

B.5.2 不可微函数

不可微函数的问题在概率论中有时会出现. 特别是当人们求解如第 2 章研究的那些函数方程时, 可微性会让一大群数学吹毛求疵者出现, 声称我们将一大类重要的潜在可能解排除在外. 但是我们注意到情况并非如此. 奥采尔证明考克斯的函数方程都可以在不假设可微性 (以更长的推导过程为代价) 的情况下求解出来, 并且得到与我们相同的解.

让我们仔细看看一般不可微函数的概念. 起初, 不可微函数概念并没有被数学家们所接受. 埃尔米特曾经写信给斯蒂尔杰斯: “我惊恐地远离了这种可怕的不可微函数的瘟疫.” 人们通常将这场瘟疫归咎于勒贝格 (1875—1941), 尽管魏尔斯特拉斯在他之前注意到了这种函数. 魏尔斯特拉斯不可微函数是

$$f(x) \equiv \sum_{n=0}^{\infty} a^n \cos(m^n x), \quad (\text{B.13})$$

其中 $0 < a < 1$ 且 m 是正奇数. 因为 m^n 总是一个整数, 它是一个周期为 2π 的普通傅里叶级数. 此外, 该级数对于所有实数 x 是一致收敛的 (因为它必须至少和 $\sum a^n$ 一样收敛), 因此它定义了一个连续函数. 但是如果 $ma > 1$, 逐项微分会产生一个严重发散的序列, 其系数在 n 中呈指数增长. 导数

$$f'(x) \equiv \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (\text{B.14})$$

对于任何 x 都不存在的证明是相当冗长的.^① 魏尔斯特拉斯函数实际上是好函数序列 (前 k 项的部分和 S_k) 的极限, 但不是一个行为良好的极限. 因为这种函数不满足条件 (B.12), 所以它们对我们没有明显的用处. 然而, 在实际应用中确实会出现这样的函数. 例如在第 7 章中, 我们尝试通过傅里叶变换方法求解积分方程 (7.49) 时, 如果核函数太宽, 就会遇到这种困难. 这样, 我们的结论是积分方程没有任何可用的解, 除非核函数 $\phi(x-y)$ 至少与 “驱动” 函数 $f(x)$ 一样尖锐.

B.5.3 臆造的不可微函数

最常被引用的不可微函数的例子是根据序列 $f(x)$ 导出的, 其中每个序列都是一个等腰直角三角形, 其斜边位于实轴上并且长度为 $1/n$. 当 $n \rightarrow +\infty$ 时, 三角形的大小缩为 0. 对于任何有限的 n , $f_n(x)$ 的斜率几乎处处为 ± 1 . 那么当 $n \rightarrow +\infty$ 时会发生什么呢? 极限 $f_\infty(x)$ 经常被漫不经心地视为不可微函数. 显然, 导数的极限 $f'_n(x)$ 不存在, 但这里讨论的是极限 $f_\infty(x) \equiv 0$ 的导数, 它当然

^① 见哈代的文章 (Hardy, 1911). 蒂奇马什 (Titchmarsh, 1939, 第 350~353 页) 只在 $ma > 1 + 3\pi/2$ 时给出了一个更短的有效证明. 一些作者指出 $f(x)$ 仅在这种情况下是不可微的, 但是据我们所知, 没有人声称哈代的证明存在错误.

是可微的. 我们可以定义任意数量的、在越来越精细尺度上具有不连续斜率的这种序列 $f_n(x)$. 基于导数的极限不存在, 将结果极限 $f_\infty(x)$ 称为不可微的错误在文献中很常见. 在许多情况下, 这样一系列坏函数的极限实际上是一个行为良好的函数 (虽然定义很笨拙), 没有理由将它排除在我们的系统之外.

勒贝格针对他的批评者这样为自己辩护: “如果有人希望总是将自己限制在行为良好的函数上, 那么就必须拒绝很久以前提出的许多简单问题的解.” 我无法举出任何已解决的具体问题, 但是可以借用勒贝格的论证来捍卫自己的立场.

拒绝好函数序列的极限就要拒绝许多当前实际问题的解. 这些极限可以而且确实用于许多有用的目的, 而当前许多数学教育与实践仍在试图消除这些目的. 事实上, 拒绝承认德尔塔函数是合法数学对象导致数学家犯了错误. 例如, 克拉默 (Cramér, 1946, 第 32 章) 给出了一个我们曾在第 17 章导出的不等式, 为参数估计量 θ^* 的抽样分布的方差设置了下限:

$$\text{var}(\theta^*) \geq \frac{(1 + \text{dB}/\text{d}\theta)^2}{n \int \text{d}x (\partial \ln f / \partial \theta)^2 f(x|\theta)}, \quad (\text{B.15})$$

其中从抽样分布 $f(x|\theta)$ 中进行了 n 次观测, $b(\theta^*) \equiv E(\theta^* - \theta)$ 是估计量的偏差.

然后克拉默指出: 如果 $f(x|\theta)$ 具有不连续性, 那么 “常规条件的条件通常不满足. 在这种情况下, 通常可以找到 ‘异常高’ 精度的无偏估计, 即方差小于常规估计的下限 (B.15)”. 既然 (B.15) 只是施瓦茨不等式, 似乎不承认例外, 他是如何得出如此惊人的结论的呢? 我们发现他使用了函数的集合论定义, 并得出导数 $\partial \ln f / \partial \theta$ 在不连续点处不存在的结论. 所以他只在 $f(x|\theta)$ 连续的那些区域上对 (B.15) 进行积分.

但是在分析中恰当的不连续函数的定义是我们对连续函数序列的极限. 当我们逼近这个极限时, 导数会出现一个更高、更尖锐的峰值. 无论我们离那个极限有多近, 尖峰都是正确函数导数的一部分, 它的贡献必须包含在精确积分中. 因此, 不连续函数 $g(x)$ 的导数必然包含不连续点 y 处的一个德尔塔函数 $[g(y^+) - g(y^-)] \delta(x - y)$, 在 $g(x)$ 的微分傅里叶级数中其贡献始终存在, 并且必须包括在内以获得正确的物理解. 如果克拉默包括这一项, (B.15) 的极限就会变为 $\text{var}(\theta^*) \geq 0$, 这几乎是一个无用的陈述, 但至少不会出现异常, 也不会违反施瓦茨不等式.

类似地, 具有有限极限的形式为

$$\int_a^b \text{d}y K(x, y) f(y) = \lambda g(x), \quad (\text{B.16})$$

积分方程的解通常涉及端点处的德尔塔函数, 如 $\delta(y - a)$ 或 $\delta'(y - b)$, 等等, 因此不相信德尔塔函数的人认为此类积分方程没有解. 但在实际物理问题中, 正是

这样的积分方程反复出现,而且必须再次包含德尔塔函数才能获得正确的物理解。米德尔顿(Middleton, 1960)给出了一些例子,它们在统计力学中对不可逆过程的预测中几乎无处不在。令人惊讶的是,很少有非物理学家意识到需要包含德尔塔函数,但是我们认为这仅仅说明了我们独立观察到的情况。那些从集合论的角度考虑基本原理的人没有看到它的局限性,因为他们几乎从不进行有用、实质性的计算。

因此,臆造的不可微函数通过越来越微小的三角形序列极限被制造出来,并且被不加批判地接受。那些这样做同时又以怀疑态度对待德尔塔函数的人承认坏函数序列的极限是合法数学对象,同时拒绝承认好函数序列的极限!在我们看来,这似乎是一种病态的策略,因为德尔塔函数在真实、实质性的计算中可以用于许多基本目的,但我们无法想象不可微函数可以提供任何有用的目的。它们的唯一用途似乎是,为捣乱分子对于任何人可以做出的任何合理及有用的数学描述提供人为设计的反例。亨利·庞加莱(Poincaré, 1909)以他特有的简洁方式指出了这一点:

过去,当人们发明一种新函数时,他们的脑海中会有一些有用的目的——现在他们发明它们,只是为了推翻祖宗的推理,而这正是他们试图摆脱的一切。

我们要指出的是,这些捣乱分子毕竟没有推翻我们祖宗的推理,病态函数的出现只是因为他们偷偷地采用与祖宗所用不同的“函数”一词的定义。如果指出这一点,就很明显没有必要修改祖宗的结论。

今天,这种人为设计数学病态函数的风潮似乎已经走到了尽头,而这正是庞加莱预见的原因,用它不能做任何有用的事情。虽然我们仍然看到劝告不要假设未知函数的可微性,但在最近的文献中很难找到不可微函数出现的具体例子——更不用说实际用于任何事情了。人们必须回到像蒂奇马什(Titchmarsh, 1939)这样的老作品才能看到它们。

因此,请注意,我们也通过以适合我们主题的方式定义术语“函数”来消除这种“瘟疫”。对某个领域“适用”的数学概念的定义是允许其定理具有最大范围的有效性和有用性,而不需要一长串例外、特殊情况和其他异常情况。在我们的工作中,术语“函数”包括好函数和好函数序列的良好行为极限,但不包含不可微函数。我们不否认包含不可微函数的其他定义的存在,正如我们不否认荧光紫色染发剂在英国存在一样。在这两种情况下,我们都只是不会用到它们。^①

^① 关于一个不同的主题,在第 17 章(第 487 页脚注^①)中,我们遵循相同的策略,通过定义术语有限时间序列的“移动平均”,使得我们的定理变得都是精确的,不需要任何混乱的“最终效果”校正。当然,这是在实际应用中最直接有用的定义,并且保存了在其他情况下会丢失的信息。

B.6 无限集合计数?

