

医薬学、生物学分野のための計量科学的手法の開発



研究ノート

高木 達也*

Development of Statistical and Machine-Learning Methods for
Medical and Biological Sciences

Key Words : Bootstrap, Multiple Comparison, Pharmacometrics, Resampling Method

はじめに

私たちの研究室では、主として、医薬学分野や生物学分野で必要とされる統計学的な、或いは、数理科学的な手法の開発、改良、熟成を行っています。今回は、主として新規多重比較法の解説をさせていただきます。

1. 多重比較の必要性

皆さんの研究室に、誕生日が同じ組み合わせはありますか？ と訪ねると、「365日もある誕生日が同じ組み合わせなんて、そんなにあるものじゃない」と仰るかもしれません。では、実際に研究室のメンバーの誕生日を調べてみてください。結構な確率で、発見できるはずです。

具体的な簡単に検証できる例として、プロスポーツのチームの中に同じ誕生日を持つ選手の組み合わせがあるかどうか探してみます。横浜FマリノスのWebページで確認してみますと、飯倉大樹選手とW. N. Lopes選手の誕生日が、6月1日で同じ、浦和レッズはいませんでした。ジュビロ磐田では、上田康太選手と西紀寛選手が5月9日で、3チーム調べたうちの2チームで、誕生日が同じ選手の組み合わせが存在しました。35人が1チームに存在すると、81.4%の確率で誕生日が同じ選手の組み合わせが存在するはずですので、2/3は、まずまず理論

通りと思われます（「誕生日のパラドックス」として有名な例です）。

このように、「稀な現象」が、たびたび観測されるのは、観測される組み合わせが多いため、俗っぽく言えば、「下手な鉄砲数打ちゃ当たる」という現象が起こっているためですが、下手な鉄砲でなく、科学的な実験結果による結論を左右するとなると、問題となります。例えば、表1のような場合、従来薬とA,B,両方の新薬との効果の比較は、1つ1つStudentのt検定を行えば、表のごとく、5%水準で有意な差が出てきます。

表1 多重比較をしなければならないデータの例

	1	2	3	4	5	6	平均	標準偏差	危険率
従来薬	0.56	0.58	0.42	0.62	0.34	0.59	0.52	0.112	
新薬A	0.78	0.58	0.74	0.69	0.52	0.62	0.66	0.099	0.0489*
新薬B	0.69	0.59	0.53	0.74	0.62	0.70	0.65	0.079	0.0485*
新薬C	0.58	0.61	0.56	0.70	0.64	0.79	0.65	0.086	0.0497*
新薬D	0.64	0.54	0.78	0.56	0.74	0.66	0.65	0.095	0.0480*

* $P < 0.05$

しかし、先の誕生日のパラドックス例を思い出してください。1人では1/365の事象（同じ誕生日の人がいる）でも、 ${}_{25}C_2=300$ 通りも観察すれば、半分近い確率で同じ誕生日の人に遭遇することになります。表1の場合も、「新薬Aが、本当は効果がないのに効果があると判断する確率」が4.9%なのですが、「新薬AまたはBが本当は効果がないのにどちらか一方が効果があると判定してしまう確率」は、それぞれ4.9%ではなく、もっと高い確率になります。通常このような場合、「多重比較」と言って、何回も繰り返すことによって「多くのものを同時に比較することによって」高くなる偶然の要素を加味した検定法を行う必要があります。



*Tatsuya TAKAGI

1956年5月生

大阪大学薬学部 (1979年)

現在、大阪大学大学院薬学研究科 生命

情報環境科学専攻 微生物生態学講座

微生物動態学分野 教授 薬学博士 計

量薬学、数理環境科学、計算薬学

TEL : 06-6879-8240

FAX : 06-6879-8243

E-mail : ttakagi@phs.osaka-u.ac.jp

2. 多重比較法の例～Bonferroni 法

何かいかめしい名前が付いていますが、極めて簡単な方法です。一般に、 α %の危険率で、 n 個のペアを比較したい場合は、 α' %、

$$\alpha' = \alpha / n \quad (1)$$

を実質の危険率として検定を行えば、上記のような間違いを冒す危険はなくなります。これが、Bonferroni 法です。

Bonferroni 法は、基本的にはほとんどすべての多重比較に応用できる汎用的な方法ですが、4つくらいの比較ならともかく、多数の比較を同時に行う場合は、不都合な点がでてきます。例えば、20 個の比較を同時にしようとしてみると、 $\alpha' = 5\% / 20 = 0.25\%$ で、0.25 %の危険率を使って検定する必要があることになり、たいていの比較は「有意でない」という結論で終わってしまいます。Bonferroni 法は実質の危険率を必要以上に低く見積もっているのです。

