

# 事故発生の偶然変動を考慮した道路区間の 事故危険度の評価手法

## EVALUATION OF ACCIDENT PRONENESS OF ROADWAY SECTIONS IN CONSIDERATION OF RANDOMNESS OF ACCIDENT OCCURRENCE

岡本 博\*・越 正毅\*\*・大蔵 泉\*\*\*・鹿島 茂\*\*\*\*  
By Hiroshi OKAMOTO, Masaki KOSHI, Izumi OHKURA  
and Shigeru KASHIMA

### 1. 本研究の目的と概要

道路構造と交通事故の発生との関係についてよく用いられる実証的な統計解析手法は、ある地域あるいはある路線の道路をいくつかの小区間に分割し、各小区間の事故件数あるいは事故率を被説明変数とし、道路構造を説明変数として重回帰分析を適用する方法である。

ところが、交通事故そのものが確率的な稀現象であるために、事故件数あるいは事故率には無視できない偶然誤差が含まれており、これが重回帰分析で得られるモデルの説明力を低下させ、またその信頼性をあいまいにしている。個々の小区間の区間長を長くすればそれらの事故率のもつ偶然誤差を小さくすることができるが、そうすれば小区間の数、つまりデータ数が少なくなり、また事故率の高い区間と低い区間とが混在して一つの小区間の中に含まれがちになって事故率の大小と道路構造との関係があいまいなサンプルになりやすい。したがって、このような偶然誤差の影響を最小限に抑えるような小区間への分割方法自体がきわめて重要な課題となる。

本研究は、事故率のもつ偶然変動の大きさと、これが重回帰分析に与える影響とを定量的に把握し、偶然性の陰に隠れた真の事故率（ここでは、危険率とよぶ）に対して有効な分析を行うための方法を提示することを目的としている。

まず交通事故を確率事象としてモデル化し、事故率のもつ偶然変動を定量的に評価した。次に事故率説明のための重回帰モデルを想定して、事故率の偶然誤差を小さく抑え、サンプル全体の精度をそろえるためのいくつか

の方法を提示した。東名・名神高速道路における実際の事故データを用いて数種のサンプルを作り、これらに重回帰モデルを適用することにより、本論文で提案する方法の有効性を検証した。

### 2. 従来の研究と本研究の位置づけ

事故の発生と道路構造との関連を分析する場合、一般に被説明変数として、次式で与えられる事故率  $p_0$ （走行台キロ当たりの事故件数）が用いられる。

$$p_0 = \frac{k_0}{ml} \dots\dots\dots (1)$$

ここに、 $k_0$ ：着目する区間内で発生した事故件数、 $m$ ：通過車両台数、 $l$ ：着目する区間の長さ

対象とする道路総延長を適当な長さをもついくつかの小区間（以後要素区間とよぶことにしよう）に分割し、おのおのの要素区間の事故率を求めて、その大小を道路構造の差異で説明しようとするのが一般に用いられる方法である。

ところが、事故件数は全体として少なく、対象道路上に確率的かつ離散的に分布しているため、要素区間のとり方によって、事故率のばらつきが異なる。従来の事故要因分析では、このようなばらつきの大きさについての吟味が十分されていなかったために、説明力の低い分析結果しか得られなかったことが多い。

森らは、危険率という概念を用いることによって、事故率のもつ偶然変動を確率論的に位置づけている<sup>1)</sup>。すなわち、ある区間の危険率  $p$  とは、この区間を自動車が行くという試行において、事故の発生する確率と定義して、事故は危険率  $p$  を母数とする二項分布に従って起こるとしている。このとき、事故率  $p_0 = k_0/m$  は、危険率  $p$  の周りに、分散  $p(1-p)/m$  で分布する確率変数となることを示している。しかしこれだけでは、事

\* 正会員 建設省北陸地方建設局道路部道路計画課  
調査第一係係長

\*\* 正会員 工博 東京大学教授 生産技術研究所

\*\*\* 正会員 工博 横浜国立大学助教授 工学部土木工学科

\*\*\*\* 正会員 工博 中央大学助教授 理工学部土木工学科

故率  $p_0$  のもつ偶然変動を意味づけることにとどまり、その大きさを定量的に評価するには至っていない。この分析に用いられている事故率データについても、その精度に対する検討はなされていない。

大蔵らは、本研究で用いたものと同じ高速道路事故を対象として一連の研究を行ってきた<sup>2),10)</sup>など。ここでは、対象となる道路を 100 m ごとに区切り、事故資料および道路構造をこの区間割りに従って整理した資料に基づいて要因分析を行っている。しかし、100 m 区間ごとの事故率では、偶然変動が大きすぎて説明力のあるモデルが得られないので、この偶然変動を取り除くために、要素区間のとり方をさまざまに変えることを試みているが、いずれも理論的根拠に基づくものではなく、試行錯誤の域を出ていない。

### 3. 交通事故の確率論的モデル化

#### (1) 危険率の定義と事故率

ある要素区間内での事故発生は、次のようにモデル化できる。

長さ  $l$  の要素区間  $S$  があり、これを通過する車両が事故を起こす確率が  $P$  であるとする。  $m$  台の車両が  $S$  を通過するとき、事故が  $k$  件発生する確率  $P_r(k|m, P)$  は、二項分布となる。

交通事故の場合には、  $P$  が十分小さく、  $m$  が十分大きいので、この二項分布はポアソン分布で近似でき、次式で表わすことができる。

$$P_r(k|\lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \cdot \lambda = mP \dots\dots\dots (2)$$

長さの異なる要素区間相互の比較の便のために、上述の  $P$  の代わりに、単位長さ当たりの事故発生確率  $p = P/l$  を考え、これを危険率と定義する。  $p$  を用いると  $\lambda = mlp$  となって、式 (2) は次のように書ける。

$$P_r(k|\lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \cdot \lambda = mlp \dots\dots\dots (3)$$

いま、着目する要素区間  $S$  で、通過車両  $m$  台に対して事故が  $k_0$  件発生したとする。最尤推定法によって  $p$  の点推定量  $\hat{p}$  を求めると、次式のようになる。

$$\hat{p} = \frac{k_0}{ml} = p_0 \dots\dots\dots (4)$$

ただし、尤度関数

$$L(k_0|p) = \frac{(mlp)^{k_0}}{k_0!} e^{-mlp} \dots\dots\dots (5)$$

つまり、事故率  $p_0$  (走行台キロ当たりの事故件数) は、危険率  $p$  の最尤不偏推定量として意味付けられる。

#### (2) 事故率のもつ偶然誤差の評価

未知の定数である危険率  $p$  の推定値として、最尤推

定量である事故率  $p_0$  を用いるとき、推定に伴う誤差  $e$  (事故件数  $k_0$  が確率変数であることによる偶然誤差、以下では偶然誤差と称する) について考える。ここに

$$e = p - p_0 \dots\dots\dots (6)$$

ネイマンの信頼区間推定法によれば、危険率  $p$  の信頼区間を推定することができる<sup>3)</sup>。式 (3) のポアソン分布は、  $\lambda \geq 5$  のとき、正規分布  $N(\lambda, \lambda)$  で近似することができ、このとき次式が成り立つ。

$$P_r\{\lambda - z_\alpha \sqrt{\lambda} < k_0 < \lambda + z_\alpha \sqrt{\lambda}\} = 1 - \alpha \dots\dots\dots (7)$$

ただし、  $1 - \alpha$  は信頼係数であり、  $z_\alpha$  は標準正規分布で  $\int_{-z_\alpha}^{z_\alpha} f(z) dz = 1 - \alpha$  となる値。

信頼区間を作るために、式 (7) の  $\{ \}$  の中の  $\lambda$  と  $k_0$  を入れ換えると、次式が得られる。

$$P_r\left\{k_0 + \frac{z_\alpha^2 - z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2} < \lambda < k_0 + \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2} \middle| \lambda\right\} = 1 - \alpha \dots\dots\dots (8)$$

ここで、  $\lambda = mlp$  の  $ml$  は定数であるから、上式の  $\{ \}$  中を  $ml$  で除して、危険率  $p = \lambda/ml$  の信頼区間とすることができる。

$$P_r\left\{\frac{k_0}{ml} + \frac{z_\alpha^2 - z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml} < p < \frac{k_0}{ml} + \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml} \middle| p\right\} = 1 - \alpha \dots\dots\dots (9)$$

すなわち、危険率  $p$  の信頼係数  $1 - \alpha$  での信頼区間は、事故率  $p_0 (= k_0/ml)$  を用いて、次式のように表わされる。

$$\left(p_0 + \frac{z_\alpha^2 - z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml}, p_0 + \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml}\right) \dots\dots\dots (10)$$