众所周知, 刘易斯·卡罗尔的儿童读物实际上是对逻辑原理的阐述, 通过以一种即使对小孩子来说也觉得可笑的形式来传达. 他的一首诗这样结尾:

他以为自己看到一种证明他是教皇的论证,

再一看, 他发现那只是一块斑纹皂.

“一个如此可怕的事实,” 他淡然说道, “让所有希望都破灭!”

事实上, 概率论中许多严肃提出的论证, 在仔细考虑后只不过是斑纹皂. 剑桥数学家哈代的著名轶事^①印证了这一想法. 麦克塔格特质疑从一个错误的命题出发可以推导出所有命题的说法, 他这样挑战哈代: “假定 $2 + 2 = 5$, 请证明我是教皇.” 哈代回答: “两边各减去 3, 这样我们就有 $1 = 2$. 现在我们认同教皇和你是两个人, 因此教皇和你是同一个人!” 但这只是文字游戏, 无限集合理论为我们提供了更高级的斑纹皂, 可以用它以更令人信服的方式证明麦克塔格特的教皇身份.

我们的假设是: 如果两个集合可以一一对应, 它们就具有相同的元素个数. 然后, 对于 $n = 1, 2, \dots$, 通过关联 $n \leftrightarrow 2n$, 我们可以将正整数与正偶数进行一一对应; 通过关联 $2n \leftrightarrow 2n - 1$, 我们同样可以将正偶数与正奇数进行一一对应. 因此, 根据这种逻辑, 我们似乎会得出以下结论:

- (A) 整数个数 = 偶数个数;
- (B) 偶数个数 = 奇数个数;
- (C) 整数个数 = $2 \times$ 偶数个数.

从 (A) 和 (C) 可以得出 $1 = 2$. 这里的推理方式与 (15.2) 和 (15.3) 中的推理方式没有太大不同.

我们的观点是, 除非作为有限集合的极限, “所有整数的集合” 是没有定义的. 如果以这种方式逼近, 通过引入具体取极限的过程, 无论我们选择什么极限过程, 都不会产生矛盾. 即使 (偶数个数)/(整数个数) 的极限比率可以是 $0 \leq x \leq 1$ 中的任意 x , (奇数个数)/(整数个数) 的极限将为 $1-x$, 我们的计数也将在极限时保持一致.

例如, 每个整数在序列 $\{1, 3, 2, 5, 7, 4, \dots\}$ 中只出现一次, 其中我们交替取两个奇数和一个偶数. 然后, 只计算由该序列的前 n 个元素组成的有限集合中的元素, 并在进行计数后取极限 $n \rightarrow +\infty$, 我们将不会得到以上不一致的命题 (A)(B)(C), 而是得到一致性的命题集合:

^① 引用自杰弗里斯的著作 (Jeffreys, 1931, 1957 版, 第 18 页).

(A') 整数个数 = $3 \times$ 偶数个数;

(B') 偶数个数 = $1/2 \times$ 奇数个数;

(C') 整数个数 = 偶数个数 + 奇数个数.

这些想法并不像人们想象的那么新颖. 伽利略在《关于两种新科学的对话》(Galileo, 1638) 中曾指出了两个奇怪的事实. 一方面, 每个整数有且只有一个平方数, 而且没有两个平方数是相同的, 从中可以看出整数的数量和平方数的数量必须相同. 另一方面, 很明显有许多整数(在某种意义上是它们中的“绝大多数”)不是平方数. 他从中得出了一个非常明智的结论:

这是当我们试图用我们有限的头脑讨论无限时出现的困难之一, 将我们赋予有限的性质分配给无限. 但是我认为这是错误的, 因为我们不能将一个无限大量说成大于或小于或等于另一个无限大量.

如下所述, 300 年后, 赫尔曼·外尔表达了几乎完全相同的判断, 可见外尔 (Weyl, 1949).

B.7 豪斯多夫球体悖论与数学病理学

上述不一致陈述在结构上几乎与关于球体上全等集合的豪斯多夫悖论相同, 除了上升到不可数集合外(这里 X, Y, Z 是几乎覆盖球体的不相交集, 因为球体的旋转使 X 与 Y 重合, 所以 X 与 Y 全等, 同样 Y 与 Z 全等. 但特别的是, X 也与 Y 和 Z 的并集全等, 尽管 $Y \neq Z$). 我们与庞加莱和外尔一样, 对数学家如何接受和发表这样的结果感到困惑. 为什么他们没有看到其中存在明显的矛盾, 使得他们所用的推理无效呢?

尽管如此, 萨维奇 (Savage, 1962) 接受这个二律背反作为完全的事实, 并将其应用到概率论中. 他说, 有人可能会很轻率地认为球体上的全等集合具有同样的可能性, 但是豪斯多夫的结果表明他的信念实际上不具有这种性质. 深入思考之后, 我被迫得出相反的结论: 我相信存在一种知识状态, 认为球体上的全等集合具有同样的可能性, 而且这比我导致豪斯多夫结果的推理可靠的信念要强得多.

豪斯多夫球体悖论和罗素的理发师悖论^①大概可以有类似的解释: 从一个具有自相矛盾性质的奇怪集合定义出发, 我们当然可以从中推导出任何荒谬的命题. 豪斯多夫为他的著作冠名“集合论”. 庞加莱曾说过一句著名的俏皮话: “后人将

^① 理发师悖论是罗素悖论中的一个典型. 有一位理发师声称: “我将为本城所有不给自己刮胡子的人刮胡子, 我也只给这些人刮胡子.” 但有一天, 这位理发师从镜子里看见自己的胡子长了, 那么他能不能给自己刮胡子呢? 如果他不给自己刮, 他就属于“不给自己刮胡子的人”, 需要给自己刮胡子; 如果他给自己刮胡子, 他又属于“给自己刮胡子的人”, 就不该给自己刮胡子. ——编者注

会把集合论视为一种已经痊愈的疾病。”但是他会为 80 年后这种疾病仍未痊愈而感到震惊。尽管如此，庞加莱的观点仍然被今天的应用数学家普遍接受。

例如，在 1983 年，我曾听过一位非常杰出的统计学家的演讲，其中报告了一项历史调查。他评论道：“我惊讶地发现，在布尔巴基时代之前，法国人实际上已经产生了一些有用的数学。”最近，诺贝尔奖获得者理论物理学家默里·盖尔曼（Murray Gell-Mann, 1992）讨论了这种情况。他认为，现代数学对物理学仍有很多价值，纯数学与科学的分歧部分是由于布尔巴基主义者的晦涩语言以及不愿详细写出任何非平凡的例子而产生的幻觉。他总结道：“纯粹数学和科学终于重聚了，幸运的是，布尔巴基‘瘟疫’正在消失。”

我们希望能如此乐观。在我们看来，这场瘟疫比单纯的语言晦涩要严重得多，它感染了纯数学的核心部分。心智健全的人不可能对其中任何部分有信心，必须找到防止出现这些荒谬悖论的规则，而这要求我们的数学教科书必须全部改写。众所周知，罗素的类型理论可以解决一些悖论，但远不能解决所有悖论。我们担心的是，即使双方都表现出最好的意愿，至少还需要另一代人的努力才能够实现纯数学和科学的和解。就目前而言，在试图将成果出口到其他领域之前，专门研究无限集合理论的人有责任整理好自己的房子。在此之前，我们这些在概率论或任何其他应用数学领域工作的人，有权要求与这种我们对之并不负有责任的疾病隔离开，并使其远离我们的领域。

