

Research Report of the Research Institute
of Industrial Safety, RIIS-RR-89, 1989
UDC 519.2:614.8-02

災害発生時間の分布に関する研究 (5)

花 安 繁 郎*

A Study on the Time Intervals between Accidents (5)

by Shigeo HANAYASU*

Abstract; The accident frequency rate has been widely used as a measurement of safety performance in many undertakings over a long period of time.

In order to explore the significant changes in accident situation in succeeding intervals of time, the time intervals between occupational accidents were utilized as a useful indicator to give expression to safety performance in undertakings having a certain accident risk.

This paper deals with the stochastic treatment of the time intervals between occupational accidents to evaluate the safety performance of working places. Emphasis was placed on the probabilistic analysis of the time intervals between accidents and the number of accidents, when the accident frequency rate varies and depends on a certain probability function.

The main features of this paper are as follows:

- (1) From the analysis of frequency distribution of the number of occurrences of accidents within a fixed interval of time, it was found that the frequency distribution of many accidents agree with the negative binomial distribution as well as the poisson distribution simultaneously. Particularly, negative binomial distributions have many good agreement with actual accident distributions rather than the poisson distribution.
- (2) Depending upon the results of frequency distributions of accidents, it was recognized that many accidents take place at random. However the accident rate itself depends on the probability distribution whose probability function can be expressed as the gamma distribution. Hence, the frequency distribution of the number of occurrences of accidents as well as the frequency distribution of time periods between accidents are represented as the mixture of probability density functions. This mixing of the probability function of accident rate to the distribution function of the number of occurrences of accidents ultimately reaches the negative binomial distribution. Whilst, the mixture of the frequency distribution of accident rate to the frequency distribution of time intervals between accidents finally reaches the composite exponential distribution.
- (3) If sufficient accidents data are available, parameters of the probability distribution function of the negative binomial distribution or the composite exponential distribution, can be estimated by making use of these accident data.
- (4) In case sufficient data are not available, the Bayesian statistical inference can be employed as an alternative method to estimate the parameters of the distribution functions. By making use of Bayes' theorem, a posterior distribution of the accident rate can be obtained as the gamma distribution whose parameters are the number of accidents and duration of time investigated. Then, the frequency distribution of the number of occurrences of accident within a given period of time, can be obtained by way of mixing the probability distribution function of accident rate

* 土木建築研究部 Construction Safety Research Division

to the poisson distribution, which yields the negative binomial distribution. Similarly by mixing the accident rate distribution with the exponential or gamma distribution, the frequency distribution of the time periods between accidents, which takes into consideration the variance of accident rate, can be represented as the composite exponential or gamma distribution.

(5) Utilizing Bayes' theorem, a renewed posterior distribution of accident rate can be estimated at the time when new accident data are acquired. Then by making use of this up-to-date accident rate distribution, the renewal of the probability distribution functions of the number of occurrences of accidents as well as the frequency of the time periods between accidents, can be achieved.

Keywords; Occupational accident, Accident frequency rate, Mixture of probability distribution, Poisson distribution, Negative binomial distribution, Composite gamma distribution, Bayes' theorem

1. まえがき

事業所において労働災害の発生危険性を評価する指標には、単位労働時間あるいは単位労働力当りの災害発生件数で示される災害発生頻度率が広く用いられている。とくに、単位労働時間数が100万時間のときの災害頻度率は、我国では災害度数率と呼ばれており、労働省による労働災害動向調査によって産業別、業種別などに分類された全国平均の観測値が毎年報告されている¹⁾。

筆者はこれまで、この災害度数率に代表される労働災害発生頻度率が、事業所等において、作業時間の経過とともに変動する過程を、労働災害が発生するまでの時間数を用いて統計的に評価する方法について考察を加えてきた²⁻⁸⁾。

これまでに行ったさまざまな調査・分析から得られた結果には、例えば、最近の建設工事における労働災害の多くがほぼランダムに発生していること、すなわち、一定期間中の災害発生数の分布がポアソン分に、また、労働災害が発生するまでの時間数の確率分布が指数分布やガンマ分布で表現されること、さらにこれらの分布式のパラメータが災害度数率と関連づけられることなどがあげられる³⁻⁵⁾。

これらの知見より、最近の労働災害の多くは、すくなくとも短期的な期間でみる限り、相互に独立にかつ単位時間当りの発生数（災害発生頻度率）はほぼ一定のもとで発生していると考えてよいと思われる。

しかしながら、災害発生頻度率はこのように時間的、空間的に一定した値であるよりも、むしろ常に変動しばらつきを有して出現する値と考えた方が合理的なことが多い。

例えば、建設工事では、一般に、作業環境や作業形態のそれぞれ異なった工程の連続的な過程を通してひとつの製品が完成されてゆく。従って、工事の進捗に応じて、労働災害発生危険性も同時に変動してゆくことは充分予想されることである。

そこで本研究では、災害発生頻度率がある確率分布に従って変動する場合を想定して、そのときの労働災害発生数や災害発生時間数などの確率分布式を求め、実際の災害事例を用いてこれら災害発生数や発生時間数の分布式を検証したり、災害発生時間数などの将来予想を行うことを試みた。本稿はそれらの検討結果をまとめたものである。

2. 災害発生頻度率の変動を考慮したときの災害発生数および災害発生時間数の分布

2.1 労働災害発生数の分布

本節では、労働災害の時間軸上での起こり方の特性を分析する手法として、一定期間中での災害発生数について、とくに災害発生率の変動するときの発生数の分布について検討を加えることとする。

まず、災害はその発生確率がどの時間帯も等しくかつ相互に独立して発生すると仮定すると、ある期間での災害発生数の分布式は、ポアソン分布で表され、また個々の災害の発生時間（間隔）は負の指数分布となることが知られている。このポアソン分布は、災害のように稀な事象の発生数を分析する際に最も基礎となる確率分布式である。

またポアソン分布と指数分布とは、同一事象を発生件数からみた場合と、発生時間からみた場合とに観点を交えて記述したものであるから、ある事象が一方の

分布式を満足していれば、他方の分布式も同時に満足していると考えてよい。

単位期間での災害発生数を災害発生頻度率 λ と定義すると、期間 t における災害発生数のポアソン分布は次式で示される。式中の $E(X)$, $V(X)$ はそれぞれ分布の期待値および分散である。

$$P(x) = \frac{(\lambda t)^x}{x!} \cdot \exp\{-\lambda t\} \quad (1)$$

$$E(X) = \lambda t, \quad V(X) = \lambda t \quad (2)$$

上式での λ は一定値であることを仮定しているが、実際には前述した通り、さまざまな要因によって変動を伴って出現する値と考えた方が合理的と思われることが多い。そこでこの λ 自身も確率変数と見なし、ある確率分布のもとで変動することを想定して発生件数の確率分布を求め、同式を用いて確率・統計的な評価を行うこととした。

この災害発生頻度率 λ にどのような確率分布を想定すべきかについては現在のところ定まった説はない。Greenwood-Wood, Greenwood-Yule は、第1次世界大戦下での英国軍需工場で作業した女子労働者に発生した労働災害を確率論的立場から分析を加えた。分析ではいくつかの確率モデルを導出したが、そのひとつとして、作業員の災害の起こし易さ(蓋然性)は、各人に一様でなく、むしろ作業員ごとに異なった大きさの蓋然性によって構成されるという考えのもとに災害発生数の分布を導出した^{9,10)}。彼らは当初蓋然性の分布として正規分布を考え、ポアソン分布との複合化を経てエルミート分布と呼ばれる分布式を導出したが、災害の蓋然性が $(-\infty \sim 0]$ までの負の値をとるのは不合理であるとして、 $[0 \sim +\infty)$ までの範囲で分布する確率分布のうちで計算上便利な Pearson III type 分布(ガンマ分布)を新たにとりあげ、ポアソン分布との複合化の結果 Greenwood-Yule の分布と呼ばれる分布式に到達した。この研究の以降、災害発生率の分布にはガンマ分布を利用することが多いようである。

従って、本論でも、災害発生頻度率の分布にはガンマ分布を仮定して以降の論議を進めることとした。

まず、災害発生率 λ の分布を次式と仮定する。

$$h(\lambda) = \frac{(c\lambda)^{k-1}}{\Gamma(K)} c \cdot \exp\{-c\lambda\} \quad (3)$$

すると λ が固定した値のときの災害発生数は(1)式で与えられているので、同式のもとで(3)式を複合化すれば災害発生数が次式で得られる。

$$\begin{aligned} P(x) &= \int_0^\infty \frac{(\lambda t)^x}{x!} \cdot \exp\{-\lambda t\} \cdot h(\lambda) d\lambda \\ &= \frac{c^k t^x}{x!} \cdot \int_0^\infty \frac{\lambda^{x+k-1}}{\Gamma(K)} \cdot \exp\{-(c+t)\lambda\} d\lambda \\ &= \frac{\Gamma(x+k)}{\Gamma(K)} \frac{1}{x!} \left(\frac{c}{c+t}\right)^k \left(\frac{t}{c+t}\right)^x \\ &= \binom{x+k-1}{x} \left(\frac{c}{c+t}\right)^k \left(\frac{t}{c+t}\right)^x \quad (4) \end{aligned}$$

$$E(X) = \frac{kt}{c}, \quad V(X) = \frac{kt}{c} \left(\frac{c+t}{c}\right) \quad (5)$$

