

組 番 氏名

1 母集団と標本

標本調査

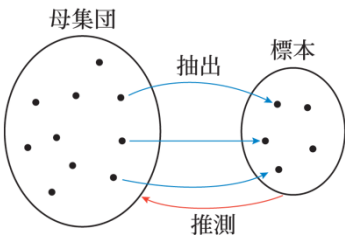
統計調査には、学校の健康診断のように、調査の対象全体をもれなく調べる全数調査と、テレビの視聴率や卵の品質調査のように、その一部を抜き出して調べ、それから全体を推測しようとする標本調査がある。

全数調査には多くの時間や費用を必要とする場合や、検査によって商品が傷つくなど現実的には無理がある場合もある。このように全数調査が難しい場合は、標本調査を行うことが多い。

標本調査では、調べようとする調査の対象全体を母集団といい、その要素の個数を母集団の大きさという。

そして、調査のために母集団から抜き出された要素の全体を標本といい、標本に含まれる要素の個数を標本の大きさという。

また、標本を抜き出すことを抽出という。



母集団の特徴を正しく推測するためには、標本に偏りがないように、確率的にみて公平な抽出をしなければならない。

このような標本の抜き出し方を無作為抽出といい、無作為抽出によって選ばれた標本を無作為標本という。

無作為抽出を行うために、乱数さいという特殊なさいころや乱数表という表を用いたり、コンピュータによって発生させた乱数を利用したりする。

乱数さいは正二十面体のさいころで、0から9までの数字がそれぞれ2面ずつ、計20の面に書いてある。

例えば、赤、青、黄の3つの乱数さいを投げて、出た目のうち、赤の目を百の位、青の目を十の位、黄の目を一の位の数とすれば、0から999までの数をもれなく選ぶことができる。



乱数表は、0から9までの数字を、各数字が現れる確率を同じにして、規則性がないように並べたものである。この本の143ページに、見やすいように2つずつの数字を組にした乱数表を載せている。

無作為抽出の方法

無作為抽出には、例えば次のような方法がある。

例 16 クラスの生徒40人から5人を無作為抽出する。まず、生徒40人に1から40までの番号をつける。

次に、乱数さいなどを用いて、乱数表のどの行、どの列から始めるかを決める。

例えば、乱数表の6行目の第1列を選ぶと、

05 33 18 08 51 51 78 57 26 17 ……

が得られる。この中から該当する生徒がいるときはその番号を抜き出し、いないときは次の組の番号に移る。

ただし、同じ数字が出てきたときは省く。

この場合は、

05 33 18 08 26

の5人が無作為抽出されたことになる。

この5人が40人の母集団から選ばれた大きさ5の無作為標本である。

乱数表

28	89	65	87	08	13	50	63	04	23	25
30	29	43	65	42	78	66	28	55	80	47
95	74	62	60	53	51	57	32	22	27	12
01	85	54	96	72	66	86	65	64	60	56
10	91	46	96	86	19	83	52	47	53	65
05	33	18	08	51	51	78	57	26	17	34
04	43	13	37	00	79	68	96	26	60	70
05	85	40	25	24	73	52	93	70	50	48
84	90	90	65	77	63	99	25	69	02	09
28	55	53	09	48	86	28	30	02	35	71

復元抽出と非復元抽出

母集団から標本を抽出するとき、抽出のたびに要素をもとに戻し、あらためて次の要素を抽出する方法を復元抽出という。これに対して、もとに戻さないで続けて次の要素を抽出する方法を非復元抽出という。

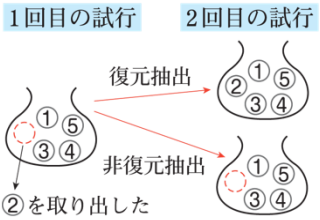
例 17 1から5までの数字が1つずつ書かれた玉の入った袋から、玉を1個ずつ取り出し大きさ2の標本を抽出する。次の場合に、標本の選び方の総数を求める。ただし、抽出する順序は区別するものとする。

(1) 復元抽出のとき、

$5^2=25$ (通り)

(2) 非復元抽出のとき、

${}_5P_2=20$ (通り)



