

第 154 講 母集団と標本

学習目標

この講では、調査したい対象全体と、実際に調べるために取り出した一部を区別できるようにする。

具体的には、次のことができるようになる。

- ・ 調査目的に応じて母集団と標本を識別できる。
- ・ 母集団の大きさ、標本の大きさ、調べる特性を区別できる。
- ・ 全数調査と標本調査を使い分ける理由を説明できる。
- ・ 標本の選び方による偏りと、無作為抽出の考え方を理解できる。

調べたい全体と、実際に調べる一部

この講で扱う製品や調査の例は、すべて架空のものである。特に断りがない限り、対象は重複なしで抽出する。また、対象数や検査結果については各設問に示された事実として扱う。

大量の製品や部品の性質を知りたいとき、すべてを調べることができる場合もあれば、一部だけを調べる場合もある。

例えば、ある日に製造された部品 1000 個の不良品割合を知りたいとする。

1000 個すべてを調べる場合、その 1000 個全体を対象として直接割合を求めることができる。ただし、これは「1000 個を漏れなく調べ、正しく判定できた」という条件のもとで成り立つ。全数を調べたからといって、測定や記録の誤りまで自動的になくなるわけではない。

一方、1000 個の中から 50 個だけを選んで調べる場合、調べた 50 個は母集団全体ではなく、母集団の特徴を考えるための標本となる。

知りたい対象全体を母集団といい、母集団から調査のために取り出した一部を標本という。

大切なのは、「調べたい対象全体」と「実際に調べた対象」を分けて考えることである。

母集団と標本

ある調査で、知りたい対象全体を母集団という。

母集団から調査のために取り出した一部の対象を標本という。

母集団に含まれる対象の個数を母集団の大きさ、標本に含まれる対象の個数を標本の大きさという。

例えば、ある製品ロット 1000 個の不良品割合を調べる場合、1000 個全体が母集団であり、その中から選んだ 50 個が標本となる。

母集団と標本は、調査の目的によって決まる相対的な関係である。ある調査では母集団であったものが、別の調査では標本になることもある。

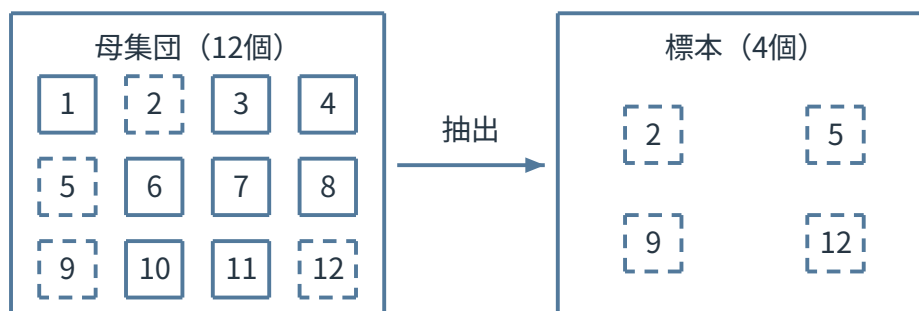


図1 架空の部品 12 個から、破線で示した 2、5、9、12 を選んだ 1 つの抽出結果の例である。同じ番号は同じ部品を表し、標本は母集団の一部である。無作為性や代表性はこの図だけでは判断せず、枠の面積も個数の比を表さない。

母集団と標本を区別する視点

標本調査で一部を選んで調べる場面では、次の順に考える。

1. 何について知りたいのかを決める。
2. 知りたい対象全体を母集団とする。
3. その母集団から実際に調べるために選んだ一部を標本とする。
4. 何の性質を調べるのかを確認する。

例えば、その日に作った架空の部品 1000 個全体の不良品割合を知るため、その中から検査用の 50 個を選んだ場合、

- ・ 母集団：その日に作った部品 1000 個
- ・ 標本：その 1000 個から選んだ部品 50 個
- ・ 調べる特性：不良品であるかどうか

となる。

ただし、全数調査では母集団全体を調べているため、調べた対象全体を自動的に一部の標本と呼ぶことはしない。

母集団と標本は、「一部を選んで調べる」という標本調査の状況で区別することが重要である。

全数調査と標本調査

母集団に含まれるすべての対象を調べる方法を全数調査という。

一方、母集団の一部だけを取り出して調べる方法を標本調査という。

全数調査では、母集団全体を直接調べるため、対象全体についての結果を得ることができる。ただし、全数調査であっても、測定方法や記録方法に誤りがあれば、その誤りが自動的になくなるわけではない。

