

100 Questions (and Answers) about Statistics

いまさら聞けない疑問に答える

統計学の

キホン

Q&A 100

ニール・J・サルキンド [著]

山田剛史・寺尾 敦・杉澤武俊・村井潤一郎 [訳]

初心者がいきなりつまづく統計学の用語。
教科書でも、ウェブで調べてもわかりにくい。
要するに、どういうこと?!

基礎から推測統計学まで、100の疑問を精選して1~2ページで解説。
辞書より詳しく、数学なしで基本概念の意味から使い方、Excelを
使った計算法まで理解できる、これまでになかったガイドブック。

新曜社

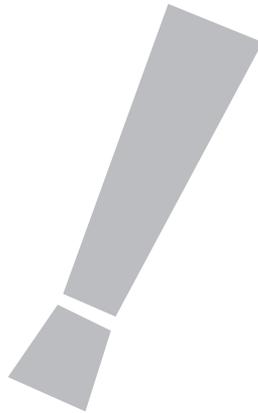


いまさら聞けない疑問に答える

統計学の キホン Q&A100

ニール・J・サルキンド [著]
山田剛史・寺尾 敦・杉澤武俊・村井潤一郎 [訳]

新曜社



100 QUESTIONS (AND ANSWERS) ABOUT STATISTICS

by Neil J. Salkind

Copyright © 2015 by SAGE Publications, Inc.

This translation is published under cooperation contract between SAGE and Shinyosha.

はじめに

今日では、経済的な成功・失敗からどのチームがワールドシリーズを制するかへの予測まで、あらゆることを説明するのに、ビッグデータという言葉がいたるところで用いられている。量的な情報とそのパターンについての理解が、これまで以上に重要になっている。野心に燃える専門家、学生、政策担当者、そして、何千という他の専門家も、データを理解する方法を求めている。基礎的な統計学、さらには高度な統計学が重要なゆえんである。

本書『統計学のキホンQ&A 100』は、基本的な統計についての簡単な理解からもっと洗練された推測統計学まで、統計学の基礎に関する、最も重要な質問を1冊の中にまとめたものである。

本書を書いたのは、私自身、長年教育に携わる中で、多くの学生だけでなく専門家も、統計学について、どのようなトピックに焦点を定める必要があるのか、詳しい情報をどこで見つけたらよいかを正しく案内してくれる、簡潔な概説を必要としているということに気づいたからである。

本書は小さな本だが、統計学に関して重要なトピックが何かを改めて確認したい人や、統計学の全くの初心者で、疑問に思っていることについて参照できる資料が必要な人に向けて書かれている。本書を、以前に学んだことを思い出す手がかり、一種の資料や知識をリフレッシュするための情報源として考えていただければと思う。本書は、学位取得に向けて総合的な試験の準備をしている大学院生や、手軽な参考書が必要な研究者、統計学の初級コースを履修していない関連分野の大学生、統計学の技法について知る必要がある政策担当者、そして、単純にこれらのツールを最も効果的に用いる方法について知りたい人のためのものである。

本書を有効に用いるコツをいくつか……。

1. 質問は、以下のように12のパートに分かれている。

- パート 1 なぜ統計学なのか
- パート 2 中心傾向の測度についての理解
- パート 3 変動の測度についての理解
- パート 4 データの視覚的表現
- パート 5 関連性についての理解
- パート 6 測定とその重要性についての理解
- パート 7 統計学における仮説の役割についての理解

- パート 8 正規分布と確率についての理解
- パート 9 有意性の概念についての理解
- パート 10 群間差についての理解
- パート 11 変数間関係を見る
- パート 12 その他の統計的方法

2. 質問と回答は、それぞれ独立に読むこともでき、率直な質問と比較的短い回答が記載されている。それぞれの質問に対して一冊の本ほどの長い回答をすることもできるだろうが、本書の目的は、読者が次の質問やトピックに進む前にある程度の知識を得られるよう、手短かに、必要な情報を示すことである。
3. これらの質問と回答のすべては互いに無関連というわけではなく、大半が補完し合っている。そのため、重要な内容について記述が強化されるとともに、参照したトピックだけでなく、関連するトピックについても確実に理解できるようになっている。
4. それぞれの質問の最後には、当該の質問に関連する3つの他の質問が参照項目として記載されている。これらは、今読んでいる質問と答えを最もよく補完すると私が考える項目である。もちろん、他にも関連する項目は多くある。

