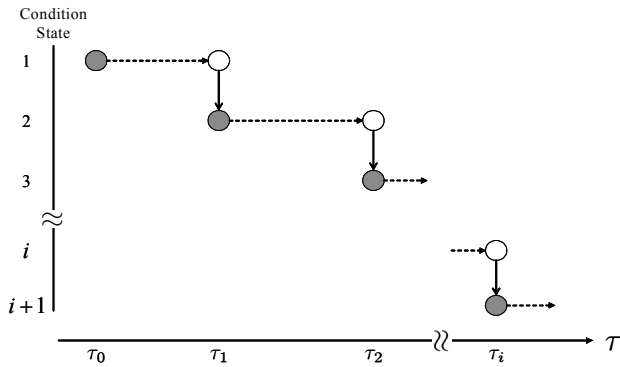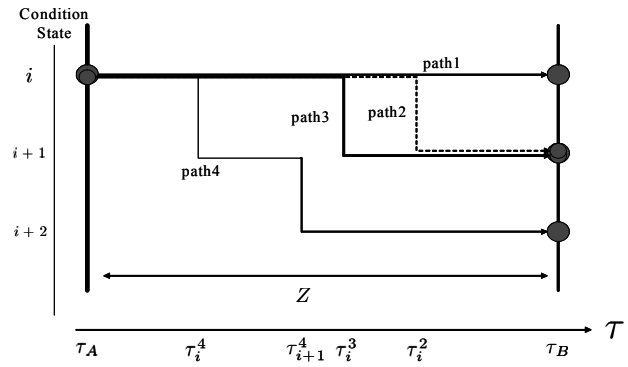
Note) In this example, the deterioration process of a bridge component is expressed in terms of the calendar time $\tau_1, \tau_2, \ldots, \tau_i$, and the condition state of the component is increased in unitary units.

Fig. 1 Timely transition of the condition state.

Note) In this example, the deterioration process of a bridge component is expressed in terms of four different sample paths. In paths 2 and 3 the condition state has advanced to one upper state condition at the calendar times τ_i^2 and τ_i^3 respectively. In path 4, the condition state has increased one state at each time τ_i^4 and τ_{i+1}^4. However, in the case of a periodical inspection carried out at times τ_A and τ_B the condition state at any point in time between inspections cannot be observed. The time length between two inspections is expressed by Z.

Fig. 2 Periodical inspection scheme of the condition state.

the deterioration progress of the component, and moreover, the condition state at each point in the time axis is restricted by the time the visual inspection was carried out. In this figure, τ represents real calendar time (the expression 'time' will be used instead throughout this paper). The deterioration of the bridge starts immediately after it is opened to the public at time τ_0. The condition state of a component is expressed by a rank J representing a state variable i ($i = 1, \cdots, J$). For a component in the sound condition, its condition state is given as $i = 1$, and increasing of condition state i describes progressing deterioration. A value of $i = J$ indicates that a component has reached its service limit. In **Fig.1** for each discrete time τ_i ($i = 1, \cdots, J - 1$) on the time-axis the corresponding condition state has increased from i to $i + 1$. Hereinafter τ_i is referred to the time a transition from a condition state i to $i+1$ occurs.

Information regarding the deterioration process of a bridge can be acquired through periodical visual inspections. However, information on the condition state based on continuous visual inspection is difficult to acquire. In this case, two periodical inspections at times τ_A and τ_B on the time-axis are considered. It is supposed that at time τ_A the condition state observed by inspection is i ($i = 1, \cdots, J - 1$). The deterioration progress in future times is uncertain. Among the infinite set of possible scenarios describing the deterioration process only one path is finally realized. **Fig.2** shows four possible sample paths. Path 1 shows no transition in the condition state i during a periodical visual inspection interval. In paths 2 and

3 the condition state has advanced to one upper state condition at the calendar times τ_i^2 and τ_i^3 respectively. The condition state of these two paths observed at time τ_B become $i + 1$. In a periodical inspection scheme, the point times τ_i^2 and τ_i^3 in which the condition state has changed from i to $i + 1$ are not determined. In addition, path 4 shows transitions in the condition state at times τ_i^4 and τ_{i+1}^4 during the inspection interval. The condition state observed at time τ_B becomes $i+2$. That is, in spite the transitions in the condition state are observable at the time of periodical inspection, it is not possible to obtain information about the times in which those transitions occur.

(3) Markov transition probability

The transition process of the condition state for a bridge component is uncertain and forecasting future states cannot be accomplished deterministically. The Markov transition probability is used to represent the uncertain transition of the condition state of a component during two points in time. As explained later in Section **3.(4)**, Markov transition probabilities can be defined for arbitrary time intervals. For simplify, Markov transition probabilities are defined to forecast the deterioration of a bridge component using the periodical inspection scheme shown in **Fig.2**. The

observed condition state of the component at time τ_A is expressed by using the state variable $h(\tau_A)$. If the condition state observed at time τ_A is i, then the state variable $h(\tau_A) = i$. A Markov transition probability, given a condition state $h(\tau_A) = i$ observed at time τ_A, defines the probability that the condition state at a future time (τ_B for example) will change to $h(\tau_B) = j$. That is

$$\text{Prob}[h(\tau_B) = j | h(\tau_A) = i] = \pi_{ij} \qquad (1)$$

The Markov transition probabilities matrix can be defined by using the transition probabilities between each pair of condition states (i, j) as

$$\mathbf{\Pi} = \begin{pmatrix} \pi_{11} & \cdots & \pi_{1J} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \pi_{JJ} \end{pmatrix} \qquad (2)$$

The Markov transition probability (1) shows the transition probability between the condition states at two given times τ_A and τ_B, therefore, it is straightforward that the values of a transition probability differ for different time intervals. Since deterioration continues as long as no repair is carried out $\pi_{ij} = 0$ $(i > j)$. From the definition of transition probability $\sum_{j=1}^{J} \pi_{ij} = 1$. Then, it holds for the Markov transition probability

$$\left.\begin{array}{l} \pi_{ij} \geq 0 \\ \pi_{ij} = 0 \quad (\text{when } i > j) \\ \sum_{j=1}^{J} \pi_{ij} = 1 \end{array}\right\} \qquad (3)$$

The highest level of deterioration is expressed by the condition state J, which remains as an absorbing state in the Markov chain as long as no repair is carried out. In this case $\pi_{JJ} = 1$.

Markov transition probabilities are defined independently of the past deterioration history. As shown in **Fig.2**, the condition state at the inspection time τ_A is i but the time in which the condition state changed from $i-1$ to i is unobservable. However, in a Markov transition probabilities model this does not have to be taken into account, on the contrary, it is assumed that the transition probability between the inspection times τ_A and τ_B is only dependent on the condition state at time τ_A so as to satisfy its Markov property. The Markov transition probabilities model is highly operative. It can be used to examine asset management strategies at the macro level in which deterioration forecasting and repair strategies of a group of bridges are attained.

3. DISAGGREGATE MARKOV TRANSITION PROBABILITIES

(1) Formulation of hazard model

The Markov transition probability (1) can be defined by using a hazard model representing the deterioration process of an individual component. The information obtained by visual inspection contains not only data on the condition state of an individual component, but also disaggregate data specific to the component such as the structural characteristics or the bridge usage conditions for example. The inspection intervals may also differ from one bridge to another. In order to estimate Markov transition probabilities based on such a variety of data, it is desirable to develop an estimation methodology that can consider specific characteristics of each individual component. In this research 1) the hazard model for the deterioration of a bridge component based on specific information is estimated, and using these results 2) a two steps methodology to estimate Markov transition probabilities is proposed. The purpose of the hazard model is to determine the transition probabilities that characterize the deterioration process of each bridge component. Markov transition probabilities determined by means of hazard models are referred as disaggregate Markov transition probabilities. The methodology to obtain the average transition probability for the whole of bridge components is later explained in Section **4.(3)**.