組	番	氏名
---	---	----

問 23 例17において、玉を1個ずつ非復元抽出する。抽出する順序を区別しないとき、大きさ2の標本の選び方は何通りあるか。

集団分布

身や体重のように、母集団の特性のうち、特に数量的に表されるものを変量という。
大きさ N の母集団において、変量 X のとり異なる値を $x_1, x_2, \dots, x_k$ とし、それぞれの値をとる個数を、それぞれ $f_1, f_2, \dots, f_k$ とする。ここで、母集団の中から要素を1つ無作為抽出するとき、各要素が抽出される確率は等しいことから、 X が x_i となる確率 p_i は、

$$p_i = \frac{f_i}{N} \quad (i=1, 2, \dots, k)$$

であり、 X は確率変数とみなすことができる。

また、確率分布は右の表のようになる。

この確率分布を、母集団分布という。

母集団分布の平均、分散、標準偏差を、それぞれ母平均、母分散、母標準偏差といい、 m, σ^2, σ で表す。

X	x_1	x_2	$\cdots$	x_k	計
P	p_1	p_2	$\cdots$	p_k	1

例 18 1から4までの数字が1つずつ書かれたカードが、それぞれ数字の枚数だけある。この10枚のカードを母集団とし、この中から1枚を無作為抽出してそこに書かれた数字を X とする。
このとき、確率変数 X の母集団分布は右の表のようになり、
母平均 m ，母分散 σ^2 ，母標準偏差 σ は、それぞれ次のようになる。

X	1	2	3	4	計
P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	1

$$m = 1 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{4}{10} = \frac{30}{10} = 3$$
$$\sigma^2 = (1 - 3)^2 \cdot \frac{1}{10} + (2 - 3)^2 \cdot \frac{2}{10} + (3 - 3)^2 \cdot \frac{3}{10} + (4 - 3)^2 \cdot \frac{4}{10} = \frac{10}{10} = 1$$
$$\sigma = 1$$

問 24 1等500円が2本，2等300円が6本，3等100円が12本の合計20本のくじを母集団とする。
このくじを1本引いたときの賞金 X の母平均 m ，母分散 σ^2 ，母標準偏差 σ をそれぞれ求めよ。

組 番 氏名

母集団から抽出された変量 X についての大きさ n の無作為標本を,

$$(X_1, X_2, \dots, X_n)$$

と表す。このとき、 $X_1, X_2, \dots, X_n$ は確率変数である。

この標本が復元抽出によって得られたものであれば、 $X_1, X_2, \dots, X_n$ は互いに独立である。

一方、この標本が非復元抽出によって得られたものであれば、 $X_1, X_2, \dots, X_n$ は互いに独立でない。しかし、母集団の大きさが n に比べて十分に大きいときには、非復元抽出でも、

$X_1, X_2, \dots, X_n$ は互いに独立であるとみなしてよい。

これからは、母集団の大きさが十分に大きい場合を考える。したがって、ある母集団から抽出される大きさ n の無作為標本は、いずれも母集団の確率分布に従う n 個の互いに独立な確率変数の組であるとみなしてよい。

母平均と標本平均

母平均と、母集団から抽出した標本の平均の関係について調べてみよう。

母集団から大きさ n の無作為標本 $(X_1, X_2, \dots, X_n)$ を抽出するとき、

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

を標本平均という。標本平均 $\bar{X}$ は確率変数である。

Approach 標本平均 $\bar{X}$ の平均と標準偏差について考えてみよう。

母平均 m 、母標準偏差 σ の母集団から大きさ n の無作為標本 $(X_1, X_2, \dots, X_n)$ を復元抽出するとき、 $X_1, X_2, \dots, X_n$ は確率変数であり、これらは母集団の確率分布に従うから、

$$E(X_1) = E(X_2) = \dots = E(X_n) = m$$

$$V(X_1) = V(X_2) = \dots = V(X_n) = \sigma^2$$

このことから、 $\bar{X}$ の平均 $E(\bar{X})$ は、

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$
$$= \frac{1}{n}E(X_1 + X_2 + \dots + X_n)$$
$$= \frac{1}{n}\{E(X_1) + E(X_2) + \dots + E(X_n)\} = \frac{1}{n} \cdot nm = m$$

$$E(aX) = aE(X)$$

また、 $X_1, X_2, \dots, X_n$ は互いに独立であるから、 $\bar{X}$ の分散 $V(\bar{X})$ は、

$$V(\bar{X}) = V\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$
$$= \frac{1}{n^2}\{V(X_1) + V(X_2) + \dots + V(X_n)\} = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

$$V(aX) = a^2V(X)$$

$$= \frac{1}{n^2}V(X_1 + X_2 + \dots + X_n)$$

$$= \frac{1}{n^2}\{V(X_1) + V(X_2) + \dots + V(X_n)\} = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

したがって、 $\bar{X}$ の標準偏差 $\sigma(\bar{X})$ は、
$$\sigma(\bar{X}) = \sqrt{V(\bar{X})} = \frac{\sigma}{\sqrt{n}}$$

