

無断使用をお断りします。日科技連出版社

ことばの式でわかる 統計的方法の 極意

竹士 伊知郎 著

日科技連

まえがき

「数式も記号もよくわからないし！ できれば数学っぽいことは勘弁してほしい。でも、近ごろ話題の統計には興味がある！」そんなあなたに「ことばの式」によって「統計の極意」を授けます！

筆者は、社会人や大学生などを相手に統計的方法に関する講義や指導を行ってきました。そこで気づいた「統計の極意」のようなものを多くの皆さんに伝授したいと考えて、本書を執筆しました。

2018年に『学びたい 知っておきたい 統計的方法』(日科技連出版社)を上梓しました。同書は大学や社会人向けのセミナーのテキストとしても、多くの場で使っていただいています。本書は、同書のいわば姉妹編として、もっと「そもそもの始まり」から話を起こしました。

統計の講義や指導をしていると、「そもそも何をしているのですか?」、「なぜそうするのですか?」、「何の役に立つのですか?」、「何のためにするのですか?」といった、What や Why に関する素朴な質問を数多くいただきます。しかし、統計に関する多くの書籍は、特に初心者が抱く素朴な疑問についていてねいに説明していません。ほとんどが、「こうすればよい」、「このようにやりなさい」という How に重点が置かれているといっても過言ではありません。

そこで本書は、思い切って What と Why に絞って書くことにしました。How に関することは多くの良書がありますのでそちらを見ていただくとよいし、難しく面倒くさい計算の多くもパソコンの統計ソフトに任せることも可能だからです。What や Why をきちんと理解していれば、パソコンの統計ソフトにデータやその他の条件を適切にインプット(入

力)できますし、アウトプット(結果)された数字の解釈に戸惑うこともありません。

本書には、ギリシャ文字やアルファベットを使った数式や数学記号というものがほとんど出てきません。その代わりに、「平均値」や「偏差」といった「ことばをそのまま使った数式」(といっても、四則演算、程度)である「ことばの式」を使います。

それは、次の2つの理由からです。1つは、 α 、 μ 、 σ 、 S 、 V 、 Σ などの記号が出てきたとたんに、数学アレルギーとでもいうべき症状を発症して、その後の学習がほとんど身につかない方が一定数おられるからです。

もう1つは、そのような記号などを使って本を書いたり講義をしたりする側は大変便利なのですが、本の読者や、講義の受講者の一部には、「 S って何でしたっけ?」、「 σ の意味をもう一度説明してほしい」、「 Σ は、どんな式でしたか?」など、話が複雑になってくると、少し前のことを忘れてしまう方が多いからです。

その点、日本語での表記を使った「ことばの式」は、何といても漢字という表意文字を使っているので、数学アレルギーは出ず、意味もわかるし思い出せるという利点があるだろうと考えました。

本書は、今さら聞けないことがすべて書いてある、知っていたらかつこいい、そして統計の専門家への道を示すために書いた本です。統計の勉強を本格的に始める前に、ぜひ手に取ってほしいと思っております。

本書を読み終えたあなたは、統計学を少し学んだ方よりも、きっと「統計的方法」のことに詳しくなっています。それは、明日から統計を使って仕事ができるということではありません。統計の本質や中身を理解して、きちんとした統計ができる素地や基礎が完璧にできたことを意味します。したがって、いい加減なデータに惑わされることもありませんし、自信をもって報告書を作成することができます。人に統計について説明

することもできます。

さらに、「品質管理検定(QC 検定)」の受検をお考えの方にも、最適な参考書の1つです。QC 検定は2005年に始まり、現在、品質管理に関する知識を客観的に評価するものとして広く定着し、1級から4級までの級が設定されています。しかし、「3級まではなんとか合格したが2級はかなり難しい」という「2級の壁」があるといわれているのも事実です。これは2級からは、「統計的方法」に関する相当の知識が求められることによります。合格のためには、統計的方法の真の意味を体系的に学習することが不可欠です。そういう意味で、本書は2級合格をめざす方々にも最適な書です。

またもう一つ、高校の数学で必修となる一連の統計学の参考書としても活用が期待されます。

面倒な手順や計算はパソコンと統計ソフトを使えばいいのです。でも本書を読んだあなたは、何も知らずに統計ソフトを使って結果だけをうのみにするような「自称専門家」では、もはやないのです。

さあ、「ことばの式」で「統計的方法の極意」を知って、本物の統計の専門家への道を踏み出しましょう。

本書の出版にあたっては、日科技連出版社の戸羽節文社長、鈴木兄宏部長、石田新係長に一方ならずお世話になりました。この場を借りて心より感謝申し上げます。

2022年3月

竹士 伊知郎

本書の読み方

本書は、以下の項目を順に説明します。できれば初めから順に読み進めてください。完走ならぬ完読を期待します。

0. 初めのはじめ
1. 知りたいことをデータの集まりすなわち母集団と考えよう
2. 知りたいことを知るために統計が必須であると知ろう
3. 母集団のばらつきと平均について知ろう
4. 母集団全体の姿である分布について知ろう(正規分布)
5. 正規分布と確率について知ろう
6. 母集団を推測するためのサンプリングの目的とその重要性を知ろう
7. サンプルから統計量を求めよう
8. サンプルの情報から母集団に関する結論を出そう(推定と検定の基本)
9. 1つの母集団, 2つの母集団の母平均, 母分散に関する結論を出そう
10. 時間によって変わる母平均に関する結論を出そう
11. 原因となるものの効果によって変わる母平均に関する結論を出そう
12. 対応する2つの母集団の関係に関する結論を出そう
13. 原因を連続的に動かしたことによる母平均の変化に関する結論を出そう

