Poisson Regression Models of Infrastructure Transition Probabilities

By Samer Madanat[1] and Wan Hashim Wan Ibrahim[2]

(Reviewed by the Highway Division)

Abstract: Markovian transition probabilities have been used extensively in the field of infrastructure management, to provide forecasts of facility conditions. However, existing approaches used to estimate these transition probabilities from inspection data are mostly ad hoc and suffer from several statistical limitations. In this paper, econometric methods for the estimation of infrastructure deterioration models and associated transition probabilities from inspection data are presented. The first method is based on the Poisson regression model and follows directly from the Markovian behavior of infrastructure deterioration. The negative binomial regression, a generalization of the Poisson model that relaxes the assumption of equality of mean and variance, is also presented. An empirical case study, using a bridge inspection data set from Indiana, demonstrates the capabilities of the two methods.

INTRODUCTION

In recent years, emphasis has shifted in the field of transportation engineering from the design and construction of new facilities to the maintenance and rehabilitation (M&R) of existing infrastructure facilities. Substantial research has been performed to develop effective and efficient infrastructure management systems. A major focus of this research has been in the area of deterioration modeling and performance analysis (Ben-Akiva and Ramaswamy 1993; Shahin et al. 1987; Butt et al. 1987; Busa 1985).

Infrastructure condition is often represented by discrete condition states (FHWA 1979; Shahin and Kohn 1981). For example, as shown in Table 1, bridge-deck condition is typically represented by condition states that can take values from 0 to 9, where 9 represents a new-deck condition (FHWA 1979). Discrete states are used instead of continuous condition indices primarily for reducing the computational complexity of the M&R decision-making process. The focus of the present paper is the development of infrastructure deterioration models with discrete condition states.

TABLE 1. Concrete Bridge Deck Condition States (FHWA 1979)

	Condition Indicators (% Deck Area)			
States (1)	Spalls (2)	Delaminations (3)	Electrical potentials (4)	Chloride content #/cu yd (5)
9	none	none	0	0
8	none	none	none > 0.35	none > 1.0
7	none	<2%	45% < 0.35	none > 2.0
6	<2% spall or sum of all deteriorated or contaminated deck concrete <20%			
5	<5% spall or sum of all deteriorated or contaminated deck concrete 20–40%			
4	>5% spall or sum of all deteriorated or contaminated deck concrete 40–60%			
3	>5% spall or sum of all deteriorated or contaminated deck concrete >60%			
2	Deck structural capacity grossly inadequate			
1	Deck repairable by replacement only			
0	Holes in deck—danger of other sections of deck falling			

The deterioration models developed are incremental models, because they predict *changes* in condition over time. Incremental models are more realistic representations of the deterioration process than models that predict condition directly because condition at a point in time is a function of both condition at a previous point in time and other explanatory variables such as age, traffic, weather, and maintenance.

Discrete incremental infrastructure deterioration models have been developed as a result of both the realistic nature of incremental models and the discrete representation of facility condition. Specifically, infrastructure management systems have used Markovian transition probabilities as inputs into discrete incremental programming algorithms that are computationally tractable (Carnahan et al. 1987; Feighan et al. 1988; Madanat 1993; Madanat and Ben-Akiva 1994).

Transition probabilities specify the likelihood that the condition of an infrastructure facility will change from one state to another in a unit time. Transitions are probabilistic in nature because infrastructure deterioration cannot be predicted with certainty due to unobserved explanatory variables, the presence of measurement errors, and the inherent stochasticity of the deterioration process. The unit of time on which the incremental models are based is the time between two consecutive inspections (also known as the inspection period). For practical reasons, the time is assumed to be constant. A typical transition probabilities matrix **P** has the following general structure:

$$\mathbf{P} = \begin{bmatrix} p_{kk} & p_{k(k-1)} & \cdot & \cdot & \cdots & \cdot & p_{k1} \\ 0 & p_{(k-1)(k-1)} & p_{(k-1)(k-2)} & \cdot & \cdots & \cdot & p_{(k-1)1} \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & p_{22} & p_{21} \\ 0 & 0 & 0 & 0 & \cdots & 0 & p_{11} \end{bmatrix} \quad (1)$$

where p_{ij} = transition probability from state i to state j; and $i = 1, \ldots, k$ and $j = 1, \ldots, k$ where k = highest condition state, and 1 = lowest condition state.