The deterioration process of a given bridge component can be formulized using the hazard model. In this section, the hazard model is briefly introduced in order to clarify some basic concepts. For more detailed descriptions readers are suggested to look at the references [9],[10]. For the deterioration process of a bridge component as illustrated in **Fig.3** it is assumed that the condition state at the calendar time τ_{i-1} has changed from $i-1$ to i. The calendar time τ_{i-1} is assumed to be the origin $y_i = 0$ of the time axis, referred in this paper as the sample time-axis. The time represented by the sample time-axis is referred from now as a 'time point' and differs from 'time' on the calendar time axis. The times τ_A and τ_B correspond to the time points y_A and y_B on the sample axis. It can be seen that $y_A = \tau_A - \tau_{i-1}$, $y_B = \tau_B - \tau_{i-1}$. Information on the condition state i at the beginning of the calendar time τ_{i-1} cannot be obtained in a periodical inspection scheme. Therefore, time points y_A and y_B on the sample time-axis cannot be correctly obtained either. For convenience of description, it is assumed that the information at the time
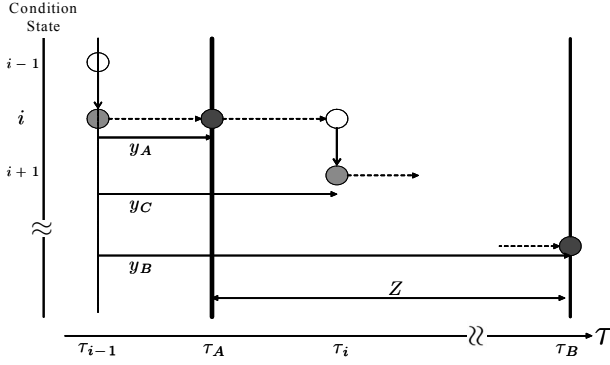
Note) In the case the condition state changes from $i-1$ to i at the calendar time τ_{i-1} the inspections carried out at times τ_A and τ_B will also correspond to the points in time y_A and y_B when using τ_{i-1} as the time origin. The figure shows a sample deterioration path in which the condition state has advanced in one unit to y_c in the interval time τ_{i-1}-y_C. However, observations at time τ_{i-1} are not possible in a periodical inspection scheme, so there is no way to obtain observations at y_A, y_B and y_C. Nevertheless, it is possible to use the information contained in $z = y_C - y_A \in [0, Z]$.

Fig. 3   Model of the deterioration process.

points is known in order to develop the model, despite this assumption is not necessarily essential. The following paragraph discusses that even without information at time points y_A and y_B an exponential hazard model can be estimated.

In the case the condition state of a bridge component at time τ_i (time point y_C) is assumed to change from i to $i+1$, the period length in which the condition state has remained at i (referred as the life expectancy of a condition state i) is represented by $\zeta_i = \tau_i - \tau_{i-1} = y_C$. The life expectancy of a condition state i is assumed to be a stochastic variable with probability density function $f_i(\zeta_i)$ and distribution function $F_i(\zeta_i)$, being ζ_i defined in the domain $[0, \infty]$. The distribution function is defined as

$$F_i(y_i) = \int_0^{y_i} f_i(\zeta_i) d\zeta_i \qquad (4)$$

The distribution function $F_i(y_i)$ represents the cumulative probability of the transition in the condition state from i to $i+1$ when i is set at the initial time point $y_i = 0$(time τ_A) and for a time interval measured along the sample time-axis until the time point y_i (time $\tau_{i-1} + y_i$). Therefore, using the cumulative probability $F_i(y_i)$, the probability $\tilde{F}_i(y_i)$ of a transition in the condition state i during the time points interval $y_i = 0$ to

$y_i \in [0, \infty]$ is defined by $\tilde{F}_i(y_i)$

$$\text{Prob}\{\zeta_i \geq y_i\} = \tilde{F}_i(y_i) = 1 - F_i(y_i) \qquad (5)$$

The conditional probability that the condition state of a component at time y_i advances from i to $i+1$ during the time interval $[y_i, y_i + \Delta y_i]$ is defined as

$$\lambda_i(y_i)\Delta y_i = \frac{f_i(y_i)\Delta y_i}{\tilde{F}_i(y_i)} \qquad (6)$$

where the probability density $\lambda_i(y_i)$ is referred as the hazard function.

(2) Exponential hazard model

In this Section, it is assumed that the deterioration process of a bridge component satisfies Markov property and that the hazard function is independent of the time point y_i on the sample time-axis. That is, for a fixed value of $\theta_i > 0$

$$\lambda_i(y_i) = \theta_i \qquad (7)$$

By using the exponential hazard function above (7) it is possible to represent a deterioration process of a bridge component that satisfies the Markov property (independency from past history). In addition, it is assumed that $\theta_i \neq \theta_j$ ($i \neq j$). By differentiating both sides of equation (5) with respect to y_i

$$\frac{d\tilde{F}_i(y_i)}{dy_i} = -f_i(y_i) \qquad (8)$$

Equation (6) then becomes

$$\lambda_i(y_i) = \frac{f_i(y_i)}{\tilde{F}_i(y_i)} = -\frac{\frac{d\tilde{F}_i(y_i)}{dy_i}}{\tilde{F}(y_i)}$$
$$= \frac{d}{dy_i}\left(-\log \tilde{F}_i(y_i)\right) \qquad (9)$$

Considering that $\tilde{F}_i(0) = 1 - F_i(0) = 1$ and by integrating equation (9):

$$\int_0^{y_i} \lambda_i(u)du = [-\log \tilde{F}_i(u)]_0^{y_i} = -\log \tilde{F}_i(y_i) \quad (10)$$

Using the hazard function $\lambda_i(y_i) = \theta_i$, the probability $\tilde{F}_i(y_i)$ that the life expectancy of the condition state i becomes bigger than y_i is expressed by

$$\tilde{F}_i(y_i) = \exp\left[-\int_0^{y_i} \lambda_i(u)du\right]$$
$$= \exp(-\theta_i y_i) \qquad (11)$$

An exponential hazard model is obtained. According to equation (8) the probability density function $f_i(\zeta_i)$ of the life expectancy of the condition state i is

$$f_i(\zeta_i) = \theta_i \exp(-\theta_i \zeta_i) \qquad (12)$$

Next, consider the condition state has changed to i at the calendar time τ_{i-1}, and remains constant

until the inspection time τ_A. That is, the condition state observed by inspection at time τ_A is i. In this case, the condition state at the time point y_A on a sample time-axis is also i, and the probability that the condition state i will remain constant at a subsequent time point measured from the time point y_A by more than z_i (≥ 0) is defined as

$$\tilde{F}_i(y_A + z_i|\zeta_i \geq y_A)$$
$$= \text{Prob}\{\zeta_i \geq y_A + z_i|\zeta_i \geq y_A\} \quad (13)$$

dividing by the probability $\tilde{F}_i(y_i)$ describing in equation (5) results

$$\frac{\text{Prob}\{\zeta_i \geq y_A + z_i\}}{\text{Prob}\{\zeta_i \geq y_A\}} = \frac{\tilde{F}_i(y_A + z_i)}{\tilde{F}_i(y_A)} \quad (14)$$

using equation (11) the right side of the above equation becomes

$$\frac{\tilde{F}_i(y_A + z_i)}{\tilde{F}_i(y_A)} = \frac{\exp\{-\theta_i(y_A + z_i)\}}{\exp(-\theta_i y_A)}$$
$$= \exp(-\theta_i z_i) \quad (15)$$

In addition, for a condition state i obtained by inspection at time point y_A the probability that the same condition state will be observed by a subsequent inspection at the time point $y_B = y_A + Z$ is

$$\text{Prob}[h(y_B) = i|h(y_A) = i] = \exp(-\theta_i Z) \quad (16)$$

where Z expresses the interval between two inspection times. The probability $\text{Prob}[h(y_B) = i|h(y_A) = i]$ is nothing but the Markov transition probability π_{ii}. That is, when an exponential hazard function is employed, the transition probability π_{ii} is dependent only on the hazard rate θ_i and the inspection interval Z. Even more, without using deterministic information on the time points y_A and y_B, it is still possible to estimate transition probabilities.

(3) Determination of Markov transition probabilities

Using an exponential hazard function the probability that the condition state at the inspection time points y_A and y_B changes from i to $i+1$ can be obtained. This transition can occur if 1) the condition state i remains constant between a time point y_A to a time point $s_i = y_A + z_i$, ($z_i \in [0, Z]$), 2) the condition state changes to $i+1$ at the time point $y_A + z_i$, and 3) it remains constant between the time interval $y_A + z_i$, y_B. Although the exact time in which the condition state transits from i to $i+1$ cannot be traced by periodical inspection, it can be temporarily assumed that the transition occurs at the time point $(y_A + \bar{z}_i) \in [y_A, y_B]$. Given a condition state i at the inspection time point y_A that remains constant until the time

point $y_A + \bar{z}_i$, the probability density that at this time point the condition state changes to $i+1$ is

$$g_i(\bar{z}_i|\zeta_i \geq y_A) = \frac{f_i(\bar{z}_i + y_A)}{\tilde{F}_i(y_A)}$$
$$= \frac{\theta_i \exp\{-\theta_i(\bar{z}_i + y_A)\}}{\exp(-\theta_i y_A)} = \theta_i \exp(-\theta_i \bar{z}_i) \quad (17)$$

With satisfying the above condition, the conditional probability density that the condition state observed at the inspection time point y_B is $i+1$ becomes

$$q_{i+1}(\bar{z}_i|\zeta_i \geq y_A)$$
$$= g_i(\bar{z}_i|\zeta_i \geq y_A) \cdot \tilde{F}_{i+1}(y_B - \bar{z}_i - y_A)$$
$$= \theta_i \exp(-\theta_i \bar{z}_i) \exp\{-\theta_{i+1}(Z - \bar{z}_i)\}$$
$$= \theta_i \exp(-\theta_{i+1} Z) \exp\{-(\theta_i - \theta_{i+1})\bar{z}_i\} \quad (18)$$

