Poisson Model of Construction Incident Occurrence

D. K. H. Chua, M.ASCE,[1] and Y. M. Goh[2]

Abstract: Construction incidents are essentially random events because they have a probabilistic component that causes their occurrence to be indeterministic. Thus, as with most random events, one of the best ways to understand and analyze construction incidents is to apply statistical methods and tools. Consequently, this paper presents a statistical framework based on the modified loss causation model (MLCM). Even though the MLCM has been used for the framework, the approach can be readily adapted for other incident causation models. The MLCM is separated into two basic components: random and systematic. The random component is represented by a probability density function (PDF), which has parameters influenced by the systematic component of the MLCM, while the systematic component is represented by the situational variables and quality of the safety management system. In particular, this paper proposes that the PDF can be represented by the Poisson distribution. Besides being a convenient and simple distribution that can be easily used in applications, the Poisson distribution had been used in various industries to model random failures or incidents. The differences in contexts and the undesirable effects of adopting an unrepresentative distribution will require formal analysis to determine the suitability of the Poisson distribution in modeling the random component of construction incident occurrence. Incident records for 14 major projects were used in the analysis. Hypothesis testing using the chi-square goodness-of-fit and dispersion tests shows that the incident occurrences can be modeled as a Poisson process characterized by some mean arrival rate. The paper also presents some applications of the proposed Poisson model to improve construction safety management, focusing on two specific concepts: the Bayesian approach and the partitioned Poisson.

DOI: 10.1061/(ASCE)0733-9364(2005)131:6(715)

CE Database subject headings: Statistical models; Construction site accidents; Occupational safety; Poisson density functions.

Introduction

In recent years there has been increased emphasis on construction safety management due to new legislation in various countries, such as the United Kingdom, Hong Kong, and Singapore, and the publication of several safety management system (SMS) standards, such as BS 8800 (BSI 1996), OHSAS 18001 (BSI 2000), and CP 79 (PSB 1999). These documents generally agree that the risk management process is a key to safety management. Furthermore, the documents also indicate that safety risk is made up of two parts: likelihood of occurrence of an incident, and severity of the incident's consequences.

An incident is defined as an unintentional and undesirable event that may or may not result in an injury, and an incident that results in an injury or fatality is defined as an accident.

This definition clearly indicates that the occurrence of a construction incident is a random event caused by such factors as window of accident opportunity, chance, and luck, which are frequently mentioned in the incident causation literature (Ramsey 1985; Sanders and Shaw 1988; Reason 1990; McKinnon 2000). However, randomness does not refer to events without cause or unaffected by human actions, but instead to the presence of variations. In the statistical sense, variation means that two situations with similar characteristics will not guarantee the same outcome (Montgomery and Runger 1999).

This random process may be modeled statistically to systematically characterize and analyze the risk posed by construction incidents. A statistical approach will allow analysis of construction incidents to be based on a stable and sound foundation provided by mathematical boundaries and reasoning, thus improving the effectiveness of safety management. However, there has been a lack of formal studies to model construction incident occurrence statistically. Most of the past statistical studies on construction incidents (Jeong 1998; Cattledge et al. 1996; Kartam and Bouz 1998; Hinze et al. 1998; Larsson and Field 2002) have been focused on summarizing incident data obtained from different sources.

In contrast, traffic safety and reliability engineering researchers have utilized numerous probability distributions to model the occurrence of incidents in their respective areas, and one of the most commonly used is the Poisson distribution (Bendell 1991; Modarres et al. 1999; Fridstrom et al. 1995). However, because of differences in the environment, scale, and nature of the processes, it would be prudent to verify that the distributions, in particular the Poisson distribution, are suitable for modeling the occurrence of construction incidents. Consequently this paper proposes a statistical interpretation of construction incident occurrence. Incident data from a number of construction sites have been used in the statistical tests to determine the goodness of fit of the Poisson

[1]Associate Professor, Dept. of Civil Engineering, National Univ. of Singapore, Singapore 119260. E-mail: cvedavid@nus.edu.sg

[2]Research Engineer, Dept. of Civil Engineering, National Univ. of Singapore, Singapore 119260. E-mail: engp0323@nus.edu.sg

Note. Discussion open until November 1, 2005. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on March 3, 2003; approved on March 2, 2004. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 131, No. 6, June 1, 2005. ©ASCE, ISSN 0733-9364/2005/6-715–722/$25.00.