在这方面，我们并不孤单，并且确实得到了许多非布尔巴基主义数学家的支持。在前言中，我们引用了莫里斯·克莱因（Morris Kline, 1980）的话，关于允许无限集合理论在应用数学中立足的危险。他反过来（在第 237 页）引用了赫尔曼·外尔的话。布劳威尔和外尔都指出，经典逻辑是被发展用于有限集合的。用外尔的话来说，将经典逻辑毫无根据地应用于无限集合的尝试是“……集合论的堕落和原罪，为此它已经受到了二律背反的惩罚。这种矛盾的出现并不令人惊讶，但是它们出现在游戏的后期阶段的确令人感到惊讶”。

但是对于这些悖论的延迟出现有一个简单的解释。在第 15 章中，我们用示例指出：如果一个错误的论证会立即导致荒谬的结果，它也将被立即放弃，我们就将从来没有听说过它。如果它在前两三次尝试中产生了合理的结果，那么在一些问题中它还会继续成功。人们会继续使用它，但一开始是保守的——用在非常相似的问题上，因此很可能会继续给出合理的结果。只是到后来，当人们变得过于自信并试图将其应用于不同类型的问题时，矛盾才会显现。^①

^① 我认识赫尔曼·外尔，曾在普林斯顿上过他的群论课程，并钦佩他是群论和广义相对论变分原理的最终权威。但是普林斯顿的布尔巴基主义数学家们嘲笑他，在背后称他为“神圣的赫尔曼”，因为他刚才的话就像《圣经》中对美德的劝勉一样。他们最好还是听他的。

同样的现象也发生在正统统计中，其中一些诸如置信区间的特定工具在很长一段时间内产生了可接受的结果。这是由于它们最初仅仅用于没有冗余参数但是具有充分统计量且没有非常重要的先验信息的简单问题。没有人注意到这样一个事实，即其数值结果与同一水平的贝叶斯后验概率区间相同（基于杰弗里斯给出的无信息先验）。置信区间被奈曼、克拉默和威尔克斯等数学家广泛应用，并被认为是相对贝叶斯方法的巨大进步。直到当人们试图将它们应用于更一般的问题，其中的矛盾才开始出现。^①最后，杰恩斯（Jaynes, 1976）证明置信区间作为推断只有在它们碰巧与贝叶斯区间一致的那些特殊情况下是令人满意的。

克莱因（Kline, 1980, 第 285 页）在这个问题上也引用了威拉德·吉布斯的话：“纯数学家可以随心所欲，但是应用数学家必须至少保持部分理智。”无论如何，理智的人不会试图在实际应用中使用诸如豪斯多夫球体悖论之类的异常理论。

最后，我们提供一些关于数学风格的更一般性的评论。

B.8 我应该发表什么？

萨维奇（Savage, 1962）曾经用这个问题表达了他的困惑。无论他选择讨论什么话题，采用什么写作风格，都肯定会因为没有做出不同的选择而受到批评。在这一点上，他并不孤单。我们希望呼吁对个体差异多一点儿宽容。

如果有人想把注意力集中在无限集合、测度论和一般的数学病理学上，他完全有权这样做。他不需要通过列举出相应的应用为之辩护或者为实际应用的缺乏而道歉。正如好久以前我们就已经意识到的那样，抽象数学有其自身的价值。

但是反过来，其他人也拥有同等的权利。如果我们选择专注于数学中那些在实际问题中 useful 并且我们能够正确地进行重要计算的方面——数学病理学家可能从来没有考虑过——我们可以自由地这样做而无须道歉。

最后，本书对数学层次和深度的选择，目的是让所有读者都能从中提取他们想要的东西。由于那些以挑剔风格为唯一目的的人总是能够这样做，我们的目标是确保那些真诚希望理解其内容的人也可以这样做。因此，我们试图给出令人信服的理由，说明为什么我们提出的想法是“显而易见的”，而我们批评的则不是，只要这样做足够简短，不会打断论证主线。这会不可避免地留下一些空白，部分由大多数章节末尾的评述填补。

在这方面，什么是或不是“显而易见”的问题是两个方面相反的技巧问题。一

① 置信区间作为关于估计量抽样性质的陈述总是正确的。然而，作为对推断参数值的陈述可能是荒谬的。例如，整个置信区间可能位于参数空间的某个区域里。通过对数据进行演绎推理，我们知道该参数不可能存在于该区域。

方面，引入经不起批判的概念——或者贬低那些站得住脚而无可辩驳的概念——的标准方式就是称其为“显而易见的”。另一方面，对显而易见的简单问题表示严重怀疑，则是将自己的深刻批判能力求全于不具备如此批判能力的其他人的标准技巧。我们试图在这两者之间选择一条中间路线，但是就像萨维奇一样，我们知道无论我们做出什么选择，都会受到其中一种类型读者的批评。

我们要避免一个常见的错误：没有什么比某个领域的作者声称数学严格性“保证了结果的正确性”更可悲的错误了。相反，经验告诉我们，越是专注于表面上的数学严格性，人们对现实世界中前提的有效性就会关注得越少，也就越有可能得出在现实世界中荒谬的最终结论。

B.9 数学礼仪

几年前，我参加了一位刚获得博士学位的年轻数学家的研讨会演讲。我知道，他得出了概率论中的一个了不起的新极限定理。他从一开始就定义他准备使用的集合，但是三块黑板还不够他写完——他一直没有写完集合定义。一小时后，当演讲不得不结束时，我们困惑地走出房间，甚至不知道他的定理是什么。

像庞加莱这样的“19 世纪数学家”也许会在几分钟内进入计算的核心，并及时完成证明，指出其后果以供进一步讨论。

这位年轻人不应该受到指责，他只是在做他被教导的“20 世纪数学家”必须做的事情。虽然他现在可能已经学会了如何更好地规划自己的演讲，但是肯定仍会浪费自己和他人的大部分时间来背诵 20 世纪数学要求的所有咒语，然后才进入实际问题。他不是遵循了更高、更严格的标准，而是研究了现代数学礼仪的受害者。

现在，如果没有指定某个集合或“空间” X 而引入一个变量 x ，你就会被指责处理的是一个未定义的问题。如果对一个函数 $f(x)$ 取微分时没有首先说明它是可微的，你就会被指责缺乏严格性。如果注意到函数 $f(x)$ 有一些对具体应用来说很自然的特殊性质，你就会被指责缺乏一般性。换句话说，你所做的每一个陈述都会受到不合礼仪的解释。

显然，如果我们的陈述中没有一些像样的精确标准，就无法传达数学结果。但是，对某种特定形式的精确性和一般性的狂热坚持可能会过头，以致我们不能达到最初的目的。20 世纪的数学常常退化为一种无聊的对抗游戏，而不是一种交流过程。

数学礼仪的狂热者根本不试图理解你的实质性信息，而只是试图挑剔你的呈现风格。只要能找到任何这样做的方法，他就会努力将你说出的内容解读为无意

义。为了自我防卫，作者们不得不把注意力集中在怎么说上，而不是说什么。而前者通常是细枝末节、不相关、吹毛求疵的东西。这会导致文字长度增加，而内容量减少。

如果我们采取不同的态度，数学交流可能会变得更加有效和愉快。对于对别人写的东西做出有礼貌解释的人来说，引入 x 作为变量这一事实已经暗示存在可能值的集合 X 。为什么每次引入变量时都需要重复该话语，从而在只需要一个符号的地方用两个符号？[实际上，值域通常会在重要的地方更清楚地表示出来，方法是在等式后添加 $(0 < x < 1)$ 等条件。]

对于有礼貌的读者来说，作者将 $f(x)$ 微分两次这一事实本身已经意味着他认为它是可微的。为什么要求他把每件事都说两遍？如果他证明命题 A 的一般性足以涵盖他的应用，为什么他必须额外说明不相关的、使 A 为真的最一般的可能条件？

与狂热主义者同样令人讨厌的是他的兄弟——强迫性的数学吹毛求疵者。我们希望作者定义他的术语，然后以与他的定义一致的方式使用它们。但是如果有其他作者曾经使用这个词表示略微不同的含义，吹毛求疵者就会指责他使用不一致的术语。我曾多次遭受这种折磨，我的同事也诉说过同样的经历。

19 世纪的数学家并不是不严格，只是理所当然地向他人提供简单的文化礼仪，并期望得到回报。这将导致人们试图解读他人所写内容的意义。这可能是在考虑到全部上下文的情况下完成的。不要将我们对每一本数学著作的阅读都变成对是否偏离官方风格的审判。

由于同情这位年轻人的困境，而且不想像他一样被奴役，我们发出以下公告。

解放宣言

我们引入的每个变量 x 都应该被理解为具有某个可能值的集合 X 。我们引入的每个函数 $f(x)$ 都应该被理解为行为足够好，因此我们用它做的事情是有意义的。我们承诺，每一个证明都足以涵盖我们对它的应用。对这个问题感兴趣的读者可以将其作为家庭作业，目的是找出结果所适用的最一般条件。

