

臨床試験における標本数の定め方

——医学統計学最近の話題 (1)*

柳井晴夫 高木廣文

はじめに

臨床医学、公衆衛生学、疫学などの諸分野においては、統計的な因果関係を立証する1つの手段として、いわゆる4分表(2×2分割表)の検定が従来から頻繁に使用されている。ところで、4分表の作成とその検定は一見単純明解にみえるが、実際はそうではなく対立仮説の設定などの統計学上の問題、さらには薬効検定におけるケースコントロール群の設定や疫学調査における前向き調査(prospective study)や後向き調査(retrospective study)においてえられる4分表の解釈上の問題など、さまざまな複雑な問題点を内包しており、サリドマイドの服用の有無と奇形児出現の有無に関するレンツの4分表に対する誤った解釈⁵⁾にみられるような4分表にまつわる誤用例は少なくない。

本稿では、これまで見落されがちであった4分表に関するもう1つの問題点である標本数の定め方に焦点をあて、2つの母集団における出現率の差を検出するのに必要とみなされる標本数をどのように定めるべきかについて考察してみることにする。

標本数の設定は本来すべての臨床試験において、はじめから考慮しなければならない問題であるが、実際には標本数を先に定めてから臨床試験を行う例は現在ではむしろまれである。したがって本稿では標本数の算出の考え方とその実際を実例を加えつつ解説しようとするものである。

1. 各種の検定方式と標本数の関連

ここで表1を眺めよう。これは架空データであるが、40歳代と50歳代の年齢階層別に喫煙量と肺癌発症の有無の関連についての4分表を示したものである。独立性の検定の統計量として知られるカイ2乗値(χ^2)はともに1%水準で有意であるから、喫煙量と肺癌発症の因果関係はある程度肯定されることになる。

しかし、 χ^2 値が大きいからといっても、50歳代の標本数が40歳代に比べて2倍になっていることから、喫煙量と肺癌発症の関連の強さは40歳代より50歳代の方が強いということにはならず、後述するように表1のデータによると、むしろ40歳代の関連の方が強くなる。

表2のような一般の4分表における2つの因子AとBの関連の強さはカイ2乗値ではなく、

$$\phi = (ad - bc) / \sqrt{(a+b)(c+d)(a+c)(b+d)} \dots (1)$$

で定義されるファイ相関係数、すなわちピアソンの積率相関係数(通常相関係数といわれるもの)において、

表1 喫煙量と肺癌発症の4分表(架空データ)

		(1) 40歳代		(2) 50歳代	
		肺癌	正常	肺癌	正常
喫煙量	20本以上/1日	80	400	190	710
	20本未満/1日	20	200	70	430
		$\chi^2 = 7.07$ (1%有意)		$\chi^2 = 10.75$ (1%有意)	

* Determination of the sample size in clinical testing — Current topics on medical statistics (1)

YANAI, Haruo: Department of psychology, Chiba University (Department of epidemiology, faculty of Medicine Tokyo University) 千葉大学人文学部 (東京大学医学部疫学教室)

TAKAGI, Hirofumi: Department of epidemiology, faculty of Medicine Tokyo University 東京大学医学部疫学教室

注：本稿は医学統計集談会における発表の結果をまとめたものである。医学統計集談会は医学統計に関する理論や実例を自由に討議することを目的とした会で、連絡先は次のとおりである。開原成允(東大病院 Tel. 815-5411 内8126)または豊川裕之(東大疫学 Tel. 812-2111 内 7717)