母集団の大きさが n に比べて十分大きいときは、非復元抽出でも上と同じことが成り立つと考えてよい。

前ページから、次のことが成り立つ。

標本平均の平均と標準偏

母平均 m 、母標準偏差 σ の母集団から大きさ n の無作為標本を抽出するとき、

1 標本平均 $\bar{X}$ の平均は、
$$E(\bar{X}) = m$$

2 標本平均 $\bar{X}$ の標準偏差は、
$$\sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

例 19 母平均60、母標準偏差15の母集団から大きさ100の標本を無作為抽出し、その標本平均を $\bar{X}$ とするとき、 $\bar{X}$ の平均と標準偏差は、

$$E(\bar{X}) = 60, \quad \sigma(\bar{X}) = \frac{15}{\sqrt{100}} = \frac{3}{2}$$

組番氏名

例題 5

赤玉と白玉が**1 : 19**の割合で多数入っている袋がある。この中から**400**個を無作為抽出するとき、***k***番目に抽出した玉が赤であれば**1**、白であれば**0**の値をとる確率変数を **$\mathbf{X}_k$** とする。この標本の標本平均 **$\mathbf{\bar{X}}$** の平均と標準偏差を求めよ。

問 25 ある国では、**10**人に**1**人の割合で左利きであることがわかっている。その国の**100**人を無作為抽出するとき、***k***番目に抽出した人が左利きであれば**1**、右利きであれば**0**の値をとる確率変数を **X_k** とする。この標本の標本平均 **$\bar{X}$** の平均と標準偏差を求めよ。

組 番 氏名

標本平均の分布

1から999までの奇数を1つずつ書いた500枚のカードから2枚を復元抽出で抜き出し、書いてある数の平均を計算する。

この操作を2000回繰り返して得た値の相対度数の分布を示したものが、右の図1である。

また、このカードを30枚抜き出して、同様の操作を2000回繰り返したときの相対度数の分布が、右の図2である。

このように操作の回数が十分に大きいとき、母集団分布がどのような分布でも、標本平均の相対度数の分布を表すヒストグラムの形は、標本の大きさを増やしていくと、正規分布曲線に近づいていくことが知られている。

一般に、次のことが知られている。

標本平均の分布

母平均 m 、母標準偏差 σ の母集団から抽出された大きさ n の標本の標本平均 $\bar{X}$ は、 n が大きいとき、正規分布 $N\left(m, \frac{\sigma^2}{n}\right)$ に従うとしてよい。

注 母集団分布が正規分布のときは、 n が大きくなっても、 $\bar{X}$ はつねに正規分布 $N\left(m, \frac{\sigma^2}{n}\right)$ に従うことが知られている。

注目 標本平均 $\bar{X}$ の標準偏差が $\frac{\sigma}{\sqrt{n}}$ となることから、 n が大きいとき、 $\bar{X}$ は母平均 m の近くに集中して分布するようになる。このことは、大数の法則と呼ばれている。

標本平均の標準化

標本平均は、標本の大きさが大きいとき正規分布に従うとしてよいことから、73ページで学んだように標本平均を標準化することによって、144ページの正規分布表を利用して標本平均の分布を調べることができる。

このことを用いて、問題を解いてみよう。

図1

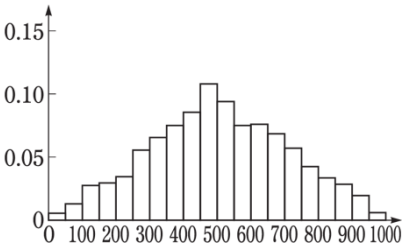
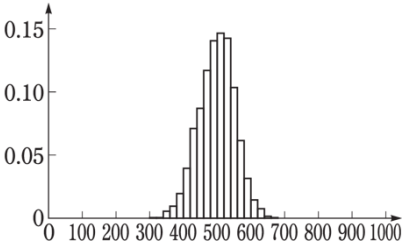


図2



例題 6

出荷前のみかん1個あたりの重さについて、母平均が100g、母標準偏差が16gであるという。ここから400個を無作為抽出するとき、この標本の標本平均が102g以上になる確率を求めよ。

問 26 電子マネーの1ヶ月間の利用金額について、母平均が24300円、母標準偏差が7000円となる母集団を考える。ここから大きさ400の標本を無作為抽出するとき、その標本平均が25000円以上になる確率を求めよ。