However, the explanation above has been applied for a fixed value $\bar{s}_i = y_A + \bar{z}_i$. The life expectancy ζ_i of a condition state i is in fact a stochastic variable, so $\bar{z}_i$ may change in range $[0, Z]$. The Markov transition probability that the condition state change from i to $i+1$ during the time points y_A and y_B is

$$\pi_{ii+1} = \text{Prob}[h(y_B) = i + 1|h(y_A) = i]$$
$$= \int_0^Z q_{i+1}(z_i|\zeta_i \geq y_A)dz_i$$
$$= \int_0^Z \theta_i \exp(-\theta_{i+1} Z) \exp\{-(\theta_i - \theta_{i+1})z_i\}dz_i$$
$$= \frac{\theta_i}{\theta_i - \theta_{i+1}}\{-\exp(-\theta_i Z) + \exp(-\theta_{i+1} Z)\} \quad (19)$$

where $\pi_{ii+1} > 0$ is indifferent to the relative size between θ_i and θ_j. The assumption $\theta_i \neq \theta_{i+1}$ implies $1 > \pi_{ii+1}$. As these characteristics are trivial in the derivation process of equation (19), the verification is omitted. The case a condition state between two inspection time points changes from i to two or more condition states $j(j \geq i + 2)$ is next considered. The distribution function and the probability density function of a period length in which a condition state j remains constant is denoted by $F_i(y_j)$ and $f_j(y_j)$ respectively. The hazard function related to the condition state j is denoted by $\lambda_j(y_j) = \theta_j$. The transition of the condition state from i to j during the time interval $[y_A, y_B]$ can occur if 1) the condition state i remains constant during the time interval y_A, $\bar{s}_i = y_A + \bar{z}_i \in [y_A, y_B]$, 2) the condition state changes to $i + 1$ at the time point $\bar{s}_i = y_A + \bar{z}_i$, 3) the condition state $i + 1$ remains constant during the time interval $\bar{s}_i = y_A + \bar{z}_i$, $\bar{s}_{i+1} = \bar{s}_i + \bar{z}_{i+1}$ ($\leq y_B$), and at this time point changes to $i + 2$. After repeating the same process 4) the condition state changed to j at the time point $\bar{s}_{j-1}$ ($\leq y_B$) remains constant until the time point y_B. The conditional probability

density that the above conditions occur simultaneously is given by

$$q_j(\bar{z}_i, \bar{z}_{i+1}, \cdots, \bar{z}_{j-1}|\zeta_i \geq y_A)$$

$$= g_i(\bar{z}_i|\zeta_i \geq y_A) \prod_{m=i+1}^{j-1} f_m(\bar{z}_m)\tilde{F}_j\left(Z - \sum_{m=i}^{j-1}\bar{z}_m\right)$$

$$= \prod_{m=i}^{j-1} \theta_m \cdot \exp\left\{ -\sum_{m=i}^{j-1}\theta_m\bar{z}_m - \theta_j(Z - \sum_{m=i}^{j-1}\bar{z}_m)\right\}$$

$$= \prod_{m=i}^{j-1} \theta_m \cdot \exp\left\{ -\theta_j Z - \sum_{m=i}^{j-1}(\theta_m - \theta_j)\bar{z}_m \right\} \quad (20)$$

where $\bar{z}_i, \cdots, \bar{z}_{j-1}$ are fixed values. Since the life expectancy ζ_i of condition states i ($i = 1, \cdots, J - 1$) is a stochastic variable, the values of $z_i \geq 0, \cdots, z_{j-1} \geq 0$ are variable under the following condition:

$$0 \leq z_i + z_{i+1} + \cdots + z_{j-1} \leq Z \quad (21)$$

Therefore, the Markov transition probabilities π_{ij} that a transition in the condition state from i to j ($j \geq i + 2$) occurs between the inspection time points y_A and y_B becomes

$$\pi_{ij} = \text{Prob}[h(y_B) = j|h(y_A) = i]$$

$$= \int_0^Z \int_0^{Z-z_i} \cdots \int_0^{Z-\sum_{m=i}^{j-2}z_m}$$

$$q_j(z_i, \cdots, z_{j-1}|\zeta_i \geq y_A)dz_i \cdots dz_{j-1}$$

$$= \sum_{k=i}^{j} \prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} \exp(-\theta_k Z) \quad (22)$$

as shown in the **APPENDIX A**. Here the following notation rule is given.

$$\begin{cases} \prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} = 1 & (\text{when} \quad k = i) \\ \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} = 1 & (\text{when} \quad k = j) \end{cases}$$

π_{iJ} is arranged using the Markov transition probabilities conditions (3)

$$\pi_{iJ} = 1 - \sum_{j=i}^{J-1} \pi_{ij} \quad (23)$$

The Markov transition probabilities based on the exponential hazard model becomes

$$\pi_{ii} = \exp(-\theta_i Z) \quad (24a)$$

$$\pi_{ii+1} = \frac{\theta_i}{\theta_i - \theta_{i+1}}\{-\exp(-\theta_i Z) + \exp(-\theta_{i+1}Z)\} \quad (24b)$$

$$\pi_{ij} = \sum_{k=i}^{j} \prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} \exp(-\theta_k Z)$$

$$(j = i, \cdots, J) \quad (24c)$$

$$\pi_{iJ} = 1 - \sum_{j=i}^{J-1} \pi_{ij} \quad (24d)$$

$$(i = 1, \cdots, J - 1)$$

(4) Time adjustment of Markov transition probability

As shown in equations (24a) - (24d), Markov transition probabilities depend on the inspection interval value Z. For clarity of presentation, the Markov transition probability is expressed as $\pi_{ij}(Z)$, so the Markov transition probabilities matrix related to the inspection time interval Z becomes

$$\boldsymbol{\Pi}(Z) = \begin{pmatrix} \pi_{11}(Z) & \cdots & \pi_{1J}(Z) \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \pi_{JJ}(Z) \end{pmatrix} \quad (25)$$

For an integer value n, two inspection interval Z and nZ are considered. The Markov transition probability matrices $\boldsymbol{\Pi}(Z)$ and $\boldsymbol{\Pi}(nZ)$ describe the same deterioration process for two different time intervals. Therefore, the Markov transition probability matrix $\boldsymbol{\Pi}(nZ)$ expressed in terms of the Markov transition probabilities matrix $\boldsymbol{\Pi}(Z)$ is

$$\boldsymbol{\Pi}(nZ) = \{\boldsymbol{\Pi}(Z)\}^n \quad (26)$$

The condition expressed above is referred as the time adjustment conditions of a Markov transition probability matrix. In order to satisfy this condition, a fixed mathematical structure between the Markov transition probabilities π_{ij} must hold. In the **APPENDIX A**, it is theoretically demonstrated that the transition probabilities (24a) - (24d), defined by using exponential hazard models, satisfy the time adjustment conditions. In other words, it is possible to find a Markov transition probabilities matrix for an arbitrary time interval Z by changing the value of the inspection interval Z contained in the transition probabilities (24a) - (24d).

4. ESTIMATION OF MARKOV TRANSITION PROBABILITIES

(1) Contents of periodical inspection data

Suppose periodical inspection data on the same kind of K bridge components is available. An inspection sample $k(k = 1, \cdots, K)$ describes two continuous periodical inspections carried out at times τ_A^k and τ_B^k and the respective condition states ratings $h(\tau_A^k)$ and $h(\tau_B^k)$ measured at those times. Differences in the inspection intervals of the samples are not inconvenient. Based on the above inspection data, the inspection interval of a sample k is defined as $Z^k = \tau_B^k - \tau_A^k$. In addition a dummy variable based on the deterioration progress patterns between two inspections times

is defined as

$$\delta_{ij}^k = \begin{cases} 1 & \text{when } h(\tau_A^k) = i \text{ and } h(\tau_B^k) = j \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

for each δ_{ij}^k $(i,j = 1, \cdots, J; k = 1, \cdots, K)$. Furthermore, the structural characteristics and usage conditions of a bridge that affect the deterioration speed of a component are represented by the vector $\boldsymbol{x}^k = (x_1^k, \cdots, x_M^k)$, where x_m^k $(m = 1, \cdots, M)$ represents the value of a characteristic variable m observed in the sample data k. The information contained in the inspection sample data k can be rearranged as $\Xi^k = (\delta_{ij}^k, Z^k, \boldsymbol{x}^k)$. On the other hand, the exponential hazard function of the deterioration process for a sample data $k(k = 1, \cdots, K)$ is