表2 一般の4分表

A \ B	1	2	計
1	a	b	a + b
2	c	d	c + d
計	a + c	b + d	N = a + b + c + d

2つの変数が1または0しかとらない場合の相関係数に対応するもので表される。 χ^2 値と ϕ 値には、標本総数を $N=(a+b+c+d)$ とおけば

$$\chi^2 = N\phi^2 \dots\dots\dots(2)$$

の関係が成立する。したがって同一の ϕ の大きさをもつ2つの4分表において、標本数が大きいほど χ^2 値は大きくなる(表1において、40歳代の場合 $\phi=0.100$ 、50歳代の場合 $\phi=0.088$ で相関は40歳代の方が高い)。つまり、 χ^2 値で示されるような統計的検定における有意性とは、与えられた標本数 N のもとで、母集団においてAとBの関連がまったくないということを、どの程度の確率で棄却できるかを示すものである。次に標本数 N と検定の有意性についてのいくつかの例を示そう。

1) 独立性の検定

(1)式で計算されるファイ相関係数 ϕ が0.4と一定の場合

- ① $N=20$ ならば $\chi^2=3.2$ (有意差なし)
- ② $N=30$ ならば $\chi^2=4.8$ (5%で有意)
- ③ $N=45$ ならば $\chi^2=7.2$ (1%で有意)

一般に ϕ が与えられている場合、5%、1%水準で有意差を検出するために必要とされる標本数はそれぞれ、 $3.84/\phi^2$ 、 $6.84/\phi^2$ となる。

2) 母相関係数の検定

一般に標本数 N で計算される標本相関係数 r_{xy} から、母相関係数 $\rho_{xy}=0$ を検定するには

$$t = \sqrt{N-2} \cdot r_{xy} / \sqrt{1-r_{xy}^2} \dots\dots\dots(3)$$

が自由度 $(N-2)$ の t 分布にしたがうことが用いられる。ここで、 $r_{xy}=0.60$ の場合

- ① $N=6$ ならば $t=1.50$ (有意差なし)
- ② $N=11$ ならば $t=2.25$ (5%で有意)
- ③ $N=18$ ならば $t=3.00$ (1%で有意)

となる。

3) 2群の平均値の差の検定

2群の分散 σ_1^2 、 σ_2^2 が等しい場合に、2群の標本平均を $\bar{X}_1$ 、 $\bar{X}_2$ 、標本分散を S_1^2 、 S_2^2 とすれば

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2}}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \dots\dots\dots(4)$$

は自由度 $(n_1 + n_2 - 2)$ の t 分布となる。ここで $n_1 = n_2 = N$

とおくと上式は

$$t = \sqrt{N-1} (\bar{X}_1 - \bar{X}_2) / \sqrt{S_1^2 + S_2^2} \dots\dots\dots(5)$$

となり、 $\bar{X}_1=16.4$ 、 $\bar{X}_2=15.6$ 、 $S_1^2=S_2^2=4$ の場合

- ① $N=25$ ならば $t=1.38$ (有意差なし)
- ② $N=50$ ならば $t=1.98$ (5%で有意)
- ③ $N=100$ ならば $t=2.56$ (1%で有意)

のように標本数の大きさによって有意水準の大きさが異なってくる。

以上述べたことは、ときおり誤って解釈されるもっとも基本的な検定の原理である。すなわち、有意差の有り無しは、医学データの背後にある本質を示すというよりは、むしろ標本数の大きさに依存するものであるといえよう。

このことを逆に考えると、(2)、(3)、(4)式で示される検定公式の有意性は、ファイ相関係数、ピアソンの標本積率相関係数 r_{xy} 、2群の平均値 $\bar{X}_1$ 、 $\bar{X}_2$ 、2群の標本比率 $\hat{p}_1$ 、 $\hat{p}_2$ が与えられている場合にどのくらいの大きさの標本数 n を集めたらよいかを示すことにあるといえよう。

なお t 分布の5%や1%の有意水準を与える限界棄却値は、自由度によって異なるが、各群の標本数 N がほぼ50以上、すなわち2群の標本数が100を越える場合には t 分布はほぼ標準正規分布に一致するから、5%水準であれば、 $t \geq 1.96$ すなわち

$$N \geq \{(S_1^2 + S_2^2) / (\bar{X}_1 - \bar{X}_2)^2\} \times 1.96^2 \dots\dots\dots(6)$$

