この本の特徴

私は、“大人のための数学教室「和」”という塾で社会人を相手に数学を教えています。ここでは統計の**検定・推定の仕組みや意味がわからないから教えて欲しい**という人が何人も門を叩きます。社会に出てから必要に迫られて、統計学を学ばなければならなくなった方たちです。

今は、AI が進化しましたから、然るべき言葉を添えて ChatGPT にデータを渡せば、検定・推定の計算方法を説明したあと、計算を実行し、結果を文章の形で回答してくれます。しかし、検定・推定の意味をぼんやりとはわかっているものの、なぜそうなるかの仕組みまではよくわからないというのです。仕組みがわからないので、ソフトが出した結果をどのようなスタンスで捉えたらよいのか態度を決めかねる人もいるようです。

そこでこの本では、統計学のなかでも、検定・推定の概念を理解してもらうことを中心テーマに据えることにしました。

この本の特徴は、

「数学が苦手な人でも、検定・推定の本質を 1 時間で理解できる」
ように書かれていることです。

そもそもこの本は、300 ページ以上もあるのに、1 時間で読むことなんてできないんじゃないの、とお考えの向きもあるでしょう。

1 時間で読んでもらうのは、第 1 章のことなんです。

この 1 章を読めば、1 時間で検定・推定の本質がわかります。しかも、難しい数学を使わないで。この 1 章の完成度は自分でもかなりイケてると思っています。この 1 章だけを切り離して、「1 時間で検定・推定がわかる本」と題して出版してもよいぐらいの内容です。表紙に鉢巻をした著者

の顔写真を載せればベストセラーになるでしょう…。

以下で、「1時間で検定・推定がわかる」ための1章での工夫を紹介しておきます。

① 中学までの数学しか使わない

まず、数式による表現は極力省きました。平均、分散・標準偏差の計算をするにも、まず公式を与えるのではなく、数値計算で例を示してから、最後に数式を使ってそのまとめをしています。ここで出てくる数学は、平方根までです。つまり、中学数学で十分にお釣りがくるというわけです。まとめでも、Σ記号を使わないようにして、ストレスを減らしました。

② 確率変数を用いない

統計学の本には、タイトルに「確率統計」とついたものがあるくらいです。本来であれば、統計学を講義するには、確率の公式や確率変数の運用が欠かせません。しかし、数学の苦手な人にとっては確率は鬼門で、統計学を学ぶ際の難所となってしまっているのです。ここで挫折してしまうので、その先にあるお宝にありつけないでいるのです。もったいないことではありませんか。

そこで1章では、確率の公式や確率変数、確率密度関数を用いないで、検定・推定を説明することにしました。検定・推定の原理は、確率の素朴な応用です。確率を形式的に表現する確率変数を用いずとも理解できることなのです。

なお、第2章では、確率の公式や確率変数、確率密度関数を用いてもう一度検定・推定を説明しています。大学で統計学の講義を受けている方、統計検定2級を目指している方は、こちらまで読んでください。

③ 統計と確率が結びついてわかる

正当な統計学の本では、平均・標準偏差などの説明をしたあと、確率変数を導入し、確率分布、検定・推定と説明が進みます。単純化すると統計→確率→統計と話題が進むわけです。とおりいっぺんの説明しかしていないと、確率の説明がどうしても浮いてしまいます。データの平均・分散と確率変数の平均・分散が結びつかないまま、検定・推定の説明へ突入するので、検定・推定をするための標本が確率試行の結果であるという捉え方ができなくなってしまうのです。

④ 本筋でないものは脇へ

検定・推定を1時間で理解するために、少しでもその理解に必要ないと判断した概念については、コラムなど脇で解説するようにしました。例えば、$V(X)=E(X^2)-\{E(X)\}^2$ の公式、モード、最頻値など記述統計に関する用語、偏差値、不偏分散はなぜ $n-1$ で割るのか、などの話題です。

また、内容的には重要であっても、検定・推定の概念を大づかみに捉えるにはむしろ邪魔であると判断したものや高校以上の数学が必要になるものについては、2章にまわしました。確率変数や正規分布以外の分布や確率密度関数の話題です。