$$\lambda_i^k(y_i^k) = \theta_i^k \quad (i = 1, \cdots, J - 1)$$

Since the condition state J is the absorption state of a Markov chain and $\pi_{JJ} = 1$ the rate of the hazard is not defined. The hazard rate θ_i^k $(i = 1, \cdots, J-1; k = 1, \cdots, K)$ characterizing the deterioration process of a bridge is considered to change in relation to the vector $\boldsymbol{x}^k$ as follows:

$$\theta_i^k = \boldsymbol{x}^k \boldsymbol{\beta}_i' \quad (28)$$

where $\boldsymbol{\beta}_i = (\beta_{i,1}, \cdots, \beta_{i,M})$ is a row vector of unknown parameters $\beta_{i,m}$ $(m = 1, \cdots, M)$ and the symbol $\prime$ indicates the vector is transposed. In order to obtain Markov transition probabilities the first step consists in estimating the exponential hazard function $\lambda_i^k(y_i^k) = \theta_i^k$ based on the inspection sample information Ξ^k $(k = 1, \cdots, K)$. In Section **4.(2)** the estimation of the exponential hazard function is described. As a second step, Markov transition probabilities are estimated using the exponential hazard functions found in the previous step. The methodology proposed by this research permits estimating Markov transition probabilities for every individual component. However, determining an optimal repair strategy for every individual component can complicate its application to the bridge management practice. For this reason, it can be more convenient in many cases to assume an average Markov transition probability analog to that of the components. The estimation of the average Markov transition probability using the exponential hazard model estimated is explained in Section **4.(3)**. In addition, using exponential hazard models, a risk management index for bridge management can also be derived. The expected elapsed period from the time the relevant rating is reached until the following rating is attained as a result of the deterioration progress (referred as the expected life expectancy of a rating) is defined by using

the survival function $\tilde{F}_i(y_i^k)$ as [9]

$$RMD_i^k = \int_0^\infty \tilde{F}_i(y_i^k) dy_i^k \quad (29)$$

By defining the survival function $\tilde{F}_i(y_i^k)$ in terms of the exponential hazard function as in equation (11), as the expected life expectancy of a rating becomes

$$RMD_i^k = \int_0^\infty \exp(-\theta_i^k y_i^k) dy_i^k = \frac{1}{\theta_i^k} \quad (30)$$

(2) Estimation of the hazard model

Information $\Xi^k = (\bar{\delta}_{ij}^k, \bar{Z}^k, \bar{\boldsymbol{x}}^k)$ can be acquired in relation to the inspection sample k, where the symbol $\bar{}$ indicates an actual measurement. The Markov transition probabilities can be expressed in terms of the hazard functions (24a)-(24d). Although the hazard rate θ_i^k $(i = 1, \cdots, J-1; k = 1, \cdots, K)$ of each state condition is contained in the Markov transition probabilities, it can be represented by equation (28) when using the vector $\bar{\boldsymbol{x}}^k$ of a bridge component. Moreover, the deterioration transition probability also depends on inspection interval $\bar{Z}^k$ in which the data was observed. For clarity of presentation, the transition probability π_{ij} is expressed as a function of the measured data $(\bar{Z}^k, \bar{\boldsymbol{x}}^k)$ obtained from visual inspection and the unknown parameters $\boldsymbol{\beta}_i$ as $\pi_{ij}(\bar{Z}^k, \bar{\boldsymbol{x}}^k : \boldsymbol{\beta}_i)$. If the deterioration progress of the bridge components in a sample K are assumed to be mutually independent, the log-likelihood function expressing the simultaneous probability density of the deterioration transition pattern for all inspection samples is [12),13)]

$$\ln[\mathcal{L}(\boldsymbol{\beta})] = \ln \left[\prod_{i=1}^{J-1} \prod_{j=i}^{J} \prod_{k=1}^{K} \left\{ \pi_{ij}(\bar{Z}^k, \bar{\boldsymbol{x}}^k : \boldsymbol{\beta}) \right\}^{\bar{\delta}_{ij}^k} \right]$$

$$= \sum_{i=1}^{J-1} \sum_{j=i}^{J} \sum_{k=1}^{K} \bar{\delta}_{ij}^k \ln \left[\pi_{ij}(\bar{Z}^k, \bar{\boldsymbol{x}}^k : \boldsymbol{\beta}) \right] \quad (31)$$

where $\bar{\delta}_{ij}^k$, $\bar{Z}^k$ and $\bar{\boldsymbol{x}}^k$ are all determined throgh inspections and $\boldsymbol{\beta}_i$ $(i = 1, \cdots, J-1)$ are parameters to be estimated. Estimations of the parameters $\boldsymbol{\beta}$ can be obtained by solving the optimality conditions

$$\frac{\partial \ln[\mathcal{L}(\hat{\boldsymbol{\beta}})]}{\partial \beta_{i,m}} = 0, \quad (32)$$

$$(i = 1, \cdots, J-1; m = 1, \cdots, M)$$

that result from maximizing the log-likelihood function (31). The optimal values $\hat{\boldsymbol{\beta}} = (\hat{\beta}_{1,1}, \cdots, \hat{\beta}_{J,M})$ are then estimated by applying a numerical iterative procedure such as the Newton Method for the $(J-1)M$ order nonlinear simultaneous equations [14)]. Furthermore, estimator for

the asymptotical covariance matrix of the parameters is given by

$$\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\beta}}) = \left[\frac{\partial^2 \ln\{\mathcal{L}(\hat{\boldsymbol{\beta}})\}}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \right]^{-1} \qquad (33)$$

The $(J-1)M \times (J-1)M$ order inverse matrix of the right-hand side of the above formula, composed by the elements $\partial^2 \ln\{\mathcal{L}(\hat{\boldsymbol{\beta}})\}/\partial\beta_{i,m}\partial\beta_{i',m'}$ results to be the inverse matrix of the Fisher information matrix.

(3) Average Markov transition probability

Given the vector $\boldsymbol{x}^k$ and the inspection interval Z^k, the Markov transition probabilities of a bridge can be estimated by using equations (24a)-(24d). Markov transition probabilities satisfying time adjustment conditions can be estimated for arbitrary inspection intervals by changing the value Z^k. As proposed in this research, a Markov transition probability matrix characterizing every bridge component can also be estimated. However, when forecasting the deterioration pattern of many bridges as a whole, in many cases, it is more convenient to search an average transition probability rather than a transition probability for every component. For that purpose, it is necessary to develop a methodology to estimate the average transition probability matrix that also satisfies at the time adjustment conditions. The methodology presented in this section pays attention to the hazard rates θ_i^k $(k = 1, \cdots, K)$. The distribution function of the bridge characteristics for the population sample of bridge components is expressed as $\Gamma(\boldsymbol{x})$. For this case, the expected value of the hazard rate $E[\theta_i]$ for the population sample is defined as

$$E[\theta_i] = \int_{\Theta} \boldsymbol{x}\boldsymbol{\beta}_i' d\Gamma(\boldsymbol{x}) \qquad (34)$$

where Θ makes reference to the population sample. A Markov transition probabilities matrix is said to satisfy the time adjustment conditions if 1) Markov transition probabilities are estimated by using the exponential hazard functions (24a) - (24d), and 2) the Markov transition probabilities matrix for each sample is defined by the hazard rates θ_i^k $(i = 1, \cdots, J-1; k = 1, \cdots, K)$. Therefore, a Markov transition probabilities matrix estimated by using the averaging equation (34) is said to also satisfy the time adjustment conditions.

Table 1   Ratings description [15].

State	Description
1	Deck is new or near new, almost no sign of deterioration
2	between 1and 3
3	Only localized areas of leakage (e.g., single longitudinal crack with leakage, or deck edges showing only spotty leakage).
4	Between 3 and 5
5	75 percent or more of the deck has leakage. Only localized spalled areas. Efflorescence along the girder top flanges.
6	Between 5 and 7
7	Heavy spalling. Heavy efflorescence. Punch through has occured or is likely. Deck saturated to point that concrete is rubble.

Note) At the level of potential risk or collapse, a slightly less deterioration level reduces the rating to the next one.

5. EMPIRICAL ANALYSIS

(1) Outline

In this section, a practical example of the model proposed above is presented. Visual inspection data on New York City's (NYC) bridges for the ten years (1987 - 1996) are used. In the U.S. carrying out visual inspections to all bridges at least once every 2 years is an obligation. The total number of bridges managed in NY is 764 giving a total of 4689 spans, and a total deck area of $1,430,000\text{m}^2$ by the year 2000. The average age of those bridges is over 75 years. In accordance to the bridge inspection manual '82 [15] of the State of New York, visual inspection is carried out for up to 25 superstructure components and 22 substructure components. The results of a visual inspection give seven levels (1-7) as a rating of the conditions of the components. In a visual inspection carried out at each span of a bridge where there are two or more similar components, the representative rating of the span is normally assumed to be the worst found. The result of these inspections is digitalized in a database for every bridge identification number. In addition, information about the bridge type and structural characteristics, location, number of spans, average traffic volume, etc. is collected.

