

『まずはこの一冊から 意味がわかる統計学』

## もくじ

はじめに	3
もくじ	8
この本の特徴	13

## 第1章

1 相対度数分布グラフ	19
1-1 資料をグラフに整理しよう	19
—度数分布表、ヒストグラム	
<b>コラム</b> データの代表値	24
2 平均、分散・標準偏差	26
2-1 資料を特徴付ける3大指標	26
—平均、分散・標準偏差	
<b>コラム</b> 分散と平均を結ぶ公式	35
2-2 平均、分散・標準偏差をヒストグラムで実感しよう	38
—平均、分散・標準偏差の意味するところ	
2-3 資料の数値に手を加えると資料の平均、分散は?	44
—平行移動、定数倍のときの平均、分散	

3 サンプル $X$ の相対度数分布グラフ	51
3-1 資料からとり出してグラフを描く	51
—サンプル $X$ の相対度数分布グラフ	
<b>コラム</b> 復元抽出と非復元抽出	59
3-2 資料の数値に手を加えると平均、分散は?	63
— $X+a$ 、 $bX$ の平均、分散	
4 正規分布	70
4-1 統計解析のなかで一番重要な度数分布	70
—正規分布	
<b>コラム</b> 偏差値	79
4-2 正規分布が持っているすばらしい特徴	82
—再生性、中心極限定理	
5 推定の考え方	88
5-1 部分を見て、全体の様子を知る方法	88
—推定の考え方	
5-2 ピンポイントでズバリ当てましょう	91
—点推定	
5-3 幅を持たせて予想しよう	95
—区間推定	

<b>6</b>	検定の考え方	110
6-1	部分から全体の特徴を判定しよう	110
	——検定の考え方	
	<b>コラム</b> 背理法と検定	123

## 第2章

<b>1</b>	確率変数	128
1-1	確率的にいろいろな数値をとるものを表すための記号	128
	——確率変数 $X$	
1-2	確率分布でも相対度数分布グラフが描ける!	132
	——確率変数 $X$ の相対度数分布グラフ	
1-3	連続する数値に対しても確率変数が定められる	136
	——連続型確率変数	
1-4	確率変数 $X$ について平均、分散を求めてみよう	145
	——確率変数 $X$ の平均、分散	
1-5	連続型確率変数でも平均、分散はある!	154
	——連続型確率変数の平均、分散	
1-6	$X + a$ 、 $bX$ 、 $X + Y$ の平均、分散は?	162
	——確率変数 $X$ の $E(X)$ 、 $V(X)$ についての公式	

<b>2</b>	二項分布	176
2-1	組合せの個数を表す記号 $C$	176
	——コンビネーションと道順	
2-2	5回投げたコインのうち、3回が表である確率	187
	——二項分布の確率	
2-3	二項分布の平均、分散は簡単に計算できる!	191
	——二項分布の平均と分散	
2-4	二項分布の極限が正規分布だ!	195
	——ラプラスの定理	
<b>3</b>	推定の応用	203
3-1	母平均、母分散、母比率を推定する	203
	——いろいろな推定	
	<b>コラム</b> 標本分散と不偏分散	222
<b>4</b>	検定の応用	226
4-1	母平均、母分散を検定しよう	226
	——いろいろな検定	
4-2	2つの母集団の母分散・母平均は等しいか	237
	——2つの母集団に関する検定	
4-3	母比率を検定しよう	252
	——適合度・独立性の検定	

5	$\chi^2$ 分布、 $t$ 分布、 $F$ 分布	260
5-1	推定・検定に使った確率分布を整理しよう	260
	—— $\chi^2$ 分布、 $t$ 分布、 $F$ 分布の定義	
	<b>コラム</b> 自由度	271
	<b>コラム</b> ポアソン分布	276
<b>第 3 章</b>		
1	2 変量の統計	282
1-1	相関係数の計算の仕方	282
	——相関係数の定義	
	<b>コラム</b> 相関係数が -1 以上 1 以下である理由	298
1-2	2 つの変数の関係を直線で表すには	301
	——回帰直線の求め方	
	<b>付録 1</b> エクセルで統計量を計算しよう	316
	——平均、分散から確率分布表まで	
	<b>付録 2</b> 分布表	323
	<b>付録 3</b> さらに学びたい人のためのブックガイド	327
	おわりに	330
	さくいん	334

### ●この本の特徴●

ぼくは、“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**推定・検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、コンピュータソフトが発達していますから、データを入力しさえすれば瞬時に機械が統計処理をし、推定・検定の結果まで出してくれます。しかし、推定・検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ます。

そこでこの本では、統計学のなかでも、推定・検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、推定・検定の本質を 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。

この第 1 章を読めば、1 時間で推定・検定の本質がわかります。しかも、難しい数学を使わないで。この第 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。この第 1 章だけを切り離して、「1 時間で推定・検定がわかる本」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時間で推定・検定がわかる」ための第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わけです。まとめでも、 $\Sigma$ 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。もったいないことはありませんか。

そこで第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、推定・検定を説明することにしました。

そのために、ぼくは、“ $X$ の相対度数分布グラフ”というものを発明しちゃいました。これは、グラフですから図として認識できるものです。数式にアレルギーがある人でも目で見て実感することができます。数学概念をビジュアル化することは、ぼくの得意とするところなのです。

このグラフの性質として、平均、標準偏差の公式、確率変数の公式の概念を全部押し込めて説明してしまったんです。これは、確率変数を用いる正式な統計学を学んできた人が読んでも、その内容がビジュアル化されているので、面白いものになっていると思います。

### 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、推定・検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。とおりに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うんです。資料の平均、分散という量と確率変数の平均、分散という量が結びつかないまま、推定・検定の説明へ突入するので、推定・検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。

そこで、第1章では、資料からサンプルをとり出したときのことを確率変数を用いないで説明していきます。このときキーとなるのが、先にも紹介した“ $X$ の相対度数分布グラフ”なんです。

### ④ 本筋でないものは脇へ

推定・検定を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、モード、最頻値など記述統計に関する用語、

$V(X) = E(X^2) - \{E(X)\}^2$ の公式、偏差値、不偏分散と標本分散の違いなどの話題です。

また、内容的には重要であっても、推定・検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、第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、第1章で本質を理解してもらった後、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、第2章に読み進んでいただきたいと思います。

第2章では、第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。

第2章では、資料からとり出したサンプルの変量が離散型の確率分布であっても面積のあるヒストグラムを描いています。これは、「面積＝確率」という統計学の基本を高校レベルで頭に叩き込んでもらうための便法です。自分でもやりすぎだと思います。この本だけで通用することですから、各種試験では注意してください。

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、

- なぜヒストグラムが曲線になるの？
- 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、全く解説してなかったり、逆に難しい解説がなされている箇所です。