J. Constr. Eng. Manage., 2005, 131(6): 715-722

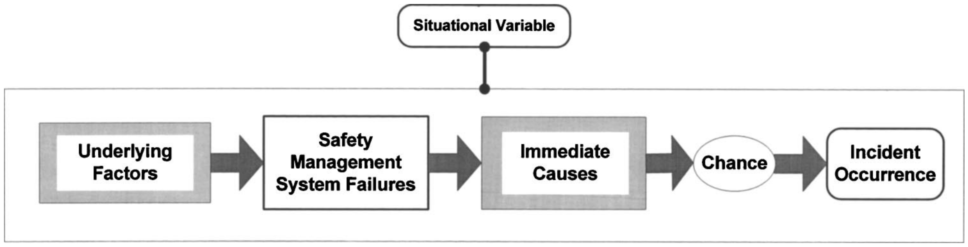

Fig. 1. Main elements of modified loss causation model (MLCM)

distribution in modeling the randomness of construction incident occurrence. Following this, the model may naturally be extended to incorporate the concept of a partitioned Poisson. Applications of the model to safety management system are also presented.

Statistical Model of Construction Incidents

A statistical model of construction incidents should highlight the random nature of construction incident occurrence and at the same time be aligned with a sound paradigm of construction incident causation. The statistical model presented here is based on the modified loss causation model (MLCM) of Chua and Goh (2004), although it may be easily applied to other incident causation models.

The main elements of the MLCM are shown in Fig. 1, comprising the situational variables, incident sequence, immediate causes, safety management system (SMS) failures, and underlying factors. The situational variables describe the context of the incident, which includes the type of project, key work activities involved, and type of equipment or machinery used. The incident sequence describes the key events during the occurrence of the incident, which may be decomposed into breakdown event, contact event, and consequences of the incident. The immediate causes are the directly attributable causes of the incident and can be classified as substandard acts and substandard conditions. SMS failures are safety measures of the SMS that failed to prevent the occurrence of the immediate causes. Underlying factors are organizationally deep rooted, relatively subjective, and contributory in nature and include broad factors such as safety culture, organization structure, and staff motivation. Furthermore, as incident occurrence is a random event, a chance factor has been added to the MLCM to reflect the probabilistic nature of incident occurrence.

From a statistical point of view, the MLCM can be reorganized and interpreted as in Fig. 2, where the model is now separated into two key components: random and systematic. The random component is inherent or objective, that is, the randomness is irremovable and uncontrollable. It is usually described by a probability density function (PDF)

$$f(\mathbf{\Phi},t) = P(X = x) \tag{1}$$

where $\mathbf{\Phi}$ is a vector representing the parameters of the PDF; t is the amount of exposure, for instance, time or man-hours worked; X is a random variable representing the number of incidents for t exposure; and x is a specific value of X, for example, a specific number of incidents for t exposure.

The systematic component, on the other hand, comprises the conditions or factors that are relatively controllable, and these would be the factors in the process system that influence the values of the parameters, $\mathbf{\Phi}$, of the PDF in Eq. (1). As depicted in Fig. 2, $\mathbf{\Phi}$ is a vector of dependent parameters influenced by a set of systematic independent variables. In the MLCM, these variables are broadly defined as situational variables and the quality of the SMS, which can be further categorized into more detailed variables. Even though immediate causes and underlying factors also influence $\mathbf{\Phi}$, they have been excluded and deemed to be represented by the SMS quality, with which they are highly correlated. This will remove multicollinearity problems in the model and retain the independence assumption required in many statistical methods. Moreover, among the three variables, quantitative measures of the SMS quality appear to be more available due to the rising trend of utilizing quantitative SMS audit checklists in the construction industry.

The PDF [function $f(\cdot)$ in Eq. (1)] that describes the random nature of construction incident occurrence is fundamental to the model. The choice of the PDF will define $\mathbf{\Phi}$ and hence the complexity of the statistical analyses. The range of possible PDFs is very wide, but ideally it should be simple to use and practical. On

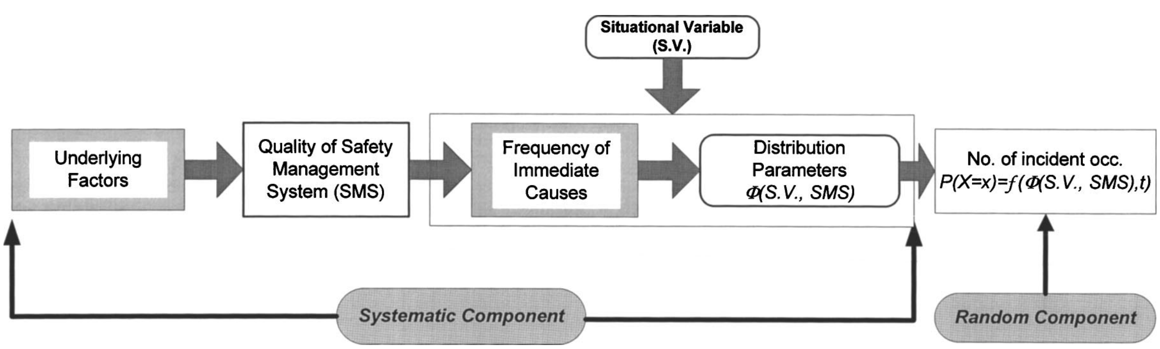

Fig. 2. Statistical interpretation of MLCM

J. Constr. Eng. Manage., 2005, 131(6): 715-722

this count, the Poisson distribution is definitely one of the most preferred PDFs, as it only has one parameter. Its suitability to model the randomness of construction incident occurrence will be verified subsequently.

