

Learning Statistics with JASP

A Tutorial for Psychology Students
and Other Beginners

Danielle J. Navarro
David R. Foxcroft
Thomas J. Faulkenberry



Learning Statistics with JASP:
A Tutorial for Psychology Students and Other Beginners

(Version $\frac{1}{\sqrt{2}}$)

Danielle Navarro
University of New South Wales
d.navarro@unsw.edu.au

David Foxcroft
Oxford Brookes University
david.foxcroft@brookes.ac.uk

Thomas J. Faulkenberry
Tarleton State University
faulkenberry@tarleton.edu

<http://www.learnstatswithjasp.com>

Overview

Learning Statistics with JASP covers the contents of an introductory statistics class, as typically taught to undergraduate psychology students. The book discusses how to get started in JASP as well as giving an introduction to data manipulation. From a statistical perspective, the book discusses descriptive statistics and graphing first, followed by chapters on probability theory, sampling and estimation, and null hypothesis testing. After introducing the theory, the book covers the analysis of contingency tables, correlation, t -tests, regression, ANOVA and factor analysis. Bayesian statistics is covered at the end of the book.

Citation

Navarro, D.J., Foxcroft, D.R., & Faulkenberry, T.J. (2019). *Learning Statistics with JASP: A Tutorial for Psychology Students and Other Beginners*. (Version $\frac{1}{\sqrt{2}}$).

日本語版について

この本は **JASP ユーザの会**の有志が翻訳を担当しています。日本語の翻訳メンバーは下記の通りです (50 音順)。

- 紀ノ定保礼 (静岡理工科大学)
- 国里愛彦 (専修大学)
- 小杉考司 (専修大学；代表者。連絡先は kosugi@psy.senshu-u.ac.jp です)
- 小林穂波 (関西学院大学)
- 五島光
- 竹林由武 (福岡県立医科歯科大学)
- 徳岡大 (高松大学)
- 難波修史 (国立研究開発法人理化学研究所)
- 北條大樹 (東京大学)
- 平川真 (広島大学)
- 武藤拓之 (京都大学 こころの未来研究センター)
- 山根嵩史 (川崎医療福祉大学)

この本は完全にオープンソースです。つまり、あなたが望む方法で自由に改変することができます (ただし著者に適切なクレジットを与える限りにおいて、です。ライセンス条項を確認してください)。

最新バージョン

この本は、翻訳の進捗に合わせて随時コンパイルされ、バージョンアップしていきます。最後にコンパイルされたのは 2024 年 4 月 25 日です。

This book is published under a Creative Commons BY-SA license (CC BY-SA) version 4.0. This means that this book can be reused, remixed, retained, revised and redistributed (including commercially) as long as appropriate credit is given to the authors. If you remix, or modify the original version of this open textbook, you must redistribute all versions of this open textbook under the same license - CC BY-SA.

<https://creativecommons.org/licenses/by-sa/4.0/>

The JASP-specific revisions to the original book by Navarro and Foxcroft were made possible by a generous grant to Tom Faulkenberry from the Tarleton State University Center for Instructional Innovation. Also, many thanks to Kristen Bowman for creating the beautiful front and back cover art for the book.

Table of Contents

Preface	ix
I Background	1
II Describing and displaying data with JASP	3
III Statistical theory	5
IV Statistical tools	13
V Endings, alternatives and prospects	15
第 1 章 References	17

第 1/ $\sqrt{2}$ 版に向けた前書き

素晴らしい姉妹本 “jamovi で学ぶ統計” や “R で学ぶ統計” のアレンジ版, “JASP で学ぶ統計” をご紹介できて嬉しく思っています。このバージョンは Dani Navarro と David Foxcroft による素晴らしい前作の上に成り立っています。前作に投入された努力がなければ, この品質は達成できなかったでしょう。このアレンジ版を出そうと思ったとき, 私はシンプルな目標を持っていました。私は Navarro と Foxcroft のテキストを私自身の授業で使いたかったのですが, 直ぐにはそうでないと思ったのは, jamovi ではなく JASP を使っていたからです。どちらも素晴らしいツールなのですが, 私は JASP のほうがちょっとばかり好きなのです。というのも, jamovi がプロジェクトとして独立する前から JASP を使っていたからです。ですから, この本を世界の JASP ユーザに提供できることを嬉しく思っています。