前式は一般に負の二項分布と呼ばれる確率分布式である。かくして、分布のパラメータ c , k を得ることが出来れば、同式を用いて任意の発生件数の確率分布を求めることが出来る。

なお、M. Greenwood, U. Yule らは観測期間として単位時間を採用し、(4)式で $t=1$ とおいた式を非均等分布(Unequal liability distribution)と名づけている。また同じく(4)式において $t=1$ とし、パラメータ c , k を新たに $h=k/c$, $b=1/c$ とおくと同式は下式のように記述出来る。

$$P(x) = \frac{\{h+(x-1)b\} \{h+(x-2)b\} \cdots \{h+b\} h}{x!} \cdot (1+b)^{-h/b-x} \quad (6)$$

上式は、壺のモデルで有名なポリヤ・エッゲンベルガー(Polya-Eggenberger)の式と呼ばれる分布式である¹¹⁾。このほか、幾何分布の和の分布として負の二項分布(この場合の分布名はパスカルの分布と呼ばれる)を導出することが出来るなど、負の二項分布は多くの成因を有している。従って、同式を用いる場合には母集団特性を十分理解しておく必要がある¹²⁾。

さて、(4)式のパラメータ c , k を推定するには、モーメント法によれば、データ平均を M , データ分散を S とすると、(5)式を用いて c , k が以下のように推定できる。

$$c = \frac{Mt}{S-M}, \quad k = \frac{M^2}{S-M} \quad (7)$$

Fig. 1 には、山陽新幹線トンネル建設工事(広島・

山口県内)で発生した労働災害について、1カ月当りの災害発生件数の分布を調べた結果を示した。同図には(4)式による負の二項分布に加えて、(1)式によるポアソン分布も併せて記した。同様に、山陽新幹線工事に引き続いて行われた上越新幹線建設工事でのトンネル建設工事における労働災害発生数の分布を Fig. 2 に示した。これらの図より、両トンネル建設工事での災害発生数は、ポアソン分布よりも負の二項分布の方がうまく適合していることがわかる。また、両トンネル工事災害を負の二項分布で表現した場合の分布のパラメータ c, k を(7)式によって求め、これらを用いて(3)式によって各プロジェクトの災害発生率の確率分布を求めた結果が Fig. 3 である。同図から、それぞれのプロジェクトにおける災害発生率は、その平均値が山陽新幹線トンネル工事災害=0.47人/月、上越新幹線トンネル工事災害=0.84人/月となっているが、全体では図に示されたような分布となっていたと考えた方が実情に近いようである。例えば山陽新幹線トンネル建設(広島・山口県内)工事では、1970年~1974年の5年間、延べ月数2157カ月もの長期間にわたって施工されたが、その間の各月の災害発生率は一定ではなく、工事の進行に応じて変動し、全体としては図に示された分布のもとで災害が発生したと考えられる。

また、他の労働災害事例として、1978年から1987年の10年間に日本全国発生した重大災害(一度に3人以上の負傷者を含んだ災害)をとりあげ、1カ月当りの発生件数の分布を調べ、その結果を図に示した。

Fig. 4 には、産業別の分析事例として、製造業と建設業における重大災害発生件数の分布を分析した結果を示した。同じく、災害の種類別に重大災害の発生数の分布を分析した事例として、Fig. 5 には土砂崩壊災害と倒壊災害の分析結果を示し、Fig. 6 には墜落災害とクレーン・デリックによる災害の分析結果を示した。さらに、都道府県別に分析した事例として、Fig. 7 には北海道と東京都での災害を示し、Fig. 8 には大阪府と神奈川県における災害の分析結果をそれぞれ示した。これらの労働災害事例の分析結果より、災害の多くが単独に負の二項分布に、あるいは負の二項分布とポアソン分布とに同時に適合していることや、とくに単独分布としてはポアソン分布よりもむしろ負の二項分布の方が災害事例をうまく表現している例が多いことが分かる。

これらのことより、災害の多くは、事象そのものはランダムに発生するが、その単位時間当りの災害発生

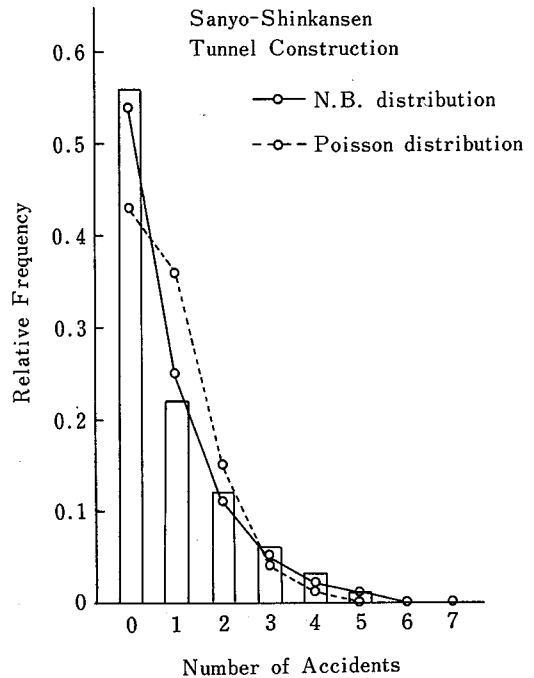


Fig. 1 Frequency distribution of accident number within a month in the Sanyo-Shinkansen tunnel construction
山陽新幹線トンネル建設工事労働災害発生数の分布

数(災害発生率)は、ガンマ分布などで示される確率分布に従って生ずる、いわゆる複合過程のもとで生じていると考えても良いと思われる。

2.2 労働災害発生時間数の分布

労働災害の時間軸上での発生特性に関する他の分析法として、本節では、労働災害が発生するまでの時間数、とくに災害発生率の変動するときの発生時間数の分布について検討する。

まず、災害事象に関する基本仮定を前節と同様とすると、個々の災害が発生するまでの時間(間隔)数は負の指数分布に、また複数件の災害が発生するまでの時間数の分布はガンマ分布に従うことが知られている。指数分布の確率密度関数を以下に示す。

$$f_1(t) = \lambda \cdot \exp\{-\lambda t\} \quad (8)$$

$$E_1(T) = 1/\lambda, \quad V_1(T) = 1/\lambda^2 \quad (9)$$

上式のパラメータは λ である。この λ は指数分布の期待値が $1/\lambda$ であることから、期待値の逆数である

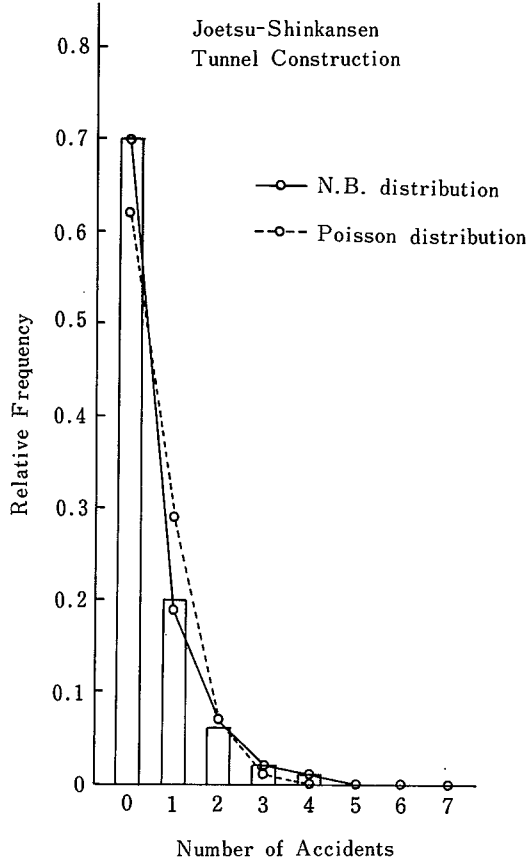


Fig. 2 Frequency distribution of accident number within a month in the Joetsu-Shinkansen tunnel construction
上越新幹線トンネル建設工事労働災害発生数の分布

単位時間当りの災害発生数，すなわち災害発生頻度率を表すものである。従って，前節で考察したように， λ 自身も確率的に変動することを想定して以下に発生時間数分布を求めることとする。

$$\begin{aligned}
 f(t) &= \int_0^{\infty} \lambda \cdot \exp\{-\lambda t\} \cdot h(\lambda) d\lambda \\
 &= \frac{c^k}{\Gamma(k)} \int_0^{\infty} \lambda^k \cdot \exp\{-(c+t)\lambda\} d\lambda \\
 &= \frac{\Gamma(k+1)}{\Gamma(k)} \cdot \frac{c^k}{(c+t)^{k+1}} \\
 &= \frac{k c^k}{(c+t)^{k+1}} = \frac{k}{c} \left(\frac{c}{c+t}\right)^{k+1} \quad (10)
 \end{aligned}$$

