

第7回

栄養調査を行うときのサンプルサイズの考え方(1)

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

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ときに、どのくらいの人件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、基本的な考え方をわかりやすく説明します。

ちょっと聞き慣れない、見慣れない、小むずかしい言葉もでてきますが、敬遠せずにぜひ一読ください。

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、どのくらいの人件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」したい場合と大きく分けて考えます(表1)。今回は、「推定」したい場合の考え方について説明しまし

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！

母集団と標本の考え方

●1回の測定でよいのか？

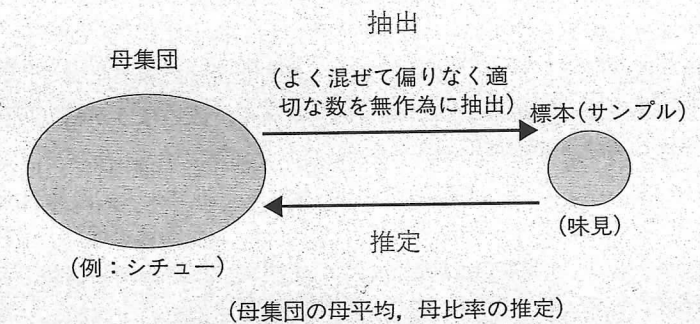
AさんとBさんとで、どちらが今日の体重は重いのか？ どちらが今日の昼食をたくさん食べるのか？ といった疑問は、「今日」(10年前でない、毎日のでもない)と限定した一つの情報を比較すれば結論がわかりますね。つまり、一回ずつ体重を測定したり、今日の昼食の食べた量を1回観察すれば、その時点で結論が出ます。

でも、栄養調査の場合には、AさんとBさんとで、どちらが不飽和脂肪酸を「通常」よくとっているのか？ AトマトとBトマトを食べている人は、どちらが血液中の抗酸化指標が高くなるの

表1 サンプルサイズの大きさの二つの決め方

- 1) 母集団のある指標(母平均, 母比率)を「推定」したい場合
→精度(推定誤差)の大きさに基づいた計算
- 2) 母集団に対して、ある仮説(差がある)を「検定」したい場合
→有意水準 α , 有意差 d , 検出力 $1-\beta$ に基づいた計算

図1 母集団と標本の関係



か？を知りたいときなど、今日1回測定すれば、結論が出るわけではないということが問題になります。

今日1回だけ測定すればよいのは、毎日同じ状態が続いている場合です。つまり多様性がない場合です。しかし、人の食事や栄養摂取には多様性がとって多くありますね。ある人(AさんかBさん)の不飽和脂肪酸の摂取量を正確にありのままに知りたい場合には、365日の摂取状況が、もとのありのままの正確な情報です。あるトマト(AトマトかBトマト)を食べている人の血液中の抗酸化物質を正確にありのままに知りたい場合には、そのトマトを食べている人全員の情報がもとの正確な情報です。

●母集団と標本

もとのありのままの情報全体を、「母集団」といいます。それは多様性があり、範囲が幾分広く、

ばらつきが大きくて、平均値や中央値の近辺に個々の値が集中しているわけではありません。

ホントのありのままの情報を知りたいときには、この「母集団」全部を対象とした調査をします。それをむずかしい言葉で、悉皆(しっかい)調査といいます。国勢調査が代表的な例ですね。しかし、「母集団」全部を調べるのは、コストがかかってとてもたいへんです。それで、この「母集団」の一部を調べて、「母集団」の特徴を推測・推定することをします。この「母集団」の一部を「標本」といいます(図1)。標本調査の代表的なものは、国民栄養調査ですね。

味見のお作法と標本抽出のしかた

ある料理をつくっているとしましょう。味がどうか、調理している最中に味見をしますね。つくっている料理全部が母集団です。全部を調べたら、

つまり全部食べてしまったら、料理がなくなってしまうので、実際には、料理の一部を箸やスプーンでとって、味見をしますね。これは、まさしく「母集団」と「標本」の関係なのです。

●「よく混ぜる」

このとき重要なのは、よく混ぜること!! 全体を代表するようなところからとること!! よく混ぜないで、偏ったところから味見用にとったものは、味の濃淡が全体と違ったりしているので、味見(標本)の味(標本解析結果)から、料理全体の出来栄(母集団)を評価(推定)することができません。この味見の例で、標本抽出の重要性を直感的におわかりいただけると思いますが、いかがでしょうか。よく混ぜていない1例や2例の事例報告やお試しレポート(?)ではなにもいえないのです。

このように、味見にはお作法が必要です。シチューの味見、つまり、まず全体をよく混ぜて(料理によってはよく混ぜるとだいなしになるものもありますが…)、シチュー全体を代表する部分(スープ、ニンジン、ブロッコリー、肉…)をとって味見することです。よく混ぜて適切にとることを、無作為抽出(作為なしに選び出す)といいます。こうしたお作法が、調査にもいえます。まず、偏りのない標本を選ぶことです。それでは、じゃあどのくらいの標本の大きさ(何人、何日、何サンプル: サンプルサイズといいます)があれば、全体について推測できるのでしょうか?