Poisson Process Model

In the context of this paper, the Poisson process can be considered a type of counting process with the random variable $X(t)$ $(t \geqslant 0)$ representing the total number of construction incidents that had occurred up to time t. In the broader sense, the variable t need not be time, but can be any continuum such as space and man-hours worked (number of workers $\times$ average number of hours worked). In comparison to conventional time intervals, man-hours worked will be able to better reflect the amount of activity on site, which in turn directly reflects the risk exposure and hence the probability of incident occurrence. Thus man-hours worked has been used for the statistical model presented here.

If construction incident occurrence follows a Poisson process, an interval of t man-hours worked can be partitioned into n subintervals of small enough length (t/n) such that at most one incident occurs within each subinterval. This innocuous condition is necessary to facilitate the derivation of the Poisson distribution based on the binomial distribution, and it is a reasonable assumption in the construction context. For instance, an appropriate subinterval would be one man-minute worked or even one man-second worked; that is, the probability that more than one incident occurs in an infinitesimally small interval is zero.

Another assumption in the Poisson process is the mutual independence of the number of incidents in disjoint intervals. This assumption is reasonable in the case of a construction project, which is composed of many workers performing diverse activities at any time.

Consequently, in a Poisson distribution, the number of incident occurrences in an interval t, $X(t)$, with $\lambda(>0)$ as the mean number of incidents in t man-hours worked, can be represented by the probability mass function (PMF), which generally can be interpreted as the PDF of discrete random variables.

$$f(x) = P[X(t) = x] = \frac{e^{-\lambda}\lambda^{x}}{x!} \tag{2}$$

Eq. (2) is based on the assumption that λ is constant. If this assumption is relaxed such that the probability of one incident in an interval is not constant but a function of independent variables (Fridstrom et al. 1995) such as situational variables (SVs) and SMS quality, then the PMF is modified as

$$f(x) = P(X(t) = x) = \frac{e^{-\lambda(\mathrm{SV},\mathrm{SMS})}[\lambda(\mathrm{SV},\mathrm{SMS})]^{x}}{x!} \tag{3}$$

In this case, the distribution is known as the nonhomogeneous Poisson distribution (Ross 2000). With reference to the model depicted in Fig. 2, function $f(\cdot)$ in the figure would refer to the nonhomogeneous Poisson distribution in Eq. (3), with $\lambda(\mathrm{SV},\mathrm{SMS})$ as the corresponding $\mathbf{\Phi}(\mathrm{SV},\mathrm{SMS})$. Note that generally the SV and SMS of construction projects can vary as construction work progresses, but these variations are usually not significant. It is shown in subsequent analyses that a homogeneous Poisson distribution would be an adequate model generally. Only one project out of the 14 cases studied showed some significant change in the SMS or SV.

Table 1. List of Contracts Chosen for Analysis

Contract	Contract description
A	Aboveground construction
B	Aboveground construction
C	Other underground construction work
D	Underground station construction
E	Underground station construction
F	Underground station construction
G	Underground station construction
H	Underground station construction with tunneling work
I	Underground station construction with tunneling work
J	Underground station construction with tunneling work
K	Underground station construction with tunneling work
L	Underground station construction with tunneling work
M	Underground station construction with tunneling work
N	Underground station construction with tunneling work

Validating Poisson Model for Construction Incidents

Data Source

The data for this study have been obtained from the Safety Department of the Land Transport Authority (LTA) of Singapore. Since 1998, the LTA has implemented the computerized Safety Information System (SITS) to capture information on incidents that occurred on LTA construction sites. The SITS contains incidents of all severity, from incidents with no injury to incidents involving fatalities, and the LTA has been nurturing a transparent and nonpenalizing culture where reporting of these incidents is greatly encouraged. This approach has helped them to collect a relatively large amount of incident information.

In all, 14 contracts with sufficient data points for statistical inference were chosen for the analysis, as depicted in Table 1. All 14 are part of the Mass Rapid Transit (MRT) or Light Rail Transit (LRT) construction projects that either have been recently completed or are still ongoing. These projects are generally large projects with an average of 340 workers on site per day and project duration of about 4 to 5 years. Most of the contracts involve construction of railway stations, and the majority of the stations are underground. Besides construction of underground stations, contracts H to N also include considerable tunneling work. Unlike the other underground construction contracts, Contract C involves the construction and installation of railway components in the underground tunnels. The above-ground construction contracts include construction of above-ground stations, a train depot for parking and maintenance of the trains, and viaducts for the railway system.