This study discusses reinforced concrete decks which are important components for maintenance and management and on which wheel loads act directly. **Table 1** shows the physical meanings

of ratings for reinforced concrete decks. In the bridge inspection manual '82 of State of New York[15], the rating evaluation is not systemized according to kinds of damages (crack, efflorescence, spalling, etc.). In Japan, on the other hand, the Periodical Bridge Inspection Manual (Draft)[16], for example, adopts a 5 point rating system (a-to-e) based upon the deterioration mechanism of mainly crack propagation. Furthermore, the width and length of the crack is given quantitatively in each rating for a more objective assessment. Regardless of the characteristics of the cracking, moreover, spalling and exposed reinforcing steel are both evaluated in another categories, and leakage, efflorescence and rust stains due to cracks are evaluated in the category of crack. In spite of these differences between both manuals, however, a comparison of both ratings of reinforced concrete decks shows that they correspond on the whole ($2 \rightarrow a$, $3 \rightarrow b$, $4 \rightarrow c$, $5 \rightarrow d$, $6\&7 \rightarrow e$).

This database is used in this research in order to estimate the parameters of the exponential hazard function. In order to consider the structural characteristics and usage conditions variables of the bridge $\boldsymbol{x}$, three variables $x_1 = 1$, x_2 and x_3 are adopted to describe an absolute term, the average traffic volume and the deck surface area (referred as deck area) respectively. Samples containing atypical data caused by misjudgement or samples in which a repair work was carried out were removed. The data in which information on average traffic and deck area is available were also filtered accordingly. The final database employed in the exponential hazard model estimation contains a total of 32,902 samples. In addition, in the 32,902 explanatory variables in each sample the traffic levels and deck areas were normalized to their maximum values to facilitate the estimation procedure.

(2) Results

Using the sample data mentioned above and previously removing the condition state rating 7 from the data 6 exponential hazard functions were estimated. For each sample, ratings describing the results of visual inspections carried out at two different points in time are recorded. The exponential hazard model (28) for this example expressed in terms of the explanation variables is defined as

$$\theta_i^k = \beta_{i,1} + \beta_{i,2}x_2^k + \beta_{i,3}x_3^k \qquad (35)$$
$$(i = 1, \cdots, 6; k = 1, \cdots, K)$$

Since there is one exponential hazard function for each of the six ratings and three unknown

Table 2 Exponential hazard model results.

State	Absolute terms $\beta_{i,1}$	Average traffic level $\beta_{i,2}$	Deck area $\beta_{i,3}$
1	0.3289	—	1.3648
	(26.144)	—	(2.547)
2	0.2071	0.0779	0.8427
	(25.432)	(2.537)	(5.099)
3	0.1334	0.1379	—
	(32.016)	(8.456)	—
4	0.0847	0.0961	0.0755
	(21.154)	(6.608)	(3.250)
5	0.0979	—	—
	(21.742)	—	—
6	0.1288	0.3842	—
	(6.951)	(4.067)	—

Note) $t-$ values are shown in parenthesis

parameters in each equation, this gives a total of 18 unknown parameters $\beta_{i,1}$, $\beta_{i,2}$, $\beta_{i,3}$ ($i = 1, \cdots, 6$). Although the exponential hazard model was firstly estimated using all 18 parameters, low values of the $t-$ and sign conditions parameters were found. Therefore, in some cases the exponential hazard model was estimated for a combination of explanation variables that do not include x_2^k and/or x_3^k. Combinations of explanatory variables were chosen so as to maximize the log-likelihood function (31) with a significance level of the $t-$ value of 95% for the explanatory variables and to satisfy sign restrictions conditions. In **Table 2** the results of the maximum likelihood estimations $\hat{\boldsymbol{\beta}}$ are shown with the respective $t-$ values of each explanatory variable. The Markov transition probabilities matrix for each sample is able to be estimated by using the exponential hazard model as proposed in this research. As explained in section **4.(3)** the average Markov transition probabilities matrix is determined in order to avoid the huge amount of individual estimations. Furthermore, in order to analyze the validity of the methodology proposed by this research a Markov transition probability matrix is also estimated by using a simplified method described as follows:

$$\bar{\pi}_{ij}(l) =$$
$$\frac{\#\{h(\tau_A^k) = i, h(\tau_B^k) = j, \tau_B^k - \tau_A^k = l, k \in [1, K]\}}{\#\{h(\tau_A^k) = i, k \in [1, K]\}}$$
$$(36)$$

where $\#\{the\ formula\}$ is the number of samples satisfying *the formula*. However, when searching for Markov transition probabilities by using a numerical method, it is necessary to fix the inspection time interval. Then, the inspection interval

Table 3 Life expectancy of a rating.

State	$E[\theta_i]$	$E[RMD_i^k]$ (year)	$E[\overline{RMD}_i^k]$ (year)
1	0.3074	3.25	3.49
2	0.2605	4.33	4.38
3	0.1653	6.27	6.29
4	0.1048	10.04	9.43
5	0.0866	11.55	7.64
6	0.2067	5.11	14.39

Fig. 4 Expected deterioration path for traffic level.

was fixed to two years, since most of the sample data were taken during this interval. The total of samples then becomes 26,748. The average values of traffic level and deck surface for this sample data are 0.2266 and 0.0431 respectively. As mentioned above, traffic levels and deck areas were normalized using the maximum values in the 32,902 samples data. The expected hazard rate $E[\theta_i]$ is obtained through equation (34), which is defined using the exponential hazard function for the database mentioned above, as shown in **Table 3**. The life expectancy of a rating, indicating the elapsed time in reaching a following rating, is expressed by equation (30). The life expectancy of a rating RMD_i^k of each of the 26,748 samples was calculated. The average values for each rating $E[RMD_i^k]$ using the proposed methodology and the average values $E[\overline{RMD}_i^k]$ obtained by the simplified method are shown in **Table 3**. The averaged life expectancies $E[\overline{RMD}_i^k]$ using Markov transition probabilities $\bar{\pi}_{ij}$ in equation (36) is given by

$$E[\overline{RMD}_i^k] = 2\bar{\pi}_{iJ}(2) + \sum_{l=1}^{\infty} \sum_{m=i}^{J-1} (2l+2)$$

$$\cdot \bar{\pi}_{im}(2l)\bar{\pi}_{mJ}(2) - \sum_{m=i+1}^{J-1} E[\overline{RMD}_m^k] \quad (37)$$

$\bar{\pi}_{ij}(2l)$ is the Markov transition probability for $2l$ years interval, which raises the transition probability matrix estimated by the simplified method to the lth power. The summation of the first and the second terms of the right hand side in the above equation means the life expectancy from the condition state i to J. The third term is the summation of $\overline{RMD}_i^k$ from the condition state $i+1$ to $J-1$, and means the life expectancy $i+1$ to J. That is to say, the life expectancy $E[\overline{RMD}_i^k]$ of the condition state I is defined as the value which subtracts the life expectancy $i + 1$ to J from the one i to J. As shown in this table the life expectancy of ratings $1\sim3$ estimated by using the methodology proposed in this research does not greatly differ from those obtained by the sim-
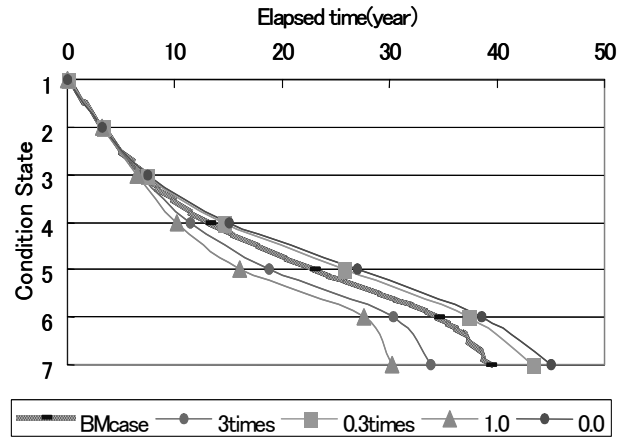
plified method. However, for higher deterioration conditions (states $4\sim6$) noticeable differences are observed. Furthermore, the average Markov transition probabilities matrix obtained by the proposed methodology is shown in **Table 4** and in **Table 5** for the simplified method. For high values of the condition state, the difference in the average transition probabilities obtained by the proposed estimations becomes large in comparison to those obtained by the simplified method because 1) the number of samples decreases relatively, and 2) the difference in the average transition probabilities becomes large due to the sample attribute.
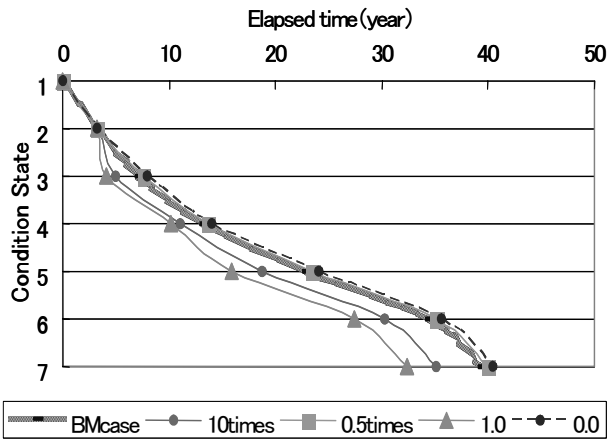
The effect of different traffic levels or deck areas in Markov transition probabilities cannot be analyzed when using the simplified method but can be handled by using the methodology proposed in this paper. The effects of different traffic levels or deck areas on the deterioration speed are analyzed. To do so, the expected deterioration path characterizing the average deterioration progress is defined in relation to the Markov transition probabilities. The expected deterioration path can be represented by a graph describing the average deterioration process during the life expectancy of a rating RMD_i^k (from the time the rating is reached to the time the next rating is attained). **Fig.4** shows the expected deterioration path for a benchmark case (BM) in which the traffic level and the deck areas are set to the sample average values 0.2266 and 0.0431 respectively. The condition state gets worse as long as the deterioration progress surpasses a fix number of years from the initial point in time. In order to observe a transition of the condition state to a subsequent ranking, in the expected deterio-

