IV-108　交通事故率の確率論的検討と事故率説明モデル

茨城県土木部　正会員　岡本 博
東京大学　正会員　鹿島 茂
東京大学　正会員　越 正毅

本研究の概要

交通事故が社会的に問題とされるようになって以来交通事故についての分析・研究は、様々な角度から試みられている。事故と道路構造との関連についての研究もその1つであり、現在までに様々な統計的手法を用いて、事故要因の分析がなされている。しかし、事故の発生が偶然性に支配されるため、データのバラツキが大きく、事故と道路構造とがうまく関連づけられていないのが現状である。本研究では、事故の発生を確率論的にモデル化することによって、事故率のもつ偶然性を定量的に評価する。さらに、偶然性に支配されることが少なく、説明力の高いモデルを得るための道路区間サンプルの作り方についての方法論を提示し、実際のデータを用いて検証する。

事故の確率論的モデル化

長さlの区間サンプルsがあり、この区間の危険率（単位長さ当りの事故発生確率）がpであるとする。m台の車輌がs内を通行するとき、事故発生件数がkである確率$P_r(k|m,l,p)$は、

$$P_r(k|m,l,p) = {}_mC_k(pl)^k(1-pl)^{m-k} \quad \cdots ①$$

で表わされる。交通事故では、pは非常に小さく、mは十分大きい。したがって、①はポアソン分布で近似できる。即ち

$$P_r(k|\lambda) = \frac{\lambda^k}{k!}e^{-\lambda} \quad , \ \lambda = mlp \quad \cdots ②$$

となる。ここで危険率pは②を支配する未知母数となる。

区間sで、m台の通過車があり、k件の事故が発生した場合を考える。sでの危険率pの最尤推定量$\hat{p}$は簡単な計算により

$$\hat{p} = k/ml \quad \cdots ③$$

であることが示される。この右辺は、一般に事故率p_0と呼ばれるものである。換言すれば、事故率p_0は、未知の危険率pの観測値であり、最尤推定量であるとして位置づけられる。他方、信頼区間推定法を用いれば②を正規分布近似することによって、pの信頼係数$1-\alpha$での信頼区間が、次式の{ }内として与えられる。

$$P_r\left\{\frac{k}{ml}+\frac{z_\alpha^2 - z_\alpha\sqrt{4k+z_\alpha^2}}{2ml} < p < \frac{k}{ml}+\frac{z_\alpha^2+z_\alpha\sqrt{4k+z_\alpha^2}}{2ml}\right\}=1-\alpha \quad \cdots ④$$

（z_αは標準正規分布で$F(z_\alpha)=1-\alpha/2$となる値）

④を $p_0 = k/ml$ を用いて変形すると

$$P_r\left\{\frac{z_\alpha^2 - z_\alpha\sqrt{4k+z_\alpha^2}}{2ml} < p-p_0 < \frac{z_\alpha^2+z_\alpha\sqrt{4k+z_\alpha^2}}{2ml}\right\}=1-\alpha \quad \cdots ⑤$$

この⑤式は、危険率pの代わりに観測値（最尤推定量）p_0を用いるときの、観測誤差$e=p-p_0$の信頼限界を与えるものと解釈できる。その信頼上限の方が、信頼下限（の絶対値）より大きいので、前者を事故率p_0に伴う観測誤差eの評価尺度として用いることにし、誤差幅e_wと呼ぶ。また、e_wの相対的な大きさを表わすため、e_wとp_0との比を誤差率e_rと呼ぶ。即ち、

$$e_w = \frac{z_\alpha^2+z_\alpha\sqrt{4k+z_\alpha^2}}{2ml} \quad \cdots ⑥$$

$$e_r = \frac{e_w}{p_0} = \frac{z_\alpha^2+z_\alpha\sqrt{4k+z_\alpha^2}}{2k} \quad \cdots ⑦$$

である。

このようにして、個々の事故率に伴う観測誤差が評価できると、サンプル全体として、事故率の精度をそろえるための基準が考えられる。即ち、観測誤差が次のような条件を満足することを、区間サンプルの要件とするのである。

i) 誤差幅基準　$e_w \leqq e_{wa}$

ii) 誤差率基準　$e_r \leqq e_{ra}$

iii) 誤差幅誤差率基準　$e_w \leqq e_{wa}$ または $e_r \leqq e_{ra}$

ここで、e_{wa}, e_{ra}は分析に必要な精度である。

事故率説明モデル

危険率は道路構造によって説明され、事故率は危険率に観測誤差を加えたものと考えられる。そこで、危険率を媒介概念として、事故率と道路構造との間に次の重回帰モデルが設定できる。

$$P_0 = X\beta + u + t = X\beta + v \qquad \cdots ⑧$$

P_0：事故率

X：道路構造　　β：回帰係数

u：方程式のはずれ　　t：観測誤差

撹乱項vについて、次のように仮定できる。

$$E(v) = \mathbf{0} \qquad \cdots ⑨$$

$$E(vv') = V = \begin{pmatrix} \sigma^2+\varphi_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2+\varphi_n^2 \end{pmatrix}$$