このようにして、1章で本質を理解してもらったあと、余力がある人、数式での表現を学んでみたい人、数式での裏付けがほしい人は、2章に読み進んでいただきたいと思います。

2章では、1章の内容を数学的に叙述して、補足していきます。といっても、すべてにきちんとした証明をつけていくわけではありません。事実だけ紹介することも多いでしょう。なにしろ、統計処理の意味がわかって使えるようになってもらうのがこの本の目的なのですから。

そして、第3章で、記述統計の1分野である相関係数や回帰直線の求め方を紹介します。ここでの工夫は、偏微分を用いず回帰直線の公式の導出をしたところです。

最後にもうひとつ、この本の目玉を。

⑤ 疑問に思うところがわかりやすく解説されている

疑問に思うところというのは具体的にいうと、

- 不偏分散はなぜ $n-1$ で割るの？
- 自由度って何？
- なぜ相関係数は-1以上1以下なの？

などです。これらを単に言葉だけでもなく、程よく数式を交えながらわかりやすく解説しています。どれも、本格的な解説書では、全く解説していなかったり、逆に難しい解説がなされている箇所です。

最後に本書を読む注意として、記号「＝」についてコメントしておきます。数学的にはおよその数を表す「≒」を使うべきところでも、「＝」を用いて表現しています。有効数字についても統一をとってありません。よろしく解釈してください。

第 1 章

UNIT

UNIT

1　度数分布表・ヒストグラム

1-1　データをグラフに整理しよう
――度数分布表、ヒストグラム

度数分布の表やグラフって、どう描くの？

これからデータ（資料）の整理の仕方とデータ（資料）の特徴を捉える数値の導き方を紹介しましょう。

ある中学校のクラスでテストをしました。20 人のテストの結果が次のようだったとします。このデータを整理してみましょう。

2人　3人　5人
43、47、52、52、54、61、67、67、68、69
6人　3人　1人
70、71、71、73、76、78、82、84、84、91

何十点台が何人いるか調べてみます。
40 点台は 2 人います。点数が 40 点台であれば、点数が 40 点以上 50 点未満として表に整理しましょう。

階級	階級値	度数
40 以上　50 未満	45	2
50 以上　60 未満	55	3
60 以上　70 未満	65	5
70 以上　80 未満	75	6
80 以上　90 未満	85	3
90 以上 100 以下	95	1

ここで、この例を参考にして、統計学の用語を紹介します。

○ サイズ（大きさ）
この場合は、調査対象の人数、データに含まれる個数、20 人のことです。本書では、分かりやすくデータの個数といったりします。

○ 変量
この場合は、テストの点数。データの各々の値です。

○ 階級（class）
40 点以上 50 点未満といった範囲のこと。

○ 階級値（midpoint）
40 点以上 50 点未満の階級であれば、40 点と 50 点の平均の 45 点です。midpoint という英語からも分かるように階級の範囲の真ん中の値です。

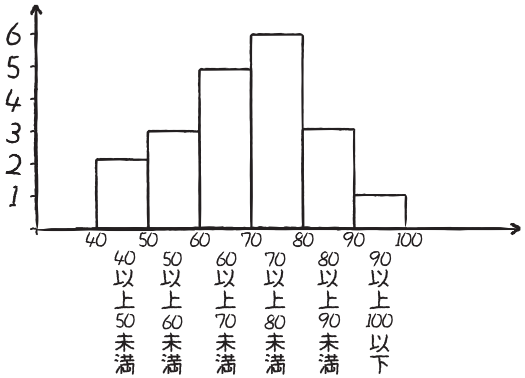
○ 階級幅（class interval）
この例の場合は 10 点。隣どうしの階級値の差ともいえます。

