

疫学研究と生物統計：現状と展望

佐藤 俊哉

疫学研究で用いられている生物統計手法のレビューを行った結果のうち、疾病発生リスクの回帰モデルを発表者の今後の研究課題として報告した。

個人の疾病発生リスクのモデル化として、generalized linear modelと同様にgeneralized risk regression model,

$$g[r(x, z)] = \alpha + \beta x + \gamma z,$$

を考えよう。ただし、 x は原因と想定されている変数（リスク要因、例えば、喫煙：1、非喫煙：0）、 z は直接研究の目的ではないが調整を必要とする変数の組（交絡変数、性、年齢など）、 $r(x, z)$ は (x, z) を持つ個人が一定期間中に疾病を発生する確率（リスク）、 g はlink functionである。我々に興味があるのはリスク要因の効果を表す β の推定である。疫学研究でもっともよく使われているモデルは、logit linkを用いたロジスティックモデル、

$$\log \frac{r(x, z)}{1-r(x, z)} = \alpha + \beta x + \gamma z,$$

である。ロジスティックモデルでは、1) リスク要因や疾病の診断に誤分類がある場合、2) 重要な交絡変数を見落としてしまった場合、3) モデルが誤っていた場合、などで β の推定にどのような影響があるか、が研究されている。しかし、ロジスティックモデルは必ずしも疫学的に意味のあるモデルではなく、むしろadditive risk regression

$$r(x, z) = \alpha + \beta x + \gamma z,$$

やrelative risk regression,

$$\log r(x, z) = \alpha + \beta x + \gamma z,$$

などの方が解釈も容易で自然なモデルである。残念ながらこれらの回帰モデルに関する研究はほとんどない。一つには疾病発生リスクは $0 \leq r(x, z) \leq 1$ でなければならないがadditive risk, relative risk regressionでは必ずしもこの条件を満たさないという問題があげられる。通常ロジスティックモデルはrelative risk regressionの近似として用いられるが、 β をoverestimateしてしまうsparse-dataの例を紹介した。このようなsparse-dataにおけるadditive risk, relative risk regressionの有効な推定方法を開発するためには、Liang and Zeger (1986)のgeneralized estimating equationアプローチをlogit link以外のlink function (identity or log link)に拡張する必要があることを述べた。

参 考 文 献

Liang, K.-Y. and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models, *Biometrika*, **73**, 13-22.

A New Look of Social Survey

吉野 諒三

近年、筆者が開発した計量心理学的モデルGeneralized High Threshold (GHT) modelと、その分析結果の多次元表示法Vectorial Angular Representation (VAR)を、日本人の国民性意識や、国民性の国際比較に関連する社会調査データに適用し、分析を進めた。その結果として特に、日本人の国民性