母集団のある指標を「推定」したい場合のサンプルサイズの考え方

●「精度」が重要

「ある地域(県や市町村)の住民の平均塩分摂取量を推定したい」「ある集団の肥満者の割合を推定したい」など、地域住民全体やあるトマトを食べている人全体の情報を「推定」したい場合には、どのくらいの「精度」(確からしさ)で推定したい

のかが、サンプルサイズを決めるうえで、重要になります。

「精度」とは、確からしさの幅をいい、その大きさを表すのに「信頼区間」という言葉を使います。「95%信頼区間」とは、真の値(ホントの母集団の平均値や比率など)が95%の確率で入っている区間をいいます(これは、5%の確率(20回に1回の確率)で、その区間に真の値(ホントの母集団の平均や比率)が入っていないことも意味しています)。つまり「95%信頼区間」の幅が代表的な精度の表わし方で、狭いほうが推定の精度が高く、広いほうが精度は低いわけです。

●サンプルサイズの計算方法

この「95%信頼区間」の幅をどの程度にするかを決めて、母集団の情報(母平均では標準偏差、母比率ではその予測値)を得れば、つぎの式で母集団の平均(母平均)や比率(母比率)を推定するためのサンプルサイズを求めることができます。

(1)母集団の平均(母平均)を求める場合に
必要なサンプルサイズ(n)の計算式
$$n \geq \frac{(\text{予想される母集団の標準偏差})^2 \times (1.96/\text{精度})^2}{\text{母比率} \times (1 - \text{母比率})}$$

<例1>ある地域(母集団)の血清総コレステロール値の平均(母平均)を95%信頼区間が $\pm 5 \text{ mg/dl}$ 以内(精度)で推定したい場合。過去のデータから母集団の血清総コレステロール値の標準偏差を30と見積もる。

$$n \geq 30^2 \times (1.96/5)^2 = 138.3$$

必要なサンプルサイズは、139人以上。

<例2>ある地域(母集団)の血清総コレステロール値の平均(母平均)を95%信頼区間が $\pm 2.5 \text{ mg/dl}$ 以内(精度)で推定したい場合。過去のデータから母集団の血清総コレステロール値の標準

偏差を30と見積もる。

$$n \geq 30^2 \times (1.96/2.5)^2 = 553.2$$

必要なサンプルサイズは、554人以上。

(2)母集団の比率(母比率)を求める場合に
必要なサンプルサイズ(n)の計算式
$$n \geq \frac{(\text{予想される母集団の比率}) \times (1 - \text{予想される母集団の比率}) \times (1.96/\text{精度})^2}{\text{母比率} \times (1 - \text{母比率})}$$

<例3>ある地域(母集団)の高脂血症者の割合(母比率)を95%信頼区間が ± 0.05 以内($\pm 5\%$ 以内)(精度)で推定したい場合。過去のデータから母集団の高脂血症者の割合を0.25(25%)と予想。

$$n \geq (0.25) \times (1 - 0.25) \times (1.96/0.05)^2 = 288.12$$

必要なサンプルサイズは、289人以上。

<例4>ある地域(母集団)の高脂血症者の割合(母比率)を95%信頼区間が ± 0.025 (2.5%)以内(精度)で推定したい場合。過去のデータから母集団の高脂血症者の割合を0.25(25%)と予想。

$$n \geq (0.25) \times (1 - 0.25) \times (1.96/0.025)^2 = 1,152.5$$

必要なサンプルサイズは、1,153人以上。

例1と例2、および例3と例4の違いは95%信

頼区間(精度)の幅だけです。精度の幅が半分になると、必要なサンプルサイズは4倍になることがわかりますね。

つぎの機会には、母集団に対してある仮説を「検定」したい場合のサンプルサイズの考え方を取り上げてみます(9月号を予定しています)。

文献

- 1) 水嶋春明: 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: 臨床研究デザイン—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.

ご質問への御礼

YSさん、いつも熱心に本連載をお読みいただき、また感想や質問も頻繁にお寄せいただきまして、誠にありがとうございます。

読者の皆さんが普段おもちの疑問に対して、わかりやすく腑に落ちるように説明できるかどうか、毎月真剣勝負を挑まれている道場主催者のような気もちでいます。もし説明がわかりにくければ、遠慮なくご指摘いただきたくお願いいたします。多くの皆様のご質問をお待ちしております。

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。すでにお送りいただいた方、ありがとうございました。質問をとりあげさせていただいた方には、掲載号をお送りします。引き続き、たくさん質問をお待ちしております。