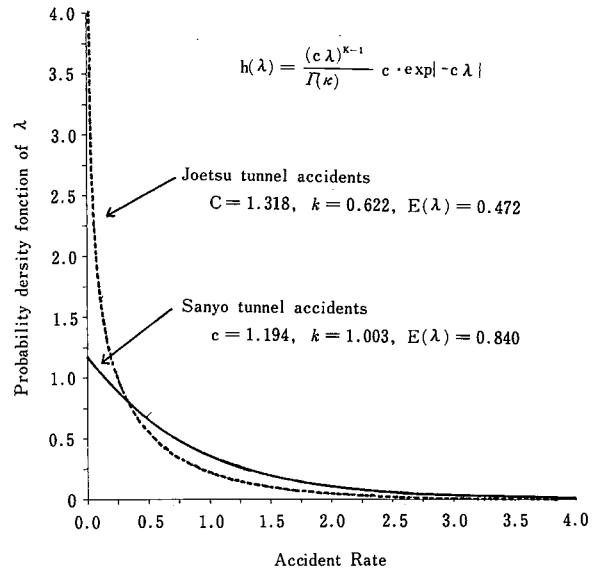


Fig. 3 Frequency distribution of accident rates (λ) in the two Shinkansen tunnel construction accidents
両新幹線トンネル工事災害の発生率 λ の分布

この λ 自身の確率分布は既に(3)式で与えられているので，同式のもとでの発生時間数は，(3)式の λ を(8)式によって複合化することによって得られる。

かくして得られた(10)式の確率分布関数および期待値，分散は次式で示される。

$$F(t) = \int_0^t f(t) dt = 1 - \left(\frac{c}{c+t}\right)^k \quad (11)$$

$$E(T) = \frac{c}{k-1}, \quad V(T) = \frac{c^2 k}{(k-2)(k-1)^2} \quad (12)$$

上式はパレート分布と呼ばれるが，指数分布とガンマ分布とを複合化したのでここでは複合化指数分布と呼ぶことにする。

前節で，ポアソン分布と指数分布とが，同一事象を発生数と発生時間数とに観点を置いて記述したものであることを述べたが，同様なことが負の二項分布と複合化指数分布との間についても言える。すなわち，ある災害事象でその発生率が(3)式によるガンマ分布で表されるとき，その発生数の分布は負の二項分布となり，また災害発生時間数の分布は複合化指数分布で示される。また，ある災害事象が一方の確率分布式を満足していれば，他方の式も同時に満足することがこの場合も成立する。

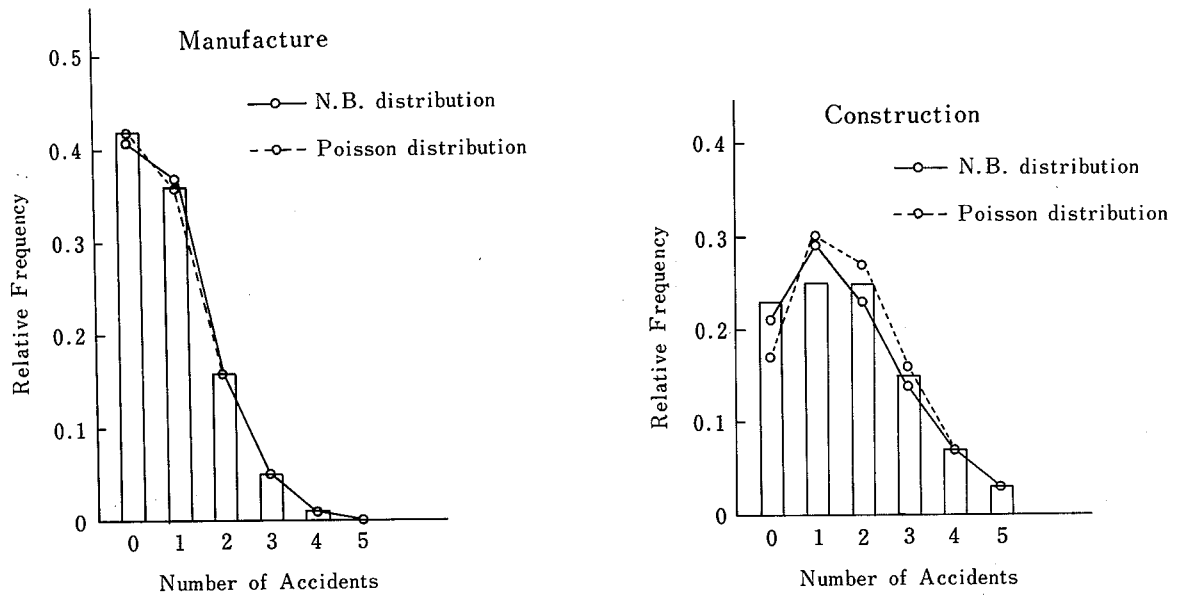


Fig. 4 Frequency distribution of serious accidents within a month classified by industrial sector from the year of 1978 to 1987
産業別重大災害発生数の分布

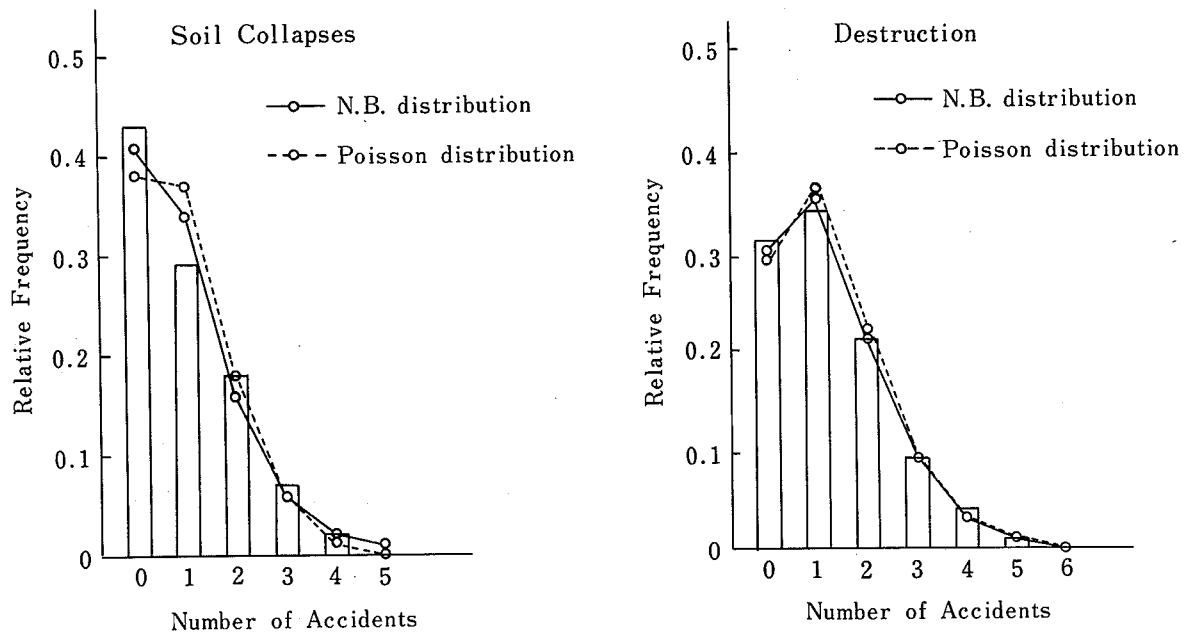


Fig. 5 Frequency distribution of serious accidents within a month classified by type of accident from the year of 1978 to 1987 (1)
災害種類別重大災害発生数の分布 (1)