例えば、架空の製品 1000 個について不良品かどうかを確認し、1000 個すべてを漏れなく正しく判定できた場合、その 1000 個全体の不良品割合を直接求めることができる。

標本調査では、一部の対象を調べることで母集団全体の特徴を考える。母集団が大きい場合、全数調査は費用や時間が大きくなるため、標本調査が利用される。

また、破壊検査では全数調査そのものが論理的に不可能になるわけではない。すべての対象を検査することはできるが、検査後にすべて販売できなくなるなど、実用上避けられる場合がある。標本調査では、調べた標本の結果と母集団全体の値が必ず一致するわけではない。そのため、標本の選び方や結果の扱い方が重要になる。

標本の大きさだけでは十分ではない

標本を大きくすれば、より多くの情報を得られる。しかし、選び方に偏りがある場合、標本を大きくしても母集団を正しく反映できないことがある。

例えば、倉庫内の箱の損傷割合を調べるために、入口近くの箱だけを大量に調べても、倉庫全体の特徴を表しているとは限らない。

重要なのは、標本の大きさだけでなく、どのように標本を選んだかである。

無作為抽出

有限の母集団から標本を選ぶとき、固定した個数を重複なしで選び、同じ個数のどの部分集合も同じ確率で選ばれるようにする方法を、この講では無作為抽出の具体例として扱う。

無作為抽出では、調査する人の判断で特定の対象だけを選ぶことを避けやすくなり、選び方による偏りを小さくすることが期待できる。

ただし、無作為に選んだとしても、偶然による標本の違いは残る。

例えば、部品ロットから無作為に選んだ標本と、母集団全体では、不良品の割合などの特性値が必ず一致する保証はない。

ここで扱う方法は、有限母集団から固定個数を選ぶ具体的な方法の一つであり、無作為抽出全体がこの方法だけで定義されるわけではない。

標本の割合について

標本の中で不良品が占める割合を調べると、その標本についての割合は分かる。

しかし、その割合が母集団全体の真の割合と必ず同じになるわけではない。

標本から母集団の特徴を考えるためには、標本の情報をどのように扱うかを考える必要がある。この講では母集団と標本の間係を理解し、以降の講で標本を利用した推測へ進む。

例題 1

ある日に作った架空の部品 1000 個について、不良品の割合を知りたい。

部品には 1 から 1000 まで番号が付いており、その中から重複しないように無作為に 50 個を選んで検査した。

次の問いに答えよ。

- (1) 母集団は何か。
- (2) 標本は何か。
- (3) 母集団の大きさと標本の大きさをそれぞれ答えよ。
- (4) 調べる特性は何か。

例題 1 の解答・解説

- (1) 知りたい対象は、その日に作った架空の部品 1000 個全体である。
したがって、母集団は「その日に作った架空の部品 1000 個」である。
- (2) 実際に検査したのは、1000 個から選んだ 50 個である。
したがって、標本は「無作為に選んだ架空の部品 50 個」である。
- (3) 母集団に含まれる部品数は 1000 個、標本に含まれる部品数は 50 個である。
したがって、母集団の大きさは 1000、標本の大きさは 50 である。
- (4) 調査の目的は不良品の割合を知ることである。
したがって、調べる特性は「不良品であるかどうか」である。
- ポイントは、対象の全体を母集団、実際に調べた一部を標本として区別することである。

例題 2

- 架空の電球 2000 個について平均寿命を知りたい。
しかし、寿命を調べるためには電球を点灯し続ける必要があり、検査した電球は販売用に戻すことができない。
そこで、2000 個から 40 個を選んで検査することにした。
次の問いに答えよ。
- (1) この調査が標本調査である理由を説明せよ。
- (2) 全数調査ではなく標本調査を行う理由を 1 つ答えよ。
- (3) 40 個の検査結果だけで、2000 個すべての平均寿命が完全に決まるといえるか。

