

## 第9回

## 栄養調査を行うときのサンプルサイズの考え方(2)

東京大学医学教育国際協力研究センター  
水嶋春朔 Mizushima, Shunsaku

独立行政法人国立健康・栄養研究所  
佐々木敏 Sasaki, Satoshi

〇〇教室の介入効果、どのトマトが抗酸化物質を増やすのに有効かなどに加えて、県民健康・栄養調査や健康づくり事業の評価のような広範囲の調査では、どのくらいの人数を対象にしたらよいのでしょうか？これを財政当局でうまく説明できると、上手な予算取りができると思うのです。(徳島県, YS)

標本の大きさ(サンプルサイズ)を決める方法については、目的によって考え方が違ってきます。ある母集団(〇〇県民や△△市住民など)の平均(母平均といいます)や比率(母比率)を「推定」したい場合と、二つの母集団(Aトマトを食べたヒト全員とBトマトを食べたヒト全員, X市とY町, 全国平均と〇〇県)の平均(母平均)や比率(母比率)に差があるのかどうかを「検定」したい場合とに大きく分けて考えます。

第7回では、手元にある標本データを利用して

第7回(7月号)に引き続いて、今回も栄養調査を行うときのサンプルサイズの考え方について取り上げてみましょう。生活習慣病の患者さんや健診で要指導となった参加者への栄養指導の効果の評価、〇〇教室の評価、あるいは住民の栄養摂取状況を知るため、健康〇〇21のベースライン調査や中間評価をしたいときに、どのくらいの人数について調べたらよいのかなど、栄養調査をするときの標本の大きさ(サンプルサイズ)に関して、基本的な考え方を説明します。

ちょっと聞きなれない、見慣れない、こむずかしい言葉も出てきますが、あまり敬遠せずにぜひ一読ください。

母集団の平均や比率などを「推定」したい場合を取り上げました。今回は、手元にある標本データを利用して、ある仮説(AとBに差があるかなど)を「検定」したい場合を取り上げましょう。

財政担当者に根拠に基づく医療や健康政策、そして栄養の必要性をきちんと説明できるように、この連載はいつも応援します！

## 母集団に対してある仮説を「検定」したい場合

仮説は「A地区とB地区で、塩分摂取量が違う」「A地区は、全国平均よりも肥満者の割合が多い」など、あるものとあるものを比較して「差」があるかどうかを問題にします。この仮説を証明するには、なぜかへそ曲りに、まず「差がない」という前提をつくります。それは、背理法という論理を使うからです。

たとえば、「地球は丸くない」と主張する人に対

表1 仮説検定のステップ

- 仮説の設定  
帰無仮説  $H_0$ : Aの母平均は、Bの母平均と等しい ( $A = B$ )  
対立仮説  $H_1$ :  
(両側検定) Aの母平均は、Bの母平均と等しくない ( $A \neq B$ )  
(片側検定) Aの母平均は、Bの母平均より大きい ( $A > B$ )  
(片側検定) Aの母平均は、Bの母平均より小さい ( $A < B$ )
- 検定統計量 (Tなど) を求める
- 設定した有意水準 ( $\alpha$  値, 0.05 など) での検定統計量を表などから得て、(2)の値と比較して、(2)が大きければ、つまり、「より分布の端にあって、とても稀、偶然に起こるとすると5%未満」であれば、帰無仮説を否定(棄却)して、対立仮説が成立することにする

表2 仮説検定の考え方

		真実(真の世界)	
		真(差がない)	偽(差がある)
検定結果	差がない(帰無仮説の採用)	正しい判断(95%) ( $1 - \alpha$ )	誤った判断(20%) 第2種の過誤( $\beta$ エラー)
	差がある(帰無仮説の棄却)	誤った判断(5%) 第1種の過誤( $\alpha$ エラー)	正しい判断(80%) ( $1 - \beta$ ):検出力(パワー)