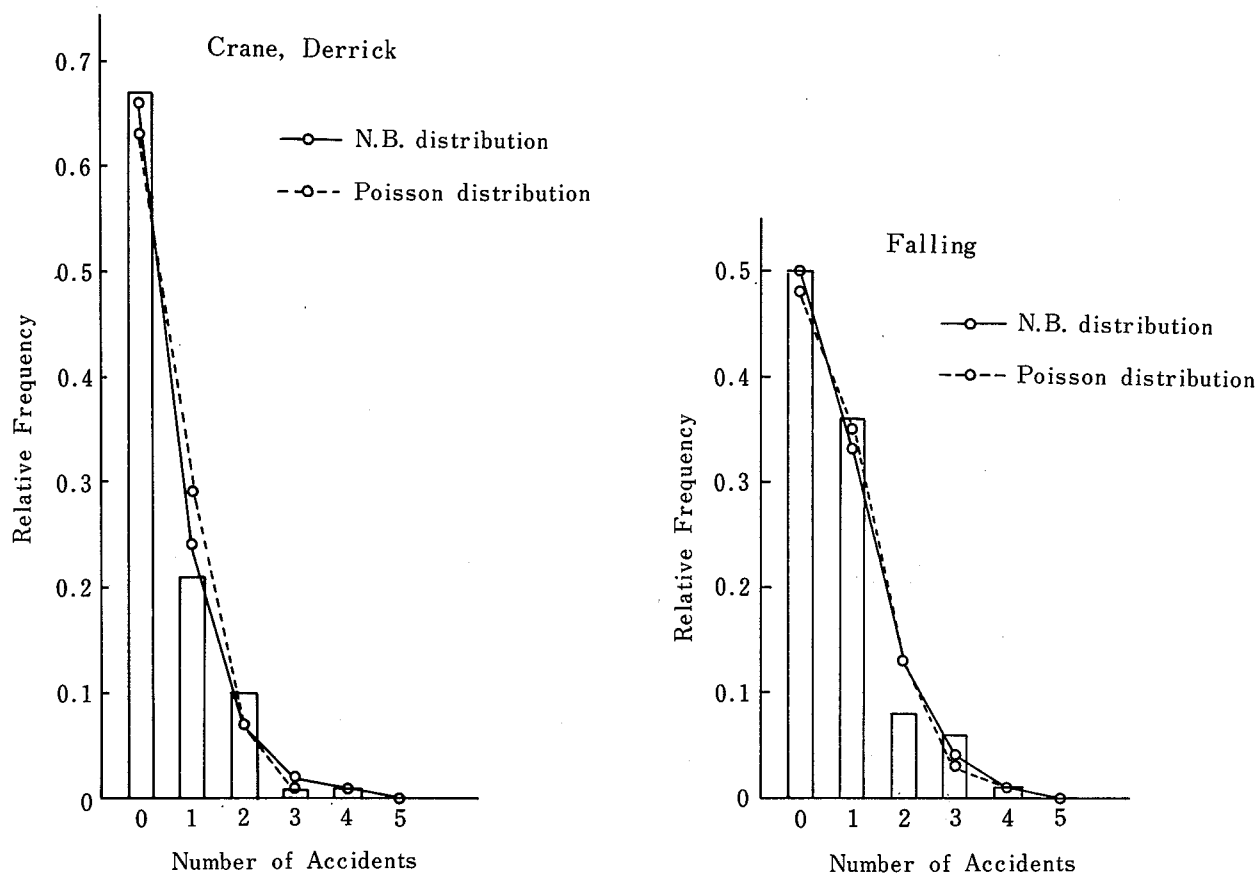


Fig. 6 Frequency distribution of serious accidents within a month classified by type of accident from the year of 1978 to 1987 (2)
 災害種類別重大災害発生数の分布 (2)

複合化指数分布(10)式のパラメータは(4)式と同じく c 、 k であり、同式による分析事例として、ここでは、前節での山陽新幹線および上越新幹線トンネル建設工事労働災害の発生数の分析結果から得られたパラメータ c 、 k を用いて複合化指数分布を計算した結果を Fig. 9 に示した。

同図より、山陽新幹線工事災害の発生時間確率分布曲線の勾配が上越新幹線工事災害のそれに比べて急になっており、山陽新幹線災害の方が上越新幹線災害よりも短期間のうちに発生し易い状態であることがわかる。これは Fig. 3 に示したように、山陽新幹線工事災害の発生率の分布が、上越工事災害のそれに比べて大きい値の領域で分布している比率が高いからである。

3. ベイズ方式による災害発生数および災害発生時間数の評価

前章では災害発生頻度率が変動することを想定して災害発生分布特性を考察した。そこでは、災害発生率 λ はガンマ分布に従う確率変数として取扱われ、ポアソン分布や指数分布との複合化を経て、負の二項分布や複合化指数分布などの複合化分布が導出され、実際の災害による検証も行われた。

しかしながら、災害発生率 λ の確率分布にガンマ分布を仮定したのは、主に複合化を行ううえでの計算上の都合による理由が大きく、同式が実体を忠実に表した最善の分布という訳では必ずしもない。

加えて、データから推定されるパラメータの信頼性についても問題がない訳ではない。前章で分析を行った災害事例のデータ数は、山陽新幹線トンネル工事災害では、延べ月数2157カ月で発生した1812件の労働災害、同じく上越新幹線トンネル工事災害では、延べ3002カ月で発生した1416件の労働災害、さらに重大災

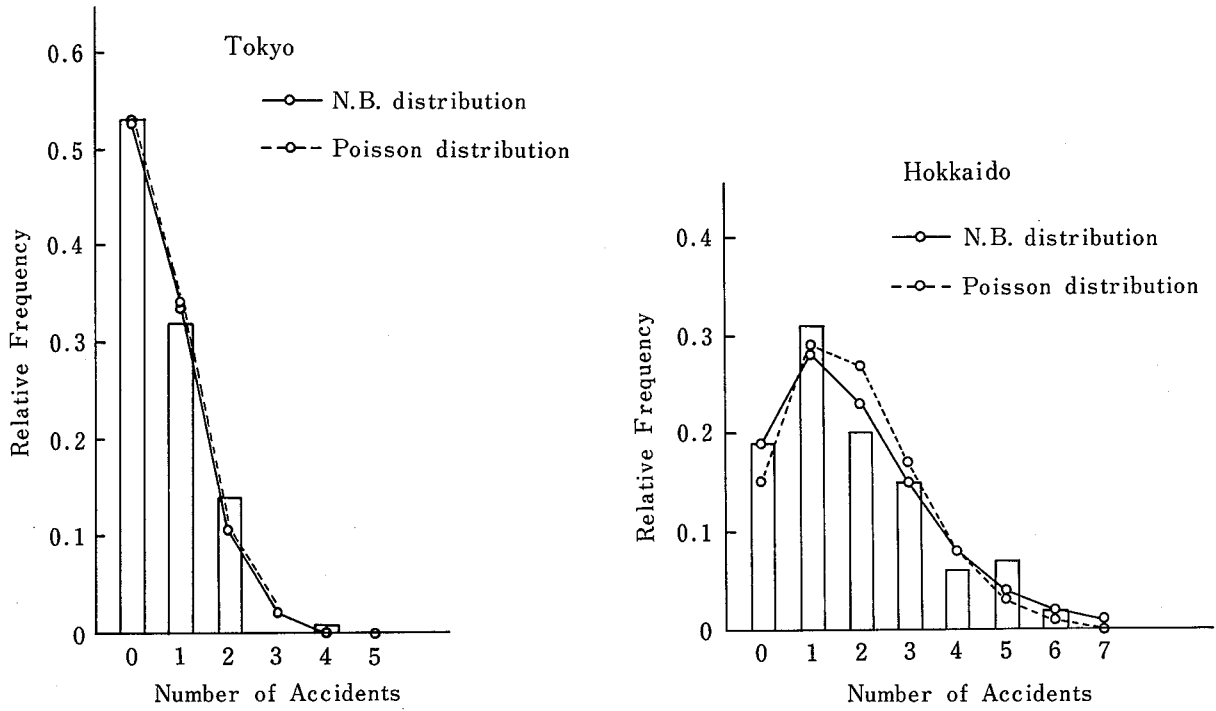


Fig. 7 Frequency distribution of serious accidents within a month classified by prefecture from the year of 1978 to 1987 (1)
都道府県別重大災害発生数の分布 (1)

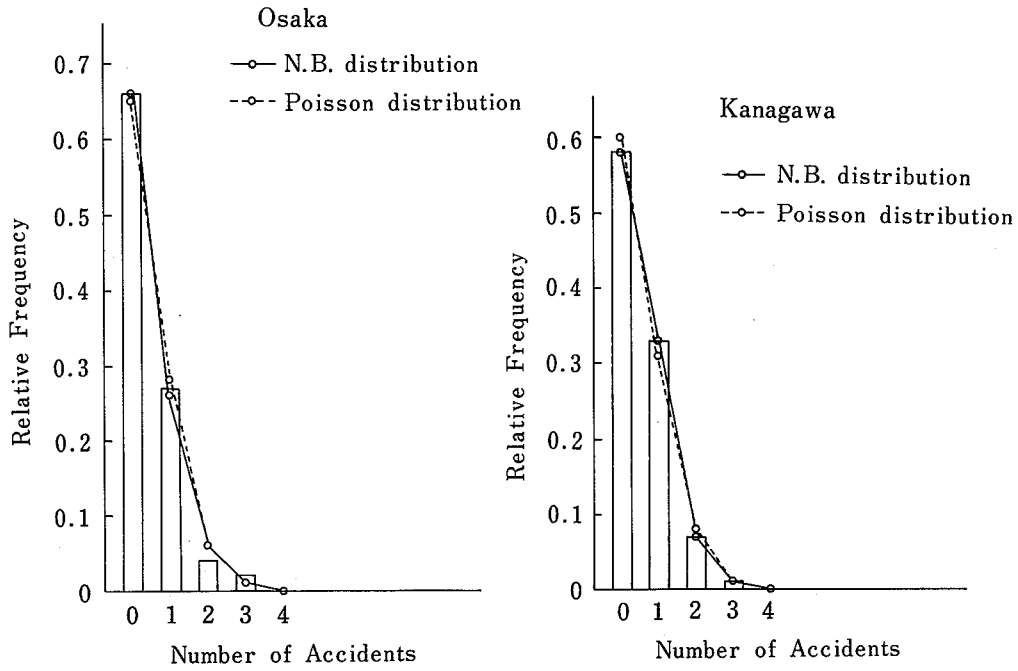


Fig. 8 Frequency distribution of serious accidents within a month classified by prefecture from the year of 1978 to 1987 (2)
都道府県別重大災害発生数の分布 (2)