This structure assumes that a facility can either stay in its current state or deteriorate to some lower state. This implies the absence of rehabilitation activities. The estimation of transition probabilities for the case where rehabilitation is per-

[1]Asst. Prof., School of Civ. Engrg., Purdue Univ., West Lafayette, IN 47907.

[2]Grad. Res. Asst., School of Civ. Engrg., Purdue Univ., West Lafayette, IN.

Note. Discussion open until November 1, 1995. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on April 29, 1994. This paper is part of the ***Journal of Transportation Engineering***, Vol. 121, No. 3, May/June, 1995. ©ASCE, ISSN 0733-947X/95/0003-0267–0272/$2.00 + $.25 per page. Paper No. 8090.

formed represents additional difficulties, which are beyond the scope of the present paper.

The most commonly used approach for the estimation of transition probabilities is the linear regression method (Carnahan et al. 1987; Jiang et al. 1988). This method starts by segmenting infrastructure facilities into groups of homogeneous explanatory variables. The purpose of this segmentation is to capture the fact that transition probabilities are a function of explanatory variables. Subsequently, for each group, a deterioration model with the condition state as the dependent variable and age as the independent variable is estimated by linear regression. Finally, a transition probabilities matrix is estimated for each group by minimizing the sum of absolute (or squared) differences between the expected value of the condition state predicted by the regression model and the theoretical expected value derived from the Markov transition probabilities (Carnahan et al. 1987; Jiang et al. 1988).

The regression-based method fails to capture the structure of the deterioration process because the change in condition within an inspection period is not explicitly modeled as a function of explanatory variables. Segmentation results in a small sample size within each segment, which restricts the number of parameters that can be estimated; as a result some probabilities in the transition matrix are arbitrarily set to zero. Finally, the use of linear regression as a deterioration modeling method is not appropriate because the dependent variable, facility condition state, is discrete rather than continuous (McKelvey and Zavoina 1975).

The remainder of this paper is organized as follows. The validity of Markov processes as models of infrastructure deterioration is investigated. Then the development of deterioration models, based on Poisson regression, and the computation of Markov transition probabilities are described mathematically. In the next section, a generalization of the Poisson regression methodology, the negative binomial model, is discussed. The next sections describe the data set used and the application of the methodology to a set of bridge decks. Finally, conclusions are presented.

TESTING VALIDITY OF MARKOVIAN ASSUMPTION

The usual approach to estimate infrastructure transition probabilities is to assume that infrastructure deterioration can be modeled as a continuous-time Markov process. A continuous-time Markov process has the property that the deterioration process of any facility in an inspection period is independent of the deterioration process in previous inspection periods. Often, the Markovian assumption is made without any verification resulting in models of questionable quality.

To investigate the validity of the Markovian assumption, the deterioration process is described in terms of change in condition states in two consecutive inspection periods. As mentioned previously, an inspection period is defined as the period between two consecutive inspections, and it is assumed to be constant. Thus, for two inspection periods, there are three condition states being observed. Let us denote the condition states observed in the first, second, and third inspections as the past state, the present state, and the future state, respectively. For two consecutive inspection periods, four deterioration sequences are possible. They are

- Sequence 1. No change in condition state occurs during the two consecutive inspection periods; the past, present, and future states are the same.
- Sequence 2. No change occurs in the first inspection period, but a change occurs in the second inspection period; the past state is the same as the present state; however, the future state is lower than both the present and the past states.
- Sequence 3. A change in condition state occurs in the first inspection period, but no change occurs in the second inspection period; the past state is higher than both the present and future states.
- Sequence 4. A change in condition state occurs in both inspection periods; the past state is higher than the present state, and the present state is higher than the future state.

From this description of the deterioration sequences, it can be seen that the deterioration process in the first inspection period for sequences 1 and 2 is the same. Similarly, the deterioration process in the first inspection period for sequences 3 and 4 is also the same. Given the deterioration process in the first inspection period, the proportion of facilities that did not deteriorate and the proportion of facilities that did deteriorate in the second inspection period are computed; these are conditional probabilities for the deterioration in the second inspection period (conditional on the process in the first time period).

