(c_1, c_2, γ) - 汚染の下での正規分布の平均のロバスト検定

南山大学理工学研究科　安 藤 周 平
南山大学理工学部　木 村 美 善

1. はじめに

ロバスト推測論においては, 標本の「汚れ」やモデル分布からの「ずれ」をモデルの分布の近傍を用いて記述することが多い. こうした近傍の 1 つに, Kakiuchi and Kimura (2012) によって提案された (c_1, c_2, γ)-近傍がある. この近傍は 3 つのパラメータの値を変えることで多様な近傍になり, よく知られている ε-汚染近傍や全変動近傍など, 有用な多くの近傍をその特殊な場合として含むものである. 本報告では, (c_1, c_2, γ)-近傍を正規分布の平均のロバスト検定問題に適用し, 2 つの近傍間の Radon-Nykodym 導関数を導出するとともに, これを用いたミニマックス検定のシミュレーション結果について述べる.

2. ロバスト検定問題

$X_1, \cdots, X_n$ を正規分布 $N(\mu, \sigma^2)$ に従う互いに独立な大きさ n の標本とし, σ^2 は既知とする. このとき, $N(\mu, \sigma^2)$ の (c_1, c_2, γ) -近傍を $\mathcal{P}_{c_1,c_2,\gamma}(\mu)$ とし, n 個の標本 $\mathbf{X} = (X_1, \cdots, X_n)$ に基づく平均 μ に関するロバスト検定問題

$$H_0: \ \mathcal{L}(\mathbf{X}) \in \mathcal{P}^n_{c_1,c_2,\gamma}(\mu_0) \quad vs. \quad H_1: \ \mathcal{L}(\mathbf{X}) \in \mathcal{P}^n_{c_1,c_2,\gamma}(\mu), \quad (\mu_0 < \mu_1 \leq \mu) \tag{1}$$

を考える. ここで $\mathcal{L}(\mathbf{X})$ は $\mathbf{X} = (X_1, \cdots, X_n)$ の確率分布であり, $\mathcal{P}^n_{c_1,c_2,\gamma}(\mu)$ は $\mathcal{P}_{c_1,c_2,\gamma}(\mu)$ の n 重積を表す. この検定問題に対する検定 φ の最大の大きさ $\alpha_\varphi(\mu_0)$ と最小検出力 $\beta_\varphi(\mu)$ は

$$\begin{aligned} \alpha_\varphi(\mu_0) &= \sup\{E_{G_n}[\varphi(\mathbf{X})] : G_n \in \mathcal{P}^n_{c_1,c_2,\gamma}(\mu_0)\} \\ \beta_\varphi(\mu) &= \inf\{E_{G_n}[\varphi(\mathbf{X})] : G_n \in \mathcal{P}^n_{c_1,c_2,\gamma}(\mu)\} \end{aligned}$$

によって定義される このとき, 問題は $\alpha_\varphi(\mu_0) \leq \alpha$ $(0 < \alpha < 1)$ を満たす水準 α 検定 φ の中で最小検出力 $\beta_\varphi(\mu)$ を最大にするミニマックス検定を求めることである.

3. 結果

$(c1, c2, \gamma)$-近傍 $\mathcal{P}_{c_1,c_2,\gamma}(\mu)$ は, c_1, c_2, γ によって定まる特殊容量 v_μ により

$$\mathcal{P}_{c_1,c_2,\gamma}(\mu) \quad = \{G \in \mathcal{M} \mid G(A) \leq v_\mu(A), \quad \forall A \in \mathcal{B}\}$$

とも表される. ここで, $\mathcal{B}$ は実数直線 $\mathcal{R}$ 上のボレル集合族, $\mathcal{M}$ は $(\mathcal{R}, \mathcal{B})$ 上の確率測度の全体からなる集合である. また, $(\mathcal{P}_{c_1,c_2,\gamma}(\mu_0),\ \mathcal{P}_{c_1,c_2,\gamma}(\mu))$ には, 最も不利な分布の対 $(Q_0,\ Q_\mu)$, すなわち, $Q_0 \in \mathcal{P}_{c_1,c_2,\gamma}(\mu_0)$, $Q_\mu \in \mathcal{P}_{c_1,c_2,\gamma}(\mu)$ であり, $\forall t \geq 0$ に対して

$$\begin{aligned} Q_0(\pi > t) &= \sup\{P(\pi > t) \ : \ P \in \mathcal{P}_{c_1,c_2,\gamma}(\mu_0)\} = v_0(\pi > t) \\ Q_\mu(\pi \leq t) &= \sup\{P(\pi \leq t) \ : \ P \in \mathcal{P}_{c_1,c_2,\gamma}(\mu)\} = v_\mu(\pi \leq t) \end{aligned}$$

を満たす分布の対が存在する. ここで, π は Q_μ の Q_0 に対する Radon-Nykodym 導関数であり, v_μ の v_0 に対する Radon-Nykodym 導関数でもある. このとき, (1) に対する水準 α ミニマックス検定は

$$\varphi^*(\mathbf{x}) = \begin{cases} 0 & \text{if} \quad \prod_{i=1}^n \pi(x_i) \ < \ \lambda_\alpha \\ \xi & \text{if} \quad \prod_{i=1}^n \pi(x_i) \ = \ \lambda_\alpha \\ 1 & \text{if} \quad \prod_{i=1}^n \pi(x_i) \ > \ \lambda_\alpha \end{cases} \tag{2}$$

により与えられる. ただし, $E_{Q_0^n}[\varphi^*(\mathbf{X})] = \alpha$ である.

π とミニマックス検定 φ^* の具体的な式, および, このミニマックス検定と一様最強力検定などのパラメトリック検定のシミュレーションによる検出力評価の詳細については当日報告する

参考文献

Kakiuchi, I. and Kimura, M. (2012). Robust nonparametric inference for the median under a new neighborhood of distributions. Technical Report of Nanzan Academic society.