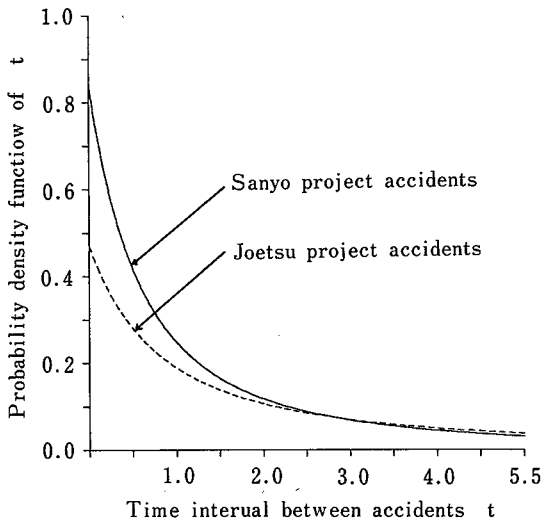


Fig. 9 Frequency distribution of time intervals between accidents in the two bullet train tunnel construction projects considering the variance of accident rates

災害発生率の変動を考慮した新幹線トンネル工事災害の発生間隔の分布

害については1978年～1987年の10年間で発生した総数1879件など、いずれの分析においても、膨大な量の災害データを用いて分布のパラメータの推定を行っている。しかし、このように大量の災害データを観察して信頼性の高いパラメータを推定出来ることは極めて限られた場合であり、実際にはその場で得られたありあわせのデータを用いて推定したり、あるいはせざるを得ないことが多い。

このように、λの確率分布を規定する際の不明確さと、少数データに起因する推定パラメータの信憑性の欠如は、労働災害のように比較的稀にしか生起しない事象を取扱う際に生ずる根本的な問題点でもある。

そこで本章では、このような問題に対処するために、ベイズの定理を用いて以下に、(1)少数のデータから災害発生率λの確率分布の導出を行う、(2)導出された確率分布式を用いて、災害発生数や発生時間数の分布を求め、災害発生数や時間数の確率的な評価や予測を行うことを試みた。

3.1 ベイズ方式による災害発生率の推定

近代統計学では、実際に観測したデータを用いてベイズの定理によって分布のパラメータの推定などの統

計的推測を行う方法論を総称してベイズ統計学と呼んでいる。このベイズ統計学の基礎であるベイズの定理は数式的には次の形式で表すことができる¹³⁾。

いまn個の事象A₁, A₂, …, A_nが互いに排反でかつ必ずどれかがおこるとする。またBを他の事象とする。このとき事象Bのもとで任意の事象A_iが起こる確率は次式で与えられる。

$$P(A_i | B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n)} \quad (13)$$

ただし、P(B|A_i)は事象A_iが与えられたときにBが起こる条件付確率である。また、ベイズの定理が利用される場合は、事象A₁, A₂, …, A_nは事象Bの原因事象、P(A_i)はA_iの事前確率、P(A_i|B)はA_iの事後確率と呼ばれている。このようにベイズの定理は、原因事象の事前確率と既知の条件付確率から事後確率を求める公式である。

以上の準備のもとに災害発生率λの事後確率をベイズの定理によって求めてみる。まず、災害発生率λのもとで期間T₀内でX₀件の災害が発生する確率は(1)式によるポアソン分布で次式の通り示される。

$$P(X_0) = \frac{(\lambda T_0)^{X_0}}{X_0!} \cdot \exp\{-\lambda T_0\}$$

ここでX₀件の災害が起こった条件のもとでλの事後確率を求める手順は以下のように示される¹⁴⁾。まず任意のε>0に対して、

$$\frac{\int_0^\Lambda (\lambda T_0)^{X_0} \cdot \exp\{-\lambda T_0\} d\lambda}{\int_0^\infty (\lambda T_0)^{X_0} \cdot \exp\{-\lambda T_0\} d\lambda} = 1 - \epsilon$$

なるλ=Λを選び、次に区間(0, Λ)をN等分して微小区間(λ_i, λ_i+Δλ)(i=1, 2, …, N, Δλ=Λ/N)を作る。さらに(13)式のベイズの定理に対応する以下の事象を考える。

B: 実現値X=X₀を得る事象

A_i: λが微小区間(λ_i, λ_i+Δλ)にある事象
するとP(A_i|B)は、災害件数が実現値X=X₀のもとで、発生率λが(λ_i, λ_i+Δλ)にある確率を示し、これはh(λ)Δλに等しい。ただしh(λ)はλの事後確率の密度関数である。一方、P(B|A_i)はλが(λ_i,

$\lambda_i + \Delta \lambda$ のもとで、 $X = X_0$ が実現する確率であり、これはポアソン分布で与えられることが分かっているので、

$$P(B|A_i) = P(X_0 | \lambda_i) = \frac{(\lambda_i T_0)^{x_0}}{X_0!} \cdot \exp\{-\lambda_i T_0\}$$

と書ける。また $P(A_i)$ は λ の区間の大きさのみに比例し、 $P(A_i) = \Delta \lambda$ で示される。よってベイズの定理により、

$$h(\lambda_i) \Delta \lambda = \frac{\Delta \lambda \frac{(\lambda_i T_0)^{x_0}}{X_0!} \cdot \exp\{-\lambda_i T_0\}}{\sum_{i=1}^n \Delta \lambda \frac{(\lambda_i T_0)^{x_0}}{X_0!} \cdot \exp\{-\lambda_i T_0\}}$$

ここで、 $N \rightarrow \infty$ 、 $\Delta \lambda \rightarrow 0$ の極限を考えると、

$$h(\lambda) = \frac{(\lambda T_0)^{x_0} \cdot \exp\{-\lambda T_0\}}{\int_0^A (\lambda T_0)^{x_0} \cdot \exp\{-\lambda T_0\} d\lambda} = \frac{(\lambda T_0)^{x_0} \cdot \exp\{-\lambda T_0\}}{(1-\varepsilon) \int_0^\infty (\lambda T_0)^{x_0} \cdot \exp\{-\lambda T_0\} d\lambda}$$

さらに ε を十分に小さく考えると、最終的に

$$h(\lambda) = \frac{(\lambda T_0)^{x_0} \cdot \exp\{-\lambda T_0\}}{\int_0^\infty (\lambda T_0)^{x_0} \cdot \exp\{-\lambda T_0\} d\lambda} = \frac{(\lambda T_0)^{x_0}}{X_0!} \cdot \exp\{-\lambda T_0\} \quad (14)$$

なる確率分布式に到達する。上式の期待値および分散は次の通りである。

$$E(\lambda) = \frac{X_0 + 1}{T_0}, \quad V(\lambda) = \frac{X_0 + 1}{T_0^2} \quad (15)$$

かくして、期間 T_0 において X_0 件の災害が発生した条件のもとでの災害発生率 λ の事後確率は、事前情報である X_0 、 T_0 をパラメータとするガンマ分布(14)式で示されることがわかった。前章において、災害発生率の分布として、主として複合化を行う計算上の都合からガンマ分布を採用したが、上にもみるとおり、ベイズ

の定理から事後分布として同分布が導出されており、ガンマ分布が計算上の利便さのみならず確率論的にも裏付けを有した分布であることが明らかとなった。

また、期間 T_0 で X_0 件の災害が発生したのに引き続き、さらに期間 T_1 で X_1 件の災害が発生したときの災害発生率 λ の事後確率分布は、(14)式の分布を新たに事前分布としてベイズの定理を適用すると、

$$h(\lambda) = \frac{\lambda^{x_0+x_1} T_0^{x_0+1} T_1^{x_1}}{X_0! X_1!} = \frac{(X_0 + X_1)!}{X_0! X_1!} \times \frac{\exp\{-\lambda(T_0 + T_1)\}}{T_0^{x_0+1} T_1^{x_1} (T_0 + T_1)^{x_0+x_1+1}} = \frac{[\lambda(T_0 + T_1)]^{x_0+x_1} (T_0 + T_1)}{(X_0 + X_1)!} \times \exp\{-\lambda(T_0 + T_1)\} \quad (16)$$

を新たな事後分布として得る。同式は、 $(X_0 + X_1)$ 、 $(T_0 + T_1)$ をパラメータとするガンマ分布式である。ベイズの定理ではこのように、データが得られる度に新規に得られた情報とそれ以前の情報とを統合し、情報を更新しながら新たな事後確率分布を求めることが出来る。

以下同様に、期間 T_i で X_i 件の災害が発生したことを (X_i, T_i) と記すと、 (X_1, T_1) 、 (X_2, T_2) 、…… (X_n, T_n) のもとでの災害発生率の事後確率分布の一般式は次式のようなになる。

$$h(\lambda) = \frac{(\lambda T)^x T}{X!} \cdot \exp\{-\lambda T\} \quad (17)$$

$$E(\lambda) = \frac{X+1}{T}, \quad V(\lambda) = \frac{X+1}{T^2} \quad (18)$$

$$\text{ここに } X = \sum_{i=0}^n X_i, \quad T = \sum_{i=0}^n T_i$$

これまで述べてきたベイズ定理を用いた災害発生率の分布を分析した事例として、ここでは、平均災害率を $\lambda = X/T = 0.1$ と一定に保ち、 $(X, T) = (10, 100)$, $(5, 50)$, $(2, 20)$ の3つの異なったケースについて災害発生率の事後分布を求めた結果を Fig. 10 に示した。同図から、同じ平均災害発生率であっても、観測期間が長いほど λ の変動が小さくなるのがみてとれる。これは分布の分散が観測時間数 T の二乗に反比例すると同時に、データのちらばり具合を相対的に表す尺度である変動係数 ($CV = \sqrt{X+1}$) が、 T が長いほど X も多くなるので相対的に変動幅も小さくなっているためである。

