

リスク曲線における確率分布パラメータの推定に関する研究

横浜国立大学 安心・安全の科学研究教育センター 正会員 花安 繁郎, 関根 和喜
独立行政法人 労働安全衛生総合研究所 正会員 大嶋 勝利

1. はじめに

リスク曲線とは、災害による被害規模とその被害規模を超える災害発生数との関係を両対数紙上に記したものである。リスク曲線は被害規模の確率分布を記述しており、同分布のパラメータはリスク曲線の形状を規定している。本稿は、リスク曲線のパラメータの推定について考察を加えたものである。

2. 産業災害による被災規模分布

産業災害による被害規模について、測定可能な最小値によって基準化した被害規模（正規化被害値）の大きさを r と記すと、その確率密度関数および上側確率分布関数は次式のパレート分布で示される。式中の n は同確率分布のパラメータである。

$$p(r) = nr^{-n-1} \quad (1)$$

$$R(r) = r^{-n} \quad (2)$$

図1には1977年～2000年の24年間に我が国の建設工事で発生した重大労働災害（一度に三人以上の受傷者を含んだ労働災害）について、災害による被害規模（一件の重大災害による受傷者数）とその被

害規模を超えた災害の発生数との関係を両対数紙上に整理したいわゆるリスク曲線を示した。同図から建設工事における重大労働災害による被災者数の規模分布が両対数紙上の直線として示されており、被害規模がパレート分布で記述されることが分かる。

3. 被害規模分布のベイズ方式によるパラメータ推定

産業災害による被害規模の分布であるパレート分布のパラメータを推定するための統計的手法には、モーメント法、最尤法、最小自乗法、ベイズ法などさまざまな手法がある。ここでは被害規模のデータが与えられたとき、ベイズの定理を用いてパラメータ n を事後分布として推定する手法を用いた。

ベイズ方式では、パラメータが確率分布として与えられる、データが得られるたびにそれまでの情報を用いてパラメータ値を更新することができる、元の確率分布と複合化して予測分布を導出することができるなど多くの解析上有利な特性を有している。

まず、 $r_i (i=1, \dots, N)$ として被害規模のデータが与えられたとき、当初はデータが無い状態から出発するので、パラメータに関する初期事前分布には $f(n)=1$ の一様分布を仮定した。そして、最初のデータ r_1 が得られたとき、 r_1 の下でのパラメータの事後分布 $f(n|r_1)$ はベイズの定理により次式として得られる。

$$f(n|r_1) = \frac{nr_1^{-n-1}}{\int_0^\infty nr_1^{-n-1} dn} = n(\ln r_1)^2 r_1^{-n} \quad (3)$$

次いで、2番目のデータ r_2 が得られたとき、 r_1, r_2 の下でのパラメータ事後分布 $f(n|r_1, r_2)$ は、 $f(n|r_1)$ を事前分布としてベイズの定理から次式となる。

$$\begin{aligned} f(n|r_1, r_2) &= \frac{nr_2^{-n-1} \cdot n(\ln r_1)^2 r_1^{-n}}{\int_0^\infty nr_2^{-n-1} \cdot n(\ln r_1)^2 r_1^{-n} dn} \\ &= \frac{n^2}{2!} [\ln(r_1 \cdot r_2)]^3 (r_1 \cdot r_2)^{-n} \end{aligned} \quad (4)$$

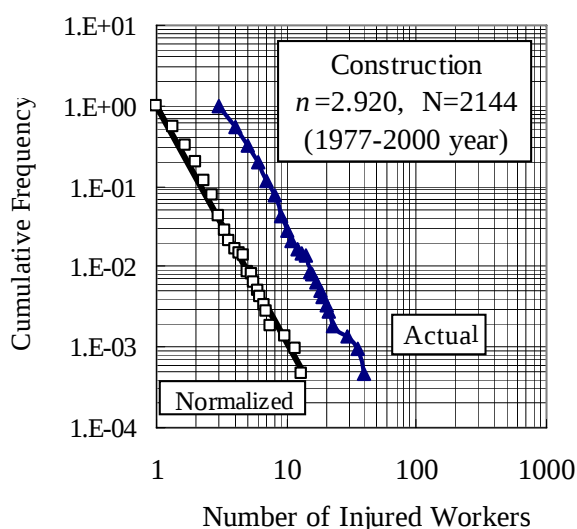


図1 被害規模と超過発生頻度との関係

Keywords : 産業災害, 被害規模分布, リスク曲線, パラメータ推定, ベイズの定理

連絡先 : 〒240-8501 横浜市保土ヶ谷区常盤台 79-5 TEL. 045-339-3778, FAX. 045-339-4294

以下、同様な繰り返しを行うと、 $r_i (i=1, \dots, N)$ のデータが得られたとき、パラメータの事後確率分布の一般式は次式で与えられる。

$$f(n|r_1, \dots, r_N) = \frac{n^N}{N!} [\ln(\prod_{i=1}^N r_i)]^{N+1} (\prod_{i=1}^N r_i)^{-n} \quad (5)$$