通过制作一个包含此公告的图章，我们可以将许多 19 世纪的数学作品转换为 20 世纪的标准。也许还有一个使用术语“ σ 代数、博雷尔场、拉东-尼科迪姆导数”的句子，印在第一页上。

现代作者可以通过在版权信息中包含这样的公告，然后以 19 世纪的风格写作，从而大大缩短他们的作品，提高可读性并且不会减少内容的信息量。也许一

些出版商看到这些话，可能会出于经济原因要求他们这样做，这将是一种对科学的服务。

在本附录中，我们提供了许多没有参考文献的简短引文。在贝尔（Bell, 1937）、菲利克斯（Félix, 1960）、克莱因（Kline, 1980）、罗和麦克利里（Rowe & McCleary, 1989）等人的著作中可以找到许多其他有趣细节的支持内容。

附录 C 卷积和累积量

首先, 我们注意到一些与概率论无关的一般数学事实. 给定一组定义在实数轴上并且未必非负的实函数 $f_1(x), f_2(x), \dots, f_n(x)$, 假设它们的积分 (零阶矩) 和第一阶、第二阶、第三阶矩存在:

$$\begin{aligned} Z_i &\equiv \int_{-\infty}^{+\infty} dx f_i(x) < +\infty, & S_i &\equiv \int_{-\infty}^{+\infty} dx x^2 f_i(x) < +\infty, \\ F_i &\equiv \int_{-\infty}^{+\infty} dx x f_i(x) < +\infty, & T_i &\equiv \int_{-\infty}^{+\infty} dx x^3 f_i(x) < +\infty. \end{aligned} \quad (\text{C.1})$$

f_1 与 f_2 的卷积定义为

$$h(x) \equiv \int_{-\infty}^{+\infty} dy f_1(y) f_2(x-y), \quad (\text{C.2})$$

或简记为 $h = f_1 * f_2$. 卷积是可结合的: $(f_1 * f_2) * f_3 = f_1 * (f_2 * f_3)$, 因此我们可以将多重卷积写为 $h = f_1 * f_2 * f_3 * \dots * f_n$ 而不会产生歧义. 这一操作下的矩会发生什么? $h(x)$ 的零阶矩是

$$Z_h = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy f_1(y) f_2(x-y) = \int dy f_1(y) Z_2 = Z_1 Z_2. \quad (\text{C.3})$$

因此, 如果 $Z_i \neq 0$, 我们可以将 $f_i(x)$ 乘以某个常数因子使得 $Z_i = 1$, 并且这个性质将在卷积下保留. 下面假设已经对于所有 i 完成操作, 那么卷积的第一阶矩为

$$\begin{aligned} F_h &= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy f_1(y) x f_2(x-y) = \int dy f_1(y) \int_{-\infty}^{+\infty} dq (y+q) f_2(q) \\ &= \int_{-\infty}^{+\infty} dy f_1(y) [y Z_2 + F_2] = F_1 Z_2 + Z_1 F_2, \end{aligned} \quad (\text{C.4})$$

所以第一阶矩在卷积下是可加的:

$$F_h = F_1 + F_2. \quad (\text{C.5})$$

对于第二阶矩, 通过类似的论证有

$$S_h = \int dy f_1(y) \int dq (y^2 + 2yq + q^2) f_2(q) = S_1 Z_2 + 2F_1 F_2 + Z_1 S_2, \quad (\text{C.6})$$

或者

$$S_h = S_1 + 2F_1 F_2 + S_2. \quad (\text{C.7})$$

减去 (C.5) 的平方, 交叉项抵消了, 我们看到还有另一个在卷积操作下的可加量:

$$[S_h - (F_h)^2] = [S_1 - (F_1)^2] + [S_2 - (F_2)^2]. \quad (\text{C.8})$$

到第三阶矩, 我们发现

$$T_h = T_1 Z_2 + 3S_1 F_2 + 3F_1 S_2 + Z_1 T_2, \quad (\text{C.9})$$

经过一些代数运算 (减去 (C.5) 和 (C.7) 的函数), 我们确认存在第三个量, 即

$$T_h - 3S_h F_h + 2(F_h)^3 \quad (\text{C.10})$$

在卷积操作下可加.

这立即可以推广到任意数量的这样的函数: 令 $h(x) \equiv f_1 * f_2 * f_3 * \cdots f_n$. 然后我们找到了可加量

$$\begin{aligned} F_h &= \sum_{i=1}^n F_i, \\ S_h - F_h^2 &= \sum_{i=1}^n (S_i - F_i^2), \\ T_h - 3S_h F_h + 2F_h^3 &= \sum_{i=1}^n (T_i - 3S_i F_i + 2F_i^3). \end{aligned} \quad (\text{C.11})$$

这些在卷积下“累加”的量称为累积量. 我们以这种方式得到它们是为了强调这个概念从根本上与概率无关.

到目前为止, 我们将第 n 个累积量定义为第 n 阶矩加上来自较低阶矩的“修正项”, 因此使结果在卷积操作下可加. 那么有两个问题需要回答: (1) 这样的修正项是否一直存在? (2) 如果存在, 如何找到构建它们的通用算法?

为了回答这些问题, 我们需要一种更强大的数学方法. 引入 $f_i(x)$ 的傅里叶变换:

$$\mathcal{F}_i(\alpha) \equiv \int_{-\infty}^{+\infty} dx f_i(x) e^{i\alpha x}, \quad f_i(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\alpha \mathcal{F}_i(\alpha) e^{-i\alpha x}. \quad (\text{C.12})$$

在卷积下, 它的行为非常简单:

$$\begin{aligned} \mathcal{H}(\alpha) &= \int_{-\infty}^{+\infty} dx h(x) e^{i\alpha x} = \int dy f_1(y) \int dx e^{i\alpha x} f_2(x-y) \\ &= \int dy f_1(y) \int dq e^{i\alpha(y+q)} f_2(q) \\ &= \mathcal{F}_1(\alpha) \mathcal{F}_2(\alpha). \end{aligned} \quad (\text{C.13})$$

换句话说, $\ln \mathcal{F}(\alpha)$ 在卷积下是可加的. 这个函数有一些与后面讨论的“倒谱”概念相关的非凡性质. 现在, 检查 $\mathcal{F}(\alpha)$ 和 $\ln \mathcal{F}(\alpha)$ 的幂级数展开式. 第一个是

$$\mathcal{F}(\alpha) = M_0 + M_1(i\alpha) + M_2 \frac{(i\alpha)^2}{2!} + M_3 \frac{(i\alpha)^3}{3!} + \cdots, \quad (\text{C.14})$$

其系数为

$$M_n = \frac{1}{i^n} \frac{d^n \mathcal{F}(\alpha)}{d\alpha^n} \bigg|_{\alpha=0} = \int_{-\infty}^{+\infty} dx x^n f(x), \quad (C.15)$$

这正好是 $f(x)$ 的第 n 阶矩. 如果 $f(x)$ 的矩达到 N 阶, 则 $\mathcal{F}(\alpha)$ 在原点可微 N 次. $\ln \mathcal{F}(\alpha)$ 也有类似的展开:

$$\ln \mathcal{F}(\alpha) = C_0 + C_1(i\alpha) + C_2 \frac{(i\alpha)^2}{2!} + C_3 \frac{(i\alpha)^3}{3!} + \dots \quad (C.16)$$

显然, 它的所有系数

$$C_n = \frac{1}{i^n} \frac{d^n}{d\alpha^n} \ln \mathcal{F}(\alpha) \bigg|_{\alpha=0} \quad (C.17)$$

在卷积下是可加的, 因此是累积量. 事实上, 术语“累积量”通常是由这种关系定义的, 而不是由累积的性质定义的. 前几个是

$$C_0 = \ln \mathcal{F}(0) = \ln \int dx f(x) = \ln Z, \quad (C.18)$$

$$C_1 = \frac{1}{i} \frac{\int dx ix f(x)}{\int dx f(x)} = \frac{\mathcal{F}}{Z}, \quad (C.19)$$

$$C_2 = \frac{d^2}{d(i\alpha)^2} \ln \mathcal{F}(\alpha) = \frac{d}{d(i\alpha)} \frac{\int dx xf(x)e^{i\alpha x}}{\int dx f(x)e^{i\alpha x}} = \frac{\int f \int x^2 f - (\int xf)^2}{(\int f)^2}, \quad (C.20)$$

或者

$$C_2 = \frac{S}{Z} - \left(\frac{\mathcal{F}}{Z} \right)^2, \quad (C.21)$$

我们认为这些正好是上面直接发现的累积量. 同样, 经过一些烦琐的计算, 可以证明 C_3 和 C_4 等于第三个和第四个累积量 (C.10). 那么我们是否在 (C.17) 中找到了一个函数的所有累积量, 还是有更多无法通过这种方式找到的累积量呢? 我们会争辩说, 如果所有 C_i 都存在 (即 $f(x)$ 具有所有阶矩, 所以 $\mathcal{F}(\alpha)$ 是整函数), 那么 C_i 唯一地确定 $cF(\alpha)$ 和 $f(x)$, 所以它们必须包括所有代数独立的累积量. 任何其他累积量必须是 C_i 的线性函数. 但是如果 $f(x)$ 没有所有阶矩, 答案就不明显了, 还需要进一步研究.

