

二標本ノンパラメトリック検定の連続化と有意確率の近似について

九州大学大学院数理学府 森山 卓
九州大学数理学研究院 前園 宜彦

1. 序

X_1, \dots, X_m を $F(x)$, Y_1, \dots, Y_n を $F(x - \theta)$ からの無作為標本とし, 帰無仮説 $H_0 : \theta = 0$ vs. 対立仮説 $H_1 : \theta > 0$ の二標本検定問題を考える. ノンパラメトリックな検定としてはメディアン検定 M やウィルコクソンの順位和検定 W_2 がよく利用されている. M は F が両側指数分布の場合の局所最強力順位検定統計量に漸近的に一致し, W_2 は F がロジスティック分布の場合の局所最強力順位検定統計量であり, 両検定統計量は理論的にもその“よさ”が保証されている.

しかし一標本検定である符号検定とウィルコクソンの符号付き順位検定に対する先行研究と同様に, 両検定間にはやや一方的な有意確率の大小の差が生じる場合があり, これにより両検定は恣意的に選択されうる問題がある. 具体的には特に標本数が小さい場合に同じ data set に対して W_2 を利用する方が M よりも有意確率は小さくなるが多くなってしまふ. さらに, 両検定統計量は離散型分布を持つためこれらは確率化検定になることから検定結果には恣意が入る余地がある. 以上の問題を検定統計量の連続化により解決を図った.

2. ノンパラメトリック検定の連続化

本講演ではカーネル型推定量を利用した, 連続化検定統計量を提案しその性質について議論する. M , W_2 の連続化として次を提案する:

$$\widetilde{M} = \frac{1}{n} \sum_{j=1}^n W^* \left(\frac{Z - Y_j}{h_N} \right) - \frac{r+1}{m+n}, \quad \widetilde{W}_2 = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n W \left(\frac{X_i - Y_j}{h_N} \right) - \frac{1}{2}$$

但し $W(t) = \int_{-\infty}^t k(u) du$ とし (\widetilde{M} に対しては $W^*(u) = 1 (u \geq 1)$ を満たす非対称カーネル k^* を選ぶ), $Z = Z_{(r+1)}$ ($Z_{(k)}$ は $(X_1, \dots, X_m, Y_1, \dots, Y_n)$ に対する順序統計量), また $N = m + n$ に対して N が奇数のとき $r+1 = \frac{N+1}{2}$ であり, N が偶数のとき $r+1 = \frac{N+2}{2}$ とする.

$h_n \rightarrow 0$ のとき $W^*(v/h_N) \rightarrow I(v \geq 0)$ (但し I は定義関数) であり, 標準化した \widetilde{M} と M は漸近的に等しく, \widetilde{W}_2 と W_2 も同様である. さらに Pitman A.R.E. も一致し, この意味でも \widetilde{M} , \widetilde{W}_2 は検定の性質を変えていない. また, 連続化検定統計量は共に distribution-free でなくなるが, 適当な条件の下で漸近正規性が成り立つ. さらにいくつかの条件の下, \widetilde{M} については Bai and Rao (1991), \widetilde{W}_2 は Maesono (1985), Bickel, Götze and van Zwet (1986) の議論を利用すると, 有意確率のより精密な評価を与えるエッジワース展開の validity を示すことができる. この展開を利用した有意確率の近似を用いると, 先の有意確率の大小の問題はシミュレーションにより大幅に改善されることが分かった.

連続化統計量 \widetilde{M} , \widetilde{W}_2 の具体的な評価, またエッジワース展開とその validity, シミュレーション結果については当日の講演で紹介する.

3. 参考文献

- [1] P. J. Bickel, F. Götze and W. R. van Zwet. (1986). The edgeworth expansion for U-statistics of degree two. *Ann. Statist.*, **14**, 1463-1484.
- [2] Maesono, Y. (1985). Edgeworth expansion for two-sample U-statistics. *Rep. Fac. Sci. Kagoshima Univ.*, (Math., Phys. & Chem.), **18**, 35-43.
- [3] Bai, Z. D. and Rao, C. R. (1991). Edgeworth expansion of a function of sample means. *Ann. Statist.*, **19**, 1295-1315.