# 解の存在確率を考慮した経験尤度

東京大・情報理工　坂東 拓馬
東京大・情報理工　清 智也

経験尤度は [1], [2] で紹介されたノンパラメトリックな統計的推論手法であり, 様々な分野で応用されている．しかしながら, 経験尤度によって構成した信頼領域の被覆確率は, 有意水準で定めたレベルを下回る傾向にあり, 小標本, 高次元, 歪度の大きな分布からサンプリングされた状況などにおいて, それは特に顕著に現れる．つまり, 帰無仮説 $H_0 : \theta = \theta_0$ に関して, $\alpha$ を有意水準, $C_\alpha$ を経験尤度で構成した信頼領域とすると, 多くの場合

$$P(\theta_0 \in C_\alpha) < 1 - \alpha \tag{1}$$

となる．この問題の軽減のために, 本発表では, 従来の経験尤度に特別な仮定を加えることなく被覆確率の近似精度を向上する方法を提案する．
経験尤度を用いた平均の検定において, データとその平均 $\mu$ から経験尤度比統計量 $\ell(\mu)$ が計算可能であり, カイ二乗分布の $1-\alpha$ パーセンタイルを $r_\alpha$ とすると, 帰無仮説 $H_0 : \mu = \mu_0$ は $\ell(\mu_0) \geq r_\alpha$ の時に棄却される．したがって, 経験尤度の $1-\alpha$ 信頼領域は

$$C_\alpha = \{\mu \mid \ell(\mu) < r_\alpha\} \tag{2}$$

と表せる．ここで閾値 $r_\alpha$ を選択する際に, 緩やかな仮定の下で $\ell(\mu_0)$ が漸近的にカイ二乗分布に従うという, [2] で示された定理を用いたが, 上記の問題 (1) はこのカイ二乗近似の精度があまり良くないことを示唆している．というのも, サンプルサイズが有限であれば, $\ell(\mu_0)$ が存在しない確率は常に正の値を取り, これが上記の問題 (1) の部分的な要因となっている．漸近的には $\ell(\mu_0)$ が存在しない確率は 0 となるので, 定理の主張は当然正しいが, 実際にサンプルサイズが有限の下で経験尤度の信頼領域を構成する場合には注意が必要である．本発表では, $\ell(\mu_0)$ が存在しない確率が被覆確率の近似精度に影響を与えることを示し, それを考慮した新たな閾値 $r'_\alpha$ を選択する方法を紹介する．そして, $r'_\alpha$ に対応する信頼領域

$$C'_\alpha = \{\mu \mid \ell(\mu) < r'_\alpha\}$$

が (2) よりも大きくなっていることを示した後, 被覆確率の向上を数値例で紹介する．最後にこの提案手法の, 経験尤度に纏わる様々な推定問題や手法への応用について議論する．

## 参考文献

[1] Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75** 237–249.

[2] Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120.