組	番	氏名
---	---	----

2 推 定

母平均などの母集団の分布の特性を示す定数の値が未知のときに、与えられた標本からその値を推定することを考えてみよう。

例えば、あるハンバーガーチェーンのMサイズのポテトを10個買い、重さを量ったところ、その平均は135.5gであった。

求めた標本平均の値から、母平均は135.5gに近いことが推測される。

そこで、母平均を幅をもって推定する方法を考えてみよう。

母平均の推定

母平均 m 、母標準偏差 σ の母集団から無作為抽出した大きさ n の標本の標本平均 $\bar{X}$ は、 n が大きいとき、正規分布 $N\left(m, \frac{\sigma^2}{n}\right)$ に従うとしてよい。

したがって、 $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ は、標準正規分布 $N(0, 1)$ に従うとしてよい。

正規分布表によると、

$P(|Z| \leq 1.96) = 0.95$ であるから、

$$P\left(|\bar{X} - m| \leq 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

よって、次の等式が成り立つ。

$$P\left(\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

すなわち、不等式

$$\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}} \quad \cdots \cdots \textcircled{1}$$

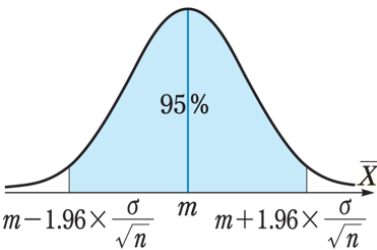
が成り立つ確率が0.95である。①で示される範囲を、母平均 m に対する信頼度95%の信頼区間といい、次のように表す。

$$\left[\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right] \quad \cdots \cdots \textcircled{2}$$

前ページのように、母平均 m について信頼度95%の信頼区間を求めることを、母平均 m を信頼度95%で推定するまたは区間推定するという。

例題 7

ある市で、**17歳**の男子**100人**を無作為抽出して身長を調べたところ、平均は**170cm**であった。この市の**17歳**の男子の身長の標準偏差を**6.5cm**として、平均身長を信頼度**95%**で推定せよ。



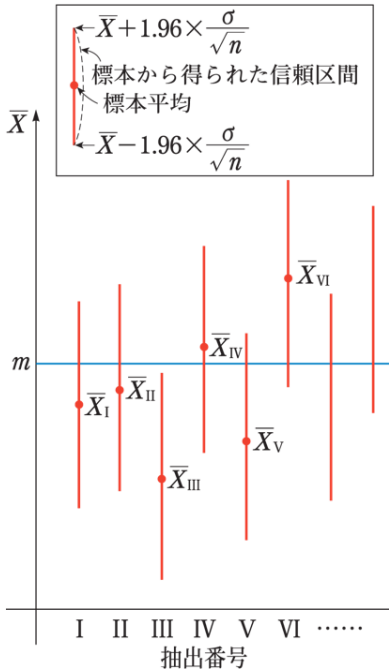
組	番	氏名
---	---	----

問 27 例題7の市において，改めて17歳の男子100人を無作為抽出して身長を調べたところ，平均は168cmであった。このとき，この市の17歳の男子の平均身長を信頼度95%で推定せよ。

標本平均 $\bar{X}$ は，抽出のたびに変わる変量であるから，信頼区間も標本抽出のたびに変わる。すなわち，右の図のⅠ，Ⅱ，……のように，標本抽出のたびに信頼区間は変わる。このとき，図の $\bar{X}_{Ⅲ}$ のように信頼区間が母平均 m を含まないこともある。

信頼度95%の信頼区間とは，95%の確率で信頼区間が m を含むという意味であり，100回の標本抽出において，95回くらいは求めた信頼区間が m を含むということである。

母標準偏差 σ の値がわかっているときは，母集団から大きさ n の標本を無作為抽出してそこから標本平均 $\bar{X}$ を求めれば，前ページの②により m の値を推定できる。



母標準偏差が未知のときの母平均の推定

ここまでは，母標準偏差 σ の値はわかっているものとしたが，実際には，わからない場合がほとんどである。しかし，標本の大きさ n が大きいときには， σ を標本の標準偏差の値 s でおき換えても，大きな違いは生じないことが知られている。このことから，次のことがいえる。

母平均の推定

標本の大きさ n が大きいとき，標本平均を $\bar{X}$ ，標本の標準偏差を s とすると，母平均 m に対する信頼度95%の信頼区間は，

$$\left[\bar{X} - 1.96 \times \frac{s}{\sqrt{n}}, \bar{X} + 1.96 \times \frac{s}{\sqrt{n}}\right]$$

同じように考えて，母平均 m に対する信頼度99%の信頼区間は，

$$\left[\bar{X} - 2.58 \times \frac{s}{\sqrt{n}}, \bar{X} + 2.58 \times \frac{s}{\sqrt{n}}\right]$$