式 (10) をさらに変形すると、

$$P_r\left\{\frac{z_\alpha^2 - z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml} < p - p_0 = e < \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml} \middle| p\right\} = 1 - \alpha \dots\dots\dots (11)$$

これは、偶然誤差  $e$  の信頼係数  $1 - \alpha$  での信頼区間を示している。

そこで、  $e$  の信頼限界のうち、絶対値の大きい方 (信頼上限がこれにあたる) を使って、危険率  $p$  の推定値として事故率  $p_0$  を用いるときの、偶然誤差の大きさを表わす指標とすることを考える。これを、本研究では誤差幅  $e_w$  とよぶことにする。すなわち

$$e_w = \text{Max}\left\{\left|\frac{z_\alpha^2 - z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml}\right|, \left|\frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml}\right|\right\} = \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml} \dots\dots\dots (12)$$

このように誤差幅  $e_w$  を定義すれば、式 (11) より、

$$P_r\{p_0 - p < e_w\} \geq 1 - \alpha \dots\dots\dots(13)$$

が成り立つ。

さらに、事故率に伴う偶然誤差の相対的な大きさを表わす尺度として、誤差率  $e_r$  を次式のように定義する。

$$e_r = \frac{e_w}{p_0} = \frac{z_a^2 + z_a \sqrt{4k_0 + z_a^2}}{2k_0} \dots\dots\dots(14)$$

上式より明らかなように、誤差率  $e_r$  は事故件数  $k_0$  だけの関数となる。

(3) 要素区間のとり方と危険率

要素区間の長さや危険率との関係を明らかにするために、着目する要素区間  $S$  をさらに短い  $t$  個の微小区間  $D_1, \dots, D_t$  に分割してみる。個々の微小区間における事故発生確率を  $P_1, \dots, P_t$  として、微小区間ごとに危険率  $p_1, \dots, p_t$  を考えると次式が成り立つ。

$$p_1 = \frac{P_1}{dl}, \dots, p_t = \frac{P_t}{dl} \quad \left( \text{ただし } dl = \frac{l}{t} \right) \dots\dots\dots(15)$$

$S$  内で、1台の通過車両が2件以上の事故を起こす確率は非常に小さいので、次式が成り立つとしてよい。

$$P = P_1 + P_2 + \dots + P_t \dots\dots\dots(16)$$

したがって、 $S$  の危険率  $p$  は、次式で与えられる。

$$p = \frac{P}{l} = \frac{P_1 + \dots + P_t}{t dl} = \frac{1}{t} (p_1 + \dots + p_t) \dots\dots\dots(17)$$

上式は、要素区間  $S$  の危険率  $p$  が、 $S$  内の微小区間の危険率  $p_i$  の平均値であることを意味している。

要素区間相互の比較によって事故要因を分析しようとするとき、このような危険率  $p$  が十分な意味をもつためには、個々の要素区間内部での事故発生確率がほぼ一様であることが望ましい。一つの要素区間内に危険率の高い微小区間と危険率の低い微小区間とが混在して、全体として中間的な危険率をもつようであれば、このような要素区間は分析の目的にそぐわないものとなる。

4. 事故率説明モデル

(1) 重回帰モデルの設定

事故率  $p_0$  は危険率  $p$  の観測値であり、危険率  $p$  は道路条件に支配されると考えることにする。そこで、危険率  $p$  を媒介概念として、事故率  $p_0$  と道路条件  $x$  との間に、重回帰モデルを適用する場合について考える。

危険率  $p$  と道路条件（道路の線形、構造、交通状況等）を表わす  $j-1$  個の説明変数との間に、次のような線形関係が成り立つと考える。

$$p = \beta_1 + \beta_2 x_2 + \dots + \beta_j x_j + u \dots\dots\dots(18)$$

$u$  は  $p$  の変動のうち  $x_2 \dots x_j$  では説明しきれない部分（以下では方程式のはずれと称する）を表わす攪乱項である。 $u$  の分布の母数および回帰係数  $\beta_1 \dots \beta_j$  は未知で

ある。 $n$  個の要素区間について式 (18) が成り立つとすると、

$$p = X\beta + u$$

$$p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_j \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{21} & \dots & x_{j1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{2n} & \dots & x_{jn} \end{pmatrix} \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \dots\dots\dots(19)$$

と表わすことができる。方程式のはずれ  $u$  について、次の仮定をおく。

$$E(u) = 0 \dots\dots\dots(20)$$

$$E(uu') = \begin{pmatrix} E(u_1^2) & E(u_1 u_2) & \dots & E(u_1 u_n) \\ \vdots & \vdots & & \vdots \\ E(u_n u_1) & \dots & \dots & E(u_n^2) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2 & 0 \\ \vdots & \vdots \\ 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2 I_n \dots\dots\dots(21)$$

ところが、各要素区間の危険率  $p$  は未知であり、その観測値（最尤不偏推定値）として、事故率  $p_0$  が得られるにすぎない。この両者の関係は偶然誤差  $e$  を用いて式 (6) より  $p_0 = p - e$  と書くことができ、次式が成り立つ。

$$E(e) = E(p - p_0) = 0 \dots\dots\dots(22)$$

また、偶然誤差相互は、互いに独立であると考えてよいであろうから次式が成立すると仮定できる。

$$E(e_i e_j) = 0 \quad (i \neq j). \quad E(e_i^2) = \varphi_i^2 \dots\dots\dots(23)$$

ここに  $\varphi_i^2$  は、観測値ごとの偶然変動の大きさを表わす。式 (6), (22), (23) を行列の形で書くと、次のようになる。

$$p_0 = p - e \quad p_0 = \begin{pmatrix} p_{01} \\ \vdots \\ p_{0n} \end{pmatrix} \quad p = \begin{pmatrix} p_1 \\ \vdots \\ p_r \end{pmatrix} \quad e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \dots\dots\dots(24)$$

$$E(e) = 0 \dots\dots\dots(25)$$

$$E(ee') = \Phi_n \dots\dots\dots(26)$$

ただし  $\Phi_n = \begin{pmatrix} \varphi_1^2 & & 0 \\ & \dots & \\ 0 & & \varphi_n^2 \end{pmatrix}$

式 (24) に式 (19) を代入すると、次式が得られる。

$$p_0 = p - e = X\beta + u - e = X\beta + w. \quad w = u - e \dots\dots\dots(27)$$

ここに  $w$  は、事故率  $p_0$  とこれを説明しようとするモデル式との間にある攪乱項である。この攪乱項の分散共分散行列を  $W$  で表わすことにすると、

$$E(w) = E(u) - E(e) = 0 \dots\dots\dots(28)$$

$$E[ww'] = W = E(uu') - E(ue') - E(eu') + E(ee') \dots\dots\dots(29)$$

ここでさらに、方程式のはずれ  $u$  と偶然誤差  $e$  とは無相関であると仮定する。すなわち、

$$E(ue') = E(eu') = 0 \dots\dots\dots(30)$$

式 (29) に対して式 (21), (26), (30) の仮定を代入すると、次式が得られる。

$$W = \sigma^2 I_n + \phi_n \dots\dots\dots (31)$$

以上をまとめると事故率を説明する重回帰モデルは、次式で表わさる。

$$p_0 = X\beta + w \quad w = u - e \dots\dots\dots (27)$$

$$E(w) = 0 \dots\dots\dots (28)$$

$$E(ww') = W = \sigma^2 I_n + \phi_n \dots\dots\dots (31)$$

(2) 偶然変動の大きさについての考察

前節で定義した偶然変動の大きさ  $\varphi^2$  について検討する。式 (6), (23) より、

$$\begin{aligned} \varphi^2 &= E(e^2) = E[(p - p_0)^2] \\ &= E\left[\left(\frac{\lambda - k_0}{ml}\right)^2\right] \quad \left(\because p_0 = \frac{k_0}{ml}, p = \frac{\lambda}{ml}\right) \\ &= \frac{1}{m^2 l^2} E[(\lambda - k_0)^2] = \frac{\lambda}{m^2 l^2} \\ &(\because k_0 \text{ は母数 } \lambda \text{ のポアソン分布に従う}) \end{aligned} \dots\dots\dots (32)$$

上式の  $\lambda$  は未知の母数であり、このままでは  $\varphi^2$  の大きさを知ることができない。そこで、フィデューシャル論によって、母数の論理的役割を逆転させ、確率変数化することにより、観測値  $k_0$  を用いて  $\varphi^2$  を評価することを考える<sup>5)</sup>。