を満たす標本数を各群ごとに集めればよい。

4) 母比率の差の検定

2つの母集団における母比率 p_1 、 p_2 の同等性、すなわち $p_1=p_2=p$ を検定する方式は、 p_1 、 p_2 、 p の推定値を標本比率 $\hat{p}_1$ 、 $\hat{p}_2$ 、 $\hat{p}$ とした場合

$$Z = (\hat{p}_1 - \hat{p}_2) / \sqrt{\hat{p}(1-\hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \dots\dots\dots(7)$$

(n_1 、 n_2 は2群の標本数 $\hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$) が、 n_1 、 n_2 が

十分大きい場合に標準正規分布 $N(0, 1)$ にしたがうことを用いて検定できる。

したがって $n_1=n_2=N$ であれば(7)式は

$$Z = \sqrt{N} (\hat{p}_1 - \hat{p}_2) / \sqrt{2\hat{p}(1-\hat{p})} \dots\dots\dots(8)$$

となる。ここで、 $\hat{p}_1=0.20$ 、 $\hat{p}_2=0.30$ の場合

- ① $N=100$ ならば $Z=1.69$ (有意差なし)
- ② $N=140$ ならば $Z=2.00$ (5%で有意)
- ③ $N=250$ ならば $Z=2.67$ (1%で有意)

となる。ここで有意水準を α とおくと、(8)式より標本数 N は

$$N \geq Z_{\alpha/2}^2 \hat{p}(1-\hat{p}) / (\hat{p}_1 - \hat{p}_2)^2 \dots\dots\dots(9)$$

となる。 $\alpha=0.05$ (5%水準) のとき $Z_{0.025}=1.96$, $\alpha=0.01$ (1%水準) のとき $Z_{0.005}=2.57$ より, それぞれの群において必要とされる最低の標本数は5%水準で $N=145$, 1%水準で $N=248$ となる。

なお, 本節で示した母比率の差の検定と1)で示した独立性の検定は, 実質的には同等となることに注意しよう。すなわち表2においてAの1と2を2つの母集団からの標本とし, それぞれの集団が因子Bの1を保有する標本比率を $\hat{p}_1=a/(a+c)$, $\hat{p}_2=b/(b+d)$ とする。ここで(8)式に $n_1=a+c$, $n_2=b+d$ を代入すると

$$Z = \frac{\sqrt{N}(ad-bc)}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \sqrt{N}\phi \dots\dots(10)$$

という関係が成立し, Z の平方が(2)式の χ^2 値に一致する。したがって母比率の差の検定で有意差がみられる場合には, 必ず独立性の検定においても有意差がみられることになるし, 理論的にも Z が標準正規分布 $N(0, 1)$ にしたがえば, $Y=Z^2$ は自由度1の χ^2 分布にしたがう。

5) 母比率が既知の場合の検定

次に4)の場合に一方の標本比率が既知の場合, すなわち標本比率を有する N 個の標本データが既知の p_0 という母比率を有する母集団からの標本と考えてよいのか否かを検定する場合を考察しよう。この場合, N が十分大ければ

$$Z = (\hat{p} - p_0) / \sqrt{\hat{p}(1-\hat{p})/N} \dots\dots\dots(11)$$

が $N(0, 1)$ にしたがうと考えてよい。したがって標本比率が $\hat{p}=X/N$ と表される場合, 有意水準を α とすれば, $p(|Z| \leq Z_{\alpha/2}) = 1-\alpha$ となる N を求めるには,

$|Z|^2 \leq Z_{\alpha/2}^2$ を変形して整頓すると

$$N \geq \frac{2X + Z_{\alpha/2}^2(1-p_0) + Z_{\alpha/2}\sqrt{4(1-p_0)X + Z_{\alpha/2}^2}}{(1-p_0)^2/2p_0} \dots\dots\dots(12)$$

となる。(×は標本中の該当者数)