○ 度数（frequency）
この場合、階級ごとの人数のこと。

○ 個体
変量を持っている調査対象。この例では試験を受けた人。

データを階級ごとに整理して度数を調べたものを**度数分布表**（frequency table）といいます。

この表をもとに各階級の度数を柱状グラフで表すと次のようになります。



このように、度数分布表を柱状グラフにして表現したものを**ヒストグラム**（histogram）といいます。柱状グラフで大きい順に並べ替えている図がありますが、これはヒストグラムではありません。ヒストグラムはヨコ軸が階級順に並んでいなくてはいけません。

ヒストグラムを描くとデータの特徴がひと目で分かり便利です。ですから、度数分布表を示すかわりに、「次のヒストグラムのように分布しています」などと示す場面が多くあります。

ただ、このままではデータどうしを比較するときに不便です。人数の異なるA中学校とB中学校の得点分布を比べたいときは、階級に含まれる度数の全体に対する割合を比較した方がよいでしょう。

データ全体に対する度数の割合を**相対度数**（relative frequency）といいます。

度数分布表をもとにして、相対度数を表に書き込むと次のようになります。このような表を相対度数分布表といいます。

$\frac{2}{20}$

階級	階級値	度数	相対度数	累積相対度数
40以上　50未満	45	2	0.1	0.1
50以上　60未満	55	3	0.15	0.25
60以上　70未満	65	5	0.25	0.5
70以上　80未満	75	6	0.3	0.8
80以上　90未満	85	3	0.15	0.95
90以上 100以下	95	1	0.05	1
		20	1	

相対度数は、データのサイズを1としたとき、階級の度数を割合で表した数です。データのサイズ（データの個数）が20で、40点以上50点未満の階級の度数は2であることから、40点以上50点未満の階級の相対度数は、

$$\frac{2}{20} = 0.1$$

となります。

これに対して、相対度数を足しあげたものを累積相対度数といいます。60点以上70点未満の累積相対度数は、表の◯のようにそれより低い階級の相対度数を足しあげて、

$$0.1 + 0.15 + 0.25 = 0.5$$

と計算します。

累積相対度数の最後の欄は1になっていますね。

これは偶然ではありません。相対度数をすべて足しあげると、つねに1になるからです。相対度数とは、データのサイズを1としたときの割合なのですから、すべての相対度数を足すと全体の割合1になるわけです。

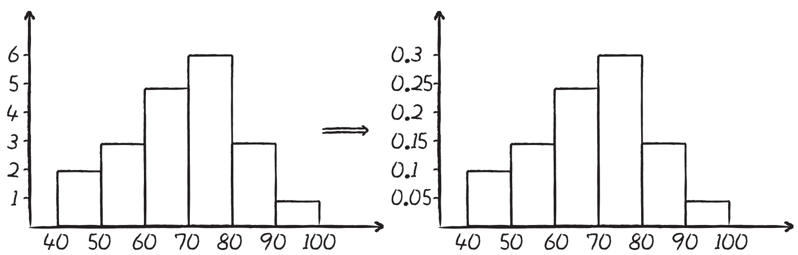
相対度数、累積相対度数のことを抽象的な書き方でまとめておきます。

データのサイズが N です。度数 f_1、f_2、…、f_n の総和が N になっています。

階級	階級値	度数	相対度数	累積相対度数
第1階級	x_1	f_1	$\frac{f_1}{N}$	$\frac{f_1}{N}$
第2階級	x_2	f_2	$\frac{f_2}{N}$	$\frac{f_1+f_2}{N}$
第3階級	x_3	f_3	$\frac{f_3}{N}$	$\frac{f_1+f_2+f_3}{N}$
⋮	⋮	⋮	⋮	⋮
第 n 階級	x_n	f_n	$\frac{f_n}{N}$	1
		N	1	

相対度数表をもとにヒストグラムを描きましょう。

度数のヒストグラムと相対度数のヒストグラムはほとんど同じではないかと思った方は勘のよい方です。タテ軸の目盛りを、度数から相対度数に振り直すだけで、相対度数のヒストグラムが完成します。



COLUMN データの代表値

本編では、データの特徴を表す重要な指標であっても、この本で説明する検定・推定には必要ないということで割愛してしまったものがあります。ここで、それらをフォローしておきます。