$k_0$  が母数  $\lambda$  のポアソン分布に従うとき、 $\lambda > 5$  ならば  $z = (k_0 - \lambda) / \sqrt{\lambda}$  は正規分布  $N(0, 1)$  に従うものとみなせる。ここで  $z$  は、観測値  $k_0$  と母数  $\lambda$  との間の規準化された誤差を表わすと考えられる。そこで、もともと考えるべき確率変数は  $z$  であるとし、 $\lambda$  は次の式 (33) の関係によって、確率変数としての性格を獲得するものとする。言い換えると、事故件数  $k_0$  が観測されたとき、 $\lambda$  は値  $k_0$  を生ずる可能性において、さまざまな値をとる確率変数であるとみなすのである。

$$z = \frac{k_0 - \lambda^*}{\sqrt{\lambda^*}} \dots\dots\dots (33)$$

( $\lambda$  を確率変数化したことを強調するため  $\lambda^*$  と記す) 変形して、 $\lambda^*$  について解き、式 (33) から  $k_0 \geq \lambda^*$  のとき  $z \geq 0$ ,  $k_0 < \lambda^*$  のとき  $z < 0$  であることを考えると、 $\lambda^*$  は次のようになる。

$$\lambda^* = k_0 + \frac{z^2 - z\sqrt{4k_0 + z^2}}{2} \dots\dots\dots (34)$$

( $z$  は  $N(0, 1)$  に従う確率変数)

上式を  $p^* = \lambda^* / ml$  を用いてさらに変形すると、次式が得られる。

$$p^* = p_0 + \frac{z^2 - z\sqrt{4k_0 + z^2}}{2ml} \dots\dots\dots (35)$$

したがって、偶然変動の大きさは、次式のようになる。

$$\begin{aligned} \varphi_p^2 &= E(e^{*2}) = E[(p^* - p_0)^2] \\ &= E\left[\left(\frac{z - z\sqrt{4k_0 + z^2}}{2ml}\right)^2\right] \\ &= \frac{1}{4m^2 l^2} E[z^4 - 2z^3\sqrt{4k_0 + z^2} + (z^2 + 4k_0)z^2] \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2m^2 l^2} [E(z^4) - E(z^3\sqrt{4k_0 + z^2}) + 2k_0 E(z^2)] \\ &\dots\dots\dots (36) \end{aligned}$$

(フィデューシャル論によったことを示すため、 $\varphi_p^2$  と記す)

ここで、 $z$  は  $N(0, 1)$  に従うから、

$$E(z^2) = 1, E(z^4) = 3, E(z^{2i-1}) = 0 \quad (i \text{ は自然数}) \dots\dots\dots (37)$$

$$\begin{aligned} E(z^3\sqrt{4k_0 + z^2}) &= \sqrt{4k_0} E\left[\sum_{i=0}^{\infty} \frac{(-1)^i (2i)!!}{i!} \left(\frac{z^2}{4k_0}\right)^i z^3\right] \\ &= 2\sqrt{k_0} \sum_{i=0}^{\infty} \frac{(-1)^i (2i)!!}{i! (4k_0)^i} E(z^{2i+3}) = 0 \end{aligned} \dots\dots\dots (38)$$

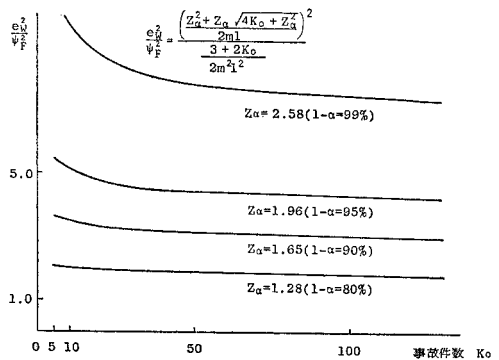
$$\therefore \varphi_p^2 = E(e^{*2}) = \frac{3 + 2k_0}{2m^2 l^2} \dots\dots\dots (39)$$

以上のように、フィデューシャル論によれば、偶然変動の大きさ  $\varphi_p^2$  は走行台キロと、観測された事故件数との関数として表わすことができる。

しかし、本研究では、危険率  $p$  を未知の定数と考え、事故率  $p_0$  を確率変数とする立場をとるものであるから、このフィデューシャル法による偶然変動の大きさの期待値  $\varphi_p^2$  は、前節および式 (32) における  $\varphi^2$  とは論理的に別のものである。フィデューシャル法をそのまま適用すると、以下に示すように、危険率  $p^*$  の期待値が、事故率  $p_0$  と一致せず、式 (32) の仮定が成り立たなくなる。

$$\begin{aligned} E(p^*) &= E\left(p_0 + \frac{z^2 - z\sqrt{4k_0 + z^2}}{2ml}\right) \\ &= p_0 + \frac{1}{2ml} \{E(z^2) - E(z\sqrt{4k_0 + z^2})\} \\ &= p_0 + \frac{1}{2ml} \end{aligned}$$

$$\begin{aligned} &(\because E(z^2) = 1, E(z^{2i+1}) = 0, E(z\sqrt{4k_0 + z^2}) \\ &= \sqrt{4k_0} E\left[\sum_{i=0}^{\infty} \frac{(-1)^i (2i)!!}{i!} \left(\frac{z^2}{4k_0}\right)^i z\right] \\ &= 2\sqrt{k_0} \sum_{i=0}^{\infty} \frac{(-1)^i (2i)!!}{i! (4k_0)^i} E(z^{2i+1}) = 0) \end{aligned}$$



図一 誤差幅とフィデューシャル法による偶然変動の期待値との比

そこで、本研究では、 $\varphi_i^2$  を  $\varphi^2$  の大きさを推定するための参考値として扱い、個々の事故率の偶然誤差の評価は、あくまでも誤差幅  $e_w$  を用いるものとする。

$\varphi_i^2$  と  $e_w^2$  との違いをみるため、比をとって図示したものが図-1である。これによれば、 $k_0$  の小さくない範囲では、 $\varphi_i^2$  と  $e_w^2$  とはほぼ比例することがわかる。

(3) 回帰係数の推定に対する偶然変動の影響

事故率説明のモデルの回帰係数の推定に、通常用いられている最小二乗法を適用することについて検討する。

$$p_0 = X\beta + w \dots\dots\dots (27)$$

において、回帰係数ベクトル  $\beta$  の推定値を  $\hat{\beta}$  とし、残差ベクトルを  $v$  とすると

$$p_0 = X\hat{\beta} + v \dots\dots\dots (40)$$

と書ける。この残差の平方和を最小とする  $\hat{\beta}$  が、求める最小二乗推定量であり、次式で与えられる<sup>9)</sup>。

$$\hat{\beta} = (X'X)^{-1}X'p_0 \dots\dots\dots (41)$$

このとき

$$\begin{aligned} E[\hat{\beta}] &= E[(X'X)^{-1}X'p_0] \\ &= E[(X'X)^{-1}X'(X\beta + w)] \\ &= \beta + (X'X)^{-1}X'E(w) = \beta \dots\dots\dots (42) \end{aligned}$$

( $\because E(w) = 0$  (式 (28)) $\dots\dots\dots (28)$ )

となり、 $\hat{\beta}$  は  $\beta$  の線形不偏推定量であることがわかる。

ここで、攪乱項  $w$  が等分散性  $E(ww') = \sigma^2 I_n$  を仮定できるのであれば、 $\hat{\beta}$  は  $\beta$  のあらゆる線形不偏推定量の中で、最も分散の小さいもの(最良線形不偏推定量)である<sup>9)</sup>。しかし、本モデルの攪乱項  $w$  は、式(31)に示されるように等分散ではないため、最小二乗推定量  $\hat{\beta}$  は最小分散性を保証されない。もし、 $W$  が既知であるならば、一般化最小二乗法が適用でき、 $\beta$  の最良線形不偏推定量  $b$  が次式で与えられる<sup>7)</sup>。

$$b = (X'W^{-1}X)^{-1}X'W^{-1}p_0 \dots\dots\dots (43)$$

ところが、ここでは、 $W$  が未知数であるため、一般化最小二乗法も適用することはできない。しかし式(43)について考察することにより、偶然変動が回帰係数の推定に与える影響を知ることができる。

式(31)より

$$W^{-1} = (\sigma^2 I_n + \Phi_n)^{-1} = \begin{pmatrix} \frac{1}{\sigma^2 + \varphi_1^2} & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma^2 + \varphi_n^2} \end{pmatrix} \dots\dots\dots (44)$$

であるから、一般化最小二乗法によって、最良推定量を求めるためには、事故率データはそれぞれ、方程式のはずれと偶然変動との和の逆数  $1/(\sigma^2 + \varphi_i^2)$  によって重み付けされなければならない。ところが、最小二乗推定法

では、すべてが同じ重みで取り扱われるため、結果的に回帰係数の推定量  $\hat{\beta}$  が、偶然変動  $\varphi_i^2$  の大きいデータに引きずられてしまい、分散の大きい、不安定なものになる。