〔例1〕

わが国において, ある疾患の有病率が $p_0=0.01$ で既知であったとしよう。ここで, その疾患の有病率がほぼ0とみられる国で N 人に調査を実施し, 1人もその疾患

に患っている人が見つからなければ, 日本とその国における有病率に有意差が判定できるものとする。この場合の N は $X=0$ を(12)式に代入してえられるもので, $N=380$ となる。もし, 1人だけその疾患に患っている人が見つければ(12)式に $X=1$ を代入してえられる $N=582$ が調査すべき最小の標本数となる。

2. 臨床試験における標本数の決め方

ところで, これまでに述べてきた検定の方式によって, たとえば4)の2群の母比率の差の検定において, 2群の標本数が等しければ ④ $p_1=0.2$, $p_2=0.3$ の場合 $N \geq 145$, ⑤ $p_1=0.2$, $p_2=0.25$ の場合 $N \geq 536$, ⑥ $p_1=0.2$, $p_2=0.21$ の場合 $N \geq 12,522$ となり, 母比率の差が小さくなるにつれて有意差検出に必要とされる標本数は著しく増加する。しかし, 十分に大きな標本を用いて有意差を検出できることがわかって, 多数の標本を集めにくい通常の医学研究には, ほとんど役にたつことはない。ところが, 通常の臨床試験などにおいては母比率の差が実験前にあらかじめ推定または規定できる場合がある。たとえば, ある薬に対し有効率の差が0.1以上あったときにはじめて新しい薬が意味があるとしようというような場合である。こうしたときには, この差を既知と考えてこの差を検出するに必要とされる最小のサンプル数を決めることができる。これが標本数決定の基本的な考え方である。

この考え方を理解するには, 第2種の過誤について理解することが必要である。

通常の検定においては第1種の誤差(仮説が正しい場合に仮説を棄却する誤まり)のみが強調化されて, 第2種の誤差(仮説が正しくない場合に仮説を選択する誤まり)の存在は軽視されがちであるが, ここで検定の基本原理にいま一度立ちかえってみることにしよう。

一般に検定する帰無仮説を

$H: \mu = \mu_0$ (母比率 μ_0 の検定と考えればよい)

これに対して, ある1つの対立仮説を

$AH: \mu = \mu_1 (\mu_1 > \mu_0)$

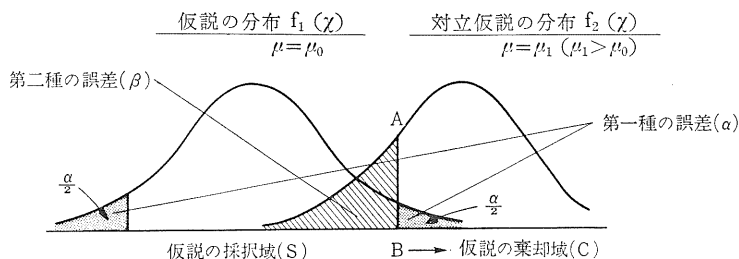


図1 単純仮説の検定

とし、仮説の棄却域を図1のCと定めると、第1種の誤差、第2種の誤差は図中の α 、 β となる（第1種の誤差 α は両側検定として考えるので、 $\frac{\alpha}{2}$ の2倍となる）。ここで対立仮説のもとでの分布 $f_2(x)$ の棄却域Cに相当する面積、すなわち

$$p(x \in C) = \int_C f_2(x) dx = 1 - \beta$$

は仮説の検出力（discriminating power）とよばれるもので、仮説が正しくない場合に仮説を棄却する確率の大きさを示すものとなる。

平均値の差の検定や母比率の差の検定においては、母平均や母比率の相等性が仮説として取り上げられるから、検出力の大きさは重要な意味をもつ。たとえば、これまでの治療法にかわって、ある種の新しい治療法が開発された場合に、新治療法に効果がみられる（治癒率が相等しくない）場合の2つの方法の治癒率の差は検出力として表現されるべきものである。

