確率変数としての p 値の挙動

ファイザー株式会社　河合統介・相澤愛奈・吉松宏樹

臨床試験の検証的な場面では p 値は重要な役割を果たし，事前に設定した有意水準よりも p 値が小さいか否かで統計的な有意性が判断される。一方，p 値は検定統計量から導出される統計量であるにも関わらず，得られた p 値のみに焦点が当たり，統計量の確率変数としての側面が軽視されがちである。帰無仮説の下で p 値は[0, 1]区間の一様分布に従うことは良く知られており[1]，本発表では，対立仮説の下での p 値の確率変数としての挙動を評価するとともに，p 値とエフェクト・サイズとの関係についても検討する。

仮説検定の帰無仮説と対立仮説をそれぞれ H_0 と H_1 とする。H_0 と H_1 の下での検定統計量の分布関数をそれぞれ F_0 と F_1 とすると，H_1 の下での p 値の分布関数は，その定義から

$$F_P(p|\mathrm{H}_1) = 1 - F_1(F_0^{-1}(1-p))$$

と表わされる[2,3]。両辺を微分することにより，p 値の確率密度関数

$$f_P(p|\mathrm{H}_1) = \frac{f_1(F_0^{-1}(1-p))}{f_0(F_0^{-1}(1-p))}$$

を得る[4]。ここで，f_0 と f_1 はそれぞれ H_0 と H_1 の下での検定統計量の確率密度関数である。

いま，正規分布に従う連続な評価項目の平均を 2 群間で比較する臨床試験の場面を想定し，試験開始前に有意水準α（片側検定），検出力$1-\beta$で標本サイズを設計したとする。このとき，2 群の平均を比較する仮説検定の p 値の H_1 の下での分布関数は

$$F_P(p|\mathrm{H}_1) = 1 - \Phi(z_{1-p} - (z_{1-\alpha} + z_{1-\beta}))$$

となり，確率密度関数は

$$f_P(p|\mathrm{H}_1) = \frac{\phi(z_{1-p} - (z_{1-\alpha} + z_{1-\beta}))}{\phi(z_{1-p})}$$

となる。ここで，Φとϕは標準正規分布の分布関数と確率密度関数，z_γは標準正規分布の$100\gamma\%$点である。H_1 の下での p 値の中央値と期待値はそれぞれ$\Phi\left(-(z_{1-\alpha} + z_{1-\beta})\right)$，$\Phi(-\left(z_{1-\alpha} + z_{1-\beta}\right)/\sqrt{2})$となる。より詳細な検討は発表当日に報告する。

参考文献

1. Hung, H.M.J., O’Neill, R.T., Bauer, P., and Kohne, K. (1997). The behavior of the p-value when the alternative hypothesis is true. *Biometrics*, 53, 11-22.
2. Sackrowitz, H., and Samuel-Cahn, E. (1999). P values as random variables - Expected p values. *The American Statistician*, 53, 326-331.
3. Bhattacharya, B., and Habtzghid, D. (2002). Median of the p value under the alternative hypothesis. *The American Statistician*, 56, 201-206.
4. Donahue R.M.J. (1999). A note on information seldom reported via the p value. *The American Statistician*, 53, 303-306.