Furthermore, to minimize effects due to instability in the reporting and recording of incidents during the early stage of implementing SITS in the projects, initial data of about 100,000 man-hours have been removed from each contract. This corresponds to between 4 and 6 months of the contracts. The data during this period have demonstrated exceptionally high variance or an exceptionally low number of incidents and would introduce unnecessary noise into the analysis if included. These contracts are generally over 3.5 years long, involving several million man-hours, so that the data discarded represent only a small portion of the project.

J. Constr. Eng. Manage., 2005, 131(6): 715-722

Table 2. Analysis Results Based on Complete Contract Data

Contract	Dispersion test P-value	Chi-square test P-value	Mean arrival rate ($\hat{\lambda}$)	Coefficient of variance	Number of intervals of 50,000 mhr
A	**0.002**	**0.005**	0.209	1.190	532
B	0.529	0.948	0.154	0.989	188
C	0.719	0.635	0.455	0.892	66
D	0.055	0.396	0.551	1.239	98
E	0.238	0.437	0.795	1.145	44
F	0.562	0.792	0.295	0.972	105
G	0.058	0.101	0.748	1.216	115
H	0.169	0.445	0.605	1.093	210
I	0.016	0.476	0.851	1.243	175
J	0.056	0.012	0.416	1.209	125
K	0.065	0.100	1.258	1.238	89
L	0.043	0.594	1.024	1.279	85
M	0.337	0.882	0.689	1.053	103
N	0.034	0.288	1.255	1.284	94

Goodness-of-Fit Test

The appropriateness of the Poisson distribution in modeling the random component of incident causation has been tested using the chi-square goodness-of-fit test (Conover 1980; Bendell 1991) and the dispersion test (Cox and Lewis 1966; Nicholson 1986; Nicholson and Wong 1993). The former is one of the most commonly used tests to determine the goodness of fit of a distribution to some observed data, in which each data point is assumed to be an independent observation of the random variable $X(t)$. The statistic, T, for this test follows the chi-square distribution and is given by

$$T=\sum_{i=1}^{c}\frac{(O_i-E_i)^2}{E_i} \tag{4}$$

where O_i is the number of observed data in class i of the data (e.g., class i may be the class with x_i incidents in the time intervals observed), and E_i the expected number of observed data in that class as given by the Poisson distribution so that

$$E_i=p_i\cdot\Sigma O_i \tag{5}$$

where p_i is the probability that $X(t)=x_i$, given by Eq. (1).

The classes for the test have been designed carefully to ensure that the assumptions of the test are not violated. For example, small values of the expected number E_i of observed data in class i can lead to a poor match between the chi-square distribution and the actual distribution of T. This problem is resolved by applying a conservative rule of thumb proposed by Cochran (1954), in which the expected number of occurrences in each class must be greater than one, and more than 80% of the expected number of occurrences in all classes must be greater than five. If the expected number of occurrences in a class is too low, the class is merged with adjacent classes to increase the E_i (Montgomery and Runger 1999).

Dispersion Test

A key characteristic of the Poisson process is that the mean rate of arrival, λ, is equal to the standard deviation. As demonstrated by Bendell (1991), the chi-square test is not able to detect whether the sample's coefficient of variation (standard deviation/mean) is significantly different from unity. Instead, the dispersion test (Nicholson 1986; Nicholson and Wong 1993; Cox and Lewis 1966) has been utilized to validate this aspect of the model. The statistic H for this test also has a chi-square distribution and is given by

$$H=\sum_{i=1}^{c}\frac{(x_i-\overline{x})^2}{\overline{x}} \tag{6}$$

where x_i is the number of incident occurrences in the time interval i, and $\overline{x}$ is the average number of incident occurrences in an interval, with c being the number of intervals in the sample.

In the analysis, the time interval size of 50,000 man-hours worked was chosen arbitrarily to ensure meaningful aggregation of incident occurrences. It is sufficiently small to prevent loss of information when incident counts are merged into large intervals. On the other hand, the intervals are not too small to cause the significance of the errors contained in the data to be amplified. In this regard, all the samples obtained for the contracts have more than an adequate number of intervals and incidents necessary for a valid test. Specifically, the number of time intervals range from 44 to over 500 (Table 2) and the minimum number of incidents in the samples is 37, well exceeding the recommended minimum of 20 and 33, respectively (Nicholson and Wong 1993).

Discussion of Results

The results of the above tests are shown in Table 2 as P-values corresponding to the probabilities for the chi-square of the respective computed T and H statistics. As is evident from Table 2, the data from all the contracts except for Contract A have P-values well exceeding 0.01, indicating that the distribution of the observed data corresponds well to that of a homogeneous Poisson distribution. Closer analysis of the observed data in Contract A also shows that the nonhomogeneous Poisson distribution would be equally valid, as will be discussed shortly.

The mean arrival rates, $\hat{\lambda}$, range from 0.154 to about 1.26 incidents per 50,000 man-hours worked at the sites. This mean arrival rate is a parameter of the Poisson process, which is dependent on the systematic factors contributed by the situational variables and the quality of the SMS, as expressed in Fig. 2.