ここで、 $\hat{\beta}$  の  $b$  に対する相対的な不安定さを示すために  $\hat{\beta}$  の効率  $E_{ff}(\hat{\beta})$  を  $b$  と  $\hat{\beta}$  の一般化分散の比によって、次式のように定義する。

$$E_{ff}(\hat{\beta}) = \frac{|\text{Var}(b)|}{|\text{Var}(\hat{\beta})|} \dots\dots\dots (45)$$

$W$  の固有値を  $r$  で表わし、 $0 \leq r_1 \leq \dots \leq r_n$  とすると、次の Watson の公式が成り立つ<sup>8)</sup>。

$$1 \geq E_{ff}(\hat{\beta}) \geq 4^j \left[ \sqrt{\frac{r_n}{r_1}} + \sqrt{\frac{r_1}{r_n}} \right]^{-2j} \dots\dots (46)$$

( $j$  は  $\beta$  の次数)

事故率の偶然変動のうち最大のものを  $\varphi^2_{\max}$ 、最小のものを  $\varphi^2_{\min}$  とすると

$$r_1 = \sigma^2 + \varphi^2_{\min} \dots\dots\dots (47)$$

$$r_n = \sigma^2 + \varphi^2_{\max} \dots\dots\dots (48)$$

である。式(46)によって与えられる効率の下限値を図示したものが図-2である。これによれば、個々の事故率のもつ偶然変動の大きさをそろえることによって、最小二乗法でも最良推定量  $b$  に近い推定量が得られることがわかる。

(4) 偶然変動の大きさとモデルの説明力

(1) で示した重回帰モデルに対して、(3) で示したように最小二乗法を用いて回帰係数  $\beta$  の推定量  $\hat{\beta}$  を求めると、事故率の偏差平方和  $S_{pp}$  は、回帰による平方和  $S_r$  と残差平方和  $S_e$  とに分解され、モデルの説明力を示す決定係数  $R^2$  は次式で与えられる。

$$S_{pp} = S_r + S_e \dots\dots\dots (49)$$

$$R^2 = S_r / S_{pp} = 1 - S_e / S_{pp} \dots\dots\dots (50)$$

この残差平方和  $S_e$  は次のように展開できる。

$$\begin{aligned} S_e &= v'v \\ &= w' \{ I_n - X(X'X)^{-1}X' \}' \{ I_n - X(X'X)^{-1}X' \} w \end{aligned}$$

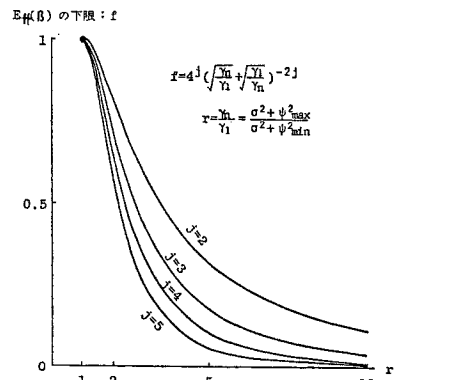


図-2 偶然変動の大きさと回帰係数推定の効率

$$= \mathbf{w}' \{I_n - X(X'X)^{-1}X'\} \mathbf{w}$$

$$= \text{trace}[\{I_n - X(X'X)^{-1}X'\} \mathbf{w}\mathbf{w}'] \dots \dots \dots (51)$$

$$(\because \mathbf{v} = \mathbf{p}_0 - X\hat{\beta} = X\beta + \mathbf{w} - X[(X'X)^{-1}X'(X\beta + \mathbf{w})])$$

$$= (I_n - X(X'X)^{-1}X')\mathbf{w}$$

ここで、残差平方和の期待値を求めると次式のようになる。

$$E(S_e) = \text{trace}[\{I_n - X(X'X)^{-1}X'\} E(\mathbf{w}\mathbf{w}')] \\ = \text{trace}[\{I_n - X(X'X)^{-1}X'\} (\sigma^2 I_n + \Phi_n)] \\ = n\sigma^2 + \sum \varphi_i^2 - [k\sigma^2 + \text{trace}\{X(X'X)^{-1}X'\Phi_n\}] \dots \dots \dots (52)$$

$$(\because \text{trace}\{X(X'X)^{-1}X'\} \\ = \text{trace}(X'X)^{-1}X'X = \text{trace} I_k = k)$$

残差平方和について、この期待値が実現しているものと仮定すれば、決定係数は次のようになる。

$$\hat{R}^2 = 1 - \frac{n\sigma^2 + \sum \varphi_i^2 - [k\sigma^2 + \text{trace}\{X(X'X)^{-1}X'\Phi_n\}]}{S_{pp}} \\ = 1 - \frac{n\sigma^2}{S_{pp}} - \frac{\sum \varphi_i^2}{S_{pp}} + \frac{k\sigma^2 + \text{trace}\{X(X'X)^{-1}X'\Phi_n\}}{S_{pp}} \dots \dots \dots (53)$$

上式の第2項は、モデル式の本来のはずれ部分を示している。第3項は、事故率のもつ偶然変動のために、モデルのもつ説明力が損なわれる部分を示している。そして第4項は、 $k$  個の説明変数を用いるために、自由度が減少して、モデルの説明力が見かけのうえで増加する部分を示している。

上式によれば、説明力の高いモデルを得るためには、単に事故率の偶然変動  $\varphi_i^2$  が小さいだけでなく、全体の偏差平方和  $S_{pp}$  の大きいことが、サンプルの重要な条件であることがわかる。なぜなら、方程式のはずれ  $\sigma^2$  は、本質的にはサンプルの作り方(要素区間の切り方)にかかわらず、一定の値として存在するはずだからである。

全体の偏差平方和  $S_{pp}$  に占める偶然変動の和  $\sum \varphi_i^2$  を偶然変動率  $R_p^2$  と定義する。

$$R_p^2 = \frac{\sum \varphi_i^2}{S_{pp}} \dots \dots \dots (54)$$

このとき

$$1 - R_p^2 = 1 - \frac{\sum \varphi_i^2}{S_{pp}} \dots \dots \dots (55)$$

この式は、式(53)で  $\sigma^2 = 0$  とし、第4項を取り除いたものである。すなわち、 $1 - R_p^2$  は、回帰モデルに方程式のはずれがない場合の、自由度調整済みの決定係数を意味している。つまり、与えられたサンプルで到達できる(到達して意味のある)決定係数の上限を示しているのである。したがって、 $R_p^2$  はサンプル全体の適性を示す一つの指標とすることができる。ただし、実際には  $\varphi^2$  は未知の母数であるので、その代用値として、(2)で示したフィデューシャル法による期待値  $\varphi_p^2$  を用いる

ことにする。

### 5. 事故率の精度の基準

#### (1) 基本的な考え方

各要素区間の事故率の偶然誤差を小さくするためには、式(12)より明らかなように、要素区間長  $l$  を長くとななければならない。しかし、これを長くすればするほど分析対象である標本の大きさ(要素区間数)が小さくなるし、また、個々の要素区間内にさまざまな危険率をもつ異質な小区間が包含されることになって、偏差平方和も小さくなり、分析の目的にそぐわなくなる。したがって、分析上許容される偶然誤差の限界を定め、かつサンプル全体の精度と偏差平方和とを保持するという妥協的な考え方が必要となる。

本章ではまず 3. で定義した誤差幅、あるいは誤差率をもとにして、サンプル全体の精度を保証するような個々の要素区間の事故件数  $k_0$ 、通過車両台数  $m$ 、区間長  $l$  が満足すべき条件についての4種類の基準を提示する。次におのおのの基準の特徴と、実際に適用するうえでの問題点とを考察する。

#### (2) 4つの基準

##### a) 誤差幅基準

すべての要素区間の誤差幅  $e_w$  をある許容誤差幅  $e_{wa}$  以下とする基準である。

すなわち式(12)より

$$e_w = \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2ml} \leq e_{wa} \dots \dots \dots (56)$$

ここで上式を、要素区間での走行台キロ  $M = ml$  に対する制約条件としてみると

$$M \geq \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2e_{wa}} \dots \dots \dots (57)$$

さらに次のように変形できる。

$$M \geq \frac{z_\alpha^2(e_{wa} + p_0)}{e_{wa}^2} \dots \dots \dots (58)$$

(ただし事故率  $p_0 = k_0/M$ )

式(58)は、事故率  $p_0$  が所要の精度を保証されるために必要な走行台キロに対する条件を示している。これを図示したものが、図-3である。この図より明らかなように、高い事故率に対しては、大きな走行台キロが必要となる。

##### b) 誤差率基準

すべての要素区間の誤差率  $e_r$  がある許容誤差率  $e_{ra}$  以下とする基準である。すなわち、式(14)より

$$e_r = \frac{z_\alpha^2 + z_\alpha \sqrt{4k_0 + z_\alpha^2}}{2k_0} \leq e_{ra} \dots \dots \dots (59)$$