2019 年夏、このオープン教育リソース (Open Educational Resource, OER) の執筆ための助成金を与えてくれたタールトン州立大学の Center for Instructional Innovation に感謝します。私の未来の学生 (そして世界中の学生諸君) に向けて, 高品質な統計テキストを (おそらく未来永劫)100/

(誤字脱字などを見つけるなどして) このテキストをよくしてくれる読者を待っています。もし何か貢献できるとしたら私にメールを送ってください (あるいは Githubg ページに参加してフォークしてください)。やっちゃおうぜ!

Thomas J. Faulkenberry

July 12, 2019

バージョン 0.70 に向けた前書き

バージョン 0.65 から今回へのアップデートでは, いくつかの新しい分析が導入されました。ANOVA の章では, 反復測定 ANOVA や共分散分析 (ANCOVA) のセクションも追加しました。因子分析やそれに関連する技術の章も導入しました。この新しい要素のスタイルは, 本書の他の章と一貫していますが, 目の肥えた読者は少し概念的で実用的な説明が強調され, 代数的要素が減っていることに気づくかもしれませんね。このことが良いことかどうか分かりませんが, 代数については少し後で追加するかもしれません。しかしそれは, 私が統計を理解して教えるときの両方のアプローチ, そして私がコース内で教えているとき学生から受け取ったフィードバックを反映したものです。これに合わせて, 私も本の残りの部分に目を通し, 代数の一部を箱や枠に入れて分離してみました。これらが重要でないとか, 役に立たないというのではなく, 学生の中にはこれらを読み飛ばしたいと思うかもしれないので, 独立したパートにすることでそうした読者の役に立てばと思うのです。

このバージョンについて, 私の学生や同僚, 特に Wakefield Morys-Carter から多くのコメントやフィードバックを受けたことに感謝しています。また世界中のみなさんから提案や修正をいただきま

したことに感謝しています！新しい試みの一つとして、この本のサンプルデータファイルが jamovi にアドオンモジュールとして読み込むことができるようになった、というのがあります。Jonathon Love がこれを援助してくれたことに感謝します。

David Foxcroft
February 1st, 2019

バージョン 0.65 に向けた前書き

本書は Danielle Navarro による 'R で学ぶ統計' の応用で、統計的なソフトウェアや分析例を jamovi に置き換えたものです。R はパワフルな統計プログラミング言語ですが、教師や学生が統計学習の最初に選択するものではありません。教師や学生によっては、ポイントしてクリックするだけで分析できるタイプのソフトウェアを好みますし、それこそ jamovi でできることです。jamovi は R を使う上で二つの側面だけに狙いを定めています。ポイントしてクリックする、グラフィカルユーザインタフェース (GUI) と、多くの機能を組み合わせた関数を提供しており、SPSS や SAS のような方法を R でプログラミングする方法を提供しています。重要なことは、jamovi はいつもフリーでオープンであること、それが中心的な価値の一つであることです。jamovi は科学コミュニティによって作られ、科学コミュニティのためのものなのですから。

このバージョンでは、多くの人に下書きを読んでもらって、幾つもの提案や訂正をもらいました。特に Dr David Emery and Kirsty Walter に感謝しています。

David Foxcroft
July 1st, 2018

バージョン 0.6 にむけた前書き

この本は 2015 年にバージョン 0.5 をリリースしてからそれほど大きく変わってはいません。それでも、前より変わったと言った方がいいと思います。私は 2016 年にアデレードからシドニーに移動し、UNSW での経歴はアデレードの頃に比べて変わってしまったので、こちらにきてから取り組む機会がずいぶん減ってしまったのです。実際に振り返ってみると、少し奇妙な感じもします。ちょっとコメントすると...