J. Constr. Eng. Manage., 2005, 131(6): 715-722

Table 3. Contract Descriptions Based on Ranked Arrival Rates

Contract	Mean arrival rate ($\hat{\lambda}$)	Contract description
B	0.154	Aboveground construction
A	0.209	Aboveground construction
F	0.295	Underground station construction
J	0.416	Underground station construction with tunneling work
C	0.455	Other underground construction work
D	0.551	Underground station construction
H	0.605	Underground station construction with tunneling work
M	0.689	Underground station construction with tunneling work
G	0.748	Underground station construction
E	0.795	Underground station construction
I	0.851	Underground station construction with tunneling work
L	1.024	Underground station construction with tunneling work
N	1.255	Underground station construction with tunneling work
K	1.258	Underground station construction with tunneling work

Generally it can be observed from Table 3 that the contracts with only aboveground construction (contracts A and B) have the lowest incidents in contrast to the contracts with tunneling works, which would have greater exposure to risks, having an average mean rate of 0.871 incidents per 50,000 man-hours worked. On average, the contracts with underground station works alone are intermediate, with an average mean rate of 0.597 incidents per 50,000 man-hours worked. The dispersion of the mean rates within each category of works may be attributed to the other systematic factors, including SMS quality. A precise correlation of the mean arrival rates to the systematic factors warrants a more detailed study, which is presently outside the scope of this paper.

With respect to Contract A, which failed the homogeneous Poisson distribution test above, an analysis of the incident rate over time shows a significant reduction in the rate of incidence after interval 120 in Fig. 3, depicting the number of incidents over time intervals of 50,000 man-hours worked. Additional tests were performed by dividing the sample into two segments, separating the difference. Three possible separation points were chosen to ensure that there are at least 34 incidents in each segment. These results are shown in Table 4, showing that the Poisson process, albeit a nonhomogeneous one, is indeed valid as well. The average mean rate for the initial stages of the project was about 0.40

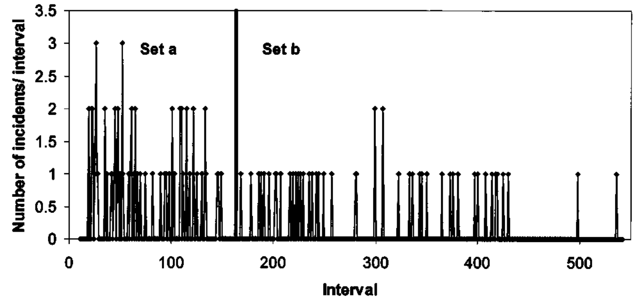

Fig. 3. Time-series plot of number of incidents per 50,000 man-hours for Contract A

incidents per 50,000 man-hours worked, compared to the significantly reduced mean rate of about 0.12 incidents per 50,000 man-hours worked. This reduction could be attributed to the difference in the nature of work of the first part involving some basement construction or due to significant improvement in the SMS following the occurrence of earlier incidents.

Application of Poisson Process Model in Construction Safety Management System

The validity of the Poisson process model presented earlier could facilitate a more statistical approach in construction safety management. Fig. 4 depicts the core elements of OHSAS 18001 (BSI 2000), which includes occupational health and safety (OH&S) policy, planning, implementation and operation, checking and corrective action, and management review. As shown in Fig. 4, the parameter of the Poisson distribution, λ, can be used as a quantitative goal or key performance indicator of the OH&S policy. This quantitative goal can be indicated as λ_G.

During the planning stage, or more specifically the risk assessment component of the planning element, the risk level of different work activities or work contexts is evaluated. Using various statistical methods, an estimate of the Poisson parameter can be determined, that is, $\hat{\lambda}_P$. Subsequently, the Bayesian approach will be elaborated to illustrate how $\hat{\lambda}_P$ can be adjusted to incorporate past experience. At this time also, additional safety measures may be devised to improve the expected incident rate.

After the plans have been implemented, on-site data will be collected and statistical analyses can be applied to compare and evaluate the differences between the actual performance, $\hat{\lambda}_O$, and λ_G or $\hat{\lambda}_P$. The λ value of other projects or companies can also be used in the comparison for benchmarking purposes. Such statis-

Table 4. Additional Test Results for Contract A

Separation point (interval number)	Subset	Dispersion test *P*-value	Chi-square test *P*-value	Mean arrival rate ($\hat{\lambda}$)	Coefficient of variance	Number of intervals of 50,000 man-hours
140	a	0.064	0.037	0.469	1.196	130
	b	0.794	0.678	0.124	0.782	402
180	a	0.025	0.021	0.388	1.225	170
	b	0.687	0.859	0.124	0.967	362
220	a	0.037	0.063	0.357	1.182	210
	b	0.477	0.955	0.112	1.002	322

