Reliability Engineering and System Safety 150 (2016) 136–146

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Estimation of expected number of accidents and workforce unavailability through Bayesian population variability analysis and Markov-based model

Márcio das Chagas Moura [a,b,*], Rafael Valença Azevedo [a,b], Enrique López Droguett [c], Leandro Rego Chaves [d], Isis Didier Lins [a,b], Romulo Fernando Vilela [e], Romero Sales Filho [f]

[a] CEERMA-Center for Risk Analysis, Reliability and Environmental Modeling, Universidade Federal de Pernambuco, Av. da Arquitetura, s/n, CEP: 50.740-550, Cidade Universitária, Recife, PE, Brazil
[b] Department of Production Engineering, Universidade Federal de Pernambuco, Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária, CEP: 50740-530, Recife, PE, Brazil
[c] Mechanical Engineering Department, University of Chile, Santiago, Chile
[d] Department of Statistics and Applied Mathematics, Federal University of Ceará, Avenida da Universidade, 2853 - Benfica, Fortaleza 60020-181, CE, Brazil
[e] CHESF-São Francisco's Hydroelectric Company, Rua Delmiro Gouvêia, 333 - San Martin, CEP: 50761-901, Recife, PE, Brazil
[f] Academic Unit of Garanhuns, Federal Rural University of Pernambuco, Garanhuns, Brazil

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form
18 January 2016
Accepted 27 January 2016
Available online 5 February 2016

Keywords:
Occupational accidents
Bayesian variability analysis
Markov model
Expected number of accidents
Workforce unavailability

ABSTRACT

Occupational accidents pose several negative consequences to employees, employers, environment and people surrounding the locale where the accident takes place. Some types of accidents correspond to low frequency-high consequence (long sick leaves) events, and then classical statistical approaches are ineffective in these cases because the available dataset is generally sparse and contain censored recordings. In this context, we propose a Bayesian population variability method for the estimation of the distributions of the rates of accident and recovery. Given these distributions, a Markov-based model will be used to estimate the uncertainty over the expected number of accidents and the work time loss. Thus, the use of Bayesian analysis along with the Markov approach aims at investigating future trends regarding occupational accidents in a workplace as well as enabling a better management of the labor force and prevention efforts. One application example is presented in order to validate the proposed approach; this case uses available data gathered from a hydropower company in Brazil.

1. Introduction

Occupational accidents pose several negative consequences to employees, employers, environment and people surrounding the locale where the event takes place. In this context, the development of quantitative models may aid the decision makers to determine adequate safety policies and operational procedures that in turn reduce the probability of occurrence and/or the severity of these undesired events.

Indeed, Moscoso et al. [31] assert that the analysis of the accident rate can be of great help to determine the safety procedures to be implemented. Yet, the investigation of the recovery time from accidents is also equally important in cases where the evaluation of unavailability and work time loss is of interest. In fact, significant research effort has been put on the quantitative analysis of occupational accidents regarding the assessment of rates of accident and recovery. For instance, Cawley and Homce [9], Chia et al. [11], Konstandinidou et al. [24] and Moscoso et al. [31] discuss occupational accidents in the electrical, construction, petrochemical and olive oil facilities, respectively, and assess the rates of accident and recovery for different causes. Camino et al. [7], Blanch et al. [6] and Lilley et al. [25] in turn analyzed the relationship between the severity of accidents, in terms of their recovery rate, and various organizational and individual variables such as age, gender, type of contract, time of accident, length of service in the company, company size, day of the week, among others. Carnero and Pedregal [8] and Freivalds and Johnson [18] evaluate accident rates in order to show the evolution and the profile of occupational accidents over the year, providing relevant information to detect periods where careful attention should be taken to improve safety. Furthermore, various researches have analyzed accident rates with the purpose to estimate the cost of

* Corresponding author at: Department of Production Engineering, Universidade Federal de Pernambuco, Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária, CEP: 50740-530, Recife, PE, Brazil. Fax: +55 81 2126 7112.
E-mail addresses: marcio@ceerma.org, marciocmoura@gmail.com (M. das Chagas Moura).

http://dx.doi.org/10.1016/j.ress.2016.01.017
0951-8320/

accidents [21,22,37,38,40]. In this context, Yakovlev and Sobel [40] showed that profit increases with the safety investment because the relationship between the prevention expenditure and the accident rates is negative.

Initiatives such as the Workgroup Occupational Risk Model (WORM) project, which has been carried out by the Ministry of Social Affairs and Employment of the Netherlands, have also given rise to a set of quantitative models that assess occupational risks. For example, Ale et al. [3] and Ale et al. [1] developed an Occupational Risk Model (ORM) that basically consists in three stages: (i) analyzing accident data; (ii) tailoring these data into a Functional Block Diagram (FBD); and (iii) optimizing solutions for risk reduction. The first step of the ORM uses a tool called "Storybuilder" that systematically classifies and analyses reports of past accidents. Ale et al. [2] also adopted the "Storybuilder" to investigate accident reports in the construction industry in The Netherlands. Papazoglou et al. [33], Aneziris et al. [4], and Aneziris et al. [5] used FBD to evaluate the occupational risk of falling from mobile ladders, for fall from height, and for crane activities respectively. Papazoglou et al. [34] in turn presented a model based on a homogeneous Poisson stochastic process to estimate the probabilities for occupational accidents per hour of exposure and for a year of average exposure considering three possible consequences: death, permanent and recoverable injuries.