ネイマン・ピアソンの仮説検定原理。図中のパーセントは、研究で習慣的に用いられるレベルを示す。

して、「その仮説が正しいとしたら、地平線の先までずっとみえるはずである。しかし実際には、みえない」、したがって、「地球は丸くない」という主張(仮説)が誤りである、と導きます。これが背理法の論理の組み立て方です。

ある地域の塩分摂取量がほかの地域や全国平均と違うことを証明するには、まず「A地区とB地区で、塩分摂取量は同じ」「A地区は、全国平均と肥満者の割合が同じ」といった文章(仮説)をつくることになります。これらを帰無(きむ)仮説といいます。そしてこの帰無仮説のもと、A地区とB地区との塩分摂取量が実際に  $Xg$  違うときに、 $Xg$ の差が偶然に起こるとするとどのくらいの確率でそれが起こるか(検定統計量)を計算します。その結果、その確率( $p$ で表わします)が5%未満( $p < 0.05$ )のときに、「偶然起こるにはその確率はとてもまれだから、もとの帰無仮説は間違っている(棄却するといいます)」ことにします。ほらやっぱりね、と「差がない」ことを否定して、「差がある」ことにするんです。表1に仮説検定のステップをまとめました。なんだかややこしいですね。

## 仮説検定に必要な3要素

母集団に対してある仮説を「検定」したい場合には、つぎの三つの要素の設定が重要になります(表2)。

(1) 有意水準  $\alpha$ : 正しい帰無仮説を誤って棄却してしまう確率。つまりホントは「差がない」(帰無仮説)のに「あわてて」間違っ「差がある」としてしまふあわてんぼう間違いの確率。第1種の過誤とか、 $\alpha$ エラーといいます。

通常  $\alpha$  は 0.05 (5%), 0.01 (1%) くらいに設定します。 $\alpha = 0.05$  に設定することは、確率として5% (20回に1回)は「差がない」のに「差がある」として間違ってしまう確率があることを前提にすることになります。有意水準  $\alpha$  が小さければ小さいほど、必要なサンプルサイズは大きくなります。

(2) 検出力  $1 - \beta$ : 正しい仮説(帰無仮説に対して対立仮説といいます)を正しく採択する確率。つまり本当に「差がある」(対立仮説)のを「差がある」と検出する力で、検出力(パワー)といっ

て  $1-\beta$  で表わします。

ちなみに  $\beta$  は、第2種の過誤といって、ホントは正しい対立仮説(差がある)を「ぼんやり」して間違えて、正しくない(「差がない」としてしまふ)ボンヤリ間違いの確率です。 $\beta$  は  $\alpha$  の4倍くらいに設定しますので、 $\alpha=0.05$  のとき、 $\beta=0.20$  となって、検出力  $1-\beta=0.80$  となります。検出力が大きければ大きいほど、必要なサンプルサイズは大きくなります。

(3) 差  $d$ : どのくらいの差  $d$  を有意な差とするか、つまり検出したい差のことです。ある母平均とほかの母平均(あるいは基準値)との差になります。検出したい差が小さければ小さいほど、必要なサンプルサイズは大きくなります。

### ある母集団の平均値や比率を、ほかの母集団や全国平均などの基準値と比較して、差があるのかどうか「検定」したい場合のサンプルサイズの考え方

「A地域(県や市町村)の住民の平均塩分摂取量が、ほかのB地域と差があるのかどうかを検定したい」「あるA集団の肥満者の割合が、全国平均値と差があるのかどうか検定したい」など、差があるのかどうか「検定」したい場合、必要なサンプルサイズは、どのくらいの「差」( $d$ )をどのくらいの「有意水準」( $\alpha$  エラー)や「検出力」( $1-\beta$ , パワー)で「検定」したいのかで決まってきます。