J. Constr. Eng. Manage., 2005, 131(6): 715-722

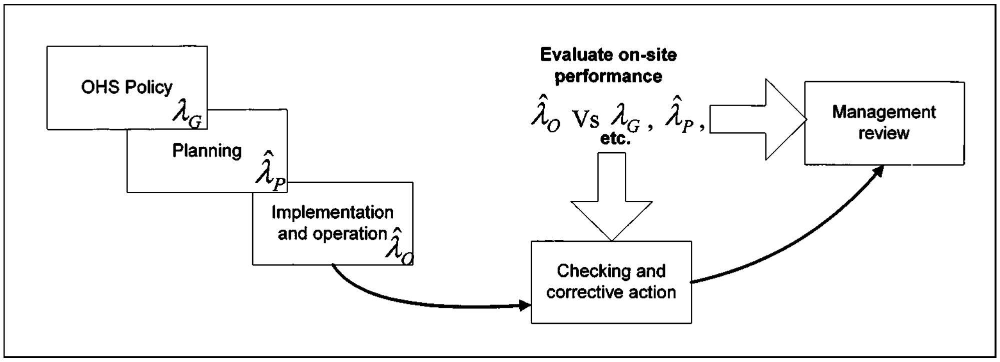

Fig. 4. Application of Poisson model and its parameter, λ, to facilitate safety management

tical tests can also help management in measuring the effectiveness of safety measures and policies by comparing the incident rate before and after implementation. Moreover, using regression analysis and a larger database set, the hazard level measured by λ can be determined for different work situations or situational variables (SVs).

Poisson Model with Bayesian Updating

The Bayesian approach (Ang and Tang 1975) is a systematic technique to incorporate objective data, such as incident occurrence data, into subjective information, such as judgment, experience, and intuition. This is especially important in construction due to the unique and "one-off" nature of its projects, where incident data are relatively scarce and there is strong dependence on engineering judgment. The Bayesian approach is unconstrained by the size of the incident data. Even in the event of no incidents, or having data based on a limited time frame, the approach can still be employed to update prior estimates.

Both sources of input for the Bayesian approach—subjective and objective data—are based on consideration of the nature of the project and safety management capability of relevant organizations in the project (as in Fig. 2). For instance, an expert assigning a likelihood incident rate value, λ, would be considering factors such as nature of work, type of equipment used, and track records of companies involved. Similarly for objective data, only data from projects with characteristics similar to those of the current project are deemed relevant in the analysis.

Accordingly, if $f'(\lambda)$ is the prior (initial estimated) distribution of the incident rate, λ, the posterior (revised) distribution $f''(\lambda)$ after incorporating incident observations may be obtained through Bayesian updating as

$$f''(\lambda) = kL(\lambda)f'(\lambda) \tag{7}$$

where $L(\lambda)$ is the likelihood of observing the incident set assuming $f'(\lambda)$, and in which k is the normalizing constant given by

$$k = \left[\int_{-\infty}^{\infty} L(\lambda)f'(\lambda)d\lambda \right]^{-1} \tag{8}$$

Since construction incidents may be modeled as a Poisson process, it is convenient to assume a gamma distribution for λ in order to form a conjugate pair with the Poisson distribution (Ang and Tang 1975; Modarres et al., 1999). A conjugate pair will permit significant mathematical simplification to Eq. (7) above, and both $f'(\lambda)$ and $f''(\lambda)$ will take the same gamma distribution, but with different values for the parameters, κ and υ. In this way, if κ′ and υ′ are the corresponding prior estimates of the parameters, the revised parameters would be

$$\kappa'' = \kappa' + x \tag{9}$$

$$\upsilon'' = \upsilon' + t \tag{10}$$

where x is the number of incidents recorded in t intervals (of 50,000 man-hours). Moreover, the parameters κ and υ of the gamma distribution are related to its mean and variance by the following relations:

$$\bar{\lambda} = \kappa/\upsilon \tag{11}$$

$$\mathrm{Var}(\lambda) = \kappa/\upsilon^2 \tag{12}$$

Thus the prior estimate of the mean rate of incident occurrence, $\bar{\lambda}'$, may be easily revised to $\bar{\lambda}''$ through the above relations. With reference to Fig. 4, $\bar{\lambda}''$ would correspond to $\hat{\lambda}_P$ of the planning component. An example of how Bayesian updating can be applied to incorporate objective incident occurrence data into subjective judgment can be found in Chua and Goh (2002).

Partitioned Poisson Processes

Another interesting application of the proposed Poisson model for construction incidents arises from the exploitation of the partitioned property of the Poisson process. With this property, the incident rate for categorized subprocesses by type of work, type of incident, or severity of incident (just to name a few) may be easily derived. In general it cannot be assumed that the distribution of incident occurrence for the categorized subprocesses will share the same distribution as the overall project, and even then, the parameters for the subprocesses are not readily available from the parameter for the overall project. The partitioned property of the Poisson model, however, implies that the subprocesses for various categorizations are also Poisson distributed and their parameters can be easily derived (Ross 2000; Wolff 1989). This is conditional on the assumption that the categorization of these incidents is random; that is, the probability of an incident being classified as a category, say category "A," is independent of its sequence and the categorization of the preceding incident. This assumption appears reasonable for the types of categories mentioned earlier since the occurrence of an incident of a particular category is generally independent of the project conditions.