ところで、一般の検定においては、対立仮説が明確に定義されず、したがって多くの場合第2種の誤差 β の大きさは定めにくい。Cohen²⁾によると、 β は α の約4倍程度、すなわち $\alpha=0.05$ であれば $\beta=0.20$ 、 $\alpha=0.01$ であれば $\beta=0.04$ に定めるのが望ましいとされている。

次に、4)で示した2つの母集団における母比率 p_1 、 p_2 の差を検出する場合に、第2種の誤差 β を考慮した場合の標本総数 N の定め方について説明しよう¹⁾。

母比率 p_1 と p_2 が相等しいという帰無仮説のもとでは標本比率の差 $\hat{d}=\hat{p}_1-\hat{p}_2$ の期待値と分散はそれぞれ0、 $pq(\frac{1}{n_1}+\frac{1}{n_2})$ となる。

したがって、有意水準 α のもとでの $\hat{d}$ の限界棄却域は

$$\hat{d}=\hat{p}_1-\hat{p}_2 \geq Z_{\alpha/2} \sqrt{pq \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad (13)$$

となる。次に、対立仮説として $\hat{d}$ の期待値を $p_1-p_2>0$ と仮定する。この場合、 $\hat{d}$ の分散は $(p_1q_1)/n_1+(p_2q_2)/n_2$ となるから（ただし $q_1=1-p_1$ 、 $q_2=1-p_2$ ）

$$Z_d = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{(p_1q_1)/n_1 + (p_2q_2)/n_2}} \quad (14)$$

は n_1 、 n_2 が十分に大きければ $N(0,1)$ にしたがうと考えてよい。

ここで図1において対立仮説 $f_2(x)$ における垂線ABの右の部分の面積は $(1-\beta)$ であるから(13)式を(14)式に代入してえられる Z_0 の値が $Z_{1-\beta}$ に等しい。すなわち

$$\frac{Z_{\alpha/2} \sqrt{\bar{p}\bar{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) - (p_1 - p_2)}}{\sqrt{(p_1q_1)/n_1 + (p_2q_2)/n_2}} = Z_{1-\beta} \quad (15)$$

となるように標本数 n_1 、 n_2 を定めればよい。

ここで、標本総数を N として、 $n_1=aN$ 、 $n_2=bN$ とおくと（ただし $a=b=1/2$ ）

$$N = \frac{\left\{ Z_{\alpha/2} \sqrt{\bar{p}\bar{q} \left(\frac{1}{a} + \frac{1}{b} \right)} - Z_{1-\beta} \sqrt{(p_1q_1)/a + (p_2q_2)/b} \right\}^2}{(p_2 - p_1)^2} \quad (16)$$

となる。 $a=b=\frac{1}{2}$ の場合、上式は

$$N = 2 \left(Z_{\alpha/2} \sqrt{2\bar{p}\bar{q}} - Z_{1-\beta} \sqrt{p_1q_1 + p_2q_2} \right)^2 / (p_2 - p_1)^2 \quad (17)$$

と変形される。なお $\bar{p}$ と $\bar{q}$ は $\bar{p}=(ap_1+bp_2)/(a+b)$ 、 $\bar{q}=1-\bar{p}$ で推定すればよい。

ところで、 $Z_{1-\beta}=Z_\beta$ であるから(16)式または(17)式の分子の第2項は正になり、検出力 $(1-\beta)$ を大きくするほど Z_β の値が大きくなり、有意差を検出するのに必要とされる標本数も大きくなる。

ところで(17)式において $\beta=0.50$ とおくと、 $Z_{1-\beta}=Z_\beta=0$ となり、(17)式は(8)式に一致するから第2種の誤差 β を考慮しない場合においてえられる1)～5)の標本数は、いわば有意差検出のための最小の標本数で、対立仮説が存在する場合の標本数に比べて少なめに評価されていることになる。