以下のような点数のデータを例にそれらを紹介していきましょう。

43、47、52、52、54、61、67、67、68、69、
71、71、71、73、76、78、82、84、84、91

○ 範囲（レンジ）range

データの最大値と最小値との差。この場合は、91 − 43 = 48

○ 最頻値（モード）mode

度数分布表から最頻値を求める場合は、度数の最も多い階級の階級値。階級幅を10点にとって、40点以上50点未満、…とするとき、一番多いのは70点台だから、階級値をとって75。

度数分布表にまとめることなく最頻値を求める場合は、データを調べると、71の度数が3で一番大きいので71です。

○ 中央値（メジアン）median

データを大きさの順に並べたときに、真ん中の順位の値。この場合はデータのサイズが20個と偶数なので、10番目の69と11番目の71、この2個の平均をとって、(69 + 71) ÷ 2 = 70

もしも、最後の91を取り除いて、データのサイズを19個とすれば、10

UNIT

4 検定の考え方

4-1 その仮説は正しいの？

——検定の原理と論法の型

こんなコトが起こるなんてウソだろ

この節では、統計学の重要な手法である「検定」の原理について例を挙げて説明します。検定という用語より、**仮説検定**（**hypothesis testing**）の方が中身を表してふさわしいですが、この本では簡潔に「検定」を使っていきます。

物語風に検定の原理について解説してみましょう。

問題 **1.13**　二項検定

あなたはC国を旅行中、市場で次のような看板を出している店を見つけました。

「コイン投げ　当てたら＄100進呈。1回＄10」

（1回＄10払って、店主の投げたコインの表裏を当てれば＄100もらえる）

すぐには参加せず様子を見ていると、6人の客がこの賭けに参加して、全員がコインの表裏を当てることができませんでした。あなたは、この状況をどう考えますか。

「公正なコイン投げであれば、
6人が連続で負けるなんてことはあり得ない。
きっと、いかさまが行なわれているんだろう」

こう思うことができれば、もうあなたは立派に検定の論法を身につけているといえます。検定で用いている論法は、

「ある仮定（前提）のもとで、
ありえそうもないことが起こったとき、
その仮定（前提）を否定する」

という論法なのです。コイン投げの例に沿って合わせて書くと、

「公正なコイン投げであれば、
ある仮定（前提）のもとで、
6人が連続で負けるなんてことはあり得ない。
ありえそうもないことが起こった
きっと、いかさまが行なわれているんだろう」
その仮定（前提）を否定する

この例に沿って、少し統計学の用語を紹介しておきましょう。

コイン投げの例において、コイン投げが公正に行なわれているということが仮定（前提）です。この仮定（前提）は、ありえそうもないことが起こった（6回連続で負ける）ので否定されました。

このようなあとで否定される仮定を、統計学では**帰無仮説**（**null hypothesis**）と呼びます。最終的に否定され、無に帰することが期待されている仮定だからです。

「ありえそうもないこと」は、数学の言葉で「起こる確率が小さいこと」と言い換えられます。起こる確率が小さいことが起こったときに、帰無仮

説を否定するわけです。「確率が小さい」では漠然としていますから、帰無仮説を疑う・疑わないの基準を数値で設定します。これを**有意水準**(**significance level**)といいます。

この本では、有意水準を一律5%に設定して説明します。検定が扱う題材によっては、有意水準をもっと小さくしなければなりません。例えば、新薬や治療の臨床実験や物理実験では有意水準を1%にする場合があります。

仮定(前提)を「否定する」ことを統計学では「棄却する(reject)」といいます。

検定で用いる論法を、統計学の用語を用いて述べてみると、

「帰無仮説を設定し、
　確率5%以下のことが起こったとき、
　帰無仮説を棄却する」

となります。

コイン投げの例では、6回連続で客が負けています。公正なコインであれば、当たる確率と当たらない確率はそれぞれ2分の1ですから、6回連続で負ける確率は、

$$\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{64}=0.015625 \qquad 約1.6\%$$