J. Constr. Eng. Manage., 2005, 131(6): 715-722

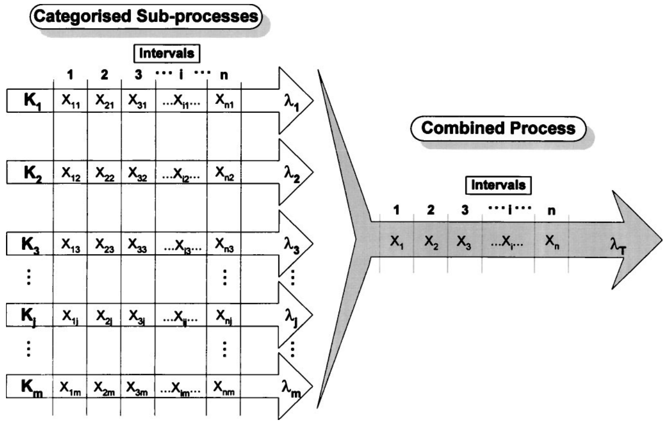

Fig. 5. Partitioning Poisson process into subprocesses

Fig. 5 illustrates the concept of a partitioned Poisson process. The combined process can be visualized to be composed of several subprocesses defined by the possible values of a categorization. For example, based on the categorization by type of incident, three subprocesses may be categorized as incidents that involve falling from a height, sudden impact with an object, and others. The random variable X_{ij} is the number of incident occurrences of subprocess j ($j=1$ to m) in interval i ($i=1$ to n). For an interval i, say $i=1$, the total number of construction incident occurrences, X_1, is the sum of X_{1j} of all the subprocesses. If the mean arrival rate of a construction incident for a project is λ_T, the distribution of a particular subprocess j will also be Poisson distributed with a mean arrival rate given by

$$\lambda_j = P(E_j|E) \times \lambda_T \tag{13}$$

where $P(E_j|E)$ is the probability of the categorization given that an incident has occurred, E_j being the categorization of the subprocess j. Then $P(E_j|E)$ can be estimated as the relative frequency of E_i, that is, the number of incidents of category E_j/total number of incidents. For a mutually exclusive and collectively exhaustive categorization of the subprocesses

$$\lambda_T = \sum_j \lambda_j \tag{14}$$

The concept of a partitioned Poisson can also be used to account for unreported incidents. This is achieved by assuming that incidents are randomly partitioned into unreported and reported incidents. If the probability of an incident being reported is p_r, then the true distribution of incident occurrence can be estimated by a Poisson distribution with parameter λ/p_r, where λ is the mean occurrence rate of reported incidents. The estimation of p_r can be based on expert opinion or statistical studies similar to the study by Alsop and Langley (2001) on traffic incidents.

Conclusions

The study has utilized the dispersion test and chi-square goodness-of-fit test to determine the suitability of the Poisson process in modeling construction incident occurrence. The results show there is no evidence to reject the Poisson distribution, which is a sound and convenient model. All except 1 of the 14 projects studied could be modeled by a homogeneous Poisson process. Even so, the remaining project could be modeled as a nonhomogeneous Poisson process comprising two homogeneous Poisson processes with the segment break occurring after some significant changes in the project.

Using the Poisson model, the incident occurrence in a project is easily characterized by the mean arrival rate, λ, and this can facilitate the use of statistical approaches in construction safety management. A λ_G can form the quantitative goal of the project, and a $\hat{\lambda}_P$ can be used to measure the risk exposure during risk assessment. On-site data can then be collected and a statistical comparison can be made to evaluate the differences between actual performance, $\hat{\lambda}_O$, and λ_G or $\hat{\lambda}_P$. The mean arrival rate, λ, also facilitates benchmarking with other projects or companies. Management can measure the effectiveness of safety measures and policies by comparing the incident rate before and after implementation. The incorporation of observed incident data into initial estimates in risk assessment exercises using the Bayesian approach has also been discussed. Furthermore, the partitioned property of the Poisson model has been exploited to extend the analysis to model incident occurrences for different categories such as different work types, incident types, and severity types. In a similar way, combined distribution of reported and unreported incidents can also be estimated based on the partitioned Poisson concept.

However, due to the present focus of the paper, the effect of the systematic factors depicted in Fig. 2 on the mean arrival rate has not been studied in detail. This will require an extensive analysis of a larger database using Poisson regression analysis with SMS scores and situational variables as independent variables. Also, the present study has been essentially based on railway construction projects. Further studies of other types of projects should help to further verify the applicability of the Poisson model for construction. Still, in the event that no other model is available, the Poisson distribution will suffice as an initial estimation.