謝 辞

出版にあたって、最高の編集制作チーム、特に、このシリーズのアイデアを熱烈に歓迎してくれた、ヴィキ・ナイトさんに感謝している。彼女は、私からの途方もない数の深夜メールにもしなやかに耐え、惜しみなく支援してくれた。そして副編集者のローレン・ハビブさんとエシカ・ミラーさんにも感謝する。彼女たちは迅速かつ効率的に仕事を進め、いつも大いに助けてくれた。制作編集者のリビー・ラーソンさんと原稿整理編集者のポーラ・L・フレミングさんにもお礼を述べたい。だが本書の誤りは、すべて私に責任がある。誤りがあつたら、この本がもっとよくなるようにEメールで知らせていただければ幸いである。

時間をとって精査してくださったレビューアの方々に感謝する。T・ジョン・アレキサンダー (テキサス・ウェズリアン大学)、ジェイミー・ブラウン (マーサー大学)、ヨンス・キム (ラスベガス、ネヴァダ大学)、リンダ・マルティネス (カリフォルニア州立大学ロングビーチ)、ジュリアナ・ラスコウスカス (カリフォルニア州立大学サクラメント)、ジェニファー・R・サーモン (エッカート大学)、ロビン・L・スペイド (モーガン州立大学)。

そして、銀河で最高の子どもたち、サラ、ミカ、テッドに。そしてもちろん、ペッパーの後継であるラッキー、根気強く、優しく、勇敢で誠実な友であるルー・Mに、ありがとう。

ニール・J・サルキンド
ローレンス, カンサス
njs@ku.edu

目次

はじめに	i	
謝辞	iii	
▶パート1 なぜ統計学なのか	1	
質問1	統計学では何を学ぶのですか？ またなぜ重要なのですか？	2
質問2	統計学はどのようにして「始まった」のですか？	3
質問3	記述統計とは何ですか？ どのように使われるのですか？	5
質問4	推測統計とは何ですか？ どのように使われるのですか？	6
質問5	私は統計学者になろうとはしていません。なぜ統計学の授業を受けるべきなのでしょう？	7
質問6	統計のソフトウェアパッケージは、どれを使ったらよいのでしょうか？ SPSS？ Excel？ 他のものでしょうか？	8
▶パート2 中心傾向の測度についての理解	11	
質問7	中心傾向の測度とは何ですか？ なぜ使われるのですか？	12
質問8	中心傾向の測度をどのように使うのですか？ その例をあげてもらえますか？	13
質問9	平均値とは何ですか？ どのように計算するのですか？	14
質問10	平均値をどのように使うのですか？ その例をあげてもらえますか？	16
質問11	中央値とは何ですか？ どのように計算するのですか？	18
質問12	中央値をどのように使うのですか？ その例をあげてもらえますか？	20
質問13	最頻値とは何ですか？ どのように計算するのですか？	22
質問14	最頻値をどのように使うのですか？ その例をあげてもらえますか？	24
質問15	中心傾向のどの測度を用いるべきか、どのように決めたらよいのですか？	26
質問16	中心傾向の測度を算出するために、Excelをどのように使ったらよいのですか？	28
▶パート3 変動の測度についての理解	31	
質問17	変動の測度とは何ですか？ それらはなぜ使われるのですか？	32
質問18	変動の測度はどのように用いられるのですか？ その例をあげてもらえますか？	33
質問19	範囲とは何ですか？ どのように計算するのですか？	34
質問20	範囲がどのように使用されるか、その例をあげてもらえますか？	36
質問21	標準偏差とは何ですか？ どのように計算するのですか？	37