Table 4   Transition probabilities matrix from estimation results.

State	1	2	3	4	5	6	7
1	0.5408	0.3485	0.0984	0.0116	0.0006	0.0000	0.0000
2	0	0.5939	0.3409	0.0606	0.0044	0.0002	0.0000
3	0	0	0.7185	0.2525	0.0273	0.0015	0.0002
4	0	0	0	0.8109	0.1731	0.0139	0.0021
5	0	0	0	0	0.8410	0.1295	0.0295
6	0	0	0	0	0	0.6614	0.3386
7	0	0	0	0	0	0	1

Table 5   Transition probabilities matrix from simplified method.

State	1	2	3	4	5	6	7
1	0.5382	0.3929	0.0595	0.0070	0.0018	0.0006	0.0000
2	0	0.6235	0.3272	0.0463	0.0022	0.0005	0.0002
3	0	0	0.7203	0.2557	0.0222	0.0018	0.0000
4	0	0	0	0.8046	0.1800	0.0126	0.0028
5	0	0	0	0	0.8413	0.1040	0.0547
6	0	0	0	0	0	0.8610	0.1390
7	0	0	0	0	0	0	1

Fig. 5   Expected deterioration path for deck area.

does not have a noticeable influence on its deterioration progress. However, for an advanced deterioration condition, the deterioration progress is notably influenced by traffic levels increments. On the other hand, **Fig.5** shows an analysis of the relation between the deterioration progress speed and the deck area for the BM case (Average traffic level 0.2266,   average deck areas 0.0431). This figure shows that the bigger the deck area the faster the deterioration progress. Unlike the case of traffic level, in condition state 2 in which deterioration is not advanced, the deck area has already affected the degradation expected value path.

(3) Sample number and estimation accuracy

The size of database used in the empirical analysis, (referred as original database) is 32,902. However, there are many cases in which a sufficient number of samples is not available. An important problem consists in defining the minimal number of samples necessary to make accurate estimations. In order to determine the accuracy of an analysis of the data used in the exponential hazard model estimations is made. Suppose that a random sample from the original database is selected. Then it is possible to analyze the influence of the sample size in the exponential hazard model estimations, and determine the minimal sample size that attains the same results obtained by using the original database. The maximum likelihood estimations of the exponential hazard function based on the original database are expressed as $\hat{\beta}$. In addition, the

ration path, it is required an elapsed time of the deterioration progress for each condition state in accordance to the life expectancy of ratings. As shown in this figure, the expected deterioration path does not necessarily have to show a convex shape. This is because the condition state is defined as a discrete variable state, and it is not continuously expressed as cardinal measure. If the standard of the condition state measures were changed, the shape of the expected deterioration path will differ as well. **Fig.4** shows the expected deterioration path for different traffic levels: a) 0 times the BM case (traffic level: 0.2266), b) 0.3 times the BM case, c) 3 times the BM case, and d) a traffic level of 1. As shown in this figure, in the case the deck does not exhibit much deterioration (condition state 1-3), the traffic level

maximum likelihood estimations based on the extracted database t are expressed as $\hat{\boldsymbol{\beta}}^t$. The exponential hazard functions based on the original database and the extracted database can be defined as

$$\hat{\theta}_i^k = \hat{\beta}_{i,1} + \hat{\beta}_{i,2}x_2^k + \hat{\beta}_{i,3}x_3^k \tag{38a}$$

$$(i = 1, \cdots, 6; k = 1, \cdots, n)$$

$$\hat{\theta}_i^{k,t} = \hat{\beta}_{i,1}^t + \hat{\beta}_{i,2}^t x_2^k + \hat{\beta}_{i,3}^t x_3^k \tag{38b}$$

$$(i = 1, \cdots, 6; k = n+1, \cdots, n+n')$$

x_2^k, x_3^k $(k = 1, \cdots, n)$, x_2^k, x_3^k $(k = n+1, \cdots, n+n')$ represent sample data of the original database and the extracted database, and n, n' denote corresponding sample numbers. That is, the corresponding sample numbers are n and n'. The exponential hazard functions based for the new pooled database k $(k = 1, \cdots, n + n')$ is defined as

$$\tilde{\theta}_i^k = (\hat{\beta}_{i,1} + \delta^k \hat{r}_{i,1}) + (\hat{\beta}_{i,2} + \delta^k \hat{r}_{i,2})x_2^k$$
$$+ (\hat{\beta}_{i,3} + \delta^k \hat{r}_{i,3})x_3^k \tag{39}$$

where

$$\begin{cases} \hat{r}_{i,1} = \hat{\beta}_{i,1}^t - \hat{\beta}_{i,1} \\ \hat{r}_{i,2} = \hat{\beta}_{i,2}^t - \hat{\beta}_{i,2} \\ \hat{r}_{i,3} = \hat{\beta}_{i,3}^t - \hat{\beta}_{i,3} \end{cases} \tag{40}$$

The dummy variable δ^k $(k = 1, \cdots, n + n')$ is defined as

$$\delta^k = \begin{cases} 0 & k = 1, \cdots, n \\ 1 & k = n+1, \cdots, n+n' \end{cases} \tag{41}$$

In order to analyze if the estimations of the exponential hazard model based on the extracted database correspond to those based on the original database, a Chow test [17] analyzing the similarity of the results $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\beta}}^t$ is applied. That is, to prove that the exponential hazard models for the condition states are similar the null hypothesis H_0 and the alternative hypothesis H_1 become

$$\begin{cases} H_0 : \hat{r}_{i,m} =_{and} 0 \ (m = 1, 2, 3) \\ H_1 : \hat{r}_{i,m} \neq_{or} 0 \ (m = 1, 2, 3) \end{cases} \tag{42}$$

The null hypothesis H_0 requires that $\hat{r}_{i,m} = 0$ simultaneously to all $m = 1, 2, 3$. The alternative hypothesis H_1, on the other hand, implies that $\hat{r}_{i,m} \neq 0$ for any arbitrary $m = 1, 2, 3$. For the estimation results based on the extraction database to be similar to those based on the original database, the log-likelihood ratio test in relation to the null hypothesis H_0 and the alternative hypothesis H_1 is

$$\xi = 2\left\{\ln[\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{r}})] - \ln[\mathcal{L}(\tilde{\boldsymbol{\beta}})]\right\} \tag{43}$$

$$\ln[\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{r}})] = \sum_{i=1}^{6} \sum_{j=i}^{7} \sum_{k=1}^{n+n'} \bar{\delta}_{ij}^k$$
$$\cdot \ln\left[\pi_{ij}(\bar{Z}^k, \bar{\boldsymbol{x}}^k : \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{r}})\right]$$

$$\ln[\mathcal{L}(\tilde{\boldsymbol{\beta}})] = \sum_{i=1}^{6} \sum_{j=i}^{7} \sum_{k=1}^{n+n'} \bar{\delta}_{ij}^k \ln\left[\pi_{ij}(\bar{Z}^k, \bar{\boldsymbol{x}}^k : \tilde{\boldsymbol{\beta}})\right]$$