です。

コイン投げの初めの推論をまとめると、

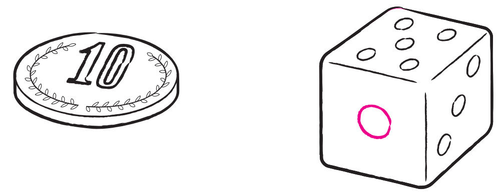
「公正なコイン投げで、
　帰無仮説：外れる確率は2分の1
　6人が連続で負けた。
　確率1.6%のことが起こった

　きっと、いかさまが行なわれているんだろう」
　帰無仮説を棄却⇒コイン投げは公正ではない

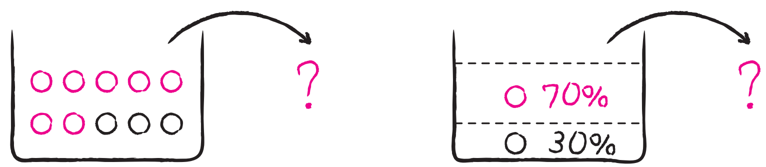
○母比率の検定

以上で、原理はお分かりいただけたと思います。
さらに、続く問題で検定を詳しく説明していきます。
その前に、簡単な例で確率のトレーニングをしておきましょう。

コインを投げて表が出る確率は2分の1です。1から6までの目が書かれている立方体のサイコロを投げて5の目が出る確率は6分の1です。



箱の中に7個の赤玉と3個の白玉、合計10個の玉が入っています。この箱の中から1個の玉を取り出すとき、取り出した玉が赤玉である確率は $\frac{7}{10}$ になります。



箱の中に何個か玉が入っています。総数は分かりませんが、70％が赤玉で、30％が白玉であることが分かっています。この中から1個の玉を取り出すとき、取り出した玉が赤玉である確率は70％、すなわち $\frac{7}{10}$ になります。

つまり、取り出した玉が赤玉である確率は、箱の中の赤玉と白玉の構成比によって決まります。

しかし、10個の玉を取り出したからといって、つねにそのうちの7個が赤玉であるわけではありません。10個の70％は7個ですが、必ずしもそうはならないことに注意しましょう。

ここまで確認したら次の問題に進みましょう。

問題 1.14　母比率の検定

ある国の大統領の支持率は70％であると発表されています。そこで、電話番号を適当に押して20人に電話を掛け、大統領を支持するか否かを聞いてみました。すると、大統領を支持すると答えた人は10人でした。支持率が70％であることを検定してください。

検定の論法は、「帰無仮説を設定し、確率5％以下のことが起こったとき、帰無仮説を棄却する」でした。

この問題の場合、帰無仮説は「支持率が70％である」です。

この仮定（帰無仮説）のもとで、「20人中10人が大統領支持」となることを確率で捉えてみましょう。

「支持率が70％」である場合、適当にかけた電話の相手が支持と回答する確率は70％$\left(\frac{7}{10}\right)$、不支持と答える確率は30％$\left(\frac{3}{10}\right)$です。

すると、20人に電話して、n 人が支持と答える確率は、

$$ {}_{20}\mathrm{C}_{n}\left(\frac{7}{10}\right)^{n}\left(\frac{3}{10}\right)^{20-n} $$

と計算できます。

${}_{20}\mathrm{C}_{n}$ が分からない人も気にせず読み進めてください。${}_{20}\mathrm{C}_{n}$ の計算の仕方、なぜこの式で確率を求めることができるかは2章で説明します。

例えば、$n=18$ のとき、

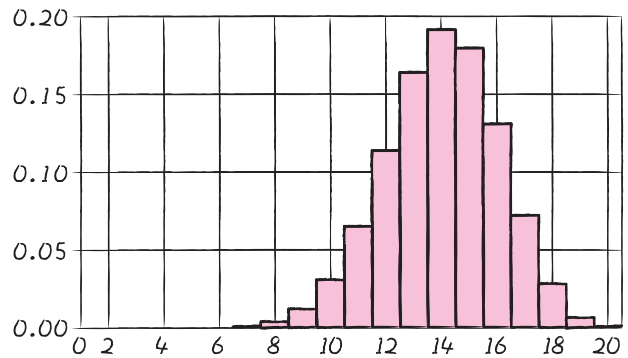
$$ {}_{20}\mathrm{C}_{18}\left(\frac{7}{10}\right)^{18}\left(\frac{3}{10}\right)^{20-18}=0.0278 \quad \rightarrow \quad 2.78\% $$