質問22	標準偏差と分散を計算する際に、なぜ、単に n ではなく、 $n-1$ を用いるのですか？	39
質問23	標準偏差はどのように使われるのですか？ その例をあげてもらえますか？	40
質問24	分散はどのように使われますか？ その例をあげてもらえますか？	41
質問25	変動の測度を計算するために、Excelをどのように用いるのですか？	42
▶パート4	データの視覚的表現	43
質問26	なぜ、データを視覚的に表現するのですか？ ほんとうに百聞は一見にしかずですか？	44
質問27	データの効果的な視覚的表現を作るためのガイドラインが、何かありますか？	46
質問28	度数分布と累積度数分布とは何ですか？ どのように作ることができますか？	47
質問29	ヒストグラムとは何ですか？ どのように手作業で作ることができますか？	49
質問30	Excelを用いて、どのようにヒストグラムを作ることができますか？	51
質問31	累積曲線とは何ですか？ Excelを用いてどのように作ることができますか？	53
質問32	縦棒グラフとは何ですか？ Excelを用いてどのように作ることができますか？	55
質問33	横棒グラフとは何ですか？ Excelを用いてどのように作ることができますか？	57
質問34	折れ線グラフとは何ですか？ Excelを用いてどのように作ることができますか？	59
質問35	円グラフとは何ですか？ Excelを用いてどのように作ることができますか？	61
質問36	散布図とは何ですか？ Excelを用いてどのように作ることができますか？	63
質問37	Excelで作成したグラフを編集するには、どうすればよいですか？	65
質問38	他の文書の中にグラフを挿入するには、どうすればよいですか？	67
質問39	いつ、どのような種類のグラフを使うべきですか？	69
▶パート5	関連性についての理解	71
質問40	相関係数とは何ですか？ どのように使うのですか？	72
質問41	相関係数をどのように使うのですか？ その例をあげてもらえますか？	73
質問42	相関係数にはどのようなタイプがありますか？	74
質問43	散布図は、相関係数の理解にどのように役立つのですか？	75
質問44	相関係数はどのように計算するのですか？	78

質問45	Excelで相関係数を計算するにはどうすればよいですか？	80
質問46	相関係数の値を解釈する、もっとも直接的な方法は何ですか？	82
質問47	決定係数とは何ですか？ どのように計算するのですか？	83
質問48	相関係数を理解し使用する上で、覚えておかなければならない重要なことは何ですか？	84
質問49	いくつかの相関係数を表示するために、行列をどのように使うことができますか？	86
質問50	相関の測度には他にどのようなものがありますか？ それらはどのように使われるのですか？	88
▶パート6	測定とその重要性についての理解	89
質問51	統計利用者にとって、なぜ、測定について理解することが重要なのですか？	90
質問52	尺度水準とは何ですか？ なぜ、それが重要なのですか？	91
質問53	信頼性とは何ですか？ どうやって信頼性を定めるのか、例をあげてください。	93
質問54	信頼性にはどんな種類がありますか？ それらをいつ用いるのですか？	95
質問55	テストの信頼性を高めるには、どのようにすればよいですか？	97
質問56	妥当性とは何ですか？ どうやって妥当性を定めるのか、例をあげてください。	99
質問57	妥当性にはどんな種類がありますか？ またそれらはどのように用いられるのですか？	101
質問58	妥当性を高めるには、どのようにすればよいのですか？	103
質問59	信頼性と妥当性の関係は何ですか？	104
▶パート7	統計学における仮説の役割についての理解	107
質問60	仮説とは何ですか？ なぜ科学的研究で重要なのですか？	108
質問61	よい仮説はどんな性質を持っていますか？	110
質問62	標本と母集団は、それぞれどのように違うのですか？	111
質問63	帰無仮説とは何ですか？ どのように使われるのですか？	112
質問64	対立仮説とは何ですか？ どのように使われるのですか？	114
質問65	帰無仮説と対立仮説はどのように違うのですか？	116
▶パート8	正規分布と確率についての理解	117
質問66	統計学の学習で、なぜ確率が重要なのですか？	118
質問67	正規曲線（ベル曲線）とは何ですか？	119
質問68	歪みと尖りとは何ですか？ これらの特徴によって、分布はどのように異なるのですか？	121