where $\hat{\boldsymbol{r}} = (\hat{r}_{1,1}, \cdots, \hat{r}_{6,3})$. In the case $\ln[\mathcal{L}(\hat{\boldsymbol{\beta}}_i, \hat{\boldsymbol{r}}_i)]$ is not restricted, (the null hypothesis H_0 is not restricted) $\ln[\mathcal{L}(\tilde{\boldsymbol{\beta}}_i)]$ express the likelihood under the restrictions of the null hypothesis H_0. In addition, $\tilde{\boldsymbol{\beta}}$ express the maximum likelihood estimations under the null hypothesis H_0. Now, as shown in **Table 2**, depending on the explanatory variable, the exponential hazard model estimations for some parameters $\hat{\beta}_{i,m} = 0$. In the exponential hazard model (39) there are 13 parameters in which $\hat{\beta}_{i,m} \neq 0$. In order to consistently compare the results of the Chow test applied to different databases, it is necessary to unify the degree of freedom of the likelihood ratio tests. For this case, for those parameters estimated by using the original database in which $\hat{\beta}_{i,m} = 0$ it is assumed that $\hat{\beta}_{i,m}^t = 0$. That is, the restriction $\hat{r}_{i,m} = 0$ in the estimation of a hazard model (39). Therefore, the degree of freedom of the likelihood ratio test for the hazard model estimations based on the original data base becomes the number of parameters for which $\hat{\beta}_{i,m} \neq 0$. If these statistical test results ξ fall in the rejection area $\xi \geq \chi^2_{(100-\alpha)}(13)$ the null hypothesis H_0 can be rejected at a $\alpha\%$ significance level. Note, that for $\chi^2_{(100-\alpha)}(13)$ expressing the χ^2 distribution for a degree of freedom 13, 13 express the number of parameters $\hat{\beta}_{i,m} \neq 0$. In the case the null hypothesis H_0 is rejected, it is considered as the exponential hazard model estimations from the extracted database are not identical with those from the original database.

The number of samples extracted from the original database can be determined. Total 7 extracted samples of which number increased from 1000 to 4000 by the step size of 500 were set and were composed by random samples taken from the original database. This procedure is repeated 1000 time for each case and 7000 extraction database are obtained. The statistics test results ξ (equation (43)) are obtained for each exponential hazard model created with each extracted databases. Furthermore, it is determined whether the null hypothesis H_0 is rejected or not for a 1 % significance level for the statistic test results calculated for 1000 extracted database respectively. Now, the fraction $\rho(w)$ of null hypothesis H_0 rejected in relation to the size of the sample w is defined as

$$\rho(w) = \frac{Number\ of\ databases\ rejecting\ H_0}{Total\ database\ size} \tag{44}$$
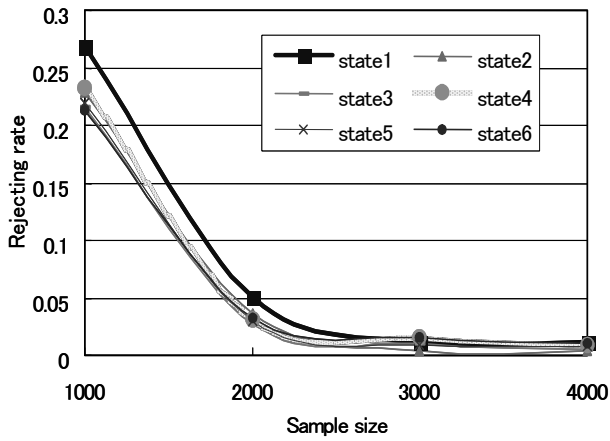
Fig. 6 Sample size and estimation accuracy.

Fig.6 shows the relation between the sample size w and null hypothesis rejecting rate $\rho(w)$. As the number of samples contained in an extraction database increases the number fraction of rejections decreases. As far as this analysis shows, for a sample size of 1000 the null hypothesis H_0 is rejected in more than 25.4% and for a sample size of 1500 it is rejected in more than 9.6%. That is, if an exponential hazard model is estimated using a sample size of 1500, there is a probability of 0.1 that the results greatly differ from those obtained by using the original database. In order to keep the rejection rate of H_0^i below 5%, it is necessary to assure a sample size of at least 2000. Although the analysis made above only applies to the database used in this research, in order to assure an estimation accuracy equivalent to the case shown in any other original database, a considerable sample size (2000 or more) obtained by visual inspection has to be accumulated

6. CONCLUSIONS

In this research a methodology to estimate Markov transition probabilities to forecast the deterioration a bridge component was proposed. The transition progress between a set of condition states representing the conditions of each bridge component were defined by using exponential hazard models. Furthermore, based on periodical inspection data, the maximum likelihood method was proposed to determine the parameters of the exponential hazard model. Heterogeneous inspection data related to specific structural characteristics and usage conditions were employed in order to estimate the deterioration of bridge components in a disaggregate way. The exponential hazard model proposed by this research permits estimating Markov transition probabilities for arbitrary time intervals. In addition, the usefulness of the Markov transition probabilities estimation method proposed by this research was positively verified through an empirical example applied to steel bridges. The methodology proposed by this research can be applied to forecast the deterioration not only of steel bridges but also to other infrastructures if the respective empirical research is accumulated. However, there are still many problems left behind 1) In actual practice of inspection, visual inspection errors arising from misjudgment for example are not deal with in this research. 2) The hazard model used in this research assumes the structure of all bridge component samples is similar. However, an error peculiar to a sample may be included in the hazard function. For this case, it is necessary to apply a different type of model like a mixture hazard model for example. 3) It is assumed that the periodic inspection data at various times is available. In many cases only inspection data at one time is available. Even in such a case, if the data obtained at the time the bridge is opened to the public, Markov transition probabilities can be estimated. An estimation method restricted to this kind of data is required. 4) It is expected that more inspection data will be accumulated in the short future. Therefore it is necessary to develop a methodology to update the parameters of the hazard model as new information is collected.

APPENDIX A. MATHEMATICAL METHOD

1) Derivation of equation (22) Let $\bar{\theta}_{m,j} = \theta_m - \theta_j$.

Equation(22)

$$= \int_0^Z \cdots \int_0^{Z - \sum_{m=i}^{j-3} z_m} \prod_{m=i}^{j-1} \theta_m$$

$$\exp(-\theta_j Z - \sum_{m=i}^{j-2} \bar{\theta}_{m,j} z_m)$$

254s

$$\int_0^{Z-\sum_{m=i}^{j-2} z_m} \exp(-\bar{\theta}_{j-1,j} z_{j-1}) dz_i \cdots dz_{j-1}$$

$$= -\frac{1}{\bar{\theta}_{j-1,j}} \int_0^Z \cdots \int_0^{Z-\sum_{m=i}^{j-3} z_m} \prod_{m=i}^{j-1} \theta_m$$

$$\exp(-\theta_{j-1} Z) \exp\left(-\sum_{m=i}^{j-2} \bar{\theta}_{m,j} z_m + \sum_{m=i}^{j-2} \bar{\theta}_{j-1,j} z_m\right)$$

$$dz_i \cdots dz_{j-2}$$

$$+\frac{1}{\bar{\theta}_{j-1,j}} \int_0^Z \cdots \int_0^{Z-\sum_{m=i}^{j-4} z_m} \prod_{m=i}^{j-1} \theta_m$$

$$\exp(-\theta_j Z) \exp\left(-\sum_{m=i}^{j-3} \bar{\theta}_{m,j} z_m\right)$$

$$\int_0^{Z-\sum_{m=i}^{j-3} z_m} \exp(-\bar{\theta}_{j-2,j} z_{j-2}) dz_i \cdots dz_{j-2}$$

by arranging the first part,

$$-\frac{1}{\bar{\theta}_{j-1,j}} \int_0^Z \cdots \int_0^{Z-\sum_{m=i}^{j-3} z_m} \prod_{m=i}^{j-1} \theta_m$$

$$\exp(-\theta_{j-1} Z) \exp\left(-\sum_{m=i}^{j-2} \bar{\theta}_{m,j-1} z_m\right) dz_i \cdots dz_{j-2}$$

$$= -\frac{\theta_{j-1}}{\bar{\theta}_{j-1,j}} \int_0^Z \cdots \int_0^{Z-\sum_{m=i}^{j-3} z_m} \prod_{m=i}^{j-2} \theta_m$$

$$\exp\left(-\theta_{j-1} Z - \sum_{m=i}^{j-2} \bar{\theta}_{m,j-1} z_m\right) dz_i \cdots dz_{j-2}$$

$$= -\frac{\theta_{j-1}}{\bar{\theta}_{j-1,j}} \pi_{ij-1}$$

and for the second part,

$$-\frac{1}{\bar{\theta}_{j-1,j} \bar{\theta}_{j-2,j}} \int_0^Z \cdots \int_0^{Z-\sum_{m=i}^{j-4} z_m} \prod_{m=i}^{j-1} \theta_m$$

$$\exp(-\theta_j Z) \exp\left(-\sum_{m=i}^{j-3} \bar{\theta}_{m,j} z_m\right)$$

$$\left[\exp\left\{-\bar{\theta}_{j-2,j}\left(Z - \sum_{m=i}^{j-3} z_m\right)\right\} - 1\right] dz_i \cdots dz_{j-3}$$

$$= -\frac{\theta_{j-1} \theta_{j-2}}{\bar{\theta}_{j-1,j} \bar{\theta}_{j-2,j}} \pi_{ij-2}$$

$$+\frac{1}{\bar{\theta}_{j-1,j} \bar{\theta}_{j-2,j}} \int_0^Z \cdots \int_0^{Z-\sum_{m=i}^{j-4} z_m} \prod_{m=i}^{j-1} \theta_m$$

$$\exp(-\theta_j Z) \exp\left(-\sum_{m=i}^{j-3} \bar{\theta}_{m,j} z_m\right) dz_i \cdots dz_{j-3}$$