$\sigma^2 = E(u^2)$：方程式のはずれの分散

$\varphi_i^2 = E(t_i)$：サンプルiの観測誤差の分散

各サンプルの観測誤差φ_i^2はそれぞれ異なるのが普通であるから、撹乱項の分散は一様でなくなる。

撹乱項の分散が一様でない場合に、回帰係数の最良線形不偏推定量を求める方法として、一般化最小二乗法（以下GLSと略記する。）がある。これは撹乱項の分散共分散行列が、未知母数αと既知行列Ωを用いて、$\alpha\Omega$と表わされる場合にしか適用できない。したがって、本モデルにGLSを直接適用することはできず、βの推定は、通常の最小二乗法（以下OLSと略記）を用いざるを得ない。けれども、GLSについて考察することにより、OLSで分散の小さいβの推定量を得るためには、各サンプルの観測誤差φ_i^2をそろえるべきであることが示される。一方、フィデューシャル法を用いてφ_i^2を推定すると、t_iがあまり小さくない範囲では、φ_i^2はe_{ri}にほぼ比例することがわかる。したがって、事故率説明モデルで、回帰係数βの推定量の分散を小さくするためには、区間サンプルの事故率の誤差幅がそろっていることが望ましいことがわかる。

ところが、⑤より、事故の多い所では誤差幅をおさえるためには、区間長を長くとらなければならない。そのため、誤差幅基準だけによると、事故の多い所が周辺の事故の少ない区間をも取り込んで1サンプルとなるため、結果的にサンプルとして意味のないものになってしまう。一方、⑥より、t_iが増すと相対的な誤差e_rは減少する。そこで誤差幅誤差率基準を用いることにより、無理のない区間サンプルが得られる。

事故率説明モデルの適用

東名・名神高速道路の西宮～厚木間（500.2km）で、昭和45年～昭和50年に発生した事故を対象として、事故率説明モデルによる重回帰分析を行った。区間サンプルは、誤差幅誤差率基準に従い、しかも区間内の事故発生状況がなるべく一様になるように、対象区間を区切って作った。又、これと比較するために、他の切り方によるサンプルを用意し、同じ説明変数を用いて分析した。表1は、サンプルの種類と分析結果の重相関係数である。誤差幅誤差率基準によるサンプルAの分析結果が図1である。表1を見ると、Aでの説明力が高いのは、説明変数のためではなく、区間サンプルの切り方が適切であったためであることがわかる。

	サンプル数	区間長[L]（100m）	4変数での重相関係数
誤差幅誤差率基準	467	21.5（7.0～51.0）	0.8078
誤差幅基準	455	22.0（7.0～82.0）	0.6425
誤差率基準	494	20.1（6.0～55.0）	0.7209
等区間長 (1)	475	21	0.7101
線形単位	822	12.1（2.0～56.0）	0.6756
等区間長 (2)	833	12	0.6705

表　1

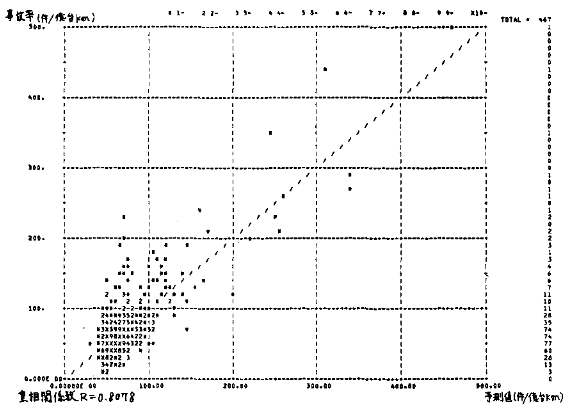

図　1

結論

事故分析において、事故率に伴う観測誤差をある範囲に抑えることにより、説明力の高いモデルが作れることがわかった。今後の同様な解析に対して、有効な方法論を提示することができたと考える。