同式の確率分布関数および期待値と分散はそれぞれ下式で示される。

$$F(n) = 1 - (\prod_{i=1}^N r_i)^{-n} \cdot \sum_{i=0}^N \frac{[n \cdot \ln(\prod_{i=1}^N r_i)]^i}{i!} \quad (6)$$

$$E(n) = \frac{(N+1)}{\ln(\prod_{i=1}^N r_i)} \quad (7)$$

$$V(n) = \frac{(N+1)}{[\ln(\prod_{i=1}^N r_i)]^2} \quad (8)$$

4. パラメータ推定のための漸化式

パラメータ n の事後確率分布の一般式は(5)式で与えられるが、式中に階乗計算 $N!$ が入っているために、計算をそのまま実行すると災害数が大きくなるに従って往々にして計算不能の状態に陥る。これを回避するためには漸化式による計算が必要となる。

そのために、まず、各被害値の積の対数（あるいは各被害値の対数の和）を P_N と記すと、 P_N は $P_1 = \ln r_1$ を初項として次の漸化式で表現される。すなわち、

$$\begin{aligned} P_N &= \ln(\prod_{i=1}^N r_i) \\ P_N &= P_{N-1} + \ln r_N \end{aligned} \quad (9)$$

また、(5)式を用いて $f_N(n)$ と $f_{N-1}(n)$ との比を取り整理すれば次の漸化式が得られる。ここで初項 $f_1(n)$ は次式で与えられる。

$$\begin{aligned} f_1(n) &= n(\ln r_1)^2 \cdot r_1^{-n} = nP_1^2 r_1^{-n} \\ f_N(n) &= \frac{n}{N} P_N \left(\frac{P_N}{P_{N-1}} \right)^N r_N^{-n} \cdot f_{N-1}(n) \end{aligned} \quad (10)$$

要するに、被害データが与えられとき、まず、各被害値の対数の和を(9)式による漸化式によって求め、次いでその値を用いて(10)式の漸化式を用いて求めるという、2段階の手続きによる計算がなされる。

別解として、 P_N が得られたとき、この値を常数として扱い、次の漸化式によって計算してもよい。

$$\begin{aligned} f_1(n) &= nP_N^2 \exp(-nP_N) = nP_N^2 (\prod_{i=1}^N r_i)^{-n} \\ f_N(n) &= \frac{n}{N} P_N \cdot f_{N-1}(n) \end{aligned} \quad (11)$$

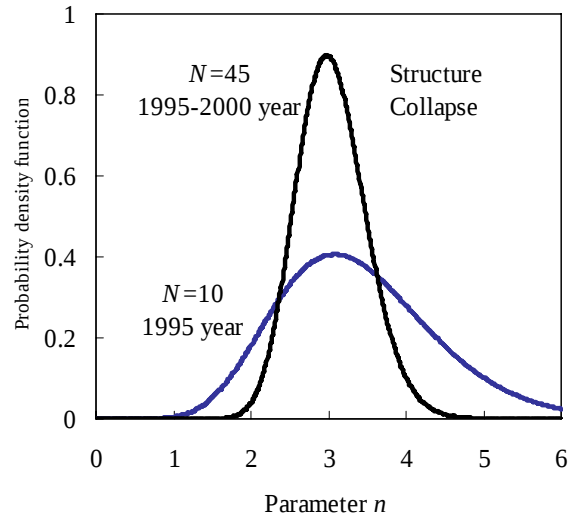


図 2 被害規模分布パラメータの分布

図2は、我が国の建設工事中での構造物の倒壊による重大労働災害について、その被災者数規模分布のパラメータの事後分布を漸化式によって求めた結果を示した。同図より1995年の単年データ($N=10$)による分布よりも、1995年～2000年の6年間データ($N=45$)による分布の分散が小さくなっていることが示されている。データ数 N が大きくなれば、最終的には次式のデルタ関数による一点分布に収束する。

$$\delta\left(n - \frac{N}{\ln(\prod_{i=1}^N r_i)}\right) \quad (12)$$

5. まとめ

産業災害による被害規模の確率分布と同分布のパラメータの推定法について考察を加えた。主たる結果は下記の通りである。

- 1) 産業災害による被害規模とその被害規模を超える災害発生数との関係を両対数紙上に記した曲線はリスク曲線と呼ばれる。リスク曲線は被害規模の確率分布を記述しており、具体的にはパレート分布で表現される。重大労働災害による被害規模（受傷者数）が同分布に従っていることが示された。
- 2) パレート分布のパラメータを、ベイズの定理を用いて被害規模データが与えられたとき、事後確率分布として推定するための一般式を導出した。
- 3) パラメータの事後確率分布は式の中に階乗計算が入っているためにデータ数が大きくなると計算が困難な状態に陥る。これを回避するために2段階の手続きによる漸化式を導出した。