質問 69	中心極限定理とは何ですか？ それはなぜ重要なのですか？	123
質問 70	z 得点とは何ですか？ どのように計算するのですか？	125
質問 71	Excel ではどのように z 得点を計算するのですか？	127
質問 72	z 得点と正規曲線にはどんな関係があるのですか？	128
質問 73	z 得点は、仮説検定とどのような関係にあるのですか？	130
▶ パート 9	有意性の概念についての理解	131
質問 74	どのようにして推測するのですか？ その具体例もあげてもらえますか？	132
質問 75	有意性の概念とは何ですか？ それはなぜ重要なのですか？	134
質問 76	第 1 種および第 2 種の誤りとは何ですか？	135
質問 77	統計的検定を研究仮説に適用するときの手順は、どのようになりますか？	137
質問 78	統計的な有意性と、有意味であることとの違いは何ですか？	138
質問 79	Excel の分析ツールとは何ですか？ 統計的検定を実行するときにどう使うのですか？	140
質問 80	Excel の関数とは何ですか？ 統計的検定を実行するときにどう使うのですか？	142
▶ パート 10	群間差についての理解	145
質問 81	どの統計的検定を使うべきかを、どうすれば知ることができますか？	146
質問 82	独立な 2 群の平均値間の t 検定とは何ですか？ どのように使われるか、例をあげてください。	148
質問 83	独立な 2 群の平均値の差の検定を行うために、Excel をどのように使えますか？	149
質問 84	対応のある 2 群の平均値間の t 検定とは何ですか？ どのように使われるか、例をあげてください。	151
質問 85	対応のある 2 群の平均値の差の検定を行うために、Excel をどのように使えますか？	152
質問 86	一要因分散分析とは何ですか？ どのように使われるか、例をあげてください。	154
質問 87	一要因分散分析の計算を行うために、Excel をどのように使えますか？	155
質問 88	多要因分散分析とは何ですか？ どのように使われるか、例をあげてください。	157
質問 89	多要因分散分析の計算を行うために、Excel をどのように使えますか？	158
質問 90	どのようにして、有意性の検定のためにノンパラメトリック検定を使うのですか？	160
質問 91	効果量とは何ですか？ なぜ重要なのですか？	162

▶パート11	変数間の関係を見る	165
質問92	相関係数を使って、関係の有意性をどのように検定しますか？ また、相関係数の使い方の例をあげてください。	166
質問93	相関係数の有意性を検定するために、Excelをどのように使ったらよいですか？	167
質問94	単回帰分析とは何ですか？ その使い方の例をあげてください。	169
質問95	単回帰式を算出するのに、Excelをどのように用いたらよいですか？	170
▶パート12	その他の統計的方法	173
質問96	多変量分散分析 (MANOVA) とは何ですか？ それはどのように用いられるのですか？	174
質問97	共分散分析 (ANCOVA) とは何ですか？ それはどのように用いられるのですか？	175
質問98	繰り返しのある分散分析とは何ですか？ それはどのように用いられるのですか？	176
質問99	重回帰分析とは何ですか？ それはどのように用いられるのですか？	177
質問100	因子分析とは何ですか？ それはどのように用いられるのですか？	178

訳者あとがき 179

索引 181

装幀＝新曜社デザイン室

パート **1**
なぜ統計学なのか



統計学では何を学ぶのですか？ またなぜ重要なのですか？

私たちはみな、ますますデータに支配されるようになった世界に生きている。データの分析が教育機関やフォーチュン500の大企業にとって重要なと同じくらい、プロスポーツチームにとっても重要なものになるうとは、誰が想像しただろうか。しかし、実際そうなのである。パターンを見つけたり、結果を予測したりするために統計学と向き合う機会が、ますます多くなっている。統計学は、複雑な結果をよりよく理解したり決定を下したりするのに役立つ、ツールの集まりから成っている。

統計学とは、数量的な情報を記述し、統合し、分析し、解釈することである。ここで情報というのは、ひとまとまりのテスト得点かもしれないし、特定のタイプの自動車に対する好みかもしれないし、あるいはバスケットボールチームが、相手チームからボールをカットしたあとで得点することがどれくらい多いかといったことかもしれない。この本を通して身につけていくさまざまなツールを使うことで、これらの問いをはじめ、その他データの分析を扱う、ほとんどどんな問いにでも答えることができる。