If the proportion of facilities in sequence 1 is not significantly different than the proportion of facilities in sequence 3, and the proportion of facilities in sequence 2 is not significantly different than the proportion of facilities in sequence 4, then it can be concluded that the deterioration process in the first inspection period is independent of the deterioration process in the second inspection period; thus, the Markovian property holds.

This test was used to investigate the validity of the Markovian assumption for our data. The data set is described in detail later. From the data, the proportions of facilities in each deterioration sequence are tabulated in Table 2. As shown at the bottom of Table 3, it can be concluded that the Markovian assumption is appropriate for the data. Note that, $L = 1$ and $L = 2$ in Table 3 refers to deterioration sequences

TABLE 2. Validation of Markovian Assumption: Proportion of Facilities in each Deterioration Sequence

Sequences (1)	Number of facilities (2)	Proportion of facilities (3)
1	123	0.651
2	66	0.349
3	120	0.652
4	64	0.348

TABLE 3. Statistical Test of Markovian Assumption

Sequences (1)	Number of facilities $(F_{L,k-1})^*$ $k = 1$ (2)	Expected proportion of facilities (3)	Expected number of facilities $(F_{L,k-2})^*$ $k = 2$ (4)	Total (5)
$L = 1$	123 (123.04)	0.652	123.23 (123.04)	246.23
$L = 2$	66 (65.96)	0.348	65.77 (65.96)	131.77
Total	189	1.00	189	378

Note: H_0: Markov assumption holds: $p_{L,k=1} = p_{L,k=2}$ for all L. Superscript * indicates expected number of facilities when H_0 holds

$$\chi^2 = \frac{(123 - 123.04)^2}{123.04} + \frac{(123.23 - 123.04)^2}{123.04} + \frac{(66 - 65.96)^2}{65.96} + \frac{(65.77 - 65.96)^2}{65.96} = 8.78e-4$$

$$\chi^2_{0.01,1} = 6.63$$

$\chi^2 < \chi_{0.01,1}$ ∴ do not reject H_0

1 and 2, respectively. Also, $k = 1$ refers to the observed number of facilities in sequence 1 or 2, whereas $k = 2$ refers to the expected number of facilities in sequence 1 or 2.

Poisson Regression Model

The Poisson regression model (Lerman and Gonzales 1980) is often a reasonable description for events that occur both randomly and independently over time. For example, the model has been adapted to describe the number of ship accidents (McCullagh and Nelder 1983), total number of patents applied by firms (Hausman et al. 1984), frequency of commuters' route change (Mannering 1989), number of switches in departure time made by commuters (Hatcher and Mahmassani 1992), and number of traffic accidents (Ben-Akiva et al. 1988; Hauer and Persaud 1990). In the present paper, the Poisson regression model is used to construct a discrete incremental deterioration model where the dependent variable is the number of drops in condition state in one inspection period. This model is in turn used to compute the elements of the transition matrix for a data set of bridge decks.

Different incremental deterioration models are estimated for different condition states because, as shown in Table 1, each drop in condition state is a realization of a different mechanistic deterioration process. For example, for new or relatively new bridge decks (state 9 or state 8), the extent of chloride content is the major indicator of deterioration. For moderate and extensive deterioration, on the other hand, the extent of spalls and delamination is the primary indicator of deterioration.

An incremental deterioration model predicts the deterioration of a facility during an inspection period. The dependent variable in an incremental deterioration model is the difference between the condition states of a facility observed in two consecutive inspections. For a facility n in condition state i, the dependent variable Z_{in} can be any integer value greater than or equal to 0, and less than or equal to i. In other words, Z_{in} is the number of drops in condition state taking place during an inspection period.

In a continuous-time Markov process, the amount of time a facility spends in each state is negative exponentially distributed, and is independent of the amount of time the facility spent in its previous state (Ross 1989). Consequently, from basic probability theory, the number of drops in condition state within an inspection period follows a Poisson distribution. Thus, if facility deterioration follows a continuous-time Markov process, then the Poisson distribution is appropriate for developing the incremental deterioration model. Specifically, for condition state i, the number of drops in condition state of facility n within an inspection period is given by the following Poisson probability mass function:

$$p(z_{in} = j) = \frac{e^{-\lambda_{in}}\lambda_{in}^{j}}{j!}, \quad j = 0, 1, 2, \ldots, i,$$

$$i = 1, 2, \ldots, k - 1 \tag{2}$$

where λ_{in} = deterioration rate of facility n in condition state i (the conditional mean of the Poisson distribution); and j = number of drops in condition state in one inspection period.