A common characteristic of most of the above-mentioned papers is that they rely on either "top-down" (national data and general statistics) or "bottom-up" (records gathered from various organizations) data-collection methods. Jallon et al. [22] state that instead of "top-down" and "bottom-up" strategies, the "local" approach, which is based on in-company data, allows the accurate assessment of occupational accidents in specific workplaces. However, "local" accident recordings are generally sparse and contain many censored records and, according to Meel et al. [28], classical statistical approaches are ineffective in this context. Therefore, the use of Bayesian methods may be more suitable for the "local" approach.

Indeed, Meel et al. [28] and Marcoulaki et al. [26] developed Bayesian approaches for the assessment of the rate of accidents. Marcoulaki et al. [27] extended these Bayesian models by considering sick leaves (time necessary for an employee to be recovered from an accident), and the associated work time loss. To this end, a prior Gamma distribution was considered to model both the rates of accident and recovery. The use of the Gamma distribution allowed an analytical solution for formulating the posterior distributions because it forms a conjugated pair with the Poisson distribution, which in turn was used in the construction of the likelihood function. From these distributions, Marcoulaki et al. [27] developed closed-form results for some occupational measures such as the accident rates, duration of recovery from an accident, and the worker unavailability. Papazoglou et al. [32] also adopted a Bayesian-based approach to assess the uncertainty in the quantification of risk rates of occupational accidents.

However, the models of Meel et al. [28], Marcoulaki et al. [27] and Papazoglou et al. [32] as well as the classical statistics based approaches have been established under the assumption that all company workers have homogeneous/identical behavior in terms of the occurrence of accidents, i.e., they assumed that rates of accident and recovery are the same for all workers. Despite that, it has been shown that due to the existence of individual factors (age, gender, experience, time, etc.) it is expected that workers have distinct rates of accident and recovery even in the situation they have similar roles in the workplace [10,36,6]. However, as pointed out by Fragola [17], the more inhomogeneous the database becomes, the tighter the uncertainty bounds due to the larger size of the aggregated population. This occurs when inhomogeneous data are aggregated as if they were homogeneous [14]. Therefore, in this case, even though the rates of accident and recovery would be better estimated in accordance with the statistical sense because the confidence intervals would be narrower, those occupational measures would be less representative of each subpopulation that composes the mixture.

Thus, it is important to analyze the variability of these rates in a population of workers in order to forecast the random behavior of the occupational accidents within a workplace. However, in a "local" perspective, we may have little or no availability of accidents data, mainly if it is desired to make a categorized analysis (for example, by type of accidents). Then, other sources of information such as data from similar facilities may be used to draw occupational measures of interest.

Therefore, this paper proposes a Bayesian method to estimate rates of accident and recovery. We consider that each worker has unique rates of accident and recovery and the Bayesian Population Variability Analysis (BPVA), also known as the first phase in a two-stage Bayesian, or hierarchical Bayes [13], is here used to estimate the variability distribution of these rates within a group of nonhomogeneous workers. Unlike Marcoulaki et al. [27], where the rates of accident and recovery are considered homogeneous for different workers and conjugate distribution pairs are used, the solution here is not analytically obtainable, and thus we resort to a Markov Chain Monte Carlo (MCMC) approach to draw the nonparametric posterior distributions. Hence, we use the outputs of the BPVA, which are the distributions of the rates of accident and recovery, to feed a two-state Markov-based model that in turn will estimate the expected number of accidents and the expected unavailability of the labor force. These quantities will provide information to investigate future trends regarding occupational accidents in the workplace as well as enabling a better management of the labor force. At the best of authors' knowledge, no article has yet adopted BPVA to model occupational accidents. Indeed, the procedures here implemented for BPVA are based on the mathematical methods developed for risk and reliability analysis, such as those in Kaplan [23], Mosleh and Apostolakis [30], Pörn [35] and Droguett et al. [13]. In this work, these methods are tailored for the context of analysis of occupational accidents.

The remainder of this paper is organized as follows. Section 2 presents the theoretical background about the BPVA in the context of occupational accidents. Section 3 presents the proposed model, illustrating the implementation of the proposed BPVA-Markov procedure. Section 4 validates the model from an example application, which uses evidence from real accident reports of a hydropower company in Brazil. Finally, Section 5 provides some concluding remarks.

2. Bayesian population variability assessment

Similarly to Singpurwalla [39] claims in the context of reliability analysis, we here argue that rates of accident λ and recovery μ are expressions of our personal uncertainty about the workplace's dynamic behavior. Given that, we can associate these quantities to each individual worker $i = 1, \ldots, m$, and then consider the variability of rates over the whole population. A representation of this variability, in the form of a probability distribution, is referred to as the Population Variability Distribution (PVD) and its assessment is named Population Variability Analysis (PVA).

As in procedures of PVA [13,20,23,35], we assume that a member of a given family of a parametric distribution may describe PVD. If we have enough data for each i-th worker, it is possible to define the parameters of PVD of λ and μ directly from the dataset. In this section, we review the BPVA based on Deely and Lindley [12], Kaplan [23], Mosleh and Apostolakis [30], Pörn [35] and Droguett et al. [13].