C.1 累积量和矩的关系

在遵守我们的约定 $Z = 1$ 的同时, 对于 $n = 0, 1, 2, \dots$, 让我们对函数的第 n 阶矩使用更一般的表示法:

$$M_n \equiv \int_{-\infty}^{+\infty} dx x^n f(x) = \frac{d^n}{d(i\alpha)^n} \int dx f(x) e^{i\alpha x} \bigg|_{\alpha=0} = i^{-n} \mathcal{F}^{(n)}(0). \quad (C.22)$$

通常也可以方便地使用符号

$$M_n = \overline{x^n} \quad (\text{C.23})$$

表示相对于函数 $f(x)$ 的 x^n 的平均值. 要强调这些通常不是概率平均值. 我们只是在指出一些一般的代数关系, 其中 $f(x)$ 不必是非负的. 对于概率平均值, 我们总是保留符号 $\langle x \rangle$ 或 $E(x)$.

如果函数 $f(x)$ 具有所有阶矩, 则其傅里叶变换具有幂级数展开

$$\mathcal{F}(\alpha) = \sum_{n=0}^{\infty} M_n (i\alpha)^n. \quad (\text{C.24})$$

显然, 第一个累积量与第一阶矩相同:

$$C_1 = M_1 = \bar{x}, \quad (\text{C.25})$$

对于第二个累积量, 我们有 $C_2 = M_2 - M_1^2$, 但这是

$$C_2 = \int dx [x - M_1]^2 f(x) = \overline{(x - \bar{x})^2} = \overline{x^2} - \bar{x}^2, \quad (\text{C.26})$$

关于其平均值的第二阶矩, 称为 $f(x)$ 的第二阶中心矩. 类似地, 第三阶中心矩是

$$\int dx (x - \bar{x})^3 f(x) = \int dx [x^3 - 3\bar{x}x^2 + 3\bar{x}^2x - \bar{x}^3] f(x), \quad (\text{C.27})$$

但这正好是第三个累积量 (C.11):

$$C_3 = M_3 - 3M_1M_2 + 2M_1^3, \quad (\text{C.28})$$

这时我们可能推测所有的累积量只是相应的中心矩. 然而, 事实证明并非如此: 我们发现第四阶中心矩是

$$\overline{(x - \bar{x})^4} = M_4 - 4M_3M_1 + 6M_2M_1^2 - 3M_1^4, \quad (\text{C.29})$$

但是第四个累积量是

$$C_4 = M_4 - 4M_3M_1 - 3M_2^2 + 12M_2M_1^2 - 6M_1^4. \quad (\text{C.30})$$

所以它们通过关系

$$\overline{(x - \bar{x})^4} = C_4 + 3C_2^2 \quad (\text{C.31})$$

相关联. 因此, 第四阶中心矩不是累积量, 也不是累积量的线性函数. 然而我们发现, 对于 $n = 1, 2, 3, 4$, 第 n 阶以下的矩和第 n 个以下的累积量是彼此唯一确定的. 我们留给读者验证这对于所有 n 是否都正确.

如果我们的函数 $f(x)$ 是概率密度, 许多有用的近似值可以通过就累积量展开前几项更有效地写出来.

C.2 示例

高斯分布的累积量是什么？令

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{[x - \mu]^2}{2\sigma^2} \right\}, \quad (\text{C.32})$$

我们发现其傅里叶变换为

$$\mathcal{F}(\alpha) = \exp\{i\alpha\mu - \alpha^2\sigma^2/2\}. \quad (\text{C.33})$$

这样，

$$\ln \mathcal{F}(\alpha) = i\alpha\mu - \alpha^2\sigma^2/2, \quad (\text{C.34})$$

所以

$$C_0 = 0, \quad C_1 = \mu, \quad C_2 = \sigma^2, \quad (\text{C.35})$$

并且所有更高阶的 C_n 都为 0. 高斯分布的特点是只有两个非平凡累积量：均值和方差.

引用文献

Abel, N. H. (1826), 'Untersuchung der Functionen zweier unabhängig veränderlichen Gröſzen x und y , wie $f(x, y)$, welche die Eigenschaft haben, dass $f[z, f(x, y)]$ eine symmetrische Function von z, x und y ist.', *J. Reine u. angew. Math. (Crelle's Journal)*, **1**, 11–15.

结合性函数方程的第一个已知示例.

Aczél, J. (1966), *Lectures on Functional Equations and their Applications*, Academic Press, New York.

Aczél, J. (1987), *A Short Course on Functional Equations*, D. Reidel, Dordrecht.

Aitken, G. A. (1892), *The Life and Works of John Arbuthnot*, Clarendon Press, Oxford.

Akaike, H. (1980), 'The interpretation of improper prior distributions as limits of data dependent proper prior distributions', *J. Roy. Stat. Soc.* **B42**, 46–52.

处理边缘化悖论的失败尝试. 作者没有意识到这个悖论对于正常先验和非正常先验是一样的, 正如本书第 15 章所讨论的那样.

Andrews, D. R. & Mallows, C. L. (1974), 'Scale mixtures of normal distributions', *J. Roy. Stat. Soc.* **B36**, 99–102.

Anscombe, F. J. (1963), 'Sequential medical trials', *JASA* **58**, 365.

宣称阿米蒂奇 (Armitage, 1960) 的序列分析是 “一场骗局”.

Arbuthnot, J. (1710), 'An argument for Divine Providence', *Phil. Trans. Roy. Soc.* **27**, pp. 186–190.

重印于肯德尔和普莱克特的著作 (Kendall & Plackett, 1977, 第 2 卷, 第 30~34 页). 这本书第 5 章讨论以数据的非概率性为理由拒绝统计假设的第一个已知示例. 约翰·阿巴思诺特 (1667—1735) 是安妮女王的医生, 也是许多主题的多产作家. 艾特肯的著作 (Aitken, 1892) 中有阿巴思诺特的传记.

Armitage, P. (1960), *Sequential Medical Trials*, Thomas, Springfield, Illinois; 2nd edn, Blackwell, Oxford (1975).

“可选停止”争议的主要源头之一. 萨维奇 (L. J. Savage, 1962) 和安斯科姆 (Anscombe, 1963) 对此进行了广泛讨论.

Barnard, G. A. (1947), 'Significance test for 2×2 tables', *Biometrika* **34**, 123–137.

Barnard, G. A. (1983), 'Pivotal inference and the conditional view of robustness (why have we for so long managed with normality assumptions?)', in Box, G. E. P., Leonard, T. & Wu, C-F., eds. (1983), *Scientific Inference, Data Analysis, and Robustness*, Academic Press Inc., Orlando, FL.

对正态分布的成功大致表达了与 145 年前奥古斯塔斯·德摩根相同的惊讶. 在第 7 章中, 我们试图解释这一点.

Bayes, Rev. T. (1763), 'An essay toward solving a problem in the doctrine of chances', *Phil. Trans. Roy. Soc.*, pp. 370–418.

莫利纳的著作 (Molina, 1963) 中有其照片. 巴纳德在《生物计量学》杂志 [*Biometrika* **45**, 293–313 (1958)] 中, 以及皮尔逊和肯德尔 (Pearson & Kendall, 1970) 对其进行了带有传记注释的重印. 发表的具体日期有不确定性, 这篇论文于 1763 年在英国皇家学会上被宣读, 但直到 1764 年才真正出版. 霍兰 (Holland, 1962) 给出了托马斯·贝叶斯 (1702—1761) 的更多传记信息. 施蒂格勒 (Stiger, 1983) 和萨贝尔 (Zabell, 1989) 提供的证据表明, 为了回应大卫·休谟 (1711—1776) 的挑战, 可能贝叶斯早在 1748 年就已经发现此结果并将其告诉朋友.

Bean, W. B. (1950), *Aphorisms from the Bedside Teaching and Writings of Sir William Osler*, (1849–1919). Henry Schumann, New York.

奥斯勒认识到医学诊断的推理形式与后来乔治·波利亚给出的形式基本相同. 卢斯特德 (L. Lusted, 1968) 将其作为贝叶斯医学诊断计算机程序的基础.