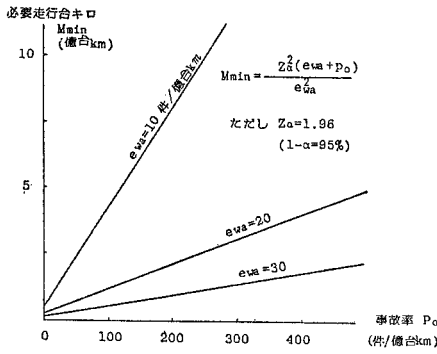


図-3 誤差幅基準での必要走行台キロ

変形すると

$$k_0 \geq \frac{(1+e_{ra})}{e_{ra}^2} z_a^2 \dots\dots\dots (60)$$

この基準では、要素区間内の事故件数だけが問題となる。式(60)は十分な事故件数があれば、走行台キロにかかわらず(すなわち事故率の値にかかわらず)、事故率のもつ相対的な誤差は小さいことを示している。

c) 誤差幅誤差率基準

許容誤差幅  $e_{wa}$  および許容誤差率  $e_{ra}$  を定め、すべての事故率が、誤差幅基準または誤差率基準のいずれかを満足することを要求する基準である。すなわち、式(56)、(59)より、

$$e_w = \frac{z_a^2 + z_a \sqrt{4k_0 + z_a^2}}{2ml} \leq e_{wa}$$

または

$$e_r = \frac{z_a^2 + z_a \sqrt{4k_0 + z_a^2}}{2k_0} \leq e_{ra} \dots\dots\dots (61)$$

あるいは、式(57)、(60)より

$$ml \geq \frac{z_a^2 + z_a \sqrt{4k_0 + z_a^2}}{2e_{wa}} \dots\dots\dots (62)$$

または

$$k_0 \geq \frac{(1+e_{ra})}{e_{ra}^2} z_a^2$$

図-4は、観測される事故率  $p_0$  と危険率  $p$  との関係を示したものである。各基準を満足する要素区間の事故率は信頼係数  $1-\alpha$  で、図中の斜線部に入っていることが、統計的に保証されることになる。

d) 走行台キロ基準

誤差幅あるいは誤差率を用いた上記の基準では、要素区間の切り方がそれぞれの要素区間内の事故件数  $k_0$  に依存することになるが、実際の分析への適用上、事故の分布状況に応じて必要走行台キロ  $ml$  を変化させながら要素区間を区切っていくことは、非常に手数のかかることである。そこでこれを簡素化して、要素区間内の事故件数にかかわらず、得られた事故率が、誤差幅誤差率基準を満足することを保証するような走行台キロの

基準を考える。これは次に示す展開から式(63)で与えられる。

$$k_0 \geq \frac{(1+e_{ra})}{e_{ra}^2} z_a^2 \text{ のとき } e_r \leq e_{ra}$$

(式(60)より、 $M=ml$ にかかわらず成立)

$$k_0 < \frac{(1+e_{ra})}{e_{ra}^2} z_a^2 (=k_c \text{ とおく})$$

のとき、 $e_w \leq e_{wa}$  となる条件を考える。式(57)右辺が  $k_0$  について単調増加であるから、 $k_0 < k_c$  であるすべての  $k_0$  に対して式(57)が成立するための必要十分条件は次のようになる。

$$M=ml \geq \frac{z_a^2 + z_a \sqrt{4k_c + z_a^2}}{2e_{wa}} = \frac{(1+e_{ra})}{e_{wa}e_{ra}} z_a^2 \text{ (} k_c \text{ を代入し整理した)}$$

以上よりすべての自然数  $k_0$  に対して、 $e_w \leq e_{wa}$  または  $e_r \leq e_{ra}$  が成立するための  $M=ml$  は次式で与えられる。

$$M=ml \geq \frac{(1+e_{ra})}{e_{wa}e_{ra}} z_a^2 \dots\dots\dots (63)$$

この基準を用いると、誤差率が許容誤差率と一致するのは、 $k_0 = (1+e_{ra})z_a^2/e_{ra}^2$  のときだけで、それ以外はすべて誤差幅または誤差率は許容値より小さな値となる。したがって同じ  $e_{wa}$ ,  $e_{ra}$  を与えるならば、誤差幅誤差率基準よりも厳しい制約条件となり、その厳しさの程度は、各要素区間内の事故件数  $k_0$  によることになる。

(3) 各基準についての考察

ここで、前章までに検討してきた個々の要素区間、あるいはサンプル全体にとっての望ましい条件を整理すると次のようになる。

- ① 個々の要素区間の内部で、事故発生頻度がほぼ均等であること(要素区間内の均質性)
- ② 事故率の偶然変動の大きさがそろっていること(偶然変動の均一性)
- ③ 事故率の全体の偏差平方和が大きいこと。すなわち、事故率の高い区間と低い区間とがうまく分離され、別々の要素区間となっていること(事故発生状

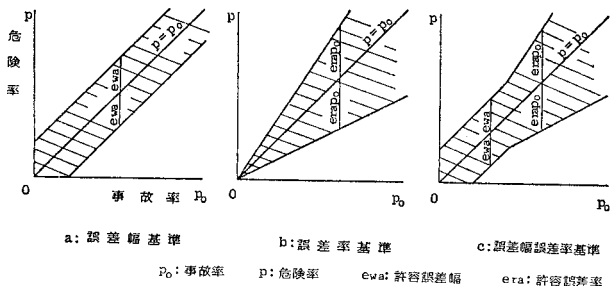


図-4 3種の誤差基準における事故率と危険率

況の弁別性)

これらの条件を念頭において、前節で示した4つの基準について検討する。

まず、誤差幅基準であるが、4.(2)で示したように、誤差幅と偶然誤差の大きさは、ほぼ比例すると考えられる。誤差幅をそろえることは、偶然変動の大きさをそろえることに相当し、②の条件を直接満足するためのものである。しかし、この基準では、高い事故率の精度を保証するためには、より大きな走行台キロが必要であるため、事故多発区間では、要素区間長を長くとらなければならない。ところが、実際には事故多発区間はそれほど長く続いていないため、その前後の事故の少ない区間をも取り入れて、一つの要素区間とせざるを得なくなる。したがって、①の要素区間内の均質性が損なわれるとともに、その事故率も小さくなって、③の弁別性の条件をも満足しないものとなってしまう。

逆に、誤差率基準では、事故件数が一定値以上あることが要件であるため、事故多発区間は、それだけで一つの要素区間として認められる。また事故の少ない区間も、同じように事故の少ない区間を長く続けて事故件数が一定値以上となったところで、一つの要素区間とできるので、①の要素区間内の均質性はおおむね満足することができる。②の偶然変動の均一性の条件については、事故率に対する相対的な偶然変動の大きさだけを制限しているため、満足されないが、著しく大きな偶然変動は排除される。しかし、事故の少ない区間に対しては、要素区間長が長くなりすぎて、事故率の小さい要素区間が少数しか得られなくなる。このため、③の弁別性条件が満足されにくくなる。

誤差幅誤差率基準では、これらの短所を補い合うことができる。すなわち、事故多発区間に対しては誤差率基準によって、事故率が高く、均質性のある要素区間が得られ、事故の少ない区間に対しては、誤差幅基準によって、回帰係数の安定性を高めるために最低限必要な偶然変動に対する条件を満たした均質な要素区間が数多く得られる。ゆえに、この基準では、①および③の条件が

満足される。②の条件については、偶然変動を均一にすることはできないが、著しく大きな偶然変動は排除されている。最後の走行台キロ基準は、誤差幅誤差率基準を簡便化したものであり、本質的な特徴は変わらないが、要素区間内の事故件数によって、誤差幅誤差率が変動する分だけ、偶然誤差に対する制約が均等でなくなっている。

以上の検討から、誤差幅誤差率基準が事故分析にとって現実的に適用しやすい基準であろうと推測できる。

## 6. 適用例（東名・名神高速道路の事故分析）

### (1) 分析対象

分析対象とした道路は、東名・名神高速道路の厚木—西宮間(500.2 km)である。事故資料は昭和45年1月～50年12月の6年間に、この区間の高速道路本線上で発生した事故の中から、次の各項に該当する1040件を除外した28983件である。

① 道路構造との関連を考えるのが不相当と思われる特殊事故、すなわち、車両火災、人の転落、転回・横断・逆行および後退車の事故、酒酔運転による事故、整備不良による事故

② 流入出車線上での事故

③ 路面凍結、積雪時の事故

道路構造データとしては100mポストごとの土工、橋梁、トンネルの構造種別、平面曲線半径、縦断勾配の3者を用いた。交通量については、前記期間の各月別日平均交通量をもとに6年間の合計通過車両台数を求めた。最小2330万台、最大7150万台、平均3850万台であった。