統計学の学習が重要な理由はいくつもある。統計学が最もよく用いられるのは、情報に基づいて決定をする手助けをすることであり、その情報は統計学を使わなければ解釈がとても難しかったり、できなかつたりする。この有用性は、ごく単純な例を見るだけでもわかる。たとえば、ある生徒の集団に対する個別のテスト得点より、生徒全体の平均点の方がより役に立つのではないだろうか。あるいは、受けたサービスについての感想を顧客集団に質問紙調査したとき、20項目の質問に1人ひとりの顧客がどう答えたかよりも、結果の平均点の方が役に立つのではないだろうか。

この2つの例は、情報を集めて記述し、分析することができれば、よりよい決定ができるという事実を示している。証拠に基づいているからである。では、まだ組織立てられていない情報の集まりが何を意味しているのかを探ることができるツールは？それが統計学なのである！

統計学は学者が行う研究、地方や国家の政策担当者が行う決定、そして、さらなる目標に向かって情報に基づいて行動しなければならないビジネスの日々の機能にとっても、計り知れない価値がある。

もっと知るには？ 質問2, 3, 4を参照。

統計学はどのようにして「始まった」のですか？

統計学の研究は、データの収集と分析にとどまらない。すなわち、情報を集めて重要な決定を行うために利用するのである。何かがどれくらいあるのか（「食べ物が底をつくまでに何日くらいあるのか」や「冬まであと何週間あるのか」）、そして、それらの数がある特定の結果（健康や安全など）にどう影響するのかということについて、人々が関心を持たなかった時代はおそらくなかっただろう。

そもそも数は、ある特定の結果と結びつけられている。もしある人が学校で勉強がよくできて、よい成績をとったなら、その後も学校で成功する可能性が高い。もしある人がよい教育を受けたなら、卒業と同時によりよい仕事が残っているだろう。そして、我々が今日人口統計学者として知っている人々（人口やその性質を研究している人々）が、多くの人が住み、働き、遊ぶ場所について数えたり分布を調べたりし始めたのは、それほど昔のことではない。

このすべては主に数学者によってなされたが、生物学や、より最近では心理学などの領域で、観測されたことの理解が強く求められ、統計学の分野が誕生した。

おそらくこの誕生の大きな節目は、フランシス・ゴルトンの研究であった。彼はチャールズ・ダーウィンのいとこで、19世紀初頭に生まれた。ゴルトンは、相関係数と呼ばれる、変数間の関係を表すのに今でも非常によく使われるツールを考案した。彼は家族間の知能に関心があった。彼の研究は（後にしばしば疑われることになったが）、家族のメンバー間の関係性を比較するための枠組みを築いた。

ゴルトンの後、社会がどんどんと複雑化し、利用可能な情報すべての複雑性を理解する必要性が増大していくにつれて、統計学は非常に多くの新しい発展を見せた。カール・ピアソン（数学者）やR・A・フィッシャー（農学者）などが、自身の研究分野で学んだことを人間行動のさまざまな側面に適用した。ここ40年のパーソナルコンピュータの出現によって、大規模データに含まれるパターンや傾向を見たいと思う人はほとんど誰でも、統計的手法の中の最も強力なものでさえ利用できるようになった。こういう分析は、現代の統計学の非常に重要な部分を占めている。大学のみならずプロスポーツチームでさえ、今では、何が役に立ち、何が役に立たないのかを特定するために、このアプローチを利用している。

統計学の歴史についてもっと知りたければ、セント・アンセルム大学のウェブサイト
を参照するとよい。

www.anselm.edu/homepage/jpitocch/biostatshist.html

もっと知るには？ 質問 1, 3, 4 を参照。

記述統計とは何ですか？ どのように使われるのですか？

記述統計は、データ（あるいは分布）の特徴をまとめ、記述するために使われる。記述統計は、データを詳しく調べるために用いられる最初のツールであり、データが「どのように見える」かを示す、いくつかの重要な指標を得る。

