# 混合ハザードモデル

京都大学大学院　医学研究科　　　　　大前　勝弘
統計数理研究所　数理・推論研究系　　江口　真透

あるイベントのハザードとそれを増減させる要因について考える．被験者が同時にリスクを増減させる要因を有する場合，それらの影響は必ずしも互いに打ち消し合うとは限らない．したがって，単一の線形予測子を用いたハザードの回帰問題を考慮するだけでは不十分な可能性がある．さらに，ハザードと共変量との関係は，異なる被験者では共通ではないかもしれない．例えば，近年の臨床医学の問題では，そのような集団間の異質性が同じ治療に対して異なる反応をもたらし得るという事実にしばしば焦点を当てている．このような複雑な関係は，もはや単一の線形関係で記述することはできない。

本発表では，このような問題に対して，線形予測子を拡張した準線形予測子 [1]

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

による Cox 回帰モデルの拡張を考える．ただし，$K$ はクラスタサイズ，$\tau\in\mathbb{R}$ は調整パラメータであり，$\pi=(\pi_1,\pi_2,\cdots,\pi_K)^\top$ は $\sum_{k=1}^{K}\pi_k=1$ を満たす混合パラメータベクトル，$\beta=(\beta_1^\top,\beta_2^\top,\cdots,\beta_K^\top)^\top\in\mathbb{R}^{pK}$ は各クラスタに対応する回帰係数の組からなるパラメータベクトルである．すなわち，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，基準ハザード関数を $h_0$ とした場合に，条件付きハザードモデル

$$h(t|X,\theta)=h_0(t)\exp\left(F(X;\theta)\right) \tag{2}$$

におけるパラメータ $\theta$ の予測関数 $F(X;\theta)$ に関して，線形予測子から準線形予測子 $F(X;\theta)=F_\tau(X;\pi,\beta)$ への拡張を考える．この拡張により得られる準線形 Cox 回帰モデルは，ハザード関数のモデルにおける混合表現を備える．この特別な場合として考察される混合ハザードモデル $h(t)=\sum_{k=1}^{K}\pi_k h_k(t)$ は，生存時間に混合分布を仮定した混合分布モデル [2] $f(t)=\sum_{k=1}^{K}\tilde{\pi}_k f_k(t)$ とは別種の混合表現を表していることがわかる．ただし，$f_k$ は混合コンポネントごとの確率（密度）関数，$\tilde{\pi}_k$ はその混合比である．実際，混合ハザードモデルと混合分布モデルから導出される生存関数 $S(t),\tilde{S}(t)$ は各コンポネントに対応する生存関数 $S_k(t),\tilde{S}_k(t)$ と次のような関係式で結ばれる：

$$\log S(t) = \sum_{k=1}^{K}\pi_k\log S_k(t) \tag{3}$$

$$\tilde{S}(t) = \sum_{k=1}^{K}\pi_k\tilde{S}_k(t) \tag{4}$$

本発表では，これら 2 つの混合モデルの違いを明らかにしながら，準線形コックス回帰モデルおよびそのパラメータ推定に関する理論的性質やシミュレーションを通じたその性能評価について報告する．

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] You, N., He, S., Wang, X., Zhu, J. and Zhang, H. (2018), Subtype classification and heterogeneous prognosis model construction in precision medicine. *Biometrics.* doi:10.1111/biom.12843