すなわち、20人中18人が支持と答える確率は2.78％と求まります。

n に0から20までの数を入れて確率を計算し、ヒストグラムを作ると次のようになります。n が0から5までは0.005％未満なので表では省略しました。

n	6	7	8	9	10	11	12	13
確率（％）	0.02	0.10	0.39	1.20	3.08	6.54	11.44	16.43

14	15	16	17	18	19	20
19.16	17.89	13.04	7.16	2.78	0.68	0.08



一番確率が大きい場合は、$n = 14$ です。これは $20 \times 0.7 = 14$ ですから納得がいきます。このヒストグラムの形は、$n = 14$ を中心にして左右に山型に広がっています。

ありそうもない確率が 5%未満になる場合を設定しましょう。ここでは、n が小さい場合と n が大きい場合で 2.5%未満ずつ設定します。

n が小さい方から設定します。

$n = 0$ から $n = 9$ までの％を足すと、

$0.02 + 0.10 + 0.39 + 1.20 = 1.71\%$

（$n = 0$ から $n = 5$ まではほぼ 0 なので $n = 6$ からしか書いていません）

$n = 0$ から $n = 10$ までの％を足すと、

$0.02 + 0.10 + 0.39 + 1.20 + 3.08 = 4.79\%$

となり 2.5%を越えますから、小さい方は $n = 0$ から $n = 9$ までとします。

次に大きい方で 2.5%未満の場合を設定します。

$n = 19$、$n = 20$ のときの％を足すと、

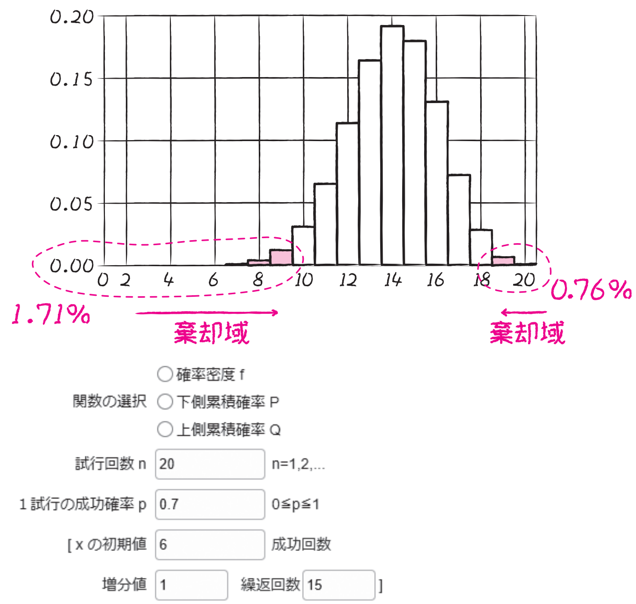
$0.68 + 0.08 = 0.76\%$

$n = 18$ から $n = 20$ までの％を足すと、

$2.78 + 0.68 + 0.08 = 3.54\%$

となり 2.5%を越えますから、大きい方は $n = 19$ から $n = 20$ までとします。

つまり、n が「0 から 9 まで、または 19 から 20 まで」になる確率は 5%未満です。「0 から 9 まで、または 19 から 20 まで」をヒストグラムに書き込むと、次のようになります。



n が 0 から 9 まで、または 19 から 20 までの数になったときに、ありえそうもないことが起こったとして、帰無仮説「支持率は 70%」を棄却することにします。「0 から 9 まで、または 19 から 20 まで」を**棄却域（rejection region）**といいます。