例題 2 の解答・解説

- (1) 2000 個すべてではなく、その中から選んだ 40 個だけを調べるため、標本調査である。
- (2) 寿命を測定すると、その電球は販売用に戻せない。また、2000 個すべてを調べるには時間や費用がかかる。
そのため、一部を調べる標本調査を行う。
- (3) 完全に決まるとはいえない。
40 個の検査結果は 2000 個全体の特徴を考えるための情報である。しかし、未調査の 1960 個の寿命は、この 40 個の結果だけでは分からない。
したがって、選んだ 40 個の平均寿命と 2000 個全体の平均寿命が一致する保証はない。

例題 3

- ある架空の倉庫には、その日に入荷した箱が 500 個ある。
箱全体の損傷割合を調べるため、入口付近に置かれている 100 個の箱だけを調べた。
次の問いに答えよ。
- (1) この標本の選び方で注意すべき点を説明せよ。
- (2) より偏りを小さくするためには、どのような選び方が考えられるか。

例題 3 の解答・解説

- (1) 入口付近の箱だけを選んでいるため、場所による選び方の偏りが生じる可能性がある。
標本の数が 100 個と多くても、選ばれ方に偏りがあれば、倉庫全体を正しく表さない可能性がある。
ただし、この情報だけから、実際に入口付近の箱の損傷割合が全体と異なるとは断定できない。
- (2) 500 個すべてに番号を付け、その中から重複なしで同じ大きさの標本が同じ確率で選ばれるようにする方法が考えられる。
このように選ぶことで、特定の場所や特徴を持つ箱だけが選ばれる偏りを小さくできる。

演習問題

練習問題 1

架空の電子部品工場で、ある日に製造された部品 1200 個の長さの特徴を調べることにした。
その部品 1200 個から重複なしで無作為に 60 個を選び、それぞれの部品の長さを 2 回ずつ測定した。
この結果、120 個の測定記録が得られた。
次の問いに答えよ。

- (1) 母集団を答えよ。
- (2) 標本を答えよ。
- (3) 母集団の大きさと標本の大きさをそれぞれ答えよ。
- (4) 測定記録の数と標本の大きさが異なる理由を説明せよ。
- (5) 調べる特性を答えよ。

練習問題 2

架空の飲料工場で、2026 年 10 月 1 日に製造された缶飲料 3000 本の内容量のばらつきを調べたい。
調査担当者は、2026 年 9 月 30 日以前に製造された商品や、10 月 2 日以降に製造された商品は調査対象に含めず、10 月 1 日の製造分だけを調べることにした。
この調査における母集団を答えよ。
また、「工場で作られるすべての缶飲料」と考える誤りの理由を説明せよ。

練習問題 3

架空の倉庫で、ある日に保管した箱 80 個すべてについて、箱を壊さずに外観を確認し、損傷の有無を記録した。
次の問いに答えよ。

- (1) この調査は全数調査か標本調査か答えよ。
- (2) この調査における母集団の大きさを答えよ。
- (3) 全数調査であるため、損傷記録の誤りは必ず 0 になるといえるか。理由を説明せよ。

練習問題 4

架空の倉庫に、ある日に入荷した箱 600 個がある。

箱全体の損傷割合を知るため、担当者は「見た目に傷がない箱」だけを選び、その中から 30 個を調べた。

次の問いに答えよ。

- (1) この標本の選び方で注意すべき点を説明せよ。
- (2) より偏りを小さくするためには、どのような選び方が考えられるか。
- (3) この調査結果だけから、倉庫全体の損傷割合や実際の偏りの大きさを断定できるか。