- 奇妙なことです，この本では一貫して私の性別を間違えています，これについては私自身に責任があると思います (笑)。12 ページにこのことについて言及した短い脚注があります。現実の生活では，私はジェンダー多様性を認める活動をしていて，この 2 年ほどはほとんど she/her の代名詞を使っています。しかし私は，面倒くさがりなので，この本での文章を訂正しようとは思ってません。
- バージョン 0.6 にむけて，私はそれほど大きく変更せず，指摘してもらったタイプミスやその他の間違いぐらいの，いくつかのマイナーチェンジだけにしました。ただ，セクション 14.4 で触れている **lsr** パッケージ (これはもうメンテナンスされていません) に含まれる `etaSquared` 関数に関する問題については注目してもらいたいと思います。この関数は，本書のようなシンプルな例ではうまく機能するのですが，見つけ切ってはないのですが確かにバグがあるんです! ですから，これについては注意しておいてください。
- 最も大きな変更はライセンスで，私はこれをクリエイティブ・コモンズライセンス (特に CC BY-SA 4.0) のもとでリリースすることにし，誰でも利用できるように全てのソースファイルを GitHub レポジトリに置きました。

おそらく **tidyverse** を使ったバージョンを誰かが書いてくれると思うのですが... これな近々 R にとってもっと重要なトピックになってくるでしょう。:-) では。Danielle Navarro

バージョン 0.5 にむけた前書き

今年もまたアップデートです。今回のアップデートは，本書の理論セクション全体に関わるものです。第 9,10,11 章は書き直しました。よくなっているといいんですが。同時に，17 章も全体的に新しくして，ベイズ統計にフォーカスしました。この変更によって本書は大きく改良されたと思います。私は常に，推測統計全体についての事実が従来型の観点から描かれていることに不満を感じていました。私もすでにベイズ流のデータ分析を自分の仕事に取り入れているのに，です。本書のどこかにベイズの手法を入れることで，本全体として良くなったなと思えるようになりました。今回のアップデートでは他にもやりたいことがいくつかあったのですが，私はいつも授業の締め切りに追われているので，アップデートが後回しになってしまいます! Dan Navarro

February 16, 2015

バージョン 0.4 に向けた前書き

前回の前書きを書いてから一年経ってしまいました。今回はいくつか重要な変更点があります。第 3,4 章は RStudio の特徴について書くのを抑えたので、読みやすくなりましたし、第 12,13 章はカイ二乗検定と t 検定を実行するための lsr パッケージの新しい関数を使うようにしたので、補正に関する議論が lsr パッケージの新しい関数を参照するように対応させました。バージョン 0.4 の電子版では、内部参照 (すなわち、セクションごとの実際のハイパーリンクです) が改良されています。これはバージョン 0.3.1 から導入されたものです。あちこちに新しいことを入れています。多くは誤字脱字の修正 (タイポを見つけてくれたひと全てに感謝します!) で、バージョン 0.3 と 0.4 が全体的に全く違うというようなことはありません。この 12 ヶ月の間、もっと中身を充実させたいと思ってきました。反復測定 ANOVA や混合モデルについての議論がないのは、全く心苦しいところです。言い訳になりますが、進捗が出ないのは私の二人目の子供が 2013 年の初めに生まれたからで、私は昨年のほとんどを生活の維持に注力したからです。結果的に、この本のように報酬のないプロジェクトは、実際に私に賃金が支払われる仕事に追いやられることになりました。今は状況が幾分改善されましたので、バージョン 0.5 ではもっと前進できるでしょう。

私を驚かせたことの一つは、この本を入手するためのダウンロード数です。ウェブサイトの基本的なトラッキング情報を、数ヶ月前から入手できたのですが、(明らかなロボットツールを除いて) この本は一日平均 90 回ダウンロードされています。これには勇気づけられます。少なくとも何人かはこの本が便利だと思ってくれてるってことですから!