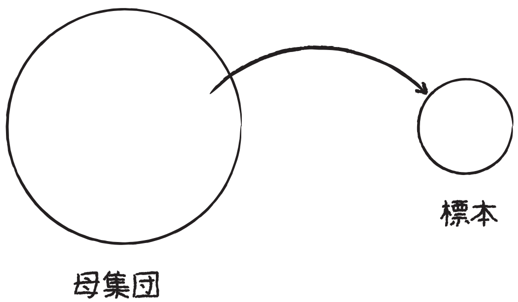
この問題の場合、$n = 10$ ですから帰無仮説「支持率は 70%」は棄却されません。このようなとき、帰無仮説を「**受容する（accept）**」または帰無仮説は「**棄却できない（fail to reject）**」といいます。

このように、検定では棄却域を設定し、n の実現値が棄却域に入るか否かで帰無仮説を棄却するか否かを決めます。

ここまでで、この問題の検定の仕組みはお分かりいただけたと思います。この問題を通して、統計学の用語を紹介していきましょう。

この問題では、ある国の大統領の支持率を調べるために、国民の一部（20人）を取り出してその支持率について検討しています。このように全体についての情報を知りたいとき、その一部を取り出して調べることは日常でもよくあることです。

例えば、料理のときの味見や交通違反取り締まりなどがそうです。全体を調べることが不可能な場合、また全体を調べるには多大なコストがかかる場合は、その一部を取り出して、全体の状況を推測することになります。



統計学では、調べたい全体を**母集団**（**population**）、調べるために取り出した一部を**標本**（**sample**）といいます。この問題の場合、ある国の国民全員が母集団、電話調査をした20人が標本です。

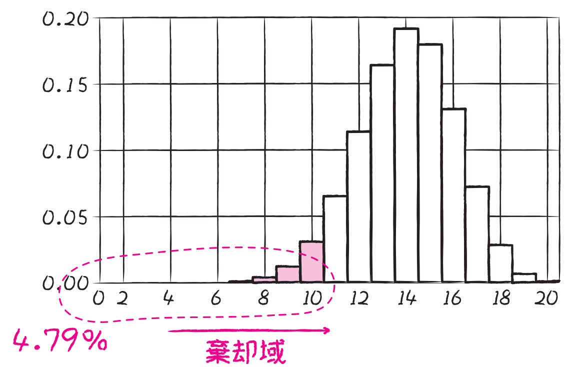
支持率は、国民全体のうち大統領を支持する国民の比率です。母集団に含まれる個体のうち、ある特徴を持つ個体の比率を**母比率**といいます。この問題1.14は、母比率を検定する問題です。

ちなみに、母集団の平均を**母平均**、分散を**母分散**といいます。

検定のように、母集団の統計的特徴を標本から把握する手法を扱う統計学を**推測統計学**（**inferential statistics**）といいます。推測統計学が扱う手法は、検定の他に標本から母集団の平均や分散・標準偏差を予測する「推定」（次節で扱う）があります。

この問題で、nの棄却域は、「0から9または19から20」というように、小さい方と大きい方に分かれていました。このように、左右両側に棄却域をとって検定することを**両側検定**（**Two-tailed test**）といいます。両側に棄却域をとった理由は、$p = 0.7$を否定したとき、真のpの値は小さい方に外れる可能性もあるし、大きい方に外れる可能性もあるからです。それで、小さい方、大きい方それぞれに2.5％未満になるように棄却域をとったのです。

もしも、支持率が$p = 0.7$よりも小さいことが確実で、小さい方だけで5％未満になるように棄却域をとるとすれば、nを0から10までにしたときの和が4.79％ですから、0から10までを確率5％未満となる棄却域としてとることができます。このように一方に棄却域をとって検定することを**片側検定**（**one-sided test**）といいます。



この問題の場合、$n = 10$ですから、nの小さい方に棄却域をとって有意水準5％の片側検定をすると帰無仮説は棄却されます。両側検定のときと異なる検定の結果となります。

両側検定にするか、片側検定にするかは実際にはナーバスな問題です。帰無仮説を棄却したいがために、あとから片側検定で判断することはフェアではありません。