目 次

まえがき *iii*

本書の読み方 *vi*

序章 初めのはじめ 1

第 1 章 知りたいことをデータの集まり
すなわち母集団と考えよう 7

第 2 章 知りたいことを知るために
統計が必須であると知ろう 11

第 3 章 母集団のばらつきと平均について知ろう 15

3.1 ばらつきについて知ろう 16

3.2 平均について知ろう 18

第 4 章 母集団全体の姿である分布について知ろう 21

4.1 正規分布 22

第 5 章 正規分布と確率について知ろう 25

5.1 分布と確率の関係 26

5.2 正規分布と確率 28

5.3 標準正規分布 30

第 6 章 母集団を推測するための サンプリングの目的とその重要性を知ろう	33
6.1 母集団を推測しよう	34
6.2 サンプルとサンプリング	34
6.3 母集団を代表するサンプルをとる	36
第 7 章 サンプルから統計量を求めよう	39
7.1 統計量	40
7.2 統計量の分布	42
インターミッション I	51
第 8 章 サンプルの情報から母集団に関する結論を出そう	53
8.1 推定と検定	54
8.2 推定とは	54
8.3 検定とは	58
インターミッション II	65
第 9 章 1 つの母集団, 2 つの母集団の母平均, 母分散に関する結論を出そう	67
9.1 1 つの母集団の母平均の推定	68
9.2 1 つの母集団の母平均の検定	71
9.3 1 つの母集団の母分散の推定	75
9.4 1 つの母集団の母分散の検定	78
9.5 2 つの母集団の母平均の差の推定	83

9.6	2つの母集団の母平均の差の検定	85
9.7	2つの母集団の母分散の比の検定	89
第10章 時間によって変わる母平均に関する結論を出そう 95		
10.1	管理図	96
10.2	\bar{X} -R 管理図による解析	99
10.3	p 管理図による解析	101
第11章 原因となるものの効果によって変わる 母平均に関する結論を出そう 105		
11.1	実験計画法	106
11.2	一元配置実験	109
11.3	二元配置実験	115
第12章 対応する2つの母集団の関係に 関する結論を出そう 123		
12.1	相関分析	124
12.2	相関分析の解析方法	129
第13章 原因を連続的に動かしたことによる 母平均の変化に関する結論を出そう 133		
13.1	回帰分析	134
13.2	単回帰分析	137
理解度テスト 144		
引用・参考文献 145		

理解度テストの解答 146

索引 147



インターミッション I

1. 統計に用いる用語の整理

ここで、これまでに出てきた用語の意味と中身を整理して、今後の学習に備えましょう。

- ① **母集団**：知りたいことをデータの集まりと考えたもの

母集団にかかわる用語：

母平均：(母集団に属するデータ)の平均→通常わからない

母分散： $\frac{((\text{データ}) - (\text{母平均}))^2}{\text{偏差}}$ の平均→通常わからない

母標準偏差： $\sqrt{\text{母分散}}$

- ② **統計量**：母集団を知るためにサンプルのデータから求めた各種の数値

統計量にかかわる用語：

平均値：(サンプルのデータ)の平均値

偏差平方和： $\frac{((\text{サンプルのデータ}) - (\text{平均値}))^2}{\text{偏差}}$ の合計

不偏分散：(偏差平方和)/(サンプルのデータの数-1)

自由度：(サンプルのデータの数)-1

標準偏差： $\sqrt{\text{不偏分散}}$

注1) 偏差平方和は、複数個の平均値とそれらの平均値から算出する場合もある。

注2) 第8章以降は、偏差平方和を「平方和」、不偏分散を「分散」と表記する。

注3) 推定・検定はもちろん、実験計画法や回帰分析でも自由度の考え方は重要であり、例えば、

$$\text{要因の自由度} = (\text{水準数}) - 1$$

などとして、しばしば登場する。

第8章では、推定と検定の基本的な考え方や留意点を学びました。さていよいよここからは、実際の場面で推定や検定がどう使われるのかです。基本的な例から話を進めます。

本章で取り上げたのは、

- ①, ② 母集団が1つの場合の母平均の推定と検定
- ③, ④ 母集団が1つの場合の母分散の推定と検定
- ⑤, ⑥ 母集団が2つの場合の母平均の差の推定と検定
- ⑦ 母集団が2つの場合の母分散の比の検定

の合計7つの例です。

たとえ母集団が同じであっても、母平均と母分散は別のものですので、それぞれ異なる方法を用います。また、2つの母集団を比べて母平均や母分散が異なるのかどうかといったことに興味があることもあります。そういった場合に適用する「母平均の差の推定と検定」と「母分散の比の検定」についても扱います。