Dan Navarro

February 4, 2014

バージョン 0.3 に向けた前書き

心の中では本当にこの本を出版したくないと思ってるんです。完成してないんですから。

私がこういうときは、その言葉通りなんです。参考文献はまた十うんではないし、章のようやぐはセクションタイトルのリストに過ぎないし、索引はないし、読者向けの練習問題はないし、構成は最適とは言えないし、トピックのカバーしている範囲は私の好みを十分に反映していません。さらに、内容的に満足していないところや、書き直さないといけない図もあり、矛盾点や誤字脱字を直す時間も十分にありませんでした。言い換えると、この本は未完成なんです。もし授業の締め切りや数週間後に予定されている赤ちゃんの存在がなければ、私は本当に後悔しなかったと思います。

つまり、もしあなたが大学での教材を探しているとか、Ph.D. の学生さんで R を勉強する方法を探しているとか、統計学の一般的な興味を持っているという人であれば、注意が必要だよと言いたい

のです。あなたが見ているのは最初の下書きで、あなたの目的に沿ったものではないかもしれないからです。もし出版にお金がかかり、インターネットが周りにない世界であれば、こんな形で公開することは決して考えられないでしょう。この本に\$80のお金を出す人がいるかと思うと(これは出版社が販売するにあたって、小売価格を申し出てくれたのです)、ちょっと申し訳なく思います。しかし今は21世紀で、フリーで私のウェブサイトにはPDFを乗せることができ、プリント・オン・デマンドサービスでハードコピーを配布すれば、出版社の教科書の半額ですみます。そして私の罪悪感を和らげるため、シェアしたいと思います!覚えておいてほしいのですが、次のサイトからみなさんは無料でソフトコピー(電子版)を入手できますし、安価なハードコピーもオンラインで入手できます。

Soft copy: <http://www.compcogscisydney.com/learning-statistics-with-r.html>

Hard copy: www.lulu.com/content/13570633

とは言え、溪谷はまだ残っています: あなたが見ているのは、作業中のバージョン0.3です。もしいつの日かバージョン1.0になれば、この仕事に責任を持って、これは誰にでも使って欲しいテキストですと言いたいです。そのとき、私はおそらくインターネットに恥ずかしげもなく後悔し、道具として活用するでしょう。しかしその日が来るまでは、私ははっきりした態度は持てずに、この仕事についてアンビバレントな状態にあるというほかありません。

これを踏まえてですが、この本を強くお勧めするあるグループがあります。2013年の学部生向け研究法(DRIPとDRIP-A)を受講する心理学の学生です。あなたにとって、この本は理想出来なものになるでしょう。というのも、あなたの統計に関する講義に合わせて書かれたものだからです。もしこのノートによる欠点が発覚した場合、直ぐにそのコメントを適用して問題を修正することができます。効果的なことに、あなたのクラスに特化されたテキストを使うことができますし、それは無料で(電子版)あるいは手数料だけで(紙版)利用できるのです。さらに良いことに、このノートはすでに検証済みです。このノートのバージョン0.1は2011年のクラスですでに使われていて、バージョン0.2は2012年のクラスで使われたのです。そして今あなたが見ているのは、新しく改良されたバージョン0.3というわけです。このノートがチタンにメッキされたスティックだというつもりはありません—あなたが学生評価フォームでそう言いたいと思ったかもしれませんが、そのときはどうぞそうしてください—というのも、実際そこまでではないからです。しかし既に何年間か検証されてきていて、うまく機能してきたんだということは言っておきたいと思います。とはいえ、何か問題が生じたときにはわたしたちが直ぐに対応しますし、少なくとも教師の先生方のうち少なくとも一人は隅々までこの本を読んでいることは間違いないのです。