### (2) 比較のためのサンプル

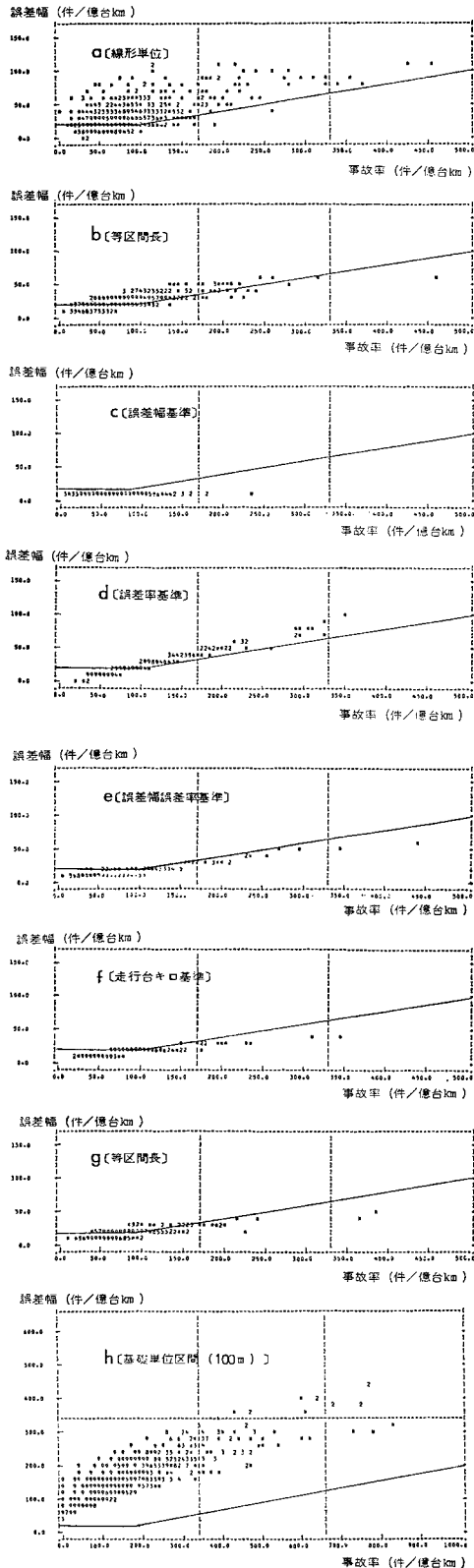
事故率のもつ偶然誤差の具体的な大きさと、これがモデルの説明力に与える影響を調べるため、7種類のサンプルを作った。おのおのサンプルの概要をまとめたもの

表-1 各サンプルの概要

サンプルの種類	要素区間数	区間長(100m単位)			事故率(件/億台km)		偶然変動率* (%)	備 考
		平均	最短	最長	平均	標準偏差		
a 線形単位	822	12.1	2	56	82.7	66.6	7.5	$e_{wa}=20$ 件/億台 km $e_{ra}=30\%$ $e_{wa}=20, e_{ra}=20$ $e_{wa}=25, e_{ra}=25$
b 等区間長(1200m)	833	12.0	12	12	75.3	43.1	9.9	
c 誤差幅基準	455	22.0	7	82	65.7	29.0	1.1	
d 誤差率基準	494	20.1	6	55	94.5	70.4	4.2	
e 誤差幅誤差率基準	467	21.5	7	51	72.8	49.3	3.9	
f 走行台キロ基準	464	21.5	12	34	75.2	39.7	7.0	
g 等区間長(2100m)	475	21.0	21	21	75.3	38.7	6.9	
h 基礎単位(100m)区間	9994	1.0	1	1			51.9	

\* 偶然変動率  $R_p^2 = \frac{\sum \sigma_{\hat{F}}^2}{S_{pp}}$  (4.(4)で定義)





図—5 各サンプルの事故率と誤差幅

が表—1 および 図—5 である。なお、誤差評価のための信頼係数は 95% とした。図—5 には、誤差幅誤差率基準で用いた許容値を実線で示した。

表—1 および 図—5 には、参考のためにデータ収集の基礎単位である 100 m 区間ごとの事故率についても記した。図—5 (h) にみられるように、個々の 100 m 区間の事故率のもつ偶然誤差は非常に大きい。また表—1 にみられるように、偶然変動率  $R_p^2=0.519$  であって、全変動の半分以上が偶然変動である。したがって、これに重回帰モデルを適用しても、得られるモデルは本当に道路の危険率を説明するものとはなりそうにないことがわかる。

7 種類のサンプルは次のようである。

a) 線形単位サンプル

これは、本論文と同じデータを用いた既往の研究<sup>9)</sup>で試みられた方法で、線形的に等質な区間を 1 要素区間としたものである。しかし、図—5 (a) をみると、サンプル全体としてかなり大きな誤差をもっており、これらの事故率が、危険率を適切に反映しているかどうかは疑問である。

b) 等区間長 (1200 m) サンプル

一般に、偶然変動を少なくするため、要素区間の区間長を長くする方法が用いられている。この方法の効果を確かめるため、1200 m を 1 要素区間とした。この区間長は、要素区間数が、a) の線形単位サンプルと同程度になるように決めたものである。図—5 (b) をみると、a) のサンプルと比べれば、全体に誤差が小さく、かつそろっている様子がわかる。

c) 誤差幅基準サンプル

許容誤差幅  $e_{wa}=20$  件/億台キロとし、誤差幅基準を満足するように、対象道路を厚木側から機械的に区切って要素区間とした。この  $e_{wa}$  は比較を容易にするために、要素区間数が e) のサンプルと同程度になるように決めたものである。(3) で述べたように、全体として、事故率の低いサンプルになっている。

d) 誤差率基準サンプル

許容誤差率  $e_{ra}=30\%$  として、c) と同様、機械的に要素区間を設定した。 $e_{ra}$  を 30% としたのは、要素区間数を c) および e) のサンプルと同程度にして、比較を容易にするためである。ただし、あまり短い要素区間をとると、事故が発生した地点と記録された地点(事故車が停止した地点)が別々の区間になってしまうおそれがあるので、着目する区間内で発生した事故のほとんどが、その区間に記録されているとの仮定があまりはずれないように、最小要素区間長を一応 600 m とした。

e) 誤差幅誤差率基準サンプル

本論文で提案する要素区間の切り方である。許容誤差

幅  $e_{wa}=20$  件/億台キロ，許容誤差率  $e_{ra}=20\%$  として，この基準を満足するように，しかも，100 m 区間ごとの事故件数分布をみながら，できるだけ区間内の事故発生頻度が一樣となるように要素区間をとった。

f) 走行台キロ基準サンプル

許容誤差幅  $e_{wa}=25$  件/億台キロ，許容誤差率  $e_{ra}=25\%$  として，e), d) と同様，機械的に要素区間をとった。必要走行台キロ  $ml$  は 0.768 億台キロとなった。この  $e_{wa}, e_{ra}$  は，要素区間数が e) のサンプルと同程度になるように決めたものである。図-5 (f) によれば，誤差幅は e) とほとんど同じか，それ以下に抑えられている。

g) 等区間長 (2100 m) サンプル

要素区間数が，e) のサンプルと同程度になるように区間長を決めた。図-5 (g) によれば，誤差幅は全体にかなり小さくなっており，最も簡単に偶然誤差を抑えるためには，要素区間長を長くそろえる方法が，ある程度有効なことがわかる。

(3) 説明変数

a) 基本的な考え方

従来の研究では，説明変数として要素区間の線形要素の平均値のほかに，要素区間内の特徴的な値 (最急勾配・最大曲率等) が用いられることもある。上流あるいは下流の影響については，まったく考慮されていないもの<sup>9)</sup>もあるし，上下の要素区間の代表値との関係が取り入れられているもの<sup>10)</sup>もある。まず，このような説明変数のとり方について，3. で設定した危険率モデルに即して検討する。

式 (17) に示されるように，ある要素区間  $S$  が  $n$  個の 100 m 基礎単位区間から成っているとき， $S$  の危険率  $p$  は各基礎単位区間の危険率  $p_i$  の平均  $\frac{1}{n} \sum_{i=1}^n p_i$  である。したがって，各  $p_i$  が，上流側および下流側のある範囲における道路特性  $x_i$  によって決まっており， $p_i = x_i \beta + u_i$  という 1 次結合で表わされるとすれば， $S$  の危険率  $p$  について  $p = x \beta + u$  と書くとき， $x$  の各要素  $x_h$  は  $x_i$  の対応する要素  $x_{ih}$  の平均  $\frac{1}{n} \sum_{i=1}^n x_{ih}$  となる。