練習問題 5

架空の製品ロット 5000 個について品質を調査する。

担当者 A は、入口付近から 100 個を選んだ。

担当者 B は、入口付近から 1000 個を選んだ。

どちらも「入口付近にある製品だけを選ぶ」という方法で調査した。

このとき、標本を 100 個から 1000 個へ増やしたことで、必ず選び方の偏りがなくなるといえるか。理由を説明せよ。

練習問題 6

架空の部品ロット 900 個について、品質を調べるために 50 個を選ぶ。

次の 2 つの方法を比較する。

- ・ A：900 個すべてに番号を付け、その中から重複なしで無作為に 50 個を選ぶ。
- ・ B：作業台の近くにある 50 個を選ぶ。

- (1) この講で扱う無作為抽出の具体例に当てはまる方法を答えよ。
- (2) B の方法を無作為抽出と呼べない理由を説明せよ。

練習問題 7

架空の部品ロット 2000 個から、無作為に 100 個を選んで不良品割合を調査した。

標本 100 個では不良品が 3 個であった。

この結果について、次の主張が正しいか答え、理由を説明せよ。

「この標本は無作為に選ばれているので、母集団 2000 個全体の不良品割合も必ず 3%である。」

練習問題 8

架空の食品保存容器ロット 1500 個について、傷の割合を調べるため、無作為に 100 個を選んだ。

選んだ 100 個には傷のある容器が 1 個もなかった。

次の主張について、正しいか答え、理由を説明せよ。

「標本 100 個で傷が 0 個だったので、ロット 1500 個全体にも傷のある容器は 1 個もない。」

練習問題 1 の解答・解説

- (1) 知りたい対象全体は、その日に製造された部品 1200 個である。
したがって、母集団は「その日に製造された部品 1200 個」である。
- (2) 実際に選んで調べた対象は、1200 個から選んだ 60 個の部品である。
したがって、標本は「無作為に選んだ部品 60 個」である。
- (3) 母集団の大きさは 1200、標本の大きさは 60 である。
- (4) 標本の単位は部品であり、1 つの部品から 2 回測定しても、標本に含まれる部品数は 1 個のままである。
今回は 60 個の部品を調べ、それぞれ 2 回測定したため、測定記録は $60 \times 2 = 120$ 個になった。
測定記録の数と標本の大きさは同じものではない。
- (5) 調べる特性は、部品の長さである。

練習問題 2 の解答・解説

母集団は「2026 年 10 月 1 日に製造された缶飲料 3000 本」である。
「工場で作られるすべての缶飲料」と考えると、調査対象ではない 9 月 30 日以前や 10 月 2 日以降の製造分まで含めてしまう。
母集団は、調査目的だけでなく対象の範囲や時点を明確にして決める必要がある。

練習問題 3 の解答・解説

- (1) 保管した箱 80 個すべてを調べているため、全数調査である。
- (2) 調査対象は、その日に保管した箱 80 個である。
したがって、母集団の大きさは 80 である。
- (3) 必ず 0 になるとはいえない。
全数調査では対象範囲のすべてを調べるが、確認方法や記録方法に誤りがあれば、記録の誤りが生じる可能性がある。
全数調査であることと、測定や記録が必ず正確であることは別のことである。

練習問題 4 の解答・解説

- (1) 見た目に傷がない箱だけを選んでいため、損傷という調べたい特性に関係する選び方の偏りが生じる可能性がある。
見た目に傷があると判断された箱が最初から選択対象から外れているため、600 個全体の特徴を表さない可能性がある。ただし、この情報だけから、見た目では確認できない傷の有無まで判断することはできない。
- (2) 600 個すべてに番号を付け、その中から重複なしで無作為に選ぶ方法が考えられる。
このように、特定の特徴を持つ箱だけが選ばれることを避けることで、選び方による偏りを小さくできる。
- (3) 断定できない。
この情報だけでは、倉庫全体の実際の損傷割合や、どの程度の偏りが生じているかは分からない。
ただし、選び方に偏りが生じる可能性があることは判断できる。

練習問題 5 の解答・解説

必ずなくなるとはいえない。

標本を大きくすると調べる情報量は増える。しかし、入口付近だけを選ぶという方法自体が変わっていないため、選び方による偏りは残る可能性がある。

標本の大きさだけでなく、標本の選び方も重要である。

練習問題 6 の解答・解説

(1) A が、この講で扱う無作為抽出の具体例に当てはまる。

番号を利用し、同じ大きさの標本となるどの組合せも同じ確率で選ばれるようにしているためである。

(2) B は作業台の近くという特定の条件で選んでいる。

そのため、すべての組合せが同じ確率で選ばれるとは限らず、無作為抽出とはいえない。

練習問題 7 の解答・解説

この主張は正しくない。

無作為に選ばれた標本であっても、標本の不良品割合が母集団全体の割合と完全に一致する保証はない。

この場合、標本 100 個では不良品 3 個だったが、それだけで母集団 2000 個全体の不良品割合が 3% であると断定することはできない。

練習問題 8 の解答・解説

この主張は正しくない。

標本 100 個で傷のある容器が 0 個だったことは、その 100 個についての結果である。

母集団 1500 個全体にも傷がないとは断定できない。標本では確認されなかった特徴が、母集団の別の部分に存在する可能性がある。

標本の結果から母集団について考えるときは、観測した部分と全体を区別する必要がある。