さてそれはさておき、この本が目指しているものが何なのかについて述べておきましょう。中心にある考えは、心理学を学ぶ人に向けて作られた統計の導入的教科書であること、です。ですから、類似の本にあなたが期待するような標準的トピックスはカバーしています。: 研究デザイン、記述統計、仮説検定の理論、 t 検定、 χ^2 検定、ANOVA、回帰などです。しかし、いくつかの章ではRの統計パッケージに言及しています。データの操作やそのほかのスクリプト、プログラミングなんかについての章も。さらにいえば、この本の目次を見たらお気づきになると思いますが、心理学の学生に統

計を教える際、これまでは無視されてきたようなことも多く含まれています。ベイズアンカ頻度主義か、という分断は確率の章で議論されますが、ネイマンとフィッシャーの仮説検定に関する不一致も扱います。確率と密度の違いについても説明します。ANOVA のアンバランスデザインにおける平方和の計算式、タイプ I, II, III の扱いについても触れます。またエピログを見ていただければ、私が追加したかったもっと発展的な要素について明らかになるでしょう。

このアプローチを取る理由は全くシンプルなものです。すなわち、学生が乗りこなせるように、そして楽しめるように、したかったのです。最近の数年間は、心理学の学部生が R を習得するのにほとんど苦労しないことに驚かされています。それは全く簡単だというわけではないですし、成績をつける基準を設定するときは少し優しめにする必要はありますが、最終的にはそこに到達できます。同様に、統計的な考え方に現れる複雑で曖昧な表現を受け入れることに対しても、学生さんはそれほど問題を感じないようです。評価基準が適切に設定されていて、それが提示されている場合は。ですから学生が習得できるのに、教えないわけにいかないでしょう？ その潜在的な能力はとても魅力的でう s。もし彼らが R を学べば、おそらく最大で最も包括的な統計ツールライブラリである CRAN にアクセスできるということでもあるのですから。そしてもし確率理論の詳細について学べば、オーソドックスな帰無仮説検定からベイズアンカの方法に乗り換えようと思ったとき、より乗り換えが容易になるのです。さらに、データ解析技術を学ぶときに高価で独自仕様になっているソフトウェアに捉われることなく、仕事に持って行くこともできます。

残念ながら、この本は全ての問題を解決する決定打ではありません。私の作業は作業中で、いつかは便利な道具になってくれると思います。数あるものの中の一つになる、そう思います。R を使う統計の基本的な導入をしようとする類書はたくさんあります。そして私の本が優れていると考えるほど、私は傲慢ではありません。でも、他の本よりも私はこの本を気に入ってますし、もしかしたら他の人もそう思ってくれるかもしれませんね。

Dan Navarro

January 13, 2013

Part I.

Background

Part II.

Describing and displaying data with JASP

Part III.

Statistical theory

Prelude to Part IV

Part IV of the book is by far the most theoretical, focusing as it does on the theory of statistical inference. Over the next three chapters my goal is to give you an introduction to probability theory (Chapter ??), sampling and estimation (Chapter ??) and statistical hypothesis testing (Chapter ??). Before we get started though, I want to say something about the big picture. Statistical inference is primarily about *learning from data*. The goal is no longer merely to describe our data but to use the data to draw conclusions about the world. To motivate the discussion I want to spend a bit of time talking about a philosophical puzzle known as the *riddle of induction*, because it speaks to an issue that will pop up over and over again throughout the book: statistical inference relies on *assumptions*. This sounds like a bad thing. In everyday life people say things like “you should never make assumptions”, and psychology classes often talk about assumptions and biases as bad things that we should try to avoid. From bitter personal experience I have learned never to say such things around philosophers!

On the limits of logical reasoning

The whole art of war consists in getting at what is on the other side of the hill, or, in other words, in learning what we do not know from what we do.

– Arthur Wellesley, 1st Duke of Wellington

I am told that quote above came about as a consequence of a carriage ride across the countryside.^{*1} He and his companion, J. W. Croker, were playing a guessing game, each trying to predict what would be on the other side of each hill. In every case it turned out that Wellesley was right and Croker was wrong. Many years later when Wellesley was asked about the game he explained that “the whole art of war consists in getting at what is on the other side of the hill”. Indeed, war is not special in this respect. All of life is a guessing game of one form or another, and getting by on a day to day basis requires us to make good guesses. So let’s play a guessing game of our own.