Bell, E. T. (1937), *Men of Mathematics*, Dover Publications, Inc., New York.

中译本 埃里克·坦普尔·贝尔. 数学大师——从芝诺到庞加莱. 徐源, 译. 宋蜀碧, 校. 上海: 上海科技教育出版社, 2012, 2018. ^①

这本传记式小品集值得阅读, 它似乎不存在替代品. 但读者应该知道, 埃里克·坦普尔·贝尔也是一位著名科幻作家 (笔名约翰·泰恩), 这种天赋在这本书中也得到展现. 我们或许可以相信其中引用的名字、日期及有文献证明的历史事实的准确性. 但是, 书中对所讨论问题的解释性陈述甚少而是包含很多关于作者的梦想和社会政治观点, 以及他对技术事实的理解水平. 例如 (第 167 页^②), 他以“社会正义”为由赞成将现代化学命名法之父拉瓦锡送上断头台. 他公然错误地谴责了拉普拉斯, 并错误地将布尔描绘成一个从不会犯错的圣人. 他展示了 (第 256 页) 对爱因斯坦工作性质的荒谬理解, 弄错了事实的时间顺序, 还告诉我们 (第 459 页) 阿基米德从不关心数学的应用!

Bellman, R. & Kalaba, R. (1956), 'On the principle of invariant imbedding and propagation through inhomogeneous media', *Proc. Natl Acad. Sci. USA* **42**, 629–632.

Bellman, R. & Kalaba, R. (1957), 'On the role of dynamic programming in statistical communication theory', *IRE Trans. PGIT-1*, 197.

Bellman, R., Kalaba, R. & Wing, G. M. (1957), 'On the principle of invariant imbedding and one-dimensional neutron multiplication', *Proc. Natl Acad. Sci.* **43**, 517–520.

Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.

中译本 詹姆斯·伯杰. 统计决策论及贝叶斯分析. 贾乃光, 译. 吴喜之, 校译. 北京: 中国统计出版社, 1998.

^① 上海科技教育出版社的中译本有 2012 年和 2018 年两种版本. ——编者注

^② 英文原书页码 167、256、459, 对应 2012 年中译本页码 194、295、524, 对应 2018 年中译本页码 180、275、495. ——编者注

Berger, J. O. & Wolpert, R. L. (1988), *The Likelihood Principle*, 2nd edn, Institute of Mathematical Statistics, Hayward, CA.

Bernardo, J. M. (1980), 'A Bayesian analysis of classical hypothesis testing', in Bernardo, J. M. *et al.*, eds., *Bayesian Statistics*, University Press, Valencia, Spain, pp. 606-47. 带有讨论.

Bernoulli, D. (1738), 'Specimen theoriae novae de mensura sortis', in Bernoulli, D. (1730-1), *Comment Acad. Sci. Imp. Petropolitanae*, **V**, pp. 175-192; English translation by Sommer, L. (1954) *Econometrica*, **22**, 23-36.

Bernoulli, D. (1777), *Mem. St Petersburg Acad.*, *Acta Acad. Petrop.*, pp. 3-33; English translation, 'The most probable choice between several discrepant observations', *Biometrika*, **45**, 1-18 (1961).

重印于皮尔逊和肯德尔的著作 (Pearson & Kendall, 1970, 第 157~167 页). 取多个观测值的平均值问题, 本书第 7 章讨论过.

Bernoulli, J. (1713), *Ars Conjectandi*, Thurnisiorum, Basel.

重印于 *Die Werke von Jakob Bernoulli*, Vol. 3, Birkhaeuser, Basel (1975), 107-286. 第一个现代极限定理. 第 IV 部分 (包含极限定理) 由宋冰 (Bing Sung) 翻译成英文, 见哈佛大学统计系技术报告 #2, 1966 年.

Bertrand, J. L. (1889), *Calcul des Probabilités*, Gauthier-Villars, Paris; 2nd edn, 1907; reprinted (1972) by Chelsea Publishing Co., New York.

人们通常仅因出现在第 4~5 页的“贝特朗悖论”而引用该书. 然而, 该书充满简洁的数学, 展示了作者良好的概念洞察力. 这两方面都往往优于一些近作. 贝特朗清楚地认识到, 我们从给定数据得出的结论在很大程度上必须依赖于先验信息, 这种理解在后来的正统统计文献中丢失了. 我们在 6.23 节引用了他的话. 然而, 我们不同意他对我们在第 7 章中给出的高斯分布的赫歇尔-麦克斯韦推导的批评. 他所认为的推导中的缺陷正是我们眼中的最大优点, 也是爱因斯坦推理的先驱. 今天非常值得了解和复兴这种推理方式.

Birnbaum, A. (1962), 'On the foundation of statistical inference (with discussion)', *J. Am. Stat. Assoc.* **57**, 269-326.

反贝叶斯主义者接受的“似然原理”的第一个证明.

Blackman, R. B. & Tukey, J. W. (1958), *The Measurement of Power Spectra*, Dover Publications, Inc., New York.

周期图被特定的平滑破坏, 抹去了其中大部分有用信息. 我们多次警告不要这样做.

Bloomfield, P. (1976), *Fourier Analysis of Time Series: An Introduction*, J. Wiley & Sons, New York.

该书作者利用布莱克曼-图基方法对来自惠特克和活森 (Whittaker & Robinson, 1924) 的所谓变星天文数据进行了荒谬的极端处理 (比噪声水平低约 50 dB!). 他未能识别出数据的虚假性 (他既不认为一个身份不明的天文台观测到连续 600 次晴朗的午夜天空有什么问题, 也不担心作者没有提及任何数据来源这一事实). 结果, 他的周期图对于变星没有给出任何信息; 其顶部显示放入模拟数据的两个正弦波, 底部仅显示数字化误差的频

谱. 这有力证明了在不合适的情况下盲目、不假思索地应用统计流程的愚蠢性. 贝叶斯主义者不得不考虑有关该现象和数据获取流程的先验信息, 所以不会犯这种错误.

Boltzmann, L. W. (1871), *Wiener Berichte* **63**, pp. 397–418, 679–711, 712–732.

在这三篇文章中首次出现 “ $p \log p$ ” 类型的熵表达式.

Boole, G. (1854), *An Investigation of The Laws of Thought*, Macmillan, London; reprinted by Dover Publications, Inc., New York (1958).

Boole, G. (1857), ‘On the application of the theory of probabilities to the question of the combination of testimonies or judgments’, *Edinburgh Phil. Trans.* **v**, xxi.

Borel, E. (1909), *Éléments de la Théorie des Probabilités*, Hermann et Fils, Paris.

详细讨论了贝特朗悖论并推测出正确解, 后来通过群不变性论证发现这种解.

Borel, E. (1924), ‘A propos d’un traité de probabilités’, *Rev. Philos.* **98**, 321–336.

对凯恩斯的著作 (Keynes, 1921) 的评论. 重印于凯伯格和斯莫克勒的著作 (Kyburg & Smokler, 1981). 和贝特朗 (Bertrand, 1889) 一样, 博雷尔非常清楚概率必须依赖于我们的先验知识状态. 遗憾的是, 两人都没有在实际应用中证明这一点的重要后果, 他们本可以避免其他人长达 50 年的错误教导.

Boring, E. G. (1955), ‘The present status of parapsychology’, *Am. Sci.*, **43**, 108–16.

作者得出结论, 超心理学家的行为是一种要研究的奇怪现象. 他指出, 无论观察到什么事实, 试图证明不存在对通灵的自然解释, 在逻辑上都是不可能的; 一个人不可能对一个普遍的否定结论加以证明 (否认因果解释存在的量子理论学家请注意这一点).

Bortkiewicz, L. V. (1898), *Das Gesetz der Kleinen Zahlen*, Teubner, Leipzig.

包含他著名的使用泊松分布拟合连续几年被马踢死的德国士兵数量的工作.

Bortkiewicz, L. V. (1913), *Die radioaktive Strahlung als Gegenstand Warscheinlichkeitstheoretischer Untersuchungen*, B. G. Teubner, Leipzig.

Box, G. E. P. (1962), ‘On the foundations of statistical inference: discussion’, *J. Am. Stat. Assoc.* **57**, (298), 311.

Box, G. E. P. & Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, MA.

乔治·爱德华·佩勒姆·博克斯和吉米·萨维奇一样, 在这个领域是异类. 博克斯是罗纳德·艾尔默·费希尔的助手, 娶了费希尔的女儿. 但是, 博克斯在推断问题上成为贝叶斯主义者, 在显著性检验问题上仍然是费希尔主义者, 因为他认为这超出了贝叶斯方法的范围. 杰恩斯 (Jaynes, 1985) 认为, 正相反, 任何合理的显著性检验都需要完整的贝叶斯方法.