図2に α と β を全部で5とおりに変えて、 $p_1=0.20$ を一定、2群の標本数を同一そして p_2 を0.25から0.50までに連続的に変化させた場合に必要とされる一方の群の標本数を示した。第1種の誤差がともに0.10である(4)と(5)を比較すると、 β が0.5である(5)の方がはるかに標本数は少なくなっていることが読みとれる。

〔例2〕

ある病院の産前外来の受診者では、未熟児出産が

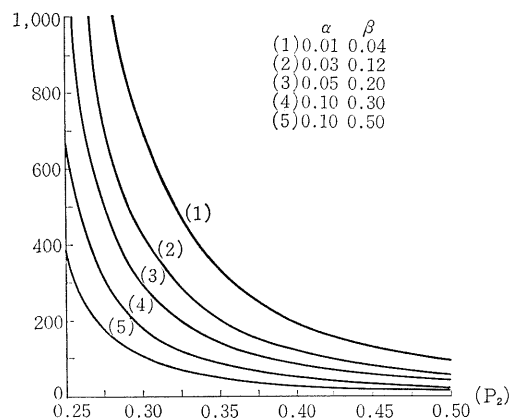


図2 $p_1=0.20$ の場合の標本数

$p_1=0.25$, 非受診者では $p_2=0.40$ と想定されている。有意水準 α を 0.01, 検出力 $(1-\beta)$ を 0.95 に指定すると, 受診者と非受診者の比を 1:1 にした場合, (17) 式より必要とされる標本数は両群ともに 342 例となり, 有意水準を変えて $\alpha=0.02$, 検出力 $1-\beta=0.90$ に下げると 253 例となる。

3. 比較すべき群が多数ある場合の母比率の検定

薬効検定においては, 患者をいくつかの群に分けて, それぞれの群に異なる薬を与え, さらにもう 1 つの群に偽薬を投与し, これを対照群として対照群と個々の薬剤投与群についての死亡率 (あるいは生存率) などを比較することがある。これは標本数の定め方でもやや特殊な場合であるが, 有名な Coronary Drug Project⁴⁾ ではこの考え方を取り入れて標本数を決めているので, ここにその方法を解説しておくこととする。

ここで, m 種の薬剤服用群の標本数を n_d (一定), 対照群の標本数を n_c , 全標本数を N とすると, $N=mn_d+n_c$ の関係が成立する。ここで第 i 種の薬剤投与群と対照群の観察期間中の死亡率を $p_i (i=1\cdots m)$, p_c とおくと標本比率 $\hat{p}_i$ と $\hat{p}_c$ の差の検定は

$$Z_i = \{\hat{p}_i - \hat{p}_c - (p_i - p_c)\} / \sqrt{(p_i q_i) / n_d + (p_c q_c) / n_c} \quad (18)$$

が $N(0, 1)$ にしたがうことを用いればよい (ただし $q_c=1-p_c$, $q_i=1-p_i$)。ここで p_i , p_c が既知の場合に Z_i を最大にするように n_d , n_c を定めるには, Z_i の分母

$$L = (p_i q_i) / n_d + (p_c q_c) / (N - m \cdot n_d) \quad (19)$$

を最小にすればよい。 L は n_d の関数であるから, これを n_d で微分し, 0 とおいて整理すると

$$\frac{n_c}{n_d} = \sqrt{\frac{m p_c q_c}{p_i q_i}} = k \quad (20)$$

となる。したがって, $N=n_c+mn_d$ より

$$n_c = kN / (m+k), \quad n_d = N / (m+k)$$

となるから, (18) 式に $a=k/(m+k)$, $b=1/(m+k)$ を代入すれば, 第 1 種の誤差を β , 検出力を $(1-\beta)$ とした場合の標本総数が算出される。

なお, この方式は Dunnett 型の多重比較の方法の範疇に属するものである。

〔例 3〕 CDP (Coronary Drug Project) における標本数の決め方⁴⁾

冠動脈性心疾患の患者での死亡と病態を軽減するための数種の薬剤の効果を判定することを目的として, 長期間にわたる臨床的な研究が行われた。薬剤投与群は, その種類と量とにより 5 群とし, 比較のために偽薬