To develop a deterioration model using this approach, the deterioration rate is expressed as a function of explanatory variables $\mathbf{X}_n$, such as age, incremental cumulative average daily traffic, wearing surface type, and environmental factors. The relationship between λ_{in} and the explanatory variables follows an exponential function as follows

$$\lambda_{in} = e^{(\boldsymbol{\beta}_i'\mathbf{X}_n)} \tag{3}$$

where $\boldsymbol{\beta}_i$ = a vector of parameters to be estimated; and $\mathbf{X}_n$ = a vector of exogeneous variables for facility n. This specification is widely used since it ensures the nonnegativity of the dependent variable (Frome et al. 1973; Lerman and Gonzalez 1980; Cameron and Trivedi 1986, 1990; Hausman et al. 1984; McCullagh and Nelder 1983; Gourieroux et al. 1984). A linear function for λ_{in} would be unrealistic, since it would not impose the positivity constraint on the Poisson parameter. It is noted that λ_{in} is a deterministic function of $\mathbf{X}_n$. Randomness in the model comes from the Poisson specification for the dependent variable.

It should also be noted that the Poisson random variable is not bounded by the preceding, but that we are using it to model a random variable Z_{in} that takes a maximum value equal to i. Although this restriction is theoretically unappealing, it is of little practical consequence, as the predicted probability of Z_{in} exceeding its allowable maximum value is insignificant. It is possible to specify a right-truncated Poisson regression model (Greene 1990). However, empirical work on our data set showed that there was no statistical difference between the models estimated using truncated and nontruncated Poisson regressions.

Maximum likelihood estimation (MLE) is used to estimate the value of the parameter vector $\boldsymbol{\beta}_i$. The likelihood function of the Poisson regression model for condition state i is the following:

$$L_i^* = \prod_{n=1}^{N_i} \frac{e^{-\lambda_{in}}\lambda_{in}^{Z_{in}}}{Z_{in}!} \tag{4}$$

Once the parameters of the incremental deterioration models are estimated by maximizing all L_i^*, they can be used to compute transition probabilities as a function of time, traffic volumes, maintenance actions, and other exogenous variables. Each of the $(k - 1)$ models is used to compute the elements of one row of the transition matrix ($k - 1$ models are needed, since the last row consists of a probability of 1 in the kk entry and 0 in all the rest). Transition probabilities from each condition state i are computed for any group of bridge decks by using a sample enumeration procedure (Ben-Akiva and Lerman 1985). The transition probabilities for each facility in the sample are computed as follows:

$$\hat{p}(Z_{in} = j|\mathbf{X}_n, i) = \frac{e^{-\lambda_{in}}\hat{\lambda}_{in}^{j}}{j!}, \quad j = 0, 1, 2, \ldots, i \tag{5}$$

where $\hat{p}(Z_{in} = j|\mathbf{X}_n, i)$ = the transition probability from state i to state $i - j$ for a facility with attribute vector $\mathbf{X}_n$.

For network-level optimization methods, average transition probabilities are used. Therefore, the individual probabilities are aggregated over the desired groups of facilities to obtain average transition probabilities. This aggregation is shown here

$$\hat{p}_{i(i-j)}^{g} = \frac{1}{N_g}\sum_{n=1}^{N_g} \hat{p}(Z_{in} = j|\mathbf{X}_n, i),$$

$$j = 0, 1, \ldots, i, \quad g = 1, \ldots, G \tag{6}$$

where $\hat{p}_{i(i-j)}^{g}$ = transition probability from state i to state $(i - j)$ for group g, where a group typically consists of facilities of the same class, structural type, geographic area, and design specifications; N_g = total number of facilities in group g; and G = total number of facility groups.

The Poisson regression model has major advantages over the commonly used linear regression approach. First, it allows us to explicitly link deterioration to relevant explanatory variables, as shown in (3), thus eliminating the need for artificially segmenting the sample as is the case for the linear regression method. Moreover, by using the entire data set in the estimation process, a full transition matrix can be estimated and better statistical results can be achieved. Finally, Poisson regression explicitly recognizes the discrete nature of the dependent variable, as opposed to linear regression.