組 番 氏名

例題 8

ある店で買い物をした客のうち、**200**人に買い物の金額を聞いたところ、平均金額は**2855**円、標準偏差は**600**円であった。この店の客**1**人当たりの平均買い物金額を信頼度**95%**で推定せよ。

問 28 ある市のガソリンスタンドのうち、**36**ヶ所でガソリン**1**リットル当たりの価格を調べたところ、平均価格は**129.7**円、標準偏差は**1.2**円であった。この市のすべてのガソリンスタンドにおけるガソリン**1**リットル当たりの平均価格を、信頼度**95%**で推定せよ。

注 例題8において、信頼度99%の信頼区間は、[2746, 2964]となる。

組 番 氏名

母比率の推定

母集団の中である性質Aをもつ要素の割合をその母比率といい，抽出された標本の中で性質Aをもつ要素の割合を標本比率という。

母比率 p を標本比率 R から推定することを考えてみよう。

性質Aの母比率が p である十分大きな母集団から，大きさ n の無作為標本 $(X_1, X_2, \cdots, X_n)$ を抽出し，それらに対して，

k 番目の標本が性質Aを持つ場合は， $X_k=1$

k 番目の標本が性質Aを持たない場合は， $X_k=0$

と定めると， $X=X_1+X_2+\cdots+X_n$ は，大きさ n の標本に含まれる性質Aをもつものの個数を表す確率変数であり，二項分布 $B(n, p)$ に従う。

さらに， n が大きいとき，75～76ページでみたように
 X は正規分布 $N(np, np(1-p))$ に従い，標本比率 $R=\frac{X}{n}$ は，

正規分布 $N\left(p, \frac{p(1-p)}{n}\right)$ に従うとしてよい。

$$E\left(\frac{X}{n}\right)=\frac{1}{n}E(X)$$
$$V\left(\frac{X}{n}\right)=\frac{1}{n^2}V(X)$$

したがって，母平均の推定と同様に考えると，母比率が p である母集団から無作為抽出した大きさ

n の標本の標本比率 R については， n が大きいとき， $P\left(|R-p|\leq 1.96\times\sqrt{\frac{p(1-p)}{n}}\right)=0.95$

が成り立つとしてよい。よって，

$$P\left(R-1.96\times\sqrt{\frac{p(1-p)}{n}}\leq p\leq R+1.96\times\sqrt{\frac{p(1-p)}{n}}\right)=0.95$$

さらに， n が大きいとき，標本比率 R は母比率 p に近いとみなしてよいことが知られている。

よって，上の根号の中に現れる p を R におき換えても差し支えないから，母比率 p に対する信頼度95%の信頼区間は，次のようになる。

$$\left[R-1.96\times\sqrt{\frac{R(1-R)}{n}}, R+1.96\times\sqrt{\frac{R(1-R)}{n}}\right]$$

例題 9

ある市の全世帯から**400**世帯を無作為抽出して，ある意見に対する賛否を調べたところ，**256**世帯が賛成であった。全世帯における賛成の母比率 p を信頼度**95%**で推定せよ。

	組	番	氏名
--	---	---	----

問 29 日本国内におけるある音楽グループを知っている人の割合を調べるために，無作為調査を行ったところ，500人の回答が得られ，そのうち50人がこの音楽グループのことを知っていた。日本国内におけるこの音楽グループを知っている人の割合を信頼度95%で推定せよ。

母平均の推定を利用した母比率の推定

性質Aの母比率が p である母集団に対して，Aをもてば1，もたなければ0をとる確率変数 X を考えると，平均 $E(X)$ は， $E(X)=0 \cdot (1 - p) + 1 \cdot p=p$ であり，母比率に等しい。よって，86ページのように，大きさ n の標本を用いて $E(X)$ を推定することで，母比率 p を推定することもできる。

X の標準偏差 σ は，

$$\sigma=\sqrt{E(X^2)-\{E(X)\}^2}=\sqrt{\{0^2 \cdot (1 - p) + 1^2 \cdot p\} - p^2}=\sqrt{p(1 - p)}$$

このことから，性質Aについての標本比率を R とすると， X について，

標本平均は， R 標本の標準偏差は， $\sqrt{R(1 - R)}$

よって，標本の大きさ n が大きいとき，88ページの母平均の推定より，性質Aの母比率 p に対する信頼度95%の信頼区間は，次のようになる。

$$\left[R - 1.96 \times \sqrt{\frac{R(1-R)}{n}}, \; R + 1.96 \times \sqrt{\frac{R(1-R)}{n}} \right]$$