このことから，各要素区間の説明変数としては，それに含まれる基礎単位区間の道路特性変数の平均値とすべきであることがわかる。この方法は，要素区間内の最大値や最小値を選び出す方法に比べて，実用的にも計算機による処理が容易であり，また説明変数値が要素区間のとり方によって大きく変動しないなどの利点もある。

本ケーススタディでは 100 m 区間ごとに道路構造種別，平面曲率および縦断勾配を用い，それぞれに上流の影響を取り入れた 100 m 区間道路特性変数を定義して，

要素区間の説明変数としては，それに含まれる 100 m 区間道路特性変数の単純平均をとった。下流の影響については考慮しなかった。

b) 用意した説明変数

上流の影響を取り入れる方法として，指数平滑を用いた。着目する第  $i$  番目の 100 m 区間の道路構造要素を  $A_i^0$  とすると，上流の影響を取り入れたこの 100 m 区間道路特性変数  $A_i^T$  は次式で与えられる。

$$A_i^T = \alpha A_i^0 + (1 - \alpha) A_i^{T-1} \dots \dots \dots (64)$$

ここに  $T$  は影響範囲 (影響半減区間) を示す添字であり，

$$(1 - \alpha)^T = \frac{1}{2} \dots \dots \dots (65)$$

となるように  $\alpha$  を設定する ( $T=5$  では 0 m ~ 500 m 上流区間の影響が  $A_i^0$  の 1/2 を占めることになる)。

本論文では，いくつかの線形要素に対して， $T$  の異なる一次説明変数を数種用意し，重回帰分析の変数選択段階で，最も説明力の高いものを選び出せるようにした。

用意した説明変数のうちおもなものは，表-2 に示すとおりである。

(4) 分析結果

上述の計 8 種類のサンプルに対して，偏回帰係数の符合の解釈その他についての吟味を経て選択された 4 つの説明変数によって重回帰分析を行った結果が表-3 である。表-3 には，変数選択の吟味を行わず，変数自動取込みのみによって選ばれた 4 変数による回帰結果も同時に示した。また，このうちの 4 種類のサンプルについて，予測値と実現値を示したものが図-6 である。

誤差幅誤差率基準サンプルに対して，これら 4 変数で重相関係数  $R=0.81$  というかなり高い説明力をもったモデルが得られた。他のサンプルに対する分析でも，これら変数の偏回帰係数は安定しており，また重相関係数も，説明変数に対する吟味を行わずに各サンプルごとに変数自動取込みによって得られた最大の値 (表-3) とあまり変わらず，さらに既往の同様な研究<sup>2)</sup>での値と比べても劣らないものである。したがって，これらの 4 変数は，誤差幅誤差率基準によるサンプルにだけ適しているのではなく，事故率をよく説明する力をもった変数と考えてよいであろう。

ここで取り上げられた 4 つの説明変数について簡単に説明すると次のようである (各変数の定義については表-2 を参照されたい)。

a) 遠心加速度指数 ( $F_0$ )

一定の機関推進力を維持している車両が，縦断勾配によって自然に加減速されながら走行するとき，曲線部で受ける横方向の加速度を示す。これは，上流の縦断勾配

表-2 用意した説明変数のうちのおもなもの

	記号	名称	定義式	影響半減区間 [T]	備考
平面線形	$C_T$	曲率	$C_0^i = 1/R_i$	0, 1, 2, 5	$R_i$ は着目する第 $i$ 番目 100 m 区間での平面曲線半径。
	$CX_T$	下限値つき曲率	$CX_0^i = \text{Max}(1/R_i, 1/1000)$	0, 1, 2, 5	
	$C_T^2$	曲率 II	$C_0^{i2} = 1/R_i^2$	0, 1, 2, 5	
	$CX_T^2$	下限値つき曲率 II	$CX_0^{i2} = \text{Max}(1/R_i^2, 1/1000^2)$	0, 1, 2, 5	
	$\Delta C_T$	曲率変化量	$\Delta C_0^i = 1/R_i - 1/R_{i-1}$	0, 1, 2, 5	
	$L_T$	直線率	$L_0^i = (\text{直線で } 1, \text{ その他で } 0)$	0, 10, 20, 50, 100	
縦断線形	$G_T$	勾配	$G_0^i = \tan \theta$ (登りが正)	0, 1, 2, 5, 10	$RV_i$ は着目する第 $i$ 番目 100 m 区間での縦断曲線半径 (凸部で正)
	$GM_T$	下り勾配	$GM_0^i = \text{Min}(\tan \theta, 0)$	0, 1, 2, 5, 10	
	$VC_T$	縦断曲率	$VC_0^i = 1/RV_i$	0, 1, 2, 5, 10	
	$SG_T$	サグ指数	$SG_0^i = \text{Min}(1/RV_i, 0)$	0, 1, 2, 5, 10	
	$CR_T$	クレスト指数	$CR_0^i = \text{Max}(1/RV_i, 0)$	0, 1, 2, 5, 10	
	$V$	速度指数	$V_i = V_{i-1} + (C_2(V_0^2 - V_{i-1}^2) - \tan \theta_i / C_3) V_{i-1}$		
	$T_T$	トンネル率	$T_0^i = (\text{トンネル部で } 1, \text{ その他で } 0)$	0, 1, 2, 5	
組合せ	$CG_T$	曲率×勾配	$CG_T^i = C_0^i \times G_T^i$	0, 1, 2, 5	平面線形と縦断線形との相乗効果を示す指数
	$CGM_T$	曲率×下り勾配	$CGM_T^i = C_0^i \times GM_T^i$	0, 1, 2, 5	
	$\Delta CG_T$	曲率変化量×勾配	$\Delta CG_T^i = \Delta C_0^i \times G_T^i$	0, 1, 2, 5	
	$\Delta CGM_T$	曲率変化量×下り勾配	$\Delta CGM_T^i = \Delta C_0^i \times GM_T^i$	0, 1, 2, 5	
	$\Delta CGM_T^2$	曲率変化量×下り勾配 II	$\Delta CGM_T^{i2} = \Delta C_0^{i2} \times (GM_T^i)^2$	0, 1, 2, 5	
	$FT$	遠心加速度指数	$FT^i = CX_0^i \times V^2 \times (1 + 0.5 T_0^i)$	0, 1, 2, 5, 10, 20	

表-3 各種サンプルに対する分析結果

	a. 線形単位	b. 等区間長 (1200 m)	c. 誤差幅基準	d. 誤差率基準	e. 誤差幅誤差率基準	f. 走行台キロ基準	g. 等区間長 (2100 m)	
重相関係数	0.676	0.670	0.643	0.721	0.808	0.716	0.710	
偏回帰係数	遠心加速度指数 $F_0$	(0.494) 0.108	(0.519) 0.103	(0.487) 0.094	(0.590) 0.135	(0.598) 0.131	(0.574) 0.106	(0.608) 0.118
	曲率変化量×下り勾配 II $\Delta CGM_0^2$	(0.302) 0.00097	(0.283) 0.00100	(0.259) 0.00140	(0.285) 0.00100	(0.383) 0.00140	(0.238) 0.00100	(0.247) 0.00090
	曲率変化量 $\Delta C_0$	(-0.065) -0.186	(-0.108) -0.256	(-0.130) -0.260	(-0.115) -0.383	(-0.143) -0.408	(-0.087) -0.192	(-0.146) -0.327
注 1)	直線率 $L_{50}$	(0.097) 0.299	(0.143) 0.272	(0.210) 0.255	(0.110) 0.315	(0.162) 0.337	(0.181) 0.272	(0.160) 0.274
	定数	-28.8	-22.9	-21.6	-42.7	-51.7	-26.9	-35.8
変数自動取込みの場合の重相関係数と説明変数注 2)	$R=0.703$ $F_0, \Delta CGM_1^2$ $\Delta CGM_0^2, L_{50}$	$R=0.674$ $F_0, \Delta CGM_1^2$ $L_{50}, \Delta F_0$	$R=0.650$ $F_0, L_{50}$ $\Delta CGM_0^2, CX_0$	$R=0.746$ $F_0, \Delta CGM_1^2$ $F_1, L_{50}$	$R=0.808$ $F_0, \Delta CGM_0^2$ $L_{50}, T_0$	$R=0.723$ $F_0, L_{50}$ $\Delta CGM_0^2, \Delta F_0$	$R=0.711$ $F_0, \Delta CGM_0^2$ $L_{50}, VC_0$	

注 1) 偏回帰係数の欄の ( ) 内は、標準偏回帰係数。

注 2) SPSS パッケージにより変数を自動取込みした結果で説明変数のうち重共線性の高いものも含まれている。この結果は、用意した変数によって重相関係数にしてどの程度まで達成できるかをみて、本論で指定した変数によるモデルの説明力の検討のための一つの目安とすることをおもな目的として併記したものである。