NEGATIVE BINOMIAL REGRESSION MODEL

One of the fundamental properties of the Poisson distribution is that the variance of the random variable is equal to the mean, which is equal to λ_{in}. However, it is not uncommon in real-world problems to encounter the situation where the variance of the data is substantially greater than the mean (the case of "overdispersion") (Dean and Lawless 1989). This is a shortcoming of the Poisson regression where the failure of this assumption has consequences similar to those for heteroscedasticity in the linear regression model: the estimated parameters are consistent, whereas the variances for the parameter estimates are inconsistent (Cameron and Trivedi 1990). Thus, the asymptotic t-test of the parameter estimates will be invalid. However, this restriction can be lifted if we introduce a disturbance term in the parameter of the Poisson distribution. Define λ^*_{in} as the new parameter of the Poisson distribution; this parameter is a random variable that can be written as a function of exogeneous variables as follows:

$$\lambda^*_{in} = e^{(\beta' X_n + \varepsilon_n)} \quad (7)$$

where ε_n = random error term representing the effect of omitted explanatory variables.

After some derivations [see Greene (1991)], a negative binomial probability mass function is obtained. It can be written as follows:

$$p(Z_{in} = j) = \frac{\Gamma\left(\frac{1}{\alpha_i} + j\right)}{\Gamma\left(\frac{1}{\alpha_i}\right) j!} \left(\frac{1}{1 + \alpha_i \lambda^*_{in}}\right)^{1/\alpha_i} \left(1 - \frac{1}{1 + \alpha_i \lambda^*_{in}}\right)^j \quad (8)$$

and

$$\frac{\mathrm{Var}(j|X_n)}{E(j|X_n)} = 1 + \alpha_i E(j|X_n) \quad (9)$$

where $\Gamma(\ )$ = gamma function; and α_i = natural rate of "overdispersion."

As with the Poisson model, it is possible to specify a right-truncated negative binomial model that accounts for the fact that the dependent variable Z_{in} is bounded in (8) by the state i.

The MLE procedure is used to estimate the value of the parameter vector $\boldsymbol{\beta}_i$ and $\boldsymbol{\alpha}_i$ simultaneously. The likelihood function of the negative binomial model for condition state i is the following:

$$L^*_i = \prod_{n=1}^{N_i} \frac{\Gamma\left(\frac{1}{\alpha_i} + Z_{in}\right)}{\Gamma\left(\frac{1}{\alpha_i}\right) Z_{in}!} \left(\frac{1}{1 + \alpha_i \lambda^*_{in}}\right)^{1/\alpha_i} \left(1 - \frac{1}{1 + \alpha_i \lambda^*_{in}}\right)^{Z_{in}} \quad (10)$$

As in the case of the Poisson regression, once the parameters of the incremental deterioration models are estimated by maximizing all L^*_1, the models can be used to compute transition probabilities as a function of explanatory variables. The transition probabilities from each condition state i are computed for any group of bridge decks by using a sample enumeration procedure.

From the results of the negative binomial regression, it is possible to check the validity of the Poisson regression model; specifically from a significance test of $\hat{\alpha}$. The Poisson regression model is appropriate for the data if $\hat{\alpha}$ is statistically close to zero. Therefore, in contrast to the Poisson regression model, the negative binomial model allows the mean of the dependent variable to be different from the variance.

DATA

The source of data for our study is the Indiana State Bridge Inventory database. This data set contains about 5,700 state-owned bridges in Indiana, and it is a subset of the National Bridge Inventory (NBI) database. The NBI database contains Federal Highway Administration required data reported by the states every year on over 550,000 bridges in the country. All federally supported bridges have been inspected every 2 yr beginning in 1978. The data set contains inspection records until year 1986. Each bridge record consists of 58 items on bridge inventory, 16 items on condition and condition appraisal, and another 16 items on the description of proposed renovations (FHWA 1979). The condition evaluation rates the condition of the major bridge components (e.g., deck, superstructure, substructure) on a scale from 0 to 9 with 0 being the worst condition and 9 being the best. The rating scheme required for the evaluation of bridge decks is shown in Table 1.

Each observation consists of two successive facility inspections. According to the federal requirements, the inspection period is every 2 yr. It should be noted that this data set does not include observations of bridge decks with condition states lower than 3; this is due to the strict standards used by the Indiana Department of Transportation.