The Poisson distribution can also be quite easily applied in practice because λ can be interpreted as the arithmetic average of the number of incidents per 50,000 man-hours, better known as the accident frequency rate (AFR). Even though in some countries the AFR is based on conventional time units instead of man-hours worked, most construction sites will have man-hours worked information that should allow the time-based AFR to be easily converted to a man-hours-worked-based AFR. Thus, the required information for the Poisson distribution is readily available.

Acknowledgments

This study would not have been possible without the support of the Safety Department of Singapore's Land Transport Authority, in particular its cooperation in facilitating access to SITS. The writers are also grateful for the time and effort of its numerous staff.

References

Alsop, J., and Langley, J. (2001). "Under-reporting of motor vehicle traffic crash victims in New Zealand." *Accid. Anal. Prev.*, 33, 353–359.

Ang, H. S., and Tang, W. H., (1975). *Probability concepts in engineering planning and design*. Vol. 1, *Basic principles*, Wiley, New York.

Bendell, A. (1991). "Analysis methodologies." *Reliability data banks*, A.

G. Cannon and A. Bendell, eds., Galliard Ltd., London.
British Standards Institution (BSI). (1996). "Guide to occupational health and safety management systems." *BS 8800*, London.
British Standards Institution (BSI). (2000). "Occupational health and safety management systems—Guidelines for the implementation of OHSAS 18001." *OHSAS 18002*, London.
Cattledge, G. H., Hendricks, S., and Stanevich, R. (1996). "Fatal occupational falls in the U.S. construction industry, 1980–1989." *Accid. Anal Prev.*, 28(5), 647–654.
Chua, D. K. H., and Goh, Y. M. (2003). "Application of case based reasoning in construction safety planning." *Computing in civil engineering*, A. D. Songer and J. C. Miles, eds., ASCE, Reston, Va., 298–307.
Chua, D. K. H., and Goh, Y. M. (2004). "Incident causation model for improving feedback of safety knowledge," *J. Constr. Eng. Manage.*, 130(4), 542–551.
Cochran, W. G. (1954). "Some methods of strengthening the common chi-square tests." *Biometrics*, 10, 417–451.
Conover, W. J. (1980). *Practical non-parametric statistics*, 2nd Ed., Wiley, New York.
Cox, D. R., and Lewis, P. A. (1966). *The statistical analysis of series of events*, Methuen and Wiley, London.
Fridstrom, L., Ifver, I., Ingebrigtsen, S., Kulmala, R., and Thomsen, L. K. (1995). "Measuring the contribution of randomness, exposure, weather and daylight to the variation in road accident counts." *Accid. Anal. Prev.*, 27(1), 1–20.
Hinze, J., Pedersen, C., and Fredley, J. (1998). "Identifying root causes of construction injuries." *J. Constr. Eng. Manage.*, 124(1), 67–71.
Jeong, B. Y. (1998). "Occupational deaths and injuries in the construction industry." *Appl. Ergon.*, 29(5), 355–360.
Kartam, N. A., and Bouz, R. G. (1998). "Fatalities and injuries in the Kuwaiti construction industry." *Accid. Anal. Prev.*, 30(6), 805–814.
Larsson, T. J., and Field, B. (2002). "The distribution of occupational injury risks in the Victorian construction industry." *Safety Sci.*, 40, 439–456.
McKinnon, R. C. (2000). *Cause, effect, and control of accidental loss with accident investigation kit*, Lewis Publishers, Boca Raton, Fla.
Modarres, M., Kaminskiy, M., and Krivtsov, V. (1999). *Reliability engineering and risk analysis—A practical guide*, Marcel Dekker, New York.
Montgomery, D. C., and Runger, G. C. (1999). *Applied statistics and probability for engineers*, 2nd Ed., Wiley, Singapore.
Nicholson, A. J. (1986). "The randomness of accident counts." *Accid. Anal. Prev.*, 18, 193–198.
Nicholson, A. J., and Wong, Y. D. (1993). "Are accidents Poisson distributed? A statistical test." *Accid. Anal. Prev.*, 25(1), 91–97.
Productivity and Standards Board (PSB). (1999). "Code of practice for safety management system for construction worksites." *CP 79*, Singapore.
Ramsey, J. (1985). "Ergonomics factors in task analysis for consumer product safety." *Journal of Occupational Accidents*, 7, 113–123.
Reason, J. (1990). *Human error*, Cambridge University Press, New York.
Ross, S. M. (2000). *Introduction to probability models*, Academic Press, New York.
Sanders, M., and Shaw, B. (1988). *Research to determine the contribution of system factors in the occurrence of underground injury accidents*, Bureau of Mines, Pittsburgh.
Wolff, R. W. (1989). *Stochastic modeling and the theory of queues*, Prentice-Hall, Englewood Cliffs, N.J.

J. Constr. Eng. Manage., 2005, 131(6): 715-722