Suppose you and I are observing the Wellesley-Croker competition and after every three hills you and I have to predict who will win the next one, Wellesley or Croker. Let’s say that *W* refers to a Wellesley victory and *C* refers to a Croker victory. After three hills, our data set looks like this:

^{*1}Source: <http://www.bartleby.com/344/400.html>.

WWW

Our conversation goes like this:

YOU: Three in a row doesn't mean much. I suppose Wellesley might be better at this than Croker, but it might just be luck. Still, I'm a bit of a gambler. I'll bet on Wellesley.

ME: I agree that three in a row isn't informative and I see no reason to prefer Wellesley's guesses over Croker's. I can't justify betting at this stage. Sorry. No bet for me.

Your gamble paid off: three more hills go by and Wellesley wins all three. Going into the next round of our game the score is 1-0 in favour of you and our data set looks like this:

WWW WWW

I've organised the data into blocks of three so that you can see which batch corresponds to the observations that we had available at each step in our little side game. After seeing this new batch, our conversation continues:

YOU: Six wins in a row for Duke Wellesley. This is starting to feel a bit suspicious. I'm still not certain, but I reckon that he's going to win the next one too.

ME: I guess I don't see that. Sure, I agree that Wellesley has won six in a row, but I don't see any logical reason why that means he'll win the seventh one. No bet.

YOU: Do you really think so? Fair enough, but my bet worked out last time and I'm okay with my choice.

For a second time you were right, and for a second time I was wrong. Wellesley wins the next three hills, extending his winning record against Croker to 9-0. The data set available to us is now this:

WWW WWW WWW

And our conversation goes like this:

YOU: Okay, this is pretty obvious. Wellesley is way better at this game. We both agree he's going to win the next hill, right?

ME: Is there really any logical evidence for that? Before we started this game, there were lots of possibilities for the first 10 outcomes, and I had no idea which one to expect. WWW WWW WWW W was one possibility, but so was WCC CWC WWC C and WWW WWW WWW C or even CCC CCC CCC C. Because I had no idea what would happen so I'd have said they were all equally likely. I assume you would have too, right? I mean, that's what it *means* to say you have "no idea", isn't it?

YOU: I suppose so.

ME: Well then, the observations we've made logically rule out all possibilities except two: WWW WWW WWW C or WWW WWW WWW W. Both of these are perfectly consistent with the evidence we've encountered so far, aren't they?

YOU: Yes, of course they are. Where are you going with this?

ME: So what's changed then? At the start of our game, you'd have agreed with me that these are equally plausible and none of the evidence that we've encountered has discriminated between these two possibilities. Therefore, both of these possibilities remain equally plausible and I see no logical reason to prefer one over the other. So yes, while I agree with you that Wellesley's run of 9 wins in a row is remarkable, I can't think of a good reason to think he'll win the 10th hill. No bet.

YOU: I see your point, but I'm still willing to chance it. I'm betting on Wellesley.

Wellesley's winning streak continues for the next three hills. The score in the Wellesley-Croker game is now 12-0, and the score in our game is now 3-0. As we approach the fourth round of our game, our data set is this:

WWW WWW WWW WWW

and the conversation continues:

YOU: Oh yeah! Three more wins for Wellesley and another victory for me. Admit it, I was right about him! I guess we're both betting on Wellesley this time around, right?

ME: I don't know what to think. I feel like we're in the same situation we were in last round, and nothing much has changed. There are only two legitimate possibilities for a sequence of 13 hills that haven't already been ruled out, WWW WWW WWW WWW C and WWW WWW WWW WWW W. It's just like I said last time. If all possible outcomes were equally sensible before the game started, shouldn't

these two be equally sensible now given that our observations don't rule out either one? I agree that it feels like Wellesley is on an amazing winning streak, but where's the logical evidence that the streak will continue?