The effect of climatic conditions is captured by dividing the state of Indiana into two regions, northern and southern, as defined in the Indiana Cost Allocation Study (Sinha et al. 1984).

The variables used in the analysis are summarized in Table 4.

TABLE 4. Description of Variables Used in Analysis

Variable number (1)	Variable name (2)	Variable description (3)
1	ws1	Wearing surface type 1: concrete without protective system (dummy variable)
2	ws2	Wearing surface type 2: asphaltic concrete without protective system (dummy variable)
3	srt1	Structure type 1: simple concrete (dummy variable)
4	reg	Climatic region (north = 1)
5	age	Bridge age: inspection year—construction year; or inspection year—major rehabilitation year; if major rehabilitation year is not zero
6	icadtpl	Incremental cumulative average daily traffic per lane
7	primary highway	Dummy variable for primary highway system classification
8	secondary highway	Dummy variable for secondary highway system classification
9	interstate	Dummy variable for interstate highway system classification
10	numspan	Number of spans

EMPIRICAL RESULTS

The results presented in this section were obtained by using the Poisson and negative binomial regression models. For each condition state, a different incremental deterioration model was estimated for all highway system classes. The estimation results of the Poisson regression model for condition state 7 are shown in Table 5, whereas the estimation results of the negative binomial model for the same state are shown in Table 6. Both models give the same coefficient estimates, and the standard errors of all coefficient estimates are not significantly different between the two models. The estimated natural rate of overdispersion, the last parameter estimated in Table 6, is significant at the 10% level (based on a one-sided test, since the natural rate of overdispersion is strictly nonnegative); therefore, the negative binomial model is the appropriate model for the data. Also, the coefficient estimates and the standard error of all coefficients remain exactly the same as in Table 5 and 6, respectively, when the right-truncated Poisson regression and right-truncated negative binomial models are used in estimating the incremental deterioration models; this was expected because the lowest state to which bridge decks in condition state 7 transfer to in our data is state 4. In the estimation results (not shown in the present paper) of condition state 4, which is the lowest condition state with sufficient observations in our data set, the difference between the estimated coefficients of the right-truncated and nontruncated models were statistically insignificant.

TABLE 5. Estimation Results for Poisson Regression Model (Condition State 7)

Variable name (1)	Condition State 7: Coefficient estimate (2)	Condition State 7: t-statistic (3)
constant	−2.331	−18.131
reg (specific to interstate and primary highway)	0.379	3.459
age	5.691e−3	1.716
icadtpl (specific to secondary highway)	1.896e−4	6.567
numspan (specific to interstate and primary highway)	0.105	3.377
srt1	0.402	3.675
ws1 (specific to interstate and primary highway)	1.098	3.825
ws2 (specific to interstate and primary highway)	1.356	4.510
interstate	−0.721	−2.332
primary highway	−0.681	−2.113

Note: Summary statistics: mean/var of the dependent variable = 0.267/0.308 = 0.867; number of transitions: 0,1,2,3; number of observations = 2,058; $L(0) = -2{,}058$; and $L(\mathbf{B}) = -1{,}188.5$.

TABLE 6. Estimation Results for Negative Binomial Model (Condition State 7)

Variable name (1)	Condition State 7: Coefficient estimate (2)	Condition State 7: t-statistic (3)
constant	−2.331	−16.223
reg (specific to interstate and primary highway)	0.379	3.159
age	5.642e−3	1.367
icadtpl (specific to secondary highway)	1.925e−4	5.190
numspan (specific to interstate and primary highway)	0.105	2.258
srt1	0.402	3.564
ws1 (specific to interstate and primary highway)	1.098	3.544
ws2 (specific to interstate and primary highway)	1.356	4.147
interstate	−0.721	−2.056
primary highway	−0.681	−1.849
alpha	0.194	1.628

Note: Summary statistics: number of transitions: 0,1,2,3; number of observations = 2,058; and $L(\mathbf{B}) = -1{,}260.131$.