そこで、もう少し検出力の高い方法として、Dunnett 法や Tukey 法が登場します。しかし、世の中うまい話はそうはなく、これらの方法には、いくつかの制限があります。例えば、両者は平均値の差の検定にしか利用できませんし、データの母集団が正規分布に従っている必要があります。そこで当研究室では、Dunnett 法などの検出力を持ち、かつ、Bonferroni 法のような汎用性を持つ新規多重比較法の開発に取り組みました。

3. 当研究室で開発された、新規多重比較法の紹介

ここでは、新規多重比較法の骨格として、以下の2点を考えました。

- 1) Bonferroni 法のように、実質の P 値を求める。こうすれば、原則としてどのような比較（分散比や比率の差など）であっても、用いることができる。
- 2) 計算機集約型手法を用いて、できる限り必要十分な実質の P 値を求める。Dunnett 法などと、検出力の点で見劣りがしないようにする。

当研究室で開発した手法のアルゴリズムは図1のようです¹⁾。

簡単のために、平均値の差の検定を例にあげます。最初に、帰無仮説 ($H_0: \mu_1 = \mu_i$) が成立する条件に基づいた（平均値に差がないという前提で）、

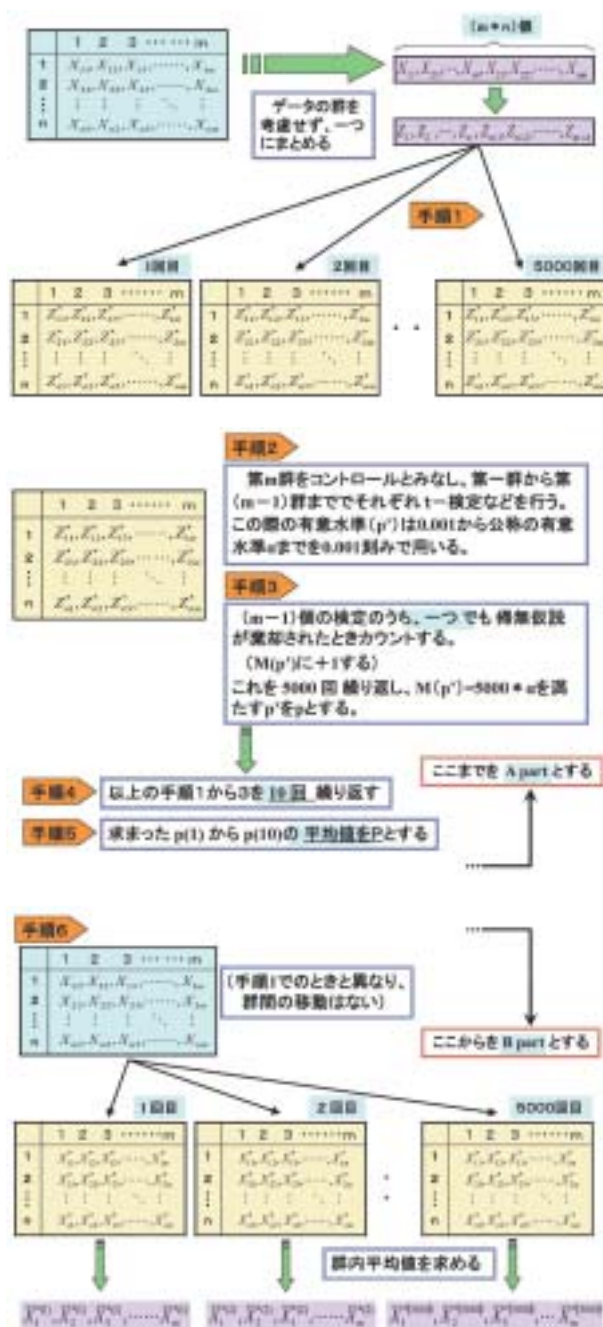


図1 新規多重比較法のアルゴリズム

Bootstrap 法を用いて、複数個の標本を作成します。Bootstrap 法は、分子進化系統樹を作成する際に用いられるのと、概念的には同じものです。ここでは、データの「所属」をバラバラにしてしまっても、標本の再抽出を行います。そもそも平均値に差がないという前提ですから、バラバラにしてしまっても、問題はないはずです。そのうえで、各データ群の平均値を求め、有意水準を少しずつ変えながら、Welch の t 検定（場合によっては、Wilcoxon の順位和検定）

を行います。前提が前提ですから、 N_{repeat} 回このようなことを繰り返せば、複数のデータ群の比較のうち、1組でも帰無仮説が棄却される回数、 N_{reject} は、用いた有意水準、 p_a の関数になるはずで

$$N_{\text{reject}} = N_{\text{repeat}} f(p_a) \quad (2)$$

設定したい公称の有意水準を p_n (例えば 0.05) とすると、

$$p_n = \frac{N_{\text{reject}}}{N_{\text{repeat}}} \quad (3)$$

を満たす N_{reject} を与える p_a が、公称の有意水準 p_n を満たす、必要な実質の有意水準となります。後は、これを用いて、通常の検定を行えばいいことになります。