表 3 CDP における各治療群の人数

治療法	実際の患者数	推定必要数
estrogen (2.5mg/日)	1,101	1,117
estrogen (5.0mg/日)	1,119	1,117
clofibrate (1.8mg/日)	1,103	1,117
dertrothyroxin (6.0mg/日)	1,110	1,117
nicotinicacid (3.0mg/日)	1,119	1,117
placebo (3.8mg/日) (lactose)	2,789	2,793
計	8,341	8,378*

* 推定値は 8,379 であるが, 桁落ちのため 1 人少ない

(lactose) を用いる対照群を他に定め, 以下の仮説のもとで必要とされる標本数を推定した。

(仮説)

- ① 5 年間の死亡率が対照群で 30.0% (p_c), 5 つの薬剤治療群でのおおの 22.5% (p_i)
- ② 5 年間の脱落率 (drop out) が 30%
- ③ 第 1 種の誤差 α が 0.01
- ④ 検出力 $(1-\beta)$ が 0.95

(20) 式に $p_c=0.3$, $p_i=0.225$, $m=5$ を代入すると

$$\frac{n_c}{n_d} = \left(\frac{5 \times 0.3 \times 0.7}{(0.225 \times 0.775)} \right)^{\frac{1}{2}} \div 2.45$$

となり, 対照群と薬剤投与群の標本数の比が求められる。CDP では, この値から $k=2.5$ を採用し, $n_d=2N/15$, $n_c=N/3$ とした。したがって $a=2/15$, $b=1/3$ を (16) 式に代入すると

$$N = \left\{ 2.326 \times \sqrt{0.21 \times \left(\frac{21}{2} \right)} + 1.645 \times \sqrt{0.174375} \right. \\ \left. \times \left(\frac{15}{2} \right) + 0.21 \times 3 \right\}^2 \div (0.075)^2 \div 5865$$

となり, 仮説②を考慮すると必要な標本総数は 8,379 人となる。これより $n_c=2,793$, $n_d=1,117$ となる。実際の研究に用いた例数は表 3 のように推定値とやや異なるが, 実用上は十分妥当な範囲のものといえよう。

おわりに

以上述べてきたように, 独立性の χ^2 検定, 平均値の差の t 検定, 母比率の差の正規分布による検定などによって求められる有意水準の大きさは, 分析に用いたデータの標本総数によって定まるものであるから, 検定方式の有意義なる使い方は有意差の検出というよりはむしろ定められた有意水準をうるに必要とされる標本数を定めることにあるといえよう。通常の方式によると対立仮説が明確にされておらず, 有意差が検出されたとしても, 仮

説とどの程度異なる対立仮説が肯定されたのか知るすべもない。そこで、本稿では対立仮説が定められる場合に、第1種の誤差だけでなく、第2種の誤差 β （または検出力 $(1-\beta)$ ）も考慮することによって、仮説を棄却し、しかも対立仮説がある程度肯定されうるために必要とされる標本数の定め方を紹介した。このような方式は対立仮説がある程度推定しうる医学データが存在する場合には、もっと積極的に導入されてしかるべきものであろう。

本論文の執筆をお推め頂き、さらに草稿を精読して有益なコメントを頂いた、東大医学部開原成允助教授に感謝する。

文献

- 1) Fleiss, J.L. 著；佐久間 昭 訳：計数データの統計学。東京大学出版会，1975.
- 2) Cohen, J.: Statistical power analysis for the behavioral sciences. Academic Press, New York, 1969.
- 3) Kramer, M. & Greenhouse, S.W.: Determination of sample size and selection of cases. pp. 356-371. In National Academy of Sciences-National Research Publication 583. 1959.
- 4) Coronary Drug Project Research Group. The Coronary Drug Project Design, methods and baseline result, Circulation, 47-48 (3) Supl, 1973.
- 5) 増山元三郎：サリドマイド。東京大学出版会。

* * *