記述統計には、大きく2つのカテゴリーがある。

第1のカテゴリーは、平均値、中央値、最頻値など、中心傾向の測度を見るものから構成される。これらはすべて「代表値」と呼ばれる。これらのいずれも、データを代表する最良の点を表すために用いることができる。たとえば、地域のあるレストランでのハンバーガーの売り上げに興味があるなら、最初の問いはおそらく、毎週平均して何個のハンバーガーが売れたかだろう。別の例をあげると、「フォードの車で最もよく売れるモデルは何か」が知りたいかもしれない。

第2のカテゴリーは、データの変動、広がり、ばらつきを見る記述ツールから構成される。これらの測度としては、範囲、標準偏差、分散などがある。これらの測度は、データの各点が互いにどれくらい離れているのかを教えてくれる。たとえば、あるグループの子どもたちが読解スキルに関してどれくらい似ているのかを知りたければ、子どもたちの読解テスト得点の標準偏差を見るだろう。この記述統計量^[訳注]の値が小さいほど、これらの子どもたち同士でばらつきが小さい（違いが少ない）。

これら2つのカテゴリー、すなわち代表値と変動の測度を組み合わせることで、あるデータの性質と、他のデータとの違いについて、非常によく表される。そして、後述するように、この2タイプの測度は、2群のデータにおける平均値差の有意性など、もっと複雑な統計的処理の土台となる。

もっと知るには？ 質問4, 7, 16を参照。

[訳注] データから求められる統計的指標、たとえば、平均や標準偏差などのこと。

推測統計とは何ですか？ どのように使われるのですか？

推測統計を使うことで、小さなグループ（しばしば標本（サンプル）と呼ばれる）での結果から、より大きなグループ（しばしば母集団と呼ばれる）へと推測を行うことができる。推測統計は、多くの場合記述統計の拡張であるが、常にそうだというわけではない。知りたい問題によっては、（記述統計を用いて）得点の平均値を知ることでも十分なこともある。たとえば、ある特定のグループの平均値が他のグループの平均値と異なるかどうかを知りたいのであれば、必ずしも推測統計を用いる必要はない。

推測統計がどのように用いられるのかの1つの例として、2つの小学生グループが受けた読解スキルのテストについて、グループ間の差を検討することを考えてみよう。一方のグループが特別な指導を受け、他方のグループは受けていないとする。この2つのグループの標本を比較して、その結果がこれらの標本が選ばれた母集団に一般化される。

小学生の母集団全体にテストをしてみたらどうだろうか。ほとんどの場合、母集団は大きすぎて、時間もお金もかかりすぎる。科学者はこの200年にわたって統計学について大いに学び、非常に小規模な観測値の集合である標本が、それよりはるかに大きな母集団をどれくらいよく代表しているかを正確に評価し、その情報を母集団について判断するために使うことができるようになった。

ある標本がその母集団をどれくらいよく代表しているかが、推測統計を使う上で鍵となるということは、想像できるだろう。代表性は、ある研究の結果が標本から母集団へ一般化できるかどうかを決めるのに役立つ。推測の手続きの結果の正確さは、その標本が母集団からどれくらいよく選ばれたかに大きく依存する。つまり、その標本が母集団をよく代表するほど、その標本から母集団への推測が信頼でき、その結果が母集団へ（そして、他の類似した母集団へも）一般化できる可能性が高くなる。

平均値の t 検定、分散分析、そして回帰分析などの推測統計は、ほとんどの統計学の入門コースで、基礎の一部となっていて、これらはとても広く使われている。要するに、記述統計はデータの特徴を記述することができる一方で、推測統計はそれらの観測結果を、より大きなグループに適用することができるのである。

もっと知るには？ 質問60, 70, 77を参照。

私は統計学者になろうとはしていません。 なぜ統計学の授業を受けるべきなのでしょう？

この答えは、最も明らかな理由 — 統計学の授業をとることが卒業の要件を満たすために必要だ、あるいは、専門の試験のために勉強していて、この分野の知識があることが求められている — だけにとどまらない。

実際、統計学の授業をとる理由は、少なくとも他に4つある。

第1に、これはあなたがこれまで受けてきた教育で経験したことがないかもしれないような、やりがいのある、知的に興味深いトピックだということである。おそらくあなたがこれまで出会ったことのない分野であり、挑戦しがいがあり、スキルを伸ばすことができる、素晴らしい活動なのである。