Referring to Tables 5 and 6, the coefficient for region is as expected: on interstate or primary highway, bridge decks in the northern region deteriorate at a faster rate than bridge decks in the southern region. This is due to the corroding effect of deicing salts, which are used more frequently in the north, especially on interstate or primary highway. Facility-age influences the rate of deterioration: as expected, bridge decks on older bridges have a higher rate of deterioration than bridge decks on new bridges; this clearly illustrates the nonstationarity of the deterioration process, at least for the facilities in this sample. The incremental average daily traffic affects the deterioration rate of bridge decks; on secondary highway, the deterioration rate of bridge decks increases with traffic. This may be because bridges on secondary highways are designed to carry smaller volumes of traffic. The number of spans affects the deterioration rate: on interstate and primary highways, the deterioration rate of bridge decks increases with the number of spans; this correlates with earlier findings (Busa 1985). The facilities' structural types also influence the rate of deterioration; simple concrete bridge decks deteriorate at a faster rate than other types of bridge-deck structures.

Wearing surface type 1 and wearing surface type 2 are wearing surfaces without any protective system. The coefficients for both types of wearing surface are as expected: on interstate or primary highway, bridge decks without protective systems have higher rates of deterioration than bridge decks that have some type of protective system.

The coefficients for interstate and primary dummy variables are as expected; bridge decks on interstate and primary highway deteriorate slower than bridge decks on secondary highway. Also the deterioration rate of bridge decks on interstate highway is relatively slower than the deterioration rate of bridge decks on primary highway. This may be because design standards are more stringent for bridge decks on interstate highways.

The sample enumeration procedure was used to compute aggregate transition probabilities for bridge decks of condition state 7. The results are compared to the observed frequencies. As can be seen in Table 7, there does not seem to

TABLE 7. Transition Probabilities for Condition State 7 (all Highway Classes)

Z_7 (1)	Observed numbers (2)	Observed frequencies (3)	Poisson regression (4)	Negative binomial (5)
0	1,602	0.778	0.775	0.781
1	385	0.187	0.190	0.181
2	49	0.024	0.031	0.032
3	22	0.011	0.004	0.006

Note: Chi-square test: H_0—estimated transition probabilities are not statistically different than transition probabilities obtained using observed frequencies

$\chi^2_{\text{poisson regression}} = 8.274$

$\chi^2_{\text{negative binomial}} = 5.401$

$\chi^2_{\text{critical}(0.01;3)} = 11.34$

be any significant difference between the transition probabilities computed using the negative binomial model and the observed frequencies. Also, the results of the Poisson regression model are relatively close to those obtained using the negative binomial model.

CONCLUSIONS

Poisson and negative binomial models for the estimation of infrastructure transition probabilities were presented. These methods were specifically developed for models in which the dependent variable takes nonnegative integer values. The Poisson regression model recognizes the Markovian nature of infrastructure deterioration and explicitly links deterioration to the relevant explanatory variables. The negative binomial model is an extension of the Poisson model that relaxes the assumption of equality between the mean and variance of the dependent variable. For our data set, this assumption is not satisfied, as shown in the results; thus, the negative binomial model is accepted. As demonstrated, the proposed methods provide accurate estimates of transition probabilities.

ACKNOWLEDGMENT

We are grateful to the Indiana Department of Transportation for allowing us to use their bridge-inspection database in this research.

