一般化平均を用いた
コックス比例ハザードモデルの拡張

総合研究大学院大学　複合科学研究科　　大前　勝弘
統計数理研究所　　　数理・推論研究系　　江口　真透

生存時間の確率変数を T，p 次元の共変量ベクトルを X とし，条件付きハザード関数を $h(t|X)$ とする．このとき，相対リスクモデルは

$$h(t|X,\theta) = h_0(t)\exp\left(F(X;\theta)\right) \tag{1}$$

と書き表せられる．ここで，h_0 は基準ハザード関数，F は予測関数，θ は予測関数のパラメータである．特に，F が線形予測関数の場合に (1) 式は Cox 比例ハザードモデルと呼ばれる．準線形予測関数を

$$F_\tau(X;\alpha,\beta) = \frac{1}{\tau}\log\left(\sum_{k=1}^{K}\exp\left(\alpha_k + \tau\beta_k^\top X\right)\right) \tag{2}$$

で定義する．ただし，$K \in \mathbb{N}$ はクラスタサイズ，$\tau \in \mathbb{R}$ はチューニングパラメータであり，$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_K)^\top \in \mathbb{R}^p$，$\beta = (\beta_1^\top, \beta_2^\top, \cdots, \beta_K^\top)^\top \in \mathbb{R}^{pK}$ は各クラスタに対応する回帰パラメータからなるパラメータベクトルである．(2) 式は一般化平均の特別な場合であり，τ を 0 に近づけた場合に線形予測関数に帰着する [1].

(2) 式を予測関数に用いた相対リスクモデル，すなわち，

$$\begin{aligned} h(t|X,\theta) &= h_0(t)\exp\left(F_\tau(X;\alpha,\beta)\right) \\ &= h_0(t)\left(\sum_{k=1}^{K}\exp\left(\alpha_k + \tau\beta_k^\top X\right)\right)^{\frac{1}{\tau}} \end{aligned} \tag{3}$$

を考える．特に，$\tau = 1$ の場合は混合ハザードモデル [2] の特別な場合とみなせる．(3) 式はパラメータに関して識別不可能であるため，適切な制限を導入する必要がある．このため，生存時間データにおける (3) 式のモデルに対する対数部分尤度関数を $l(\theta)$ としたとき，目的関数として

$$P(\theta) = l(\theta) - \lambda\sum_{\ell\neq m}\sum_{j=1}^{p}|\beta_{\ell j}\beta_{mj}| \tag{4}$$

を考える．本発表では，(3) 式のハザードモデルに関する解析的な考察と，(4) 式による適切なパラメータ学習法およびその応用について議論する．また，シミュレーションおよび実データ解析を通じて，準線形判別関数により拡張されたコックス比例ハザードモデルの性能を評価する．

参考文献

[1] Omae, K., Komori, O. and Eguchi, S. (2017) Quasi-linear score for capturing heterogeneous structure in biomarkers. *BMC Bioinformatics.* 18:308
[2] Louzada-Neto, F., Mazucheli, L. and Achcar, J.A. (2002) Mixture hazard models for lifetime data. it Biometrical journal. 44:1 pp3-14.