組	番	氏名
3	仮説検定	

仮説検定

Approach 母集団についての仮説を検定することを考えてみよう。

全国の高校2年生を対象とする数学の試験の第1回の結果は、得点の平均が60点であった。第2回の平均はまだ出ていないが、受験した生徒100人を無作為抽出して得点を調べたところ、その平均は58.3点であった。第2回の母標準偏差を7.9点とすると、第2回の平均は、第1回の平均60点から変化したといえるだろうか。

これを直接証明するのは難しいから、背理法のようにまず、「第2回の平均は60点である」と仮定する。

86ページで学んだように、標本平均 $\bar{X}$ について、

$$|\bar{X} - 60| \leq 1.96 \times \frac{\sigma}{\sqrt{n}} \quad \cdots \cdots \textcircled{1}$$

が成り立つ確率は、0.95である。

ここで、 $n=100$ 、 $\sigma=7.9$ 、 $\bar{X}$ を58.3とすると、

$$(\textcircled{1} \text{の左辺}) = |58.3 - 60| = 1.7, \quad (\textcircled{1} \text{の右辺}) = 1.96 \times \frac{7.9}{\sqrt{100}} \cong 1.5$$

となり、 $\textcircled{1}$ が成り立たないから、無作為抽出によって確率が5%未満のことが起こったと考えられる。

これは非常に低い確率であるから、「平均は60点である」という仮定が否定できる。

すなわち、平均は60点から変化したと判断できる。

上ではじめに仮定した「第2回の平均は60点である」を帰無仮説といい、それを否定した「第2回の平均は60点でない」を対立仮説という。上のように、帰無仮説が否定できるかどうかを調べて、対立仮説を受け入れるかどうかを判断する統計的な手法を仮説検定という。

また、帰無仮説は H_0 、対立仮説は H_1 で表す。

帰無仮説は、標本平均や標本比率などの分布が扱いやすいものに設定するとよい。

標本平均や標本比率について、帰無仮説を否定してもよいと考えられる値の範囲をあらかじめ定めておく必要がある。この範囲を棄却域という。棄却域は、基準となる数値 α をもとにして、標本平均や標本比率が棄却域に入る確率が α となるように定める。この α を有意水準または危険率といい、0.05や0.01のような小さい値にとる。例えば、 $\alpha=0.05$ のとき、その仮説検定を有意水準5%の仮説検定という。

母集団から抽出した標本の標本平均や標本比率の値が棄却域にあれば、帰無仮説は否定できると判断する。これを帰無仮説を棄却するという。このとき、対立仮説を受け入れる。

標本平均や標本比率の値が棄却域になれば、帰無仮説は否定できないと判断する。これを帰無仮説を棄却しないという。

注 「帰無仮説を棄却しない」とは、帰無仮説を否定するだけの根拠が得られなかったという意味であり、帰無仮説が肯定されるということではない。

正規分布を利用した母平均の仮説検定

86ページで学んだように、母平均 m 、母標準偏差 σ の母集団から無作為抽出した大きさ n の標本の標本平均 $\bar{X}$ について、

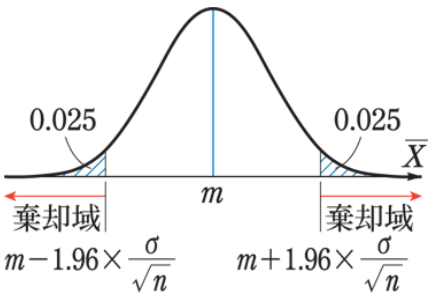
$$|\bar{X} - m| \leq 1.96 \times \frac{\sigma}{\sqrt{n}}$$

が成り立つ確率は0.95である。

よって、帰無仮説として母平均が m であると仮定したとき、標本平均 $\bar{X}$ についての有意水準5%の棄却域は、

$$|\bar{X} - m| > 1.96 \times \frac{\sigma}{\sqrt{n}}$$

を満たす範囲とする。このように、棄却域を分布の両側に設定するような検定を両側検定という。



88ページと同様に、標本の大きさ n が大きいとき、母標準偏差 σ を標本の標準偏差 s でおき換えても大きな違いは生じないため、棄却域は、 $|\bar{X} - m| > 1.96 \times \frac{s}{\sqrt{n}}$ を満たす範囲と考えることができる。

組 番 氏名

例題 10
出荷予定の卵から**100**個を無作為抽出し，重さを調べたところ，平均は**59.6g**，標準偏差は**2.0g**であった。この卵全体の**1**個当たりの重さの平均は**60g**といえるか。帰無仮説**H₀**を「平均は**60g**である」，対立仮説**H₁**を「平均は**60g**でない」として，有意水準**5%**で検定せよ。