YOU: I think you're being unreasonable. Why not take a look at *our* scorecard, if you need evidence? You're the expert on statistics and you've been using this fancy logical analysis, but the fact is you're losing. I'm just relying on common sense and I'm winning. Maybe you should switch strategies.

ME: Hmm, that is a good point and I don't want to lose the game, but I'm afraid I don't see any logical evidence that your strategy is better than mine. It seems to me that if there were someone else watching our game, what they'd have observed is a run of three wins to you. Their data would look like this: YYY. Logically, I don't see that this is any different to our first round of watching Wellesley and Croker. Three wins to you doesn't seem like a lot of evidence, and I see no reason to think that your strategy is working out any better than mine. If I didn't think that WWW was good evidence then for Wellesley being better than Croker at *their* game, surely I have no reason now to think that YYY is good evidence that you're better at *ours*?

YOU: Okay, now I think you're being a jerk.

ME: I don't see the logical evidence for that.

Learning without making assumptions is a myth

There are lots of different ways in which we could dissect this dialogue, but since this is a statistics book pitched at psychologists and not an introduction to the philosophy and psychology of reasoning, I'll keep it brief. What I've described above is sometimes referred to as the riddle of induction. It seems entirely *reasonable* to think that a 12-0 winning record by Wellesley is pretty strong evidence that he will win the 13th game, but it is not easy to provide a proper logical justification for this belief. On the contrary, despite the *obviousness* of the answer, it's not actually possible to justify betting on Wellesley without relying on some assumption that you don't have any logical justification for.

The riddle of induction is most associated with the philosophical work of David Hume and more recently Nelson Goodman, but you can find examples of the problem popping up in fields as diverse as literature (Lewis Carroll) and machine learning (the "no free lunch" theorem). There really is something weird about trying to "learn what we do not know from what we do know". The critical

point is that assumptions and biases are unavoidable if you want to learn anything about the world. There is no escape from this, and it is just as true for statistical inference as it is for human reasoning. In the dialogue I was taking aim at your perfectly sensible inferences as a human being, but the common sense reasoning that you relied on is no different to what a statistician would have done. Your “common sense” half of the dialog relied on an implicit *assumption* that there exists some difference in skill between Wellesley and Croker, and what you were doing was trying to work out what that difference in skill level would be. My “logical analysis” rejects that assumption entirely. All I was willing to accept is that there are sequences of wins and losses and that I did not know which sequences would be observed. Throughout the dialogue I kept insisting that all logically possible data sets were equally plausible at the start of the Wellesely-Croker game, and the only way in which I ever revised my beliefs was to eliminate those possibilities that were factually inconsistent with the observations.

That sounds perfectly sensible on its own terms. In fact, it even sounds like the hallmark of good deductive reasoning. Like Sherlock Holmes, my approach was to rule out that which is impossible in the hope that what would be left is the truth. Yet as we saw, ruling out the impossible *never* led me to make a prediction. On its own terms everything I said in my half of the dialogue was entirely correct. An inability to make any predictions is the logical consequence of making “no assumptions”. In the end I lost our game because you did make some assumptions and those assumptions turned out to be right. Skill is a real thing, and because you believed in the existence of skill you were able to learn that Wellesley had more of it than Croker. Had you relied on a less sensible assumption to drive your learning you might not have won the game.

Ultimately there are two things you should take away from this. First, as I’ve said, you cannot avoid making assumptions if you want to learn anything from your data. But second, once you realise that assumptions are necessary it becomes important to make sure you *make the right ones!* A data analysis that relies on few assumptions is not necessarily better than one that makes many assumptions, it all depends on whether those assumptions are good ones for your data. As we go through the rest of this book I’ll often point out the assumptions that underpin a particular statistical technique, and how you can check whether those assumptions are sensible.

Part IV.

Statistical tools

Part V.

Endings, alternatives and prospects

1. References