APPENDIX. REFERENCES

Ben-Akiva, M., Ceder, A., Cheng, L., and Liss, M. (1988). "Estimation of intersection traffic safety improvements." *Final Rep., Prepared for Massachusetts Department of Public Works*, Ctr. for Transp. Studies, MIT, Cambridge, Mass.
Ben-Akiva, M., and Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand. The MIT Press*, Cambridge, Mass.
Ben-Akiva, M., and Ramaswamy, R. (1993). "An approach for predicting latent infrastructure facility deterioration." *Transp. Sci.*, 27(2), 174–193.
Busa, G. D. (1985). "Modelling concrete bridge deck deterioration," MSc thesis, Massachusetts Institute of Technology, Cambridge, Mass.
Butt, A. A., Feighan, K. J., Shahin, M. Y., and Carpenter, S. (1987). "Pavement performance prediction model using the Markov process." *Transp. Res. Record 1123*, Transp. Res. Board, National Research Council, Washington, D.C.
Cameron, A. C., and Trivedi, P. K. (1986). "Econometric models based on count data: comparisons and applications of some estimators and tests." *J. Appl. Econometrics*, 1, 29–53.
Cameron, A. C., and Trivedi, P. K. (1990). "Regression-based tests for overdispersion in the Poisson model." *J.Econometrics*, 46, 347–364.
Carnahan, J. V., Davis, W. J., and Shahin, M. Y. (1987). "Optimal maintenance decisions for pavement management." *J. Transp. Engrg.*, ASCE, 113(5), 554–572.
Dean, C., and Lawless, J. F. (1989). "Tests for detecting overdispersion in Poisson regression models." *J. Am. Statistical Assoc.*, 84(406), 467–472.
Federal Highways Administration (FHWA). (1979). "Recording and coding guide for structure inventory and appraisal of the nation's bridges." *Rep.*, U.S. Department of Transportation, Washington, D.C.
Feighan K. J., Shahin, M. Y., Sinha, K. C., and White T. D. (1988). "Application of dynamic programming and other mathematical techniques to pavement management systems." *Transp. Res. Record 1200*, Transp. Res. Board, National Research Council, Washington, D.C., 90–98.
Frome, E. L., Kutner, M. H., and Beauchamp, J. J. (1973). "Regression analysis of Poisson-distributed data." *J. Am. Statistical Assoc.*, 68(344), 935–940.
Gourieroux, A., Monfort, A., and Trognon, A. (1984). "Pseudo maximum likelihood methods: applications to Poisson models." *Econometrica*, 52(3), 701–720.
Greene, W. H. (1990). *Econometric analysis*, MacMillan Publishing Co., New York, N.Y.
Greene, W. H. (1991). "LIMDEP version 6.0." *User's manual and reference guide*. Econometric Software, Inc., Bellport, N.Y.
Hatcher, S. G., and Mahmassani, H. S. (1992). "Daily variability of route and trip scheduling decisions for the evening commute." *Transp. Res. Record 1357*, Transp. Res. Board, National Research Council, Washington, D.C.
Hauer, E., and Persaud, B. (1990). "How to estimate the safety of rail-highway grade crossings and the safety effect of warning devices." *Transp. Res. Record 1114*, Transp. Res. Board, National Research Council, Washington, D.C., 131–140.
Hausman, J., Hall, B., and Griliches, Z. (1984). "Econometric models for count data with an application to the patents—R&D relationship." *Econometrica*, 52(4), 909–938.
Jiang, Y., Saito, M., and Sinha, K. C. (1988). "Bridge performance prediction model using the Markov Chain." *Transp. Res. Record 1180*, Transp. Res. Board, Washington, D.C., 25–32.
Lerman, S. R., and Gonzales, S. L. (1980). "Poisson regression analysis under alternate sampling strategies." *Transp. Sci.*, 14(4), 346–364.
Madanat, S. (1993). "Optimal infrastructure management decisions under uncertainty." *Transp. Res. C*, 1(1), 77–88.
Madanat, S., and Ben-Akiva, M. (1994). "Optimal inspection and repair policies for infrastructure facilities." *Transp. Sci.*, 28(1), 55–62.
Mannering, F. (1989). "Poisson analysis of commuter flexibility in changing route and departure times." *Transp. Res.*, 23B(1), 53–60.
McCullagh, P., and Nelder, J. A. (1983). *Generalized linear models*. Chapman and Hall, London, England.
McKelvey, R. D., and Zavoina, W. (1975). "A statistical model for the analysis of ordinal level dependent variables." *J. Math. Sociology*, 4(2), 103–120.
Ross, S. M. (1989). "Introduction to probability models." 4th Ed., Academic Press, Inc., New York, N.Y.
Shahin, M., and Kohn, S. (1981). "Pavement maintenance management for roads and parking lots." *Tech. Rep. M-294*, Constr. Engrg. Res. Lab., U.S. Army Corps of Engineers, Champaign, Ill.
Shahin, M. Y., Nunez, M. M., Broten, M. R., Carpenter, S. H., and Sameh, A. (1987). "New techniques for modelling pavement deterioration." *Transp. Res. Record 1123*, Transp. Res. Board, National Research Council, Washington, D.C.
Sinha, K. C., Fwa, T. F., Sharaf, E. A., Tee, A. B., and Michael, H. L. (1984). "Indiana highway cost allocation study." *Final Rep. FHWA/IN/JHRP-84-20*, Purdue University, West Lafayette, Ind.