問 30 ある工場で作られているボルトの直径は**7mm**であるという。ある日，作られたボルト**256**個を調べたところ，直径の平均は**7.01mm**，標準偏差は**0.02mm**であった。この工場で作られているボルトの直径は**7mm**といえるか。帰無仮説**H₀**を「直径は**7mm**である」，対立仮説**H₁**を「直径は**7mm**でない」として，有意水準**5%**で検定せよ。

注 前ページと同様に，88ページで学んだことから，有意水準**1%**の棄却域は， $|\bar{X} - m| > 2.58 \times \frac{s}{\sqrt{n}}$

を満たす範囲とする。

よって，例題10において，有意水準を**1%**とすると，棄却域は， $|\bar{X} - 60| > 2.58 \times \frac{2}{\sqrt{100}} = 0.516$

を満たす範囲である。 $\bar{X}$ に**59.6**を代入すると， $|59.6 - 60| = 0.4 < 0.516$ であるから，帰無仮説は棄却されなくなる。

組 番 氏名

正規分布を利用した母比率の仮説検定

89ページで学んだように、母比率が p である母集団から無作為抽出した大きさ n の標本の標本比率 R について、 n が大きいとき、

$$|R - p| \leq 1.96 \times \sqrt{\frac{p(1-p)}{n}}$$

が成り立つ確率は0.95である。

よって、帰無仮説として母比率が p であると仮定したとき、標本比率 R についての有意水準5%の棄却域は、

$$|R - p| > 1.96 \times \sqrt{\frac{p(1-p)}{n}}$$

を満たす範囲とする。

このことを利用して、仮説検定を行ってみよう。

例題 11

ある植物の種子は、発芽率が**45%**であるといわれている。この種子を**400**個まいたところ、**160**個が発芽した。この種子の発芽率は**45%**であるといえるか。帰無仮説 H_0 を「発芽率は**45%**である」、対立仮説 H_1 を「発芽率は**45%**でない」として、有意水準**5%**で検定せよ。

問 31 ある農場で生産されるメロンは**20%**が規格外であるといわれている。このメロン**100**個を無作為抽出して調べたところ、**12**個が規格外であった。このメロンの規格外の比率は**20%**であるといえるか。帰無仮説 H_0 を「比率は**20%**である」、対立仮説 H_1 を「比率は**20%**でない」として、有意水準**5%**で検定せよ。

組 番 氏名

応用例題 12

あるさいころを**100**回投げたところ、**1**の目が**15**回出た。このさいころは、**1**の目の出る確率が $\frac{1}{6}$ であるといえるか、有意水準**5%**で検定せよ。

問 32 ある検定の合格率は**40%**だといわれている。ある回の受検者から無作為抽出した**400**人に合否を確認したところ、**150**人が合格していた。この検定の合格率は**40%**であるといえるか、有意水準**5%**で検定せよ。

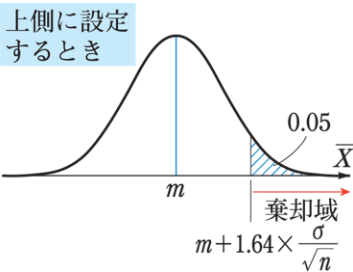
組 番 氏名

研究 片側検定

検定を行う際、棄却域を分布の片側に設定することもある。このような検定を片側検定という。片側検定では、92ページの標本平均 $\bar{X}$ について有意水準5%の場合、

$$|\bar{X} - m| > 1.64 \times \frac{\sigma}{\sqrt{n}}$$

を満たす範囲のうち、棄却域を上側に設定するときは m の右側の部分、下側に設定するときは m の左側の部分とする。



例題

あるハンバーガーチェーンでは、**M**サイズのポテトの重さが**135g**であると公表している。そこで、このポテトを**64**個買い、重さを調べたところ、平均は**136.6g**、標準偏差は**7.2g**であった。このポテトの重さの平均は、**135g**より重いといえるか。帰無仮説 H_0 を「平均は**135g**である」、対立仮説 H_1 を「平均は**135g**より重い」として、有意水準**5%**で片側検定せよ。

問題 ある宅配サービスの利用者**100**人を無作為抽出し、配達時間を調べたところ、平均は**25**分、標準偏差は**18**分であった。この宅配サービスの配達時間の平均は**30**分より短いといえるか。帰無仮説 H_0 を「平均は**30**分である」、対立仮説 H_1 を「平均は**30**分より短い」として、有意水準**5%**で片側検定せよ。