4. 他の、当研究室で開発された、機械学習法

当研究室で開発した手法は、多重比較法だけではありません。他にも数多くありますので、一部を簡単に紹介いたします。

4.1 決定木法における記述子の重要性の解析²⁾

決定木などの機械学習法は、学習効率に優れている反面、例えば教師つき学習では、記述子が重要ななどの情報に乏しく、要因解析法としては今ひとつの感がありました。そこで、当研究室では、Bootstrap 法の概念を流用して、決定木法で記述子の予測に対する重要性の解析を行うことに成功しました。簡単に言えば、すべての記述子を用いた場合の Bootstrap サンプルの予測性と、どれか一つ記述子を削除した場合の予測性との比較を行い、あまり予測性に変化がないのであれば、その記述子は重要でないということになり、予測性が格段に低下すれば、

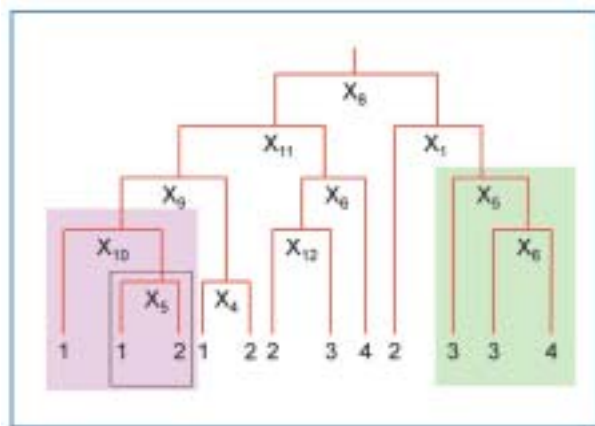


図2 決定木 (二分木) の例。

予測になくなくてはならない記述子ということになります。小児悪性脳腫瘍の予後予測に対する、要因の重要性解析の結果を、図3に示しました。

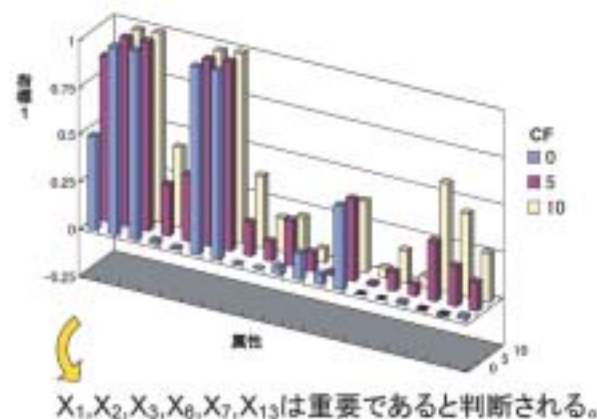


図3 小児悪性脳腫瘍の予後を左右する因子の決定木及び新規アルゴリズムによる解析。
X1-X3: 年齢、X6, X7: Grade Code、X13: 原始神経外胚葉性腫瘍。

4.2 ロジスティックPLS回帰を用いた医薬品の3次元定量的構造活性相関解析(3D-QSAR)³⁾

3-DQSARは、以前より行われてきました。しかしながら、キーポイントとなるPLS回帰は、応答変数に計量データを想定しているため、活性値が順序カテゴリーデータで与えられるような医薬品に関しては、適用に制限がありました。当研究室では、Logistic PLSRを用いたCoMFA⁴⁾法により、順序カテゴリーデータでも適切に3-DQSAR解析が行えるようにしました。

以上のように、当研究室では化学、医薬学に用いる、様々な計量科学的手法の開発、改良、熟成を行っています。結果は、差し支えない限り、やはり当研究室開発になる、WWWベースの統計的手法のプログラムパッケージ、MEPHAS⁵⁾で公開する予定にしています。

参考文献

- 1) T. Takagi, K. Okamoto, Y. Yokogawa, M. Yokota, K. Kurokawa, and T. Yasunaga, *J. Comput. Aided Chem.*, 4, 27-34 (2003).
- 2) T. Takagi, K. Okamoto, M. Yokota, and T. Yasunaga, *J. Comput. Aided Chem.*, 5, 35-46 (2004).

3) T. Ohgaru, R. Shimizu, K. Okamoto, M. Kawase, Y. Shirakuni, R. Nishikiori, and T. Takagi, *J. Chem. Inf. Modeling*, **48** (1), 207 - 212 (2008).

4) R. D. Cramer, D. E. Patterson, and J. D. Bunce, *J. Am. Chem. Soc.*, **110**, 5959-5967 (1988).

5) <http://www.gen-info.osaka-u.ac.jp/MEPHAS/>