$\alpha$  値はだいたい 0.05, 0.01, 0.001 に設定しますが、この値が小さいほど、あわてて間違える確率(差がないのに差があるといってしまう確率: 第1種の過誤,  $\alpha$  エラー)が小さくなり、その分、大きなサンプルサイズが必要になります。また、ぼんやりして間違える確率(差があるのに差がないといってしまう確率: 第2種の過誤,  $\beta$  エラー)は、0.10 や 0.20 にとります。 $1-\beta$  は検出力(パワー)といって、対立仮説が正しいときに、正しいことを証明する力をいいます。つまり、「差があ

表3 有意水準 ( $\alpha$ ) と検出力 (パワー;  $1-\beta$ ) に対応する Z 値

有意水準 ( $\alpha$ )		$Z_\alpha$	検出力	
片側検定	両側検定		(パワー: $1-\beta$ )	$Z_\beta$
0.5%	1.0%	2.576	50%	0.000
1.0%	2.0%	2.326	70%	0.524
2.5%	5.0%	1.960	80%	0.840
5.0%	10.0%	1.645	90%	1.282
10.0%	20.0%	1.282	95%	1.645

下線部分は通常使う設定値

るときに差がある」と証明する力ですね。この  $1-\beta$  が大きいほうが、サンプルサイズは大きくなります。こうしたサンプルサイズの求め方を、検出力ベースのサンプルサイズの算出方法といえます。サンプルサイズを求める計算式には、標準正規分布(平均が0で、分散が1の分布。Z分布と呼びます)を利用して、 $\alpha$  と  $\beta$  の確率を面積として与える値 ( $Z_\alpha$ ,  $Z_\beta$ ) を使います。ちょっとややこしいですが、表3にあげたような数字をよく使います。

(1) ある集団の平均(平均値  $m_1$ , 分散  $\sigma^2$ ) が基準値(平均値  $m_0$ ) と差があるかどうかを、 $\alpha=0.05$ ,  $\beta=0.20$  で検討する場合に必要なサンプルサイズ( $n$ )の計算式

$$n \geq (Z_\alpha + Z_\beta)^2 \times \text{分散 } \sigma^2 / (\text{平均値 } m_1 - \text{平均値 } m_0)^2$$

<例1>ある地域集団の血清総コレステロール値が全国平均と差があるかどうか検討したい場合。

全国平均  $m_0=185$ , この地域集団の平均  $m_1=195$ , 分散  $\sigma^2=300$  とする(検出したい差は10)。

$$n \geq (1.96 + 0.84)^2 \times 300 / (195 - 185)^2 = 23.5$$

必要なサンプルサイズは、24人以上。

<例2>ある地域集団の血清総コレステロール値が全国平均と差があるかどうか検討したい場合。

全国平均  $m_0=185$ , この地域集団の平均  $m_1=190$ , 分散  $\sigma^2=300$  とする(検出したい差は5)。

$$n \geq (1.96 + 0.84)^2 \times 300 / (190 - 185)^2 = 94.1$$

必要なサンプルサイズは、95人以上。

(2) 二つの集団 A (平均値  $m_1$ , 分散  $\sigma_1^2$ ) と集団 B (平均値  $m_2$ , 分散  $\sigma_2^2$ ) の平均に差があるかどうかを、 $\alpha=0.05$ ,  $\beta=0.20$  で検討する場合に必要なサンプルサイズ( $n$ )の計算式

$$n \geq (Z_\alpha + Z_\beta)^2 \times (\text{分散 } \sigma_1^2 + \text{分散 } \sigma_2^2) / (\text{平均値 } m_1 - \text{平均値 } m_2)^2$$