Thus, π_{ij} becomes

$$\pi_{ij} = -\frac{\theta_{j-1}}{\bar{\theta}_{j-1,j}} \pi_{ij-1} - \frac{\theta_{j-1} \theta_{j-2}}{\bar{\theta}_{j-1,j} \bar{\theta}_{j-2,j}} \pi_{ij-2} - \cdots$$

$$+ \prod_{m=i+1}^{j-1} \frac{1}{\bar{\theta}_{m,j}} \int_0^Z \prod_{m=i}^{j-1} \theta_m \exp(-\theta_j Z)$$

$$\exp(-\bar{\theta}_{i,j} z_i) dz_i$$

$$= -\sum_{k=i+1}^{j-1} \prod_{m=k}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \pi_{ik}$$

$$-\prod_{m=i}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \{\exp(-\theta_i Z) - \exp(-\theta_j Z)\}$$

$$= -\sum_{k=i+1}^{j-1} \prod_{m=k}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \pi_{ik}$$

$$-\prod_{m=i}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \pi_{ii} + \prod_{m=i}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \exp(-\theta_j Z)$$

$$= -\sum_{k=i}^{j-1} \prod_{m=k}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \pi_{ik} + \prod_{m=i}^{j-1} \frac{\theta_m}{\bar{\theta}_{m,j}} \exp(-\theta_j Z)$$

Furthermore by considering $\pi_{i,i} = \exp(-\theta_i Z)$, the next equation is obtained by the equation(19)

$$\pi_{ii+1} = \frac{\theta_i}{\bar{\theta}_{i,i+1}} \{\pi_{ii} + \exp(-\theta_{i+1} Z)\}$$

Now, $\pi_{ij}(i = 1, \cdots, J-1, j = i, \cdots, J-1)$ is a linear combination of $\pi_{i,i}, \pi_{i,i+1}, \cdots, \pi_{i,j-1}$, so the above equation can be expressed as

$$\boldsymbol{\pi}_i = -\boldsymbol{\pi}_i \mathbf{A}_i + \boldsymbol{C}_i$$

However, $\boldsymbol{\pi}_i = (\pi_{i,i}, \pi_{i,i+1}, \cdots, \pi_{i,J-1}).(J-i) \times (J-i)$ order matrix is expressed by

$$\begin{pmatrix} 0 & \frac{\theta_i}{\theta_{i,i+1}} & \frac{\theta_i \theta_{i+1}}{\theta_{i,i+2} \theta_{i+1,i+2}} & \cdots & \prod_{m=i}^{J-2} \frac{\theta_m}{\theta_{m,J-1}} \\ 0 & 0 & -\frac{\theta_{i+1}}{\theta_{i+1,i+2}} & \cdots & \prod_{m=i+1}^{J-2} \frac{\theta_m}{\theta_{m,J-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

For $\alpha_i(p) = i + p - 1$, $\alpha_j(q) = i + q - 1$, the component (p,q) of $\mathbf{A}_i$ becomes

$$\mathbf{A}_i(p,q) = \begin{cases} \prod_{m=\alpha_i(p)}^{\alpha_j(q)-1} \frac{\theta_m}{\theta_{m,\alpha_j(q)}} & \text{when } p < q \\ 0 & p \geq q \end{cases}$$

and for $(J-i)$th order vector, $\boldsymbol{C}_i$ for the element q becomes

$$\boldsymbol{C}_i(q) = \prod_{m=i}^{\alpha_j(q)-1} \frac{\theta_m}{\bar{\theta}_{m,\alpha_j(q)}} \exp(-\theta_{\alpha_j(q)} Z)$$

Let $\mathbf{B}_i$ be the inverse matrix of $\mathbf{A}_i + \mathbf{E}$, so for the component (p,q):

$$\mathbf{B}_i(p,q) = \begin{cases} \prod_{m=\alpha_i(p)}^{\alpha_j(q)-1} \frac{\theta_m}{\theta_{m+1,\alpha_i(p)}} & \text{when } p \leq q \\ 0 & p > q \end{cases}$$

Thus, $\boldsymbol{\pi}_i = \boldsymbol{C}_i \mathbf{B}_i$ and

$$\pi_{ij} = \sum_{l=1}^{J-i} \boldsymbol{C}_i(l) \mathbf{B}_i(l, j-i+1)$$

$$= \sum_{l=1}^{j-i+1} \prod_{m=i}^{i-2+l} \frac{\theta_m}{\bar{\theta}_{m,i-1+l}} \exp(-\theta_{i-1+l}Z)$$

$$\cdot \prod_{m=i-1+l}^{j-1} \frac{\theta_m}{\bar{\theta}_{m+1,i-1+l}}$$

$$= \sum_{k=i}^{j} \prod_{m=i}^{k-1} \frac{\theta_m}{\bar{\theta}_{m,k}} \prod_{m=k}^{j-1} \frac{\theta_m}{\bar{\theta}_{m+1,k}} \exp(-\theta_k Z)$$

2)Time adjustment prove Here, it is shown that the Markov transition probability matrix obtained by means of the exponential hazard model satisfies the time adjustment conditions. When $n = 1$ is self contained. When $n = k-1$, it is assumed $\boldsymbol{\Pi}((k-1)Z) = \{\boldsymbol{\Pi}(Z)\}^{k-1}$. Thus, $\pi_{ii}(kZ) = \pi_{ii}((k-1)Z)\pi_{ii}(Z)$, $\pi_{ij}(kZ) = \sum_{h=i}^{j} \pi_{ih}((k-1)Z)\pi_{hj}(Z)$.

REFERENCES

1) Kaito, K., Abe, M., Fujino, Y. : Prediction of deterioration of bridge structures based on actual inspection data, *JSCE Journal*, No.744/IV-61, pp.29-38, 2003.(in Japanese)
2) Kagimioto, H., Sato, M., Kawamura, M. : Evaluation of the degree of deterioration in ASR damaged concretes and analyses of their pore solutions, *JSCE Journal*, No.641/V-46, pp.241-251, 2002.(in Japanese)
3) Kobayashi, K., Miyagawa, T. : Study on estimation of corrosion rate of reinforcing steel in concrete by measuring polarization resistance, *JSCE Journal*, No.669/V-50, pp.173-186, 2001.(in Japanese)
4) Tsutsumi, T., Shirai, S., Yasuda, N., Matsusima, M. : Evaluation on parameters of chloride induced damage based on actual data in-situ, *JSCE Journal*, No.544/V-32, pp.33-41, 1996.(in Japanese)
5) Ohno, T., Uomoto, K. : Prediction of occurrence of cracks due to autogeneous shrinkage and drying shrinkage, *JSCE Journal*, No.662/V-49, pp.29-44, 2000.(in Japanese)
6) Qi, L., Seki, H., Takagi, M. : Study on corrosion of reinforcing bar due to concrete neutralization under alternate drying and wetting conditions, *JSCE Journal*, No.697/V-54, pp.1-11, 2002.(in Japanese)
7) Saeki, T., Otsuki N., Nagataki, S. : Quantiative estimation of steel corrosion in mortar due to carbonation, *JSCE Journal*, No.532/V-30, pp.55-66, 1996.(in Japanese)
8) Lee, T. C., Judge, G.G., and Zellner, A.: *Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data*, Amsterdam, North-Holland, 1970.
9) Lancaster, T.: *The Econometric Analysis of Transition Data*, Cambridge University Press, 1990.
10) Gourieroux, C.: *Econometrics of Qualitative Dependent Variables*, Cambridge University Press, 2000.
11) Shin, H.C., and Madanat, S.M. : Development of stochastic model of pavement distress initation, *JSCE Journal*, No.744/IV-61, pp.61-67, 2003.
12) Tobin, J.: Estimation of relationships for limited dependent variables, *Econometrica*, 26, pp.24-36, 1958.
13) Amemiya, T. and Boskin, M.: Regression analysis when the dependent variables is truncated lognormal, with an application to the determination of the duration of welfare dependency, *International Economic Review*, Vol.15, p. 485, 1974.
14) Isota, K., Ono, Y.: Handbook of Numerical Methods, Ohmsha Ltd., 1990.
15) State of New York, Department of Transportation: *Bridge Inspection Manual '97*, 1997.
16) Japan Bridge Engineering Center: *Periodical Bridge Inspection Manual (Draft)*, 2005.
17) Morimune, K. : Econometrics, Toyo Keizai INC., 1999.