Box, J. F. (1978), *R. A. Fisher: The Life of a Scientist*, Wiley, New York.

琼·费希尔·博克斯是罗纳德·艾尔默·费希尔的小女儿, 她讲述了其他人不知道的关于费希尔的许多个人轶事, 其中穿插着对他所从事工作的叙述.

Bredin, J.-D. (1986), *The Affair: The Case of Alfred Dreyfus*, G. Braziller, New York.

这本书第 5 章讨论了意见分歧心理现象的一个著名例子.

Bretthorst, G. L. (1988), *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture Notes in Statistics, Vol. 48, Springer-Verlag, Berlin.

该书作者 1987 年博士论文的修订版和扩充版。这是计量经济学趋势分析和核磁共振数据频谱分析的重要新进展。

Buck, B. & Macaulay, V. A., eds. (1991), *Maximum Entropy in Action*, Clarendon Press, Oxford.

作者在牛津大学举办了八场讲座, 涵盖核磁共振、光谱学、等离子体物理学、X 射线晶体学和热力学方面的基本概念与应用。这是迄今为止对初学者来说最好的入门材料, 并且其中足够的技术细节对实践科学家也有用。请注意, 这里所谓的“最大熵”被“窗口化”或数据“预滤波”等特定方法——由于会破坏数据中的某些信息为我们所谴责——扭曲。如果应用得当, 概率论完全有能力从未删减的原始数据中提取所有相关信息, 并且在允许这样做时以最少的总计算量取得最优效果。

Cantril, H. (1950), *The 'Why' of Man's Experience*, Macmillan, New York.

Carnap, R. (1952), *The Continuum of Inductive Methods*, University of Chicago Press.

Chadwick, J. (1958), *The Decipherment of Linear B*, Cambridge University Press.

Cheeseman, P. (1988), 'An inquiry into computer understanding', *Comput. Intell.* 4, 58–66.

另请参阅其后 76 页的讨论。作者尝试向人工智能社区解释贝叶斯原理, 但遭到不理解者在说什么的讨论者的强烈反对。杰恩斯 (Jaynes, 1990b) 描述了这种境况。

Chernoff, H. & Moses, L. E. (1959), *Elementary Decision Theory*, J. Wiley & Sons, Inc., New York.

该书首次出版时, 被描述为“唯一不落后于时代 20 年的统计学教科书”。现在该书已经落后大约 40 年, 因为作者无法接受概率不是频率的概念, 没有意识到直接贝叶斯方法可以简单得多地得到所有相同结果。尽管如此, 它仍是对沃尔德原始思想有趣和引人入胜的阐述, 比沃尔德的著作 (Wald, 1950) 容易阅读得多。

Copi, I. M. (1994), *Introduction to Logic*, 9th edn, Macmillan, New York.

中译本 欧文·M·柯匹, 卡尔·科恩. 逻辑学导论 (第 13 版). 张建军, 潘天群, 顿新国, 译. 北京: 中国人民大学出版社, 2014.

Cournot, A. A. (1843), *Exposition de la Theorie des Chances et des Probabilités*, L. Hachette, Paris.

重印于 Oeuvres Complètes, J. Vrin, Paris (1984). 对拉普拉斯的第一批攻击之一, 埃利斯、布尔、维恩、贝尔和其他人紧随其后, 直到如今。

Cox, R. T. (1946), 'Probability, frequency, and reasonable expectation', *Am. J. Phys.* 14, 1–13.

在我们看来, 这篇文章是自拉普拉斯以来对概率论概念 (相对于纯数学) 阐述的最重要进步。

Cox, R. T. (1961), *The Algebra of Probable Inference*, Johns Hopkins University Press, Baltimore MD.

考克斯的文章 (Cox, 1946) 的扩展, 有额外的结论和更多的讨论. 杰恩斯 (Jaynes, 1963) 发表了评论.

Craig, J. (1699), *Theologiae Christianae Principia Mathematica*, Timothy Child, London.

带评论的重印见 Daniel Titus, Leipzig (1755). 更多评论见施蒂格勒的文章 (Stigler, 1986a).

Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton University Press.

中译本 H. 克拉美. 统计学数学方法. 魏宗舒, 郑朴, 吴锦, 译. 北京: 高等教育出版社, 1966.

这标志着相对于贝叶斯方法对置信区间高度信心的鼎盛时期, 基于纯粹理念的理由将其断言为通常做法, 没有注意到这两种方法产生的实际结果. 对它的评论见附录 B 和杰恩斯的著作 (Jaynes, 1976, 1986a).

Crick, F. (1988), *What Mad Pursuit*, Basic Books, Inc., New York.

中译本 弗朗西斯·克里克. 狂热的追求: 科学发现回忆录. 傅贺, 译. 长沙: 湖南科学技术出版社, 2020.

克里克生活和工作的回忆录, 其中充满了对一般科学行为的重要观察和建议, 以及关于他在生物学中具有决定性意义的重要工作的精彩的技术细节——其中大部分发生在著名的克里克-沃森发现 DNA 结构几年之后. 几乎同样重要的是, 这是对沃森一书 (Watson, 1968, 见参考文献) 的纠正. 我们在这里看到了克里克在 1974 年记录的 DNA 双螺旋结构发现故事的另一面, 对事件的回忆有所不同. 在我们看来, 这项工作作为重要科学发现的案例是有历史价值的, 这些发现没有我们的数学形式的概率推断的帮助, 但是——至少在克里克看来——严格遵守以波利亚给出的定性形式的规则. 我们希望理论物理学家也能如此推理.

Crow, E. L., Davis, F. A. & Maxfield, M. W. (1960), *Statistics Manual*, Dover Publications, Inc., New York.

有很多有用的表和图, 但阐述的方法很传统, 从不考虑信息内容, 因此从来没有觉察到他们从数据中提取信息的缺陷. 在杰恩斯的著作 (Jaynes, 1976) 中对此有一些有趣的讨论.

Dawid, A. P., Stone, M. & Zidek, J. V. (1973), 'Marginalization paradoxes in Bayesian and structural inference', *J. Roy Stat. Soc.* **B35**, 189-233.

在第 15 章中讨论.

Dawkins, R. (1987), *The Blind Watchmaker*, W. W. Norton & Co., New York.

中译本 理查德·道金斯. 盲眼钟表匠: 生命自然选择的秘密. 王道还, 译. 北京: 中信出版社, 2016.

对不了解达尔文理论的宗教原教旨主义者不断攻击达尔文理论的回应. 理查德·道金斯是牛津大学动物学教授. 他像 120 年前的查尔斯·达尔文一样耐心地详细解释为什么可以将自然事实解释为自然法则的运作, 而无须援引目的论. 我们对此完全同意. 不幸的是, 道金斯的热情似乎超出了他的逻辑范围. 在书的封面上, 他声称这也解释了一个非常不同的事情: “为什么进化证据揭示了一个没有设计的宇宙.” 我们看不出有任何证据能证明这一点; 基本逻辑警告我们试图证明一个普遍的否定结论是愚蠢的.

道金斯与原教旨主义者的斗争仍在继续. 1993 年, 剑桥大学神学院设立了星桥神学

与自然科学讲座。道金斯在全国报刊上发声对此表示谴责, 并强调神学与科学价值相比空洞无用。这促使剑桥诺贝尔奖获得者化学家马克斯·佩鲁茨发表了反驳意见, 他说道: “科学教给我们自然法则, 但宗教教导我们应该如何生活。……道金斯博士将科学家描绘成破坏宗教信仰的团体, 这损害了公众对科学家的认知。”在我们看来, 道金斯谴责的是武断的神学体系, 而不是其道德教义, 这两者是非常不同的事情。我们在第 20 章末尾给出了对此的看法。

de Finetti, B. (1937), 'La prevision: ses lois logiques, ses sources subjectives', *Ann. Inst. H. Poincaré*, **7**, 1-68.

英文翻译: 'Prevision, its logical laws, its subjective sources', 收录于凯伯格和斯莫克勒的著作 (Kyburg & Smokler, 1981).

de Finetti, B. (1972), *Probability, Induction and Statistics*, John Wiley & Sons, New York.

de Finetti, B. (1974a), 'Bayesianism', *Intern. Stat. Rev.* **42**, 117-130.

de Finetti, B. (1974b), *Theory of Probability*, 2 vols., J. Wiley & Sons, Inc., New York.