組	番	氏名
---	---	----

4 標本調査の活用

標本調査の例として、テレビの視聴率を取り上げてみよう。

テレビの視聴率には、世帯視聴率と個人視聴率があり、それぞれテレビを所有する世帯の標本調査と、その世帯の4歳以上の個人を対象とする標本調査を行っている。ここでは、番組ごとの世帯視聴率や個人視聴率を、標本の中でその番組を視聴していた世帯や人数の割合とみなして問題を考えたり、自分で標本調査を設計したりしてみよう。

例 20 ある番組の関西地区での世帯視聴率が10%のとき、関西地区全体の世帯視聴率について推定する。

関西地区の標本世帯数は1200世帯である。

標本比率は0.1であるから、母比率に対する信頼度95%の信頼区間は、

$$\left[0.1 - 1.96 \times \sqrt{\frac{0.1(1-0.1)}{1200}}, 0.1 + 1.96 \times \sqrt{\frac{0.1(1-0.1)}{1200}}\right]$$

よって、 [0.083, 0.117]

すなわち、信頼度95%で、関西地区全体の世帯視聴率は8.3%以上11.7%以下と推定される。

問 33 関東地区の標本世帯数は2700世帯である。ある番組の関東地区での世帯視聴率が10%のとき、関東地区全体の世帯視聴率を信頼度95%で推定せよ。

問＋34 例20において、世帯視聴率が1%上がるとき、例20の結果と比較せよ。

Aさんのまわりでは、あるテレビ番組が流行っており、Aさんの知り合いのおよそ半分がその番組を見ている。Aさんは、自分の通う学校で、全校生徒1500人の半数がその番組を見ているかを、全校生徒数は十分に大きいと仮定して検定することにした。

	組	番	氏名
--	---	---	----

問 35 Aさんは，全校生徒から100人を無作為抽出し，その番組を見ているかを聞いたところ，53人が見ていると回答した。この結果から，全校生徒の50%がその番組を見ているといえるか，有意水準5%で検定せよ。

応用例題 13

前ページの番組について，全校生徒の個人視聴率を信頼度**95%**で推定する。標本比率 **R** が**50%**であると仮定するとき，全校生徒から何人以上を無作為抽出して調べれば，信頼区間の幅が**0.1**，すなわち**10%**以下になるかを求めよ。ただし，信頼区間[**a** , **b**]の幅とは， **$b - a$** のことである。

	組	番	氏名
--	---	---	----

問 36 応用例題13において，信頼区間の幅を0.1からその半分の0.05にすると，抽出する必要がある人数はおよそ何倍になるか。また，さらに半分の0.025にすると人数はおよそ何倍になるか。

問+37 Aさんは，全校生徒の中から400人を無作為抽出して，ある番組を見ているか調べたところ，216人が見ていると回答した。一方で，新聞で公表されたその番組の個人視聴率は16%であった。このようなことが起こる原因を考えよ。

いろいろな番組の視聴率を，インターネットなどを使って調べてみよう。そして，標本調査を設計し，自分のクラスや学年，学校などでの視聴率を調べて比べてみよう。

また，視聴率以外にも標本調査が活用できるものがないか探してみよう。

組 番 氏名

1 1, 2, 3の数字を1つずつ書いたカードが, それぞれ3枚, 2枚, 5枚ある。この10枚のカードを母集団とするととき, カードの数字について次の問いに答えよ。

数字	1	2	3	計
度数	3	2	5	10

- (1) 母平均 m と母標準偏差 σ を求めよ。
- (2) この母集団から, 大きさ4の標本 X_1, X_2, X_3, X_4 を復元抽出するとき, その標本平均 $\bar{X}$ の平均 $E(\bar{X})$ と標準偏差 $\sigma(\bar{X})$ を求めよ。

- 2 ある高校で, 全校生徒から144人を無作為抽出して身長を測定したところ, 平均は169.0cm, 標準偏差は7.0cmであった。この高校の全校生徒の平均身長を信頼度95%で推定せよ。

組

番

氏名

3 ある工場で作った製品**400**個のうち、**12**個が不良品であった。この工場で作った全製品における不良品の比率を、信頼度**95%**で推定せよ。

4 ある店で売られているメダルは重さが**10g**であるという。このメダルを**196**個購入し、重さを調べたところ、平均は**10.05g**、標準偏差は**0.04g**であった。この店のメダルの重さの平均は**10g**といえるか、有意水準**5%**で検定せよ。

組	番	氏名
---	---	----

- 5 ある一定の飼育条件に置いたテントウムシの卵の 孵 化率は，55%といわれている。この飼育条件のもとで，テントウムシの卵100個のうち50個が孵化した。このとき，孵化率は55%であるといえるか，有意水準5%で検定せよ。