

2002 年 01 月 11 日

分割表の最小度数が 5 より小さい場合の解析方法について
(2001-2BC 講義「計数値データの解析 ①」での質問表より)

文責：稲葉太一（神戸大学）

0. 質問内容

- ① 母不良率の違いの検定以外に、不良率の違いを検証する手法はあるのでしょうか。
② 検定の性質上、「不良率に違いがあるとはいえない」とは言えても、「不良率は同じである」とは言えないと思いますが、「不良率は同じである」といえる手法はあるのでしょうか。

表 1。データ表

	良品	不良品	合計
抜取	47	3	50
全数	1784	16	1800

(全数には、抜取結果も含みます。)

1. 質問 ① についての回答

まず、1 つ目の質問についてですが、色々あります。まず、分割表 (BC テキスト第 5 章 5.7 節「分割表による解析」) で計算し、検定することができます。表 1 を以下の表 2 のように表現すると分かり易いので、慣例として表 2 の表現が採用されています。

表 2。データ表

	良品	不良品	合計
抜取	47	3	50
その他	1737	13	1750
全数	1784	16	1800

ここで、BC テキスト p.5-68 によると

$$\chi_0^2 = \frac{(ad - bc)^2 T}{T_1 T_2 T_{\cdot 1} T_{\cdot 2}} = \frac{(47 \times 13 - 3 \times 1737)^2 \times 1800}{50 \times 1750 \times 1784 \times 16} = 15.25$$

$$> \chi^2(1, 0.05) = 3.84$$

となるので、有意水準 5% で有意差ありという結論になります。ただし、この方法は 2 群のデータでは、BC テキスト第 5 章で紹介されている母不良率の差の検定の「直接法」と常に同じ結論となります。つまり、抜取での不良品数が「3」であるので近似条件を (完全には) 満たしておらず、近似が良くないという状況には変わりありません。したがって、結論に若干の不安が残ります。

次に、近似を使わずに確率を直接求める「フィッシャーの直接確率法」という方法を紹介します。資料を添付しましたので、参考にしてください。（「毒性・薬効データの統計解析」吉村功編著，サイティスト社，p.76-77）今回のデータは、この解析で良いでしょう。

この手法の考え方は、周辺度数が与えられたとして、今回の度数がどの程度珍しいかを、超幾何分布の確率を用いて求めるやり方です。「並べ替え検定」とも呼ばれています。ただし、階乗の値が大きいのので以下のような若干の工夫が必要です。

具体的には、以下のように P_1, P_2, P_3, P_4 を求めて、 $P_2 + P_3 + P_4 = 0.9912$ であることから、求める確率は $1 - 0.9912 = 0.0088 < 0.05$ と分かり、行と列は独立ではない、即ち、有意水準 1% でも、抜取検査の結果と全体の結果には有意差があると結論されます。

$$\begin{aligned} P_1 &= \frac{f_{1.}! f_{2.}! f_{.1}! f_{.2}!}{f_{11}! f_{12}! f_{21}! f_{22}! n!} = \frac{50! 1750! 1784! 16!}{47! 3! 1737! 13! 1800!} = \frac{1750!}{1737!} \times \frac{1784!}{1800!} \times \frac{50!}{47!} \times \frac{16!}{13! 3!} \\ &= \frac{1750 P_{13} \times 50 \times 49 \times 48 \times 16 \times 15 \times 14}{1800 P_{16} \times 3!} \\ &= \frac{1750 \times 1749 \times \cdots \times 1738 \times 50 \times 49 \times 48 \times 16 \times 15 \times 14}{1800 \times 1799 \times \cdots \times 1785 \times 3!} = 0.0080 \end{aligned}$$

ここで、添付資料では $f_{11} = 47 < e_{11} = \frac{50 \times 1784}{1800} = 49.56$ のときは、 f_{11} を $f_{11} - 1$ 等と置き換えて P_2 を計算し、式中に $0!$ が出てくるまで計算して和を求めるように指示されているが、これだと $f_{22} = 13$ が 1 つずつ減って行って 0 になるまでの 14 通りを計算して合計しなければなりません。それに対して、今回は $f_{12} = 3$ であるので、求める事象の残りの事象（排反事象）の確率を求めるのが楽です。そこで、指示されているのと反対の方向に置き換えて計算することにします。

$$\begin{aligned} P_2 &= \frac{f_{1.}! f_{2.}! f_{.1}! f_{.2}!}{f_{11}! f_{12}! f_{21}! f_{22}! n!} = \frac{50! 1750! 1784! 16!}{48! 2! 1736! 14! 1800!} = \frac{1750!}{1736!} \times \frac{1784!}{1800!} \times \frac{50!}{48!} \times \frac{16!}{14! 2!} \\ &= \frac{1750 \times 1749 \times \cdots \times 1737 \times 50 \times 49 \times 16 \times 15}{1800 \times 1799 \times \cdots \times 1785 \times 2!} = 0.0621 \end{aligned}$$

同様に、 P_3, P_4 も求める。

$$\begin{aligned} P_3 &= \frac{50! 1750! 1784! 16!}{49! 1! 1735! 15! 1800!} = \frac{1750!}{1735!} \times \frac{1784!}{1800!} \times \frac{50!}{49!} \times \frac{16!}{15! 1!} \\ &= \frac{1750 \times 1749 \times \cdots \times 1736 \times 50 \times 16}{1800 \times 1799 \times \cdots \times 1785 \times 1!} = 0.2932 \\ P_4 &= \frac{50! 1750! 1784! 16!}{50! 0! 1734! 16! 1800!} = \frac{1750!}{1734!} \times \frac{1784!}{1800!} \times \frac{50!}{50!} \times \frac{16!}{16!} \\ &= \frac{1750 \times 1749 \times \cdots \times 1735}{1800 \times 1799 \times \cdots \times 1785} = 0.6359 \end{aligned}$$

2. 質問 ② についての回答

おっしゃる通り、「母平均が等しい」ことを示す方法はありません。ただ、薬の新製品の許認可の段階で「生物学的同等性の検定」という手法があります。新薬が副作用が少ないとか患者の苦痛が少ないといった利点がある時には、従来からある最も良く効く薬（標準薬）と同程度の効き目があれば十分だろうという考えです。

これについて、先ほど紹介した本「毒性・薬効データの統計解析」の p.220 ~ 226 の次のような記述があります。ご参考までに転記します。

「新製剤の平均値が標準製剤の 20% ずれたときの検出力が 80% 以上となる試験で、有意差がなければ、BE（生物学的同等性）を認める。」（厚生省の研究班）