の影響を取り入れ、線形によって車両に加えられる物理的な不安定性を表わす指標として取り上げたものである。この変数は、非常に大きな説明力をもっていることが表-3 に示されている。

b) 曲率変化量×下り勾配 II ( $\Delta CGM_0^2$ )

局地的な線形の厳しさを表わす量である。

c) 直線率 ( $L_{50}$ )

線形要素の中で直線の使われ方の指標である。分析結果から、直線の多いところで、事故率が高くなることが示されている。これは直線が多用されている名神で、東名よりも事故率が高いという従来の知見と一致するものである。

d) 曲率変化量 ( $\Delta C_0$ )

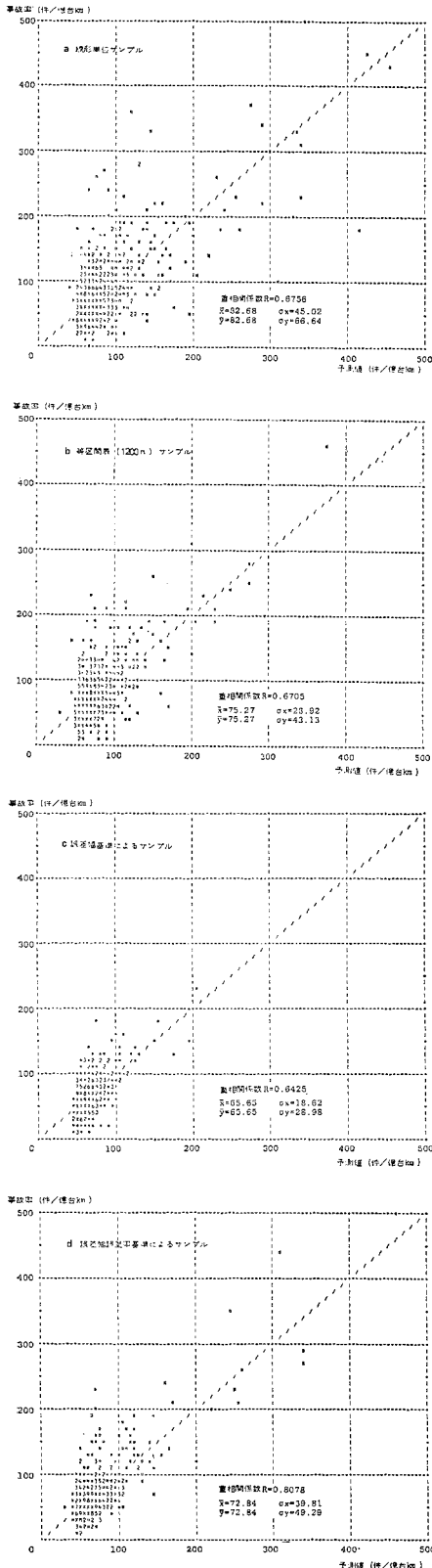
着目する地点に至るまでのハンドル操作の多少を表わ

す指標である。分析結果では、上流でハンドル操作の多かった地点で、おそらくは運転者の緊張度が高まっているために、事故が少なくなることが示された。

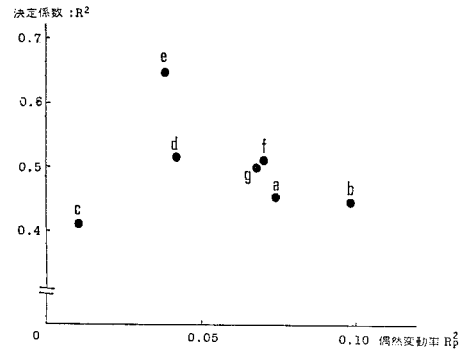
## 7. 考察と結論

これらの説明変数による説明力の大小に基づいて、要素区間の切り方の適否について考察する。図-7 は 7 通りのサンプルについて、4 変数による決定係数 (重相関係数 ( $R$ ) の 2 乗) と式 (54) で定義したサンプルの偶然変動率との関係を示したものである。

線形単位サンプルでは、等区間長サンプルに比べれば、偶然変動率はかなり小さくなっているが、図-5 (b) によると、個々の事故率にはかなり大きな偶然誤差を合



図—6 各サンプルの事故率と計算値



- a : 線形単位
- b : 等区間長 (1200 m)
- c : 誤差幅基準
- d : 誤差率基準
- e : 誤差幅誤差率基準
- f : 走行台キロ基準
- g : 等区間長 (2100 m)

図—7 偶然変動率と決定係数

むものも少なくない。これらが危険率を適切に反映していないため、モデルに十分な説明力が得られないものと考えられる。等区間長 (1200 m) サンプルでは、偶然変動率が高く、全体に偶然誤差が大きいためモデルの説明力は低い。走行台キロ基準および等区間長 (2100 m) サンプルでは、個々の偶然誤差は 図—5 からわかるように、誤差幅誤差率基準と同程度に抑えられている。しかし、表—1 にみられるように、事故率の分散がかなり小さくなってきている。このため偶然変動率があまり小さくならず、モデルの説明力が小さくなってきている。誤差幅基準サンプルでは、偶然変動率が非常に小さいにもかかわらずモデルの説明力は低い。これは、事故多発区間で区間長が著しく長くなり、事故率の高い要素区間を抽出できなくなって、4. (3) で示したように、方程式のはずれに比べて、事故率の分散が小さくなりすぎているためである。誤差率基準サンプルでは、偶然変動率は、誤差幅誤差率基準のそれにほぼ等しく、さらに、事故率の分散はかなり大きい。それにもかかわらず、説明力が低いのは、許容誤差率が大きいため、事故率の高いデータが大きな偶然誤差をもっており、危険率を適切に反映していないことによるのであろう。このように許容誤差率を大きくとらなければならない理由は、小さな許容誤差率では事故の少ない区間に対する制約条件が厳しくなって、その要素区間が長くなり、その結果誤差幅誤差率基準と同程度の要素区間数が得られなくなるからである。

以上から、本研究で検討したいくつかのサンプリングの方法の中では、要素区間内の均質性に配慮しつつ誤差幅誤差率基準に基づいて区切ったサンプリングが優れており、実用的に最も適した方法であるといえる。

これらより本研究をまとめると次のようになる。

本研究では、事故率に含まれる偶然誤差を確率論的に評価する方法を提示し、重回帰分析に対して偶然変動が

与える影響を考察した。事故分析に際しては個々の事故率の精度の検討が必要で、そのための基準として誤差幅誤差率基準が適切であることが知られた。この基準を満足するように要素区間を区切って事故率を求めることにより、確率的な偶然変動に隠された真の危険率に対して分析が可能となるのである。

適用例を通じて、事故率に含まれる偶然誤差の具体的な大きさを示すとともに、同じ説明変数を用いても要素区間のとり方によって重回帰モデルの説明力に大きな差が生じることを示し、本研究で提案した方法が有効であることを実証した。適用例の中で実際の事故率について計算した結果によると、偶然誤差は通常考えられているよりはるかに大きいことが知られた。東名・名神という交通量の多い道路を対象にし、かつ6年という長期間のデータを用いてさえ偶然誤差はこのように大きいのであるから、一般道路で必ずしも十分の交通量がない場合や道路・交通条件の変化があるために短期間のデータしか適用できないような場合には、統計的分析に耐えられないことも十分考えられる。分析にあたっては、分析しよ

うとする精度とデータから得られる精度とについて、十分に検討する必要がある。

適用例中の許容誤差幅  $e_{wa}=20$  件/億台キロ、許容誤差率  $e_{ra}=20\%$  は事故の発生状況および過去の分析例を参考にして適当に選んだものである。分析に際してどのような許容値 ( $e_{wa}$ ,  $e_{ra}$ ) を選ぶのが適切であるか、および許容値の重回帰分析結果に与える影響については今後の検討課題としたい。

#### 参 考 文 献

- 1) 森 尚雄・植松俊夫：交通環境要因から見た事故の統計分析，科学警察研究所報告・交通編，Vol. 8, No. 1, 1967.
- 2) 交通工学研究会：事故要因分析，昭和 50 年 2 月。
- 3)~5) 藤本 照：統計数理の基礎と応用，日刊工業新聞社。
- 6), 7) J. ジョンストン (竹内 啓 ほか訳)：計量経済学の方法 (上) (下)，東洋経済新報社。
- 8) 佐和隆充：計量経済学の基礎，東洋経済新報社。
- 9) 日本道路公団：道路構造に関する高速道路の事故解析 (東名・名神高速道路)，昭和 51 年 3 月。
- 10) 交通工学研究会：事故要因分析，昭和 53 年 2 月。  
(1981.7.24・受付)