<例3>A集団とB集団の血清総コレステロール値に差があるかどうか検討したい場合。

A集団の血清総コレステロールの平均値  $m_1=185$ , 分散  $\sigma_1^2=300$ , B集団の平均値  $m_2=195$ , 分散  $\sigma_2^2=310$  とする(検出したい差は10)。

$$n \geq (1.96 + 0.84)^2 \times (300 + 310) / (185 - 195)^2 = 47.8$$

必要なサンプルサイズは、48人以上。

<例4>A集団とB集団の血清総コレステロール値に差があるかどうか検討したい場合。

A集団の血清総コレステロールの平均値  $m_1=185$ , 分散  $\sigma_1^2=300$ , B集団の平均値  $m_2=190$ , 分散  $\sigma_2^2=310$  とする(検出したい差は5)。

$$n \geq (1.96 + 0.84)^2 \times (300 + 310) / (185 - 190)^2 = 191.3$$

必要なサンプルサイズは、192人以上。

(3) ある集団 A の比率(有病率など,  $p_1$ ) が基準値 ( $p_0$ ) と差があるかどうかを、 $\alpha=0.05$ ,  $\beta=0.20$  で検討する場合に必要なサンプルサイズ( $n$ )の計算式

$$n \geq \{Z_\beta \sqrt{p_1(1-p_1)} + Z_\alpha \sqrt{p_0(1-p_0)}\}^2 / (p_1 - p_0)^2$$

<例5>ある地域(母集団)の高脂血症者の割合(母比率)を0.25(25%)と予想し、全国平均0.20(20%)と差があるかどうか検討したい場合。

$$n \geq \{0.84 \sqrt{0.25(1-0.25)} + 1.96 \sqrt{0.20(1-0.20)}\}^2 / (0.25 - 0.20)^2 = 526.9$$

必要なサンプルサイズは、527人以上。

(4)二つの集団 A (有病率など,  $p_1$ ) と集団 B ( $p_2$ ) の比率に差があるかどうかを、 $\alpha=0.05$ ,  $\beta=0.20$  で検討する場合に必要なサンプルサイズ

( $n$ )の計算式

$$n \geq \{Z_\beta \sqrt{p_1(1-p_1)} + Z_\alpha \sqrt{2p(1-p)}\}^2 / (p_1 - p_2)^2$$

( $p$  は、 $p_1$  と  $p_2$  の平均)

<例6> A地域の高脂血症者の割合を0.25(25%), B地域の高脂血症者の割合を0.20(20%)として、差があるかどうか検討したい場合。

$$n \geq \{0.84 \sqrt{0.25(1-0.25)} + 0.20(1-0.20)\}^2 / (0.25 - 0.20)^2 = 765.2$$

必要なサンプルサイズは、776人以上。

ちょっとややこしいですが、一度自分で計算機を使って計算してみると、案外簡単に求めることができることがわかると思います。

実際には、無作為な抽出はむずかしく、参加率が高くないことも多いでしょう。次回以降、さらにみなさんのお悩みにお答えしていきたいと思えます。

#### 文献

- 1) 水嶋春朔: 地域診断のすすめ方—根拠に基づく健康政策の基盤—医学書院, 2000.
- 2) 丹後俊郎: 新版医学への統計学. 朝倉書店, 1993.
- 3) 折笠秀樹: 臨床研究デザイナー—医学研究における統計入門. 真興交易医学出版部, 1995.
- 4) 日本疫学会: 疫学—基礎から学ぶために—. 南江堂, 1996.
- 5) 中村好一: 基礎から学ぶ楽しい疫学. 医学書院, 2002.

#### 執筆者より

第49回日本栄養改善学会学術総会(沖縄)の自由集会「第5回EBN自由集会—超基礎講座: 人間栄養学研究におけるデータの読み方・作り方」の翌日、同自由集会の交流会を企画いたしました。詳しくは、次ページをご覧ください。沖縄で読者の多くの皆さんとお目にかかれることを楽しみにしております。(水嶋・佐々木)

#### 編集部より

このコーナーでは読者からの質問を受けつけています。水嶋先生、佐々木先生に質問のある方は、巻末のアンケートはがきか、E-mail(jcn@ishiyaku.co.jp)で編集部宛にお寄せください。質問をとりあげさせていただきます方には、掲載号をお送りします。