第2に、データや他の証拠に基づいた意思決定がますます求められる社会にあつて、統計学の理解はかけがえのないものである。統計学を理解すると、データに含まれるパターンや傾向を巧みに理解し、それらを用いて情報に基づいた決定をすることができるようになるだろう。あなたは、どの治療法が最もよいか、数学の指導のあるアプローチ法がなぜ他の方法よりよいのか、あるいは、電子教科書が紙に印刷されたものよりも効果的であるかどうかを理解するために、どんなデータが利用可能かを知らたいかもしれない。統計学は、あなたが問いたいと思った研究に基づくどんな疑問に対しても、どのように問題を提起し、答えたらよいかを教えてくれるツールを集めたものである。

第3に、あなたは同級生や教師、あるいは他の人と、いつでも知的かつ有能に交流できる知識をそなえた市民になるだろう。あるジャーナルの論文の結果が議論される時、あるいは、社会学、医学、行動科学などの興味深い結果についてある結論が提示されたとき、あなたは何が言われているか理解し、鋭敏に、情報に基づいた反応ができるだろう。

最後に、統計学の学習課題は卒業研究、あるいは、もしあなたがすでに大学院生であれば今後の研究のさらなる教育的・専門的な機会のための、素晴らしい方法である。学校教育をすべて終えていたら？ そうならば、統計学の理解は、あなたの専門領域における発展のための、より多くの潜在力を与えてくれる。

もっと知るには？ 質問1, 2, 6を参照。

統計のソフトウェアパッケージは、 どれを使ったらよいでしょうか？ SPSS？ Excel？ 他のものですか？

多くの統計分析ソフトウェアパッケージが利用できる。無料のものも、シェアウェア（そのソフトウェアに価値があると考えたらお金を払う）も、市販の商品もある。どれを選ぶかは、さまざまな要因による。

SPSSは長い間作られている非常に人気のあるソフトウェアパッケージで、主に社会科学や行動科学の研究者たちに使われている。高価であるが、大部分の高等教育機関が学内のサーバにインストールしている。したがって、あなたが学生なら、追加の費用なしで使うことができる。その上、機能は制限されているものの、比較的安価な学生版もある。

Excelは世界で最もよく使われている表計算プログラムで、数値やテキストの行や列の操作ができる。統計分析に特化したプログラムではないが、SPSSなどの製品とほぼ置き換え可能な、非常に多くの関数や内蔵の分析ツールが含まれている。Excelは強力なグラフ作成ツールも備えていて、Wordを含むMicrosoft Officeパッケージの一部なので、Excelで分析した結果をWordの文書に非常に簡単に挿入できる。SPSSと同様に、Excelは多くの教育機関でいつでも使えるようになっている。本書では、例示のためにExcelを使う。Excelはほとんどの初級・中級のユーザにとってすぐに利用可能で、もっとも簡単に使えるソフトウェアだからである。

それではどれを選ぶべきだろうか。SPSSでもExcelでも、本書で議論されているほぼすべてのことを実行できる（それほど普及していない、他の多くのプログラムでも同様である）。しかし、あなたがこれらの1つ、あるいは他のどれかのパッケージを使うと決める際には、以下の諸点を考慮するとよい。

- そのプログラムはどれくらい高価か。あなたの所属機関でお金を請求されずに使えるか。
- 学生版があるか、そして、あなたはそれを購入する資格があるか。
- そのプログラムはモジュール式か。もしそうなら、どのモジュールが必要か。
- 電話やメールでのテクニカルサポートが、いつでも利用可能か。
- そのソフトウェアを使うための適切なハードウェア（メモリやストレージの容量など）を持っているか。

- あなたのコンピュータには、そのソフトウェアを実行できるOSが入っているか。

そのソフトウェアがあなたのニーズに適合するか、購入を決める前に確認すること。しかし、ほぼどんなプログラムでも、基本的なニーズを満たしていることを知っておいてほしい。

もっと知るには？ 質問16, 79, 80を参照。