3.2 ベイズ方式による災害発生数の分布

期間 T で X 件の災害が発生したときの災害発生率 λ の事後分布が(17)式で表されることが分かったので、同式のもとで災害発生数の分布を求めるには、(1)式のポアソン分布を(17)式で複合化すればよい。

$$\begin{aligned}
 P(k) &= \int_0^\infty \frac{(\lambda t)^k}{K!} \cdot \exp\{-\lambda t\} \\
 &\quad \times \frac{(\lambda T)^X T}{X!} \cdot \exp\{-\lambda T\} d\lambda \\
 &= \frac{(X+K)!}{X!K!} \cdot \frac{T^{X+1} t^k}{(T+t)^{X+K+1}} \\
 &\quad \times \int_0^\infty \frac{[\lambda(T+t)]^{X+K} (T+t)}{(X+K)!} \\
 &\quad \quad \times \exp\{-\lambda(T+t)\} d\lambda \\
 &= \frac{(X+K)!}{X!K!} \cdot \frac{T^{X+1} t^k}{(T+t)^{X+K+1}} \\
 &= \binom{X+K}{K} \left(\frac{T}{T+t}\right)^{X+1} \left(\frac{t}{T+t}\right)^k
 \end{aligned} \tag{19}$$

$$\begin{aligned}
 E(K) &= \frac{(X+1)t}{T}, \\
 V(K) &= \frac{(X+1)(T+t)t}{T^2}
 \end{aligned} \tag{20}$$

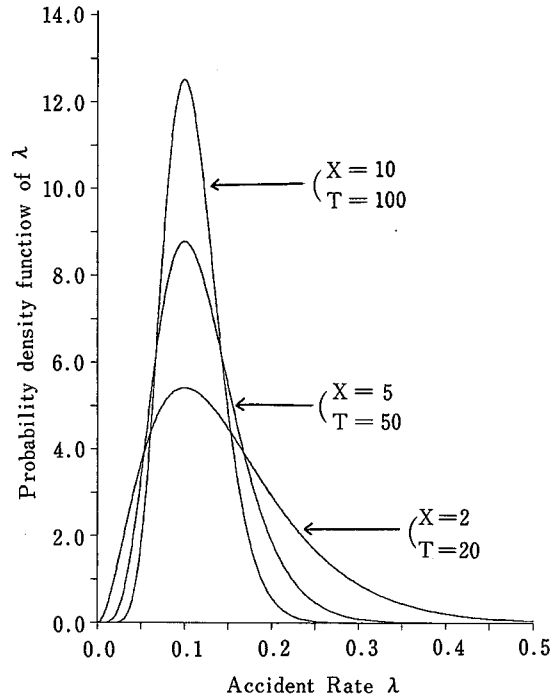


Fig. 10 Posterior distribution of accident rate based on the Bayes' theorem

ベイズの定理による災害発生率の事後分布

結果は(4)式と同じ負の二項分布となるが、(4)式と異なる点は分布のパラメータ X, T が既に与えられていることにある。また(19)式の計算を行うには以下の漸化式を利用する方法が便利である。

$$\begin{aligned}
 P(0) &= \left(\frac{T}{T+t}\right)^{X+1} \tag{21} \\
 P(1) &= \frac{(X+1)!}{X!1!} \left(\frac{T}{T+t}\right)^{X+1} \left(\frac{t}{T+t}\right)^1 \\
 &= (X+1) \left(\frac{t}{T+t}\right) P(0) \\
 P(2) &= \frac{(X+2)!}{X!2!} \left(\frac{T}{T+t}\right)^{X+1} \left(\frac{t}{T+t}\right)^2 \\
 &= \frac{(X+2)}{2!} \left(\frac{t}{T+t}\right) \\
 &\quad \times \frac{(X+1)!}{X!1!} \left(\frac{T}{T+t}\right)^{X+1} \left(\frac{t}{T+t}\right)^1 \\
 &= \frac{(X+2)}{2} \left(\frac{t}{T+t}\right) P(1)
 \end{aligned}$$

以下要するに災害発生件数 k 件の発生確率 $P(K)$ は、初項 $P(0)$ から出発して次式によって求めることができる。

$$P(K) = \left(\frac{X+K}{K} \right) \left(\frac{t}{T+t} \right) P(K-1) \quad (22)$$

同式による計算事例として、前節で求めた3種類の災害発生率 λ の事後分布 (Fig. 10) のもとで、期間を $t=100$ のときの災害発生数の分布を計算した結果を Fig. 11 に示した。同図に示される通り、事前情報として与えられる平均災害発生率 ($\lambda = X/T$) が同一であっても、事前情報の観測時間数の長さに応じて災害発生数の分布形状が異なることがわかる。Fig. 10 と同じく観測期間が長いほど災害発生数の分布の変動幅が小さくなっているが、理由は λ の事後確率分布のときと同じである。

ここで、Fig. 10, Fig. 11 において、 $\lambda = X/T = \text{const.}$ に保ったままで観測期間 $T \rightarrow \infty$ (従って $X \rightarrow \infty$) の極限を考えてみよう。まず λ の分布については式 (18) より、

$$\lim_{t \rightarrow \infty} \{E(\lambda)\} = X/T \quad (23)$$

$$\text{および } \lim_{t \rightarrow \infty} \{V(\lambda)\} = 0$$

が得られ、最終的には $h(\lambda) = X/T$ の一点分布となることが分かる。

一方災害発生数の極限分布は、災害発生率 λ の極限分布が $\lambda = X/T$ の一点分布となることから、

$$P(K) = \frac{(t \cdot X/T)^k}{K!} \cdot \exp\{-t \cdot X/T\} \quad (24)$$

なるポアソン分布となる。このようにポアソン分布は負の二項分布の極限分布であるので、ポアソン分布を負の二項分布の近似分布として用いる場合には、災害発生率 λ の信頼性を十分に検討しておく必要がある。

3.3 ベイズ方式による災害発生時間の分布

期間 T で X の災害が発生した条件のもとでの災害発生時間数の分布を求めるには、災害発生時間数の分布 (8) 式と災害発生率 λ の事後分布 (17) 式とを複合化すればよい。

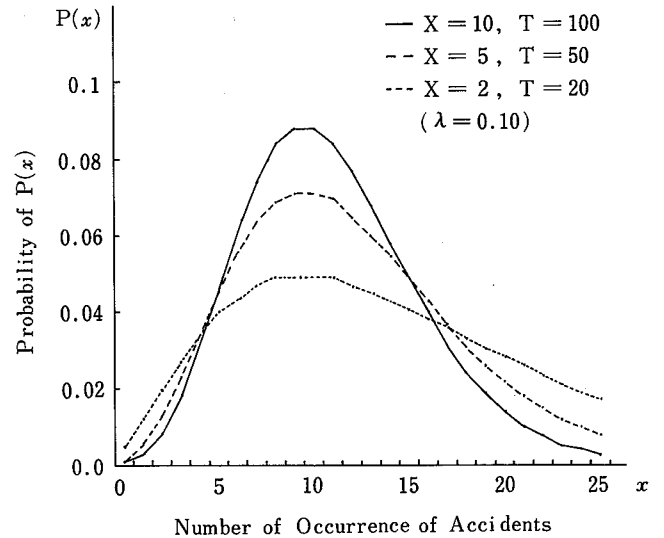


Fig. 11 Frequency distribution of the number of accidents within 100 million hour based on the Bayes' theorem

ベイズ方式によって災害発生率の変動を考慮したときの100万時間当りの災害発生数の分布

$$\begin{aligned} f(t) &= \int_0^\infty \lambda \cdot \exp\{-\lambda t\} \\ &\quad \times \frac{(\lambda T)^x T}{X!} \cdot \exp\{-\lambda T\} d\lambda \\ &= \frac{T^{x+1}}{X!} \int_0^\infty \lambda^{x+1} \cdot \exp\{-\lambda(T+t)\} d\lambda \\ &= \frac{(X+1)!}{X!} \cdot \frac{T^{x+1}}{(T+t)^{x+2}} \\ &= \frac{(X+1)}{T} \left(\frac{T}{T+t} \right)^{x+2} \quad (25) \end{aligned}$$

また同式の確率分布関数および期待値、分散は次式で示される。

$$F(t) = \int_0^t f(t) dt = 1 - \left(\frac{T}{T+t} \right)^{x+1} \quad (26)$$

$$E(t) = \frac{T}{X}, \quad V(t) = \frac{T^2(X+1)}{X^2(X-1)} \quad (27)$$

また上式で $\lambda = X/T$ を一定に保ったままで観測期間を $T \rightarrow \infty$ (従って $X \rightarrow \infty$) としたときの極限分布は、

$$\lim_{t \rightarrow \infty} \frac{(X+1)}{T} = \frac{X}{T}$$

および

$$\lim_{t \rightarrow \infty} \left[\log \left(\frac{T}{T+t} \right)^{x+2} \right] \doteq - (X+2) \frac{t}{T}$$