ここで、「3つ以上の母集団を同時に扱う場合はどうするの?」という疑問を抱かれた方もあるかと思いますが、ご安心を。第11章以降に答えがあります。

9.1 1つの母集団の母平均の推定

こんな場面で使う

自動車部品を製造しており、お客様からの要望によってある部分の寸法が重要である。

- 現状の工程で製造している部品の寸法 S の母平均を知りたい。

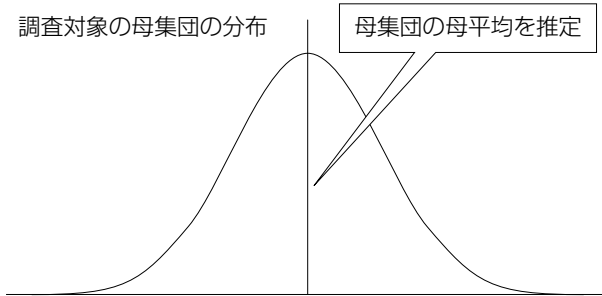


図 9.1 母平均の推定

(1) 調べたいこと

知りたい1つの母集団の母平均を推定したい(図 9.1).

例：お客様の要望に対して現状の工程で製造している寸法 S の母平均を調べたい。

(2) まずやること

- 1) 母集団からサンプルの数を決めてサンプルをランダムに採取してデータを得る。
- 2) 信頼率を決める。

例：現状の工程で生産されている部品からランダムに部品を 30 個採取して寸法 S を測定しました。信頼率は 95% としました。

(3) どう調べる

- 1) サンプルのデータの平均値, サンプルのデータの不偏分散を求める。
- 2) 母平均の点推定値はサンプルのデータの平均値となる。
- 3) サンプルのデータの平均値はサンプルの数とサンプルのデータの不偏分散とで決まる t 分布に従うので, 信頼率に応じた t 分布の値

を求めて信頼区間を求める。

信頼率 95% の信頼区間の求め方は以下のようになる。

検定統計量 t は、

$$\text{検定統計量 } t = \frac{(\text{サンプルのデータの平均値}) - \text{母平均の値}}{\sqrt{(\text{サンプルのデータの不偏分散}) / (\text{サンプルの数})}}$$

であり、この値が、 t 分布の上側 2.5% 点と下側 2.5% 点の間にある確率は 95% であるので、

$$\text{下側 2.5\% 点} < \frac{(\text{サンプルのデータの平均値}) - (\text{母平均の値})}{\sqrt{(\text{サンプルのデータの不偏分散}) / (\text{サンプルの数})}} < \text{上側 2.5\% 点}$$

から、

$$\begin{aligned} & (\text{サンプルのデータの平均値}) + \text{下側 2.5\% 点} \times \sqrt{(\text{サンプルのデータの不偏分散}) / \text{サンプルの数}} \\ & < \text{母平均の値} < \\ & (\text{サンプルのデータの平均値}) + \text{上側 2.5\% 点} \times \sqrt{(\text{サンプルのデータの不偏分散}) / (\text{サンプルの数})} \end{aligned}$$

となる。

例：得られた 30 個のデータから平均値と不偏分散を計算して母平均の点推定値を求め、さらに区間推定を行いました。

(4) わかること

知りたい母集団の母平均の点推定値と信頼率に応じた信頼区間がわかる。

例：信頼区間からお客様の要望を満足していると確認できました。

著者紹介

竹士 伊知郎 (ちくし いちろう)

1979年 京都大学工学部卒業、(株)中山製鋼所入社。

金沢大学大学院自然科学研究科博士後期課程修了、博士(工学)。

現在 QMビューローちくし 代表、関西大学化学生命工学部 非常勤講師、(一財)日本科学技術連盟 嘱託。

日本科学技術連盟等の団体、大学、企業において、品質管理・統計分野の講義、指導、コンサルティングを行っている。

主な品質管理・統計分野の著書に、『学びたい 知っておきたい 統計的方法』(単著、日科技連出版社)、『QC 検定受検テキストシリーズ』、『QC 検定対応問題・解説集シリーズ』、『QC 検定模擬問題集シリーズ』、『速効! QC 検定シリーズ』、『TQM の基本と進め方』(いずれも共著、日科技連出版社)がある。

ことばの式でわかる統計的方法の極意

2022年4月27日 第1刷発行

著者 竹士 伊知郎

発行人 戸羽 節文

発行所 株式会社 日科技連出版社

〒151-0051 東京都渋谷区千駄ヶ谷5-15-5

DSビル

電話 出版 03-5379-1244

営業 03-5379-1238

検印
省略

Printed in Japan

印刷・製本 壮光舎印刷

© Ichiro Chikushi 2022

ISBN 978-4-8171-9757-3

URL <https://www.juse-p.co.jp/>

本書の全部または一部を無断でコピー、スキャン、デジタル化などの複製をすることは、著作権法上での例外を除き禁じられています。本書を代行業者等の第三者に依頼してスキャンやデジタル化することは、たとえ個人や家庭内での利用でも著作権法違反です。