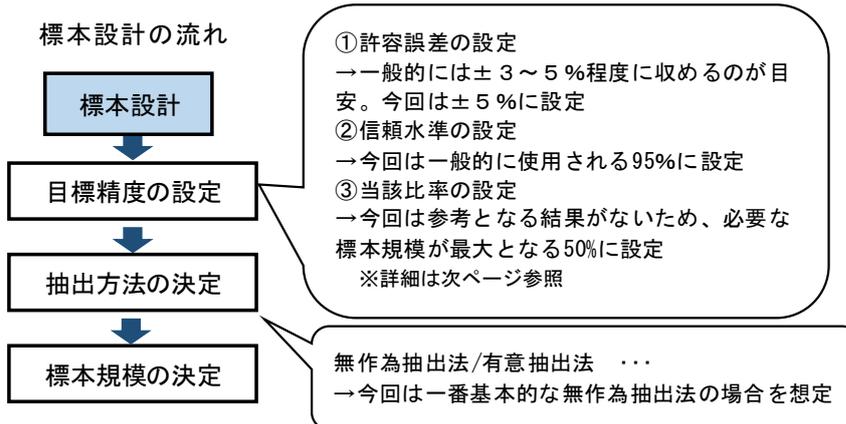


## 第2章-7 アンケート調査を行いたい

### 例えば

県民に対して、調査対象者を無作為に抽出してアンケート調査（設問「毎日朝食を食べるかどうか」⇒回答「はい」または「いいえ」の二者択一形式）を行いたい。どのくらいの人数の調査対象者が必要か？

### アンケート調査に必要な人数の一例



上記のとおり、一般的によく用いられる精度で、比率を調べる（設問に対して「はい」と回答した人の割合◆%など）アンケート調査を行う場合、おおむね **400 人分のサンプル** が必要となる。（※単純無作為抽出法で抽出）

→400 人分のサンプルで調査した場合、調査結果の誤差が±5%となる確率が95%になる。

⇒ただし…

400 人に調査して必ずしも全員から回答が得られるとは限らない。調査方法や調査対象により回収率には差が生じることが考えられるため、過去の調査や類似の調査を参考に最低限必要な400 人分のサンプルが得られるように、対象者数を決定する。

（例）二者択一で比率を調べるアンケートの回収率を20%で設定した場合

⇒アンケート対象者数は2,000人以上必要となる。（2,000人×20%=400人）

### ！注意！

・年代、地域ごとなど、さらに細かい区分での分析を行う場合は、それぞれの区分ごとにサンプル数が必要なので、区分ごとの対象者数を設定する必要がある。

（例えば、年齢～20歳、21～60歳、61歳～の3つの年代区分で集計したいとき、上記の設定の場合、必要な対象者数400人がそれぞれの区分に必要⇒全体で約1,200人分必要となる。）

## 解説

○必要サンプル数 400 人とは？

※標本規模の求め方の式（抽出法の中で、一番基本的な単純無作為抽出法の場合）

<b>【計算式】</b>	$n = \frac{\lambda^2 p(1-p)}{d^2}$	$n$ : (必要な) 標本規模	$d$ : 標本誤差
		$\lambda$ : 信頼水準によって定まる値	$p$ : 当該比率

### ①標本誤差の設定

(? 標本誤差とは…調査対象者全員 (=母集団：設問例の場合、全県民) に調査を実施した結果と、調査対象者の中から一部の人を抽出した調査 (標本調査) の結果では、結果は一致せず、いくらかの誤差が生じる。この誤差のことを標本誤差という。)

: 誤差が小さい方が、アンケート結果の精度は高くなる。調査実施者において、調査の目的、用途、予算などにに基づき、どの程度の誤差まで許容するか (どのくらいの精度を必要とするか) を決定する。一般的には 3 ~ 5 % 程度におさめるのが 1 つの目安である。

→今回は 5 % ( $d = 0.05$ ) に設定。

(ただし、この誤差には回答誤り、集計ミスなど調査の過程で生じる誤差は含まれない)

### ②信頼水準によって定まる値の設定

(? 信頼水準とは…標本調査を行った場合には、全数調査の結果から大きく外れた値が生じる可能性があるが、そのような可能性を減らすため、設定する値のこと。)

:  $\lambda = 1.96$  (信頼水準 95% のときの値) とするのが統計的処理としては一般的である。

(※調査対象によっては信頼水準を 95% 以外に設定する場合もある)

→ $\lambda = 1.96$  に設定。…標準正規分布における設定値

### ③当該比率の設定

(? 当該比率とは…調査の結果得られる比率 (設問例の場合、「はい」と回答した人の割合))

: 直近の同様の調査や事前調査などから設定する。事前に参考となる結果がない場合や、過去の当該比率の変動が大きい場合などは、必要な標本規模が最大となる  $p = 0.5$  とするのが標本誤差が小さくなるためよい。

→今回  $p = 0.5$  に設定

⇒設定した数値を上記計算式に代入すると、 $n = 384.16 \approx 400$  人 となる

※【参考】母集団が少ない場合：母集団がおよそ 1 万人以下の場合、上記計算式で求めた標本規模 ( $n$ ) に修正係数をかけた数をサンプル数とする。

<修正係数 = (母集団の人数) ÷ (母集団の人数 + 標本規模( $n$ ) - 1) >

## 参考

○サンプル数を増やしたときまたは減らしたときの誤差の目安は？

⇒前述の計算式の標本誤差（d）を変えると、必要なサンプル数も変化する。

標本誤差の目安については、下表のとおり。

サンプル数が多いほど調査の精度が上がり、少ないほど精度が下がることが分かる。サンプル数が少ない方が調査コストは減るが、調査精度の観点から慎重に検討する必要がある。

サンプル数別標本誤差早見表（回収率 100%、比率を調査する場合）

	サンプル数（人）								
	100	400	500	1,000	2,000	3,000	5,000	7,000	10,000
標本誤差の目安	±9.8%	±4.9%	±4.4%	±3.1%	±2.2%	±1.8%	±1.4%	±1.2%	±0.98%

※信頼水準 95%、当該比率 50%に設定した場合

※単純無作為抽出法による調査の場合

※他の抽出法の場合、複数回答の場合、記述式の設問の場合などは、上記表の数字によらない

### ！注意！

・本項目で解説したサンプル数は、

①単純無作為抽出法により、

②ある単一回答（一つの選択肢を選ぶ）の設問に関して、特定の回答をした人の比率を調べるアンケート調査をしたい

場合のみに目安とすることができる。

⇒他の抽出法の場合（二段抽出法、層化抽出法など）や、複数回答の場合（選択肢の中から2つ以上選ぶことができる場合）、記述式の設問の場合（「1か月の消費支出額」など実数を調査する場合）などは、サンプル数の計算の仕方が異なるので注意。

※設問の選択肢が多い場合（例えば、三択調査（選択肢A、B、C））であっても、単一回答形式かつ、「Aを選択した人の比率」について調査したいという場合であれば、今回の式を利用してサンプル数を求めることは可能。（「Aを選択した人」と「それ以外」に分けて当該比率を設定できる場合。）

※アンケート内に複数の設問がある場合には、調査の中で最も注目する設問について、目標精度等の設定から得られた必要サンプル数をアンケート調査の全体のサンプル数とするとよい。