より

$$\begin{aligned} \lim_{t \rightarrow \infty} \left(\frac{T}{T+t} \right)^{x+2} &= \lim_{t \rightarrow \infty} \left[\exp \left\{ - \frac{(X+2)t}{T} \right\} \right] \\ &= \exp \left\{ - \frac{X}{T} t \right\} \end{aligned}$$

などの関係から次式の指数分布がその極限分布として得られる。

$$f(t) = \left(\frac{X}{T} \right) \cdot \exp \left\{ - \frac{X}{T} t \right\} \quad (28)$$

上式による計算事例としてFig. 12に、平均災害発生率 $\lambda = X/T = 0.1$ のときで $(X, T) = (1, 10), (2, 20), (5, 50), (10, 100)$ の4つのケースについて計算した結果を示した。同図には、 $\lambda = X/T$ を一定に保ちながら $T \rightarrow \infty$ としたときの極限分布である指数分布を破線で示した。観測期間 T が長いほど指数分布に漸近してゆく様子がうかがえる。

以上是个々の災害が発生する時間数(時間間隔)の分布であるが、以下ではより一般的な複数件の災害が発生するまでの時間数の分布について考察を加える。

災害発生率を λ に固定したときの個々の災害発生時間(間隔)数の分布は(8)式の指数分布で与えられているが、このとき複数個 k 件の災害が発生するまでの時間数の分布は、指数分布の和の分布として、たたみ込み積分によって次式のガンマ分布となることが知られている²⁾。

$$f_k(t) = \frac{(\lambda t)^{k-1} \lambda}{(K-1)!} \exp\{-\lambda t\} \quad (29)$$

上式は確率密度関数であるが、その分布関数並びに期待値および分散はそれぞれ次の式で示される。

$$F_k(t) = 1 - \sum_{i=0}^{k-1} \frac{(\lambda t)^i}{i!} \cdot \exp\{-\lambda t\} \quad (30)$$

$$E_k(t) = K/\lambda, \quad V_k(t) = K/\lambda^2 \quad (31)$$

よって、 T 時間で X 件の災害が発生した条件のもとで、以降に複数件の災害が発生する時間数の分布は、

(29)式を事後分布(17)式で複合化すればよい。結果は、 K 件目の災害発生時間数を t_k と書くと、その確率密度関数として、

$$f_k(t_k) = \binom{X+K}{K} \frac{K}{t_k} \left(\frac{T}{T+t_k} \right)^{x+1} \left(\frac{t_k}{T+t_k} \right)^k \quad (32)$$

が得られ、またその確率分布関数および期待値、分散はそれぞれ次式で示される。

$$F_k(t_k) = 1 - \left(\frac{T}{T+t_k} \right)^{x+1} \cdot \sum_{i=0}^{k-1} \binom{X+i}{i} \left(\frac{t_k}{T+t_k} \right)^i \quad (33)$$

$$E(T_k) = K \frac{T}{X}, \quad V(T_k) = \frac{KT^2(X+K)}{X^2(X-1)} \quad (34)$$

上に示した式は特に定まった分布式名はないので、ここでは複合化ガンマ分布と呼ぶことにする。上記の各式で $K=1$ とおくと、複合化指数分布の(25)~(27)式となることが分かる。

また複合化ガンマ分布の分布関数は負の二項分布の加算値を1から引いた値となっているが、これと同様な関係が複合化する前の分布でも成立している。すなわち、(29)式の複合化前のガンマ分布の分布関数は、ポアソン分布の和を1から引いた値となっている。

また、(32)式と(19)式を比べてみると、複合化ガンマ分布は、負の二項分布に K/t_k を掛けたものであることが分かる。従って(32)式の計算を簡便に行うには、まず(22)式の漸化式から所定の災害件数 K 件の $P(K)$ を求め、この $P(K)$ を用いて次式で計算すればよい。

$$f_k(t) = \frac{K}{t} P(K) \quad (35)$$

ただし

$$f_{k=1}(0) = \frac{X+1}{T}, \quad f_{k \neq 1}(0) = 0 \quad (36)$$

また、(32)式で $\lambda = X/T$ を一定に保ち、観測期間を $T \rightarrow \infty$ にした極限の分布は、複合指数分布のときと同様な手続きを経て、最終的に次の式に到達する。

$$f_k(t) = \frac{(X/T \cdot t)^{k-1}}{(K-1)!} \frac{X}{T} \cdot \exp\left\{-\frac{X}{T} t\right\} \quad (37)$$

上式で $K=1$ とおくと指数分布の(28)式となる。

上で述べた複合化ガンマ分布による計算事例として、ここでは $K=3$ の場合について、Fig. 12と同じく、平均災害発生率を $\lambda = X/T = 0.1$ として (X, T) のいくつかのケースについてその発生時間数の確率密度関数を求めた結果をFig. 13に示した。

事業所等において、例えば期間 $T=10$ までに $X=1$ の災害が起こったとすると、その情報のもとでの災害発生時間数の分布は、同図の $(X, T)=(1, 10)$ の分布曲線で示される。同分布曲線は正方向にひずんだ分布であることがわかる。また、同じ事業所で引き続いて次の $T=10$ の期間で災害も同じく $X=1$ 件の災害を得たとすると、当初からの災害情報を合計すると、 $T=20$ において $X=2$ を得たことになる。この情報のもとでの災害発生時間数の分布は $(X, T)=(2, 20)$ の分布曲線で表されることとなる。以下同様にして、観測期間長 T が延びて災害情報も増すに従って、増加した災害情報を用いて災害発生率の新たな事後分布を推定し、この新しい事後分布のもとでの災害発生時間数の分布を更新するという手続きを繰回して評価を行うことが出来る。

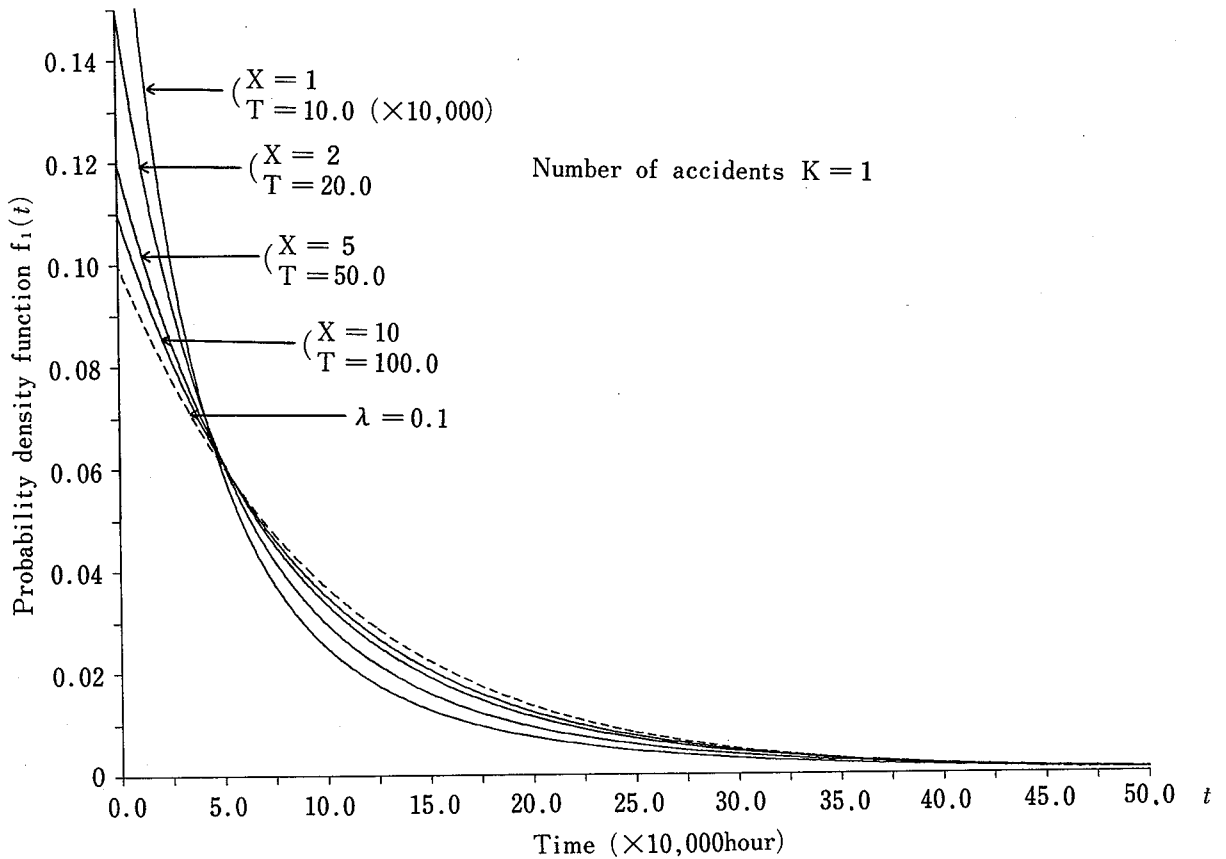


Fig. 12 Frequency distribution of time intervals between accidents considering the variance of accident rates based on the Bayes' theorem ($K=1$)

災害発生率の変動をベイズ方式で考慮したときの
災害発生時間分布 ($k=1$)

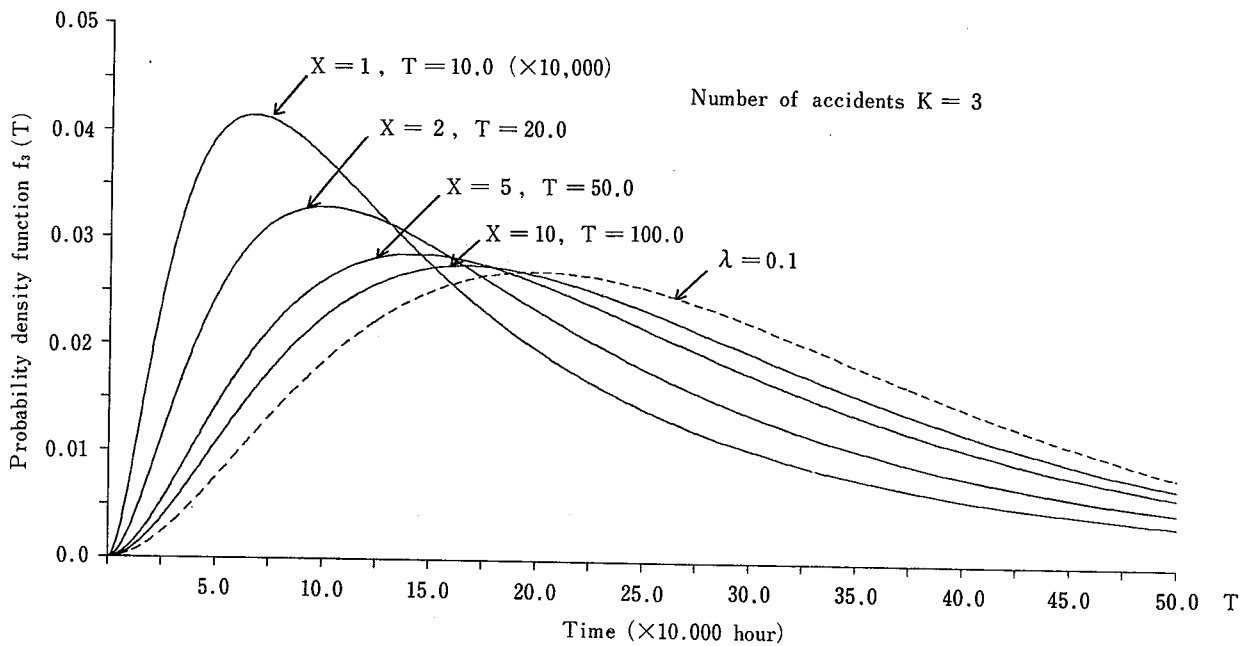


Fig. 13 Frequency distribution of time intervals between accidents considering the variance of accident rates based on the Bayes' theorem ($K=3$)

災害発生率の変動をベイズ方式で考慮したときの
災害発生時間分布 ($k=3$)

Fig. 12, Fig. 13には、このようにして観測期間 T が長くなるに従って災害発生時間数の分布が極限分布の指数分布あるいはガンマ分布に漸近してゆく様子が示されている。このように、指数分布あるいはガンマ分布を複合化ガンマ分布の近似分布として用いるには、事前情報を得るための観測期間が十分大きいことが必要であることが分かる。

以上本章では、ベイズの定理を利用して、ある時点までに得られた災害情報をもとに災害発生率の確率分布を推定し、得られた災害発生率の確率分布をポアソン分布やガンマ分布と複合化して発生数の分布である負の二項分布や、災害発生時間数の分布である複合化ガンマ分布を導出し、これらの確率分布式を用いて災害発生数の評価や発生時間数の予測を行うことを試みた。

災害データが数多くあるときには、これらのデータから信頼度の高い災害発生率を推定することが出来るが、一般の事業所では必ずしも十分なデータが蓄積されている訳ではない。したがって、少数のデータであっても災害発生率を推定することが出来るベイズ方

式の分析法は、このような場合には極めて有効な手法と思われる。さらにベイズ定理による推定法は、観測期間が長くなり災害情報が多く得られるに従って、推定される災害発生率の確信度が高まる特性を有しているので、少数データ段階での推定から大量データによる推定まで、幅広い段階での評価に適用出来る利点も有している。今後この手法が安全性評価で果たす役割は大きいと思われる。

4. ま と め

以上本論文では、労働災害発生頻度率がある確率分布に従って変動する場合を想定し、そのときの労働災害発生数や災害発生時間数などの確率分布式を求めたのち、実際の災害事例を用いてこれらの式の妥当性を検討し、ついで、これらの結果をふまえて、災害データが与えられたときに災害発生数や発生時間数の予測を行う研究を行った。それらの結果をまとめると以下のようなものである。

(1) 一定期間中における災害発生数の分析結果から、多くの災害が、単独に負の二項分布に、あるいは

負の二項分布とポアソン分布とに同時に適合していることを示した。とくに単独の分布としては、ポアソン分布よりも、むしろ負の二項分布の方が災害事例をうまく表現している場合が多いことを明らかにした。

(2) このことより、多くの災害現象は、災害事象そのものはランダムに発生するが、その単位時間当りの災害発生数(災害発生率)はある確率分布(具体的にはガンマ分布)に従って生ずる、いわゆる複合過程のもとで災害が起こっていることを明らかにした。

(3) 災害事例が数多くあれば、災害データを利用して、負の二項分布やポアソン分布のパラメータである災害発生頻度率の分布を推定し、これらを用いて災害発生数や発生時間数の確率分布式を規定することが出来ることを示した。

(4) 災害事例が少ないときには、ヘイズ方式によってこれらの分布のパラメータを推定することが出来ることを明らかにした。すなわち、分布のパラメータ(災害発生率)の分布は、ベイズの定理から災害件数 X と発生時間数 T をパラメータとするガンマ分布となることを導き出し、ついでこの分布とポアソン分布と複合化して得られる災害発生数の分布は負の二項となることや、また、災害発生時間数の分布は、発生時間数のガンマ分布と災害発生率のガンマ分布とを混合した、いわゆる複合化ガンマ分布となることを示した。さらに、これらの複合化分布は、災害発生率を固定したままで発生時間数や災害件数を十分大きくしたときは、ポアソン分布やガンマ分布がその極限分布として得られることを示し、ポアソン分布やガンマ分布が複合化分布の一種の近似分布であることを明らかにした。

(5) ベイズ方式を利用すれば、災害データが得られる度に災害発生率のパラメータの分布を更新して新しい事後分布を求め、この分布を用いて新たな複合化分布を導き出し、同確率分布式によってそれまでの災害発生数や災害発生時間数の確率的な予測や評価を行うことが出来ることを示した。

(平成2年3月30日 受理)

参 考 文 献

- 1) たとえば、労働災害動向調査報告(昭和63年)、労働大臣官房政策調査部、1989
- 2) 花安繁郎：災害発生間隔の分布に関する研究、労働省産業安全研究所研究報告、RIIS-RR-26-3, 1977
- 3) 花安繁郎：災害発生時間による安全水準の評価について、土木学会論文報告集、第301号、pp.105~113, 1980
- 4) 花安繁郎：災害発生時間の分布に関する研究(2)、労働省産業安全研究所研究報告、RIIS-RR-32-4, 1984
- 5) 花安繁郎：労働災害発生系列の均質性に関する研究、建築学会構造系論文報告集、第352号、pp.1~9, 1985
- 6) 花安繁郎：災害発生時間の分布に関する研究(3)、労働省産業安全研究所研究報告、RIIS-RR-86-6, pp.69~92, 1986
- 7) 花安繁郎：災害発生時間の分布に関する研究(4)、労働省産業安全研究所研究報告、RIIS-RR-88-3, pp.23~40, 1988
- 8) 花安繁郎、五十嵐日出夫：逐次確率比検定法による労働災害発生率の統計評価に関する研究、土木学会論文集、第409号/VI-11, pp.47~56, 1989
- 9) M. Greenwood, H. M. Wood: The Incidence of Industrial Accidents upon Individuals with Special Reference to Multiple Accidents, 1919, In: W. Haddon Jr., E. A. Suchman and D. A. Klein, Accident Research, Harper and Row, New York, pp.389~396, 1964 (Reprinted from Report No. 4, 1919, Industrial Fatigue Research Board, Medical Research Committee, Great Britain)
- 10) M. Greenwood, G. U. Yule: An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular References to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents, J. of Roy. Statis. Soc., Vol.83, pp.255~279, 1920
- 11) 印東太郎：確率および統計、コロナ社、pp.104~113, 1967
- 12) 増山元三郎：少数例のまとめ方Ⅱ、竹内書店新社、pp.590~595, 1976
- 13) 芝祐順、渡部洋、石塚智一編：統計用語辞典、新曜社、pp.239~241, 1984
- 14) 大村平、今田直孝：推測統計のFORTRAN、オーム社、pp.92~97, 1972