阿德里安·史密斯的英译本同样展示了这部作品的机智和幽默。布鲁诺·德菲内蒂写这本书时充满了快乐, 但是他几乎写出两句话, 就要加上括号讨论突然出现在他脑海中的一个不同的主题, 这一点忠实地反映在译文中。书中充满了所有认真学习该领域的学生都应该知道的有趣信息, 但是由于其混乱无序而无法总结。任何一个话题的讨论都可能分散在多个不同章节中, 没有任何交叉引用, 所以人们最好随机地阅读这本书的各页。

de Groot, M. H. (1970), *Optimal Statistical Decisions*, McGraw-Hill Book Co., New York.

de Groot, M. H. (1975), *Probability and Statistics*, Addison-Wesley Publishing Co., Reading, MA; 2nd edn. (1986).

这本教科书中充满了有用的结果, 代表了正统统计和现代贝叶斯推断的中间过渡阶段。莫里斯·德格罗特 (1931—1989), 过渡贝叶斯主义者萨维奇的博士生, 清楚地看到了贝叶斯方法的技术优势, 并且是我们每年两次的 NSF-NBER 贝叶斯研讨会的常客和演讲者, 但是他仍然保留了正统的术语、符号和一般的绝对主义心态。因此, 他仍然区分“真实概率”和“估计概率”, 仿佛前者真实存在, 并且把“概率论”和“统计推断”严格区别开来, 就好像它们是不同的主题一样。这并不妨碍他获得标准的有用结果, 通常是通过沿用发明特定工具的正统习惯而不是应用概率论规则。(我们目前的相对主义理论认识到不存在“绝对”概率这样的东西, 因为所有概率都表达了使用者的信息状态, 因此也必然依赖于使用者的信息状态。这使得概率论一般原则统一适用于所有推断问题, 而不需要任何特定工具。)莫里斯·德格罗特的传记和参考书目可以在 *Statistical Science*, vol 6, pp. 4-14 (1991) 中找到。

de Groot, M. H. & Goel, P. (1980), 'Only normal distributions have linear posterior expectations in linear regression', *J. Am. Stat. Assoc.* **75**, 895-900.

另一种由高斯 (Gauss, 1809) 首次发现的联系, 在第 7 章中有讨论。

de Moivre, A. (1718), *The Doctrine of Chances: or, A Method of Calculating the Probability of Events in Play*, W. Pearson, London; 2nd edn, Woodfall, London (1738); 3rd edn, Millar, London (1756); reprinted by Chelsea Publishing Co., New York (1967).

- de Moivre, A. (1733), *Approximatio ad Summam Terminorum Binomii $(a + b)^n$ in Seriem expansi*; Photographic reproduction in Archibald, R. C., *Isis* **8**, 671–683, (1926).
- de Morgan, A. (1838), *An Essay on Probabilities*, Longman & Co., London.
- de Morgan, A. (1847), *Formal Logic: or the Calculus of Inference Necessary and Probable*, Taylor & Watton, London.
对拉普拉斯观点的热情阐述.
- de Morgan, A. (1872), *A Budget of Paradoxes*, 2 Vols, de Morgan, S., ed., London; 2nd edn, Smith, D. E., ed. (1915).
Dover Publications, Inc., New York (1954) 以单卷重印. 奥古斯塔斯·德摩根 (1806—1871) 是一位数学家和逻辑学家, 1828~1866 年间在伦敦大学学院工作. 他的笔记不仅涉及逻辑, 还涉及逻辑悖论. 后者保存在这本书中, 令人愉快地描述了 19 世纪英国大量存在的方圆论者、反哥白尼派、反牛顿派、宗教狂热者、命理学家和精神错乱者的活动. 它生动地描绘了严肃的学者为了在科学上取得任何进步而必须克服的困难, 其中有很多有趣轶事.
- de Morgan, S. (1882), *Memoir of Augustus de Morgan*, Longman, Green, London.
关于德摩根的更多传记和轶事材料.
- Devinatz, A. (1968), *Advanced Calculus*, Holt, Rinehart and Winston, New York.
- Dyson, F. J. (1958), 'Review of Lighthill (1957)', *Phys. Today* **11**, 28.
- Edwards, A. W. F. (1974), 'The history of likelihood', *Int. Stat. Rev.* **42**, 9–15.
- Edwards, A. W. F. (1991), *Nature* Aug. 1, p. 386.
- Edwards, H. M. (1989), 'Kronecker's philosophical views', in Rowe, D. E. & McCleary, J., eds., *The History of Modern Mathematics*, Vol. 1, pp. 67–77.
- Einstein, A. (1905a), 'On the electrodynamics of moving bodies', *Ann. Phys. Leipzig* **17**, 891–921.
- Einstein, A. (1905b), 'Does the inertia of a body depend on its energy content?', *Ann. Phys. Leipzig* **18**, 639–641.
- Elias, P. (1958), 'Two famous papers', Editorial, *IEEE Trans. IT-4*, p. 99.
- Ellis, R. L. (1842), 'On the foundations of the theory of probability', *Camb. Phil. Soc.*, vol. viii.
埃利斯是英国的库尔诺, 他发起了反拉普拉斯运动, 使得科学推断的进程倒退了一个世纪.
- Ellis, R. L. (1863), *The Mathematical and Other Writings of Robert Leslie Ellis M. A.*, Wm. Walton, ed., Deighton, Bell, Cambridge.
- Erickson, G. J. & Smith, C. R. (1988), eds., *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Vol. 1, *Foundations*; Vol. 2, *Applications*, Kluwer Academic Publishers, Dordrecht, Holland.
- Euler, L. (1749), *Recherches sur la Question des Inégalitiés du Mouvement de Saturne et de Jupiter, Sujet Proposé pour le Prix de l'Année 1748 par l'Académie Royale des Sciences de Paris*; reprinted in *Leonhardi Euleri, Opera Omnia*, ser. 2, Vol. 25, Turici, Basel, (1960).

欧拉放弃了从 75 个相互矛盾的观察数据中估计 8 个未知参数的问题,但还是赢得了奖金。

Feinberg, S. E. & Hinkley, D. V. (1980), *R. A. Fisher: An Appreciation*, 2nd edn, Lecture Notes in Statistics #1, Springer-Verlag, Berlin.

这是该书于 1979 年出版后的第 2 次印刷。如果将其视为历史文献而不是对当前统计原理的说明,则是宝贵的资料来源。书中包含了费希尔最重要推导的丰富技术细节,其中提供了他的大量作品列表,包括 4 本书和 294 篇已发表的文章。但是在对费希尔的赞誉中,本书未能注意到在 1979 年已经确立的一些东西:费希尔强烈拒绝的杰弗里斯的更简单、更统一的方法实际上可以完成费希尔方法所做的一切,具有相同或更好的结果,并且几乎总是更简单。此外,这些方法可以轻松处理费希尔永远无法克服的技术难题(例如冗余参数或缺乏充分统计量),因此这本书也倾向于延续有害的费希尔传奇。

Félix, L. (1960), *The Modern Aspect of Mathematics*, Basic Books, Inc., New York.

布尔巴基主义观点,相反的观点见克莱因的著作(Kline, 1980)。

Feller, W. (1950), *An Introduction to Probability Theory and its Applications*, Vol. 1, J. Wiley & Sons, New York; 2nd edn, 1957; 3rd edn, 1968.

[中译本] 威廉·费勒. 概率论及其应用·卷 1, 第 3 版. 胡迪鹤, 译. 北京: 人民邮电出版社, 2021.

Feller, W. (1966), *An Introduction to Probability Theory and its Applications*, Vol. 2, J. Wiley & Sons, New York; 2nd edn, 1971.

[中译本] 威廉·费勒. 概率论及其应用·卷 2, 第 2 版. 郑元禄, 译. 北京: 人民邮电出版社, 2021.

Fine, T. L. (1973), *Theories of Probability*, Academic Press, New York.

Fisher, R. A. (1922), 'On mathematical foundations of theoretical statistics', *Phil. Trans. Roy. Soc. Lond. Ser. A*, **222**, 309–368.

“充分统计量”这一术语的引入。

Fisher, R. A. (1925), *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh. 有 12 个后续版本。

Fisher, R. A. (1930b), *The Genetical Theory of Natural Selection*, Oxford University Press; 2nd (rev) edn, Dover Publications, Inc., New York (1958).

在这里,费希尔表明孟德尔遗传学与达尔文进化论并不冲突,正如孟德尔在 20 世纪初所假设的那样;相反,孟德尔遗传的“粒子”或“离散”性质清除了达尔文理论的一些突出困难,这是由于混合遗传假设导致的,大多数生物学家——包括达尔文本人——在 19 世纪 60 年代认为这是理所当然的。回想一下,孟德尔的工作,以及对显性和隐性基因等的了解,比达尔文晚;这是达尔文(1809—1882)从来不知道的,直到 1900 年之后才被大家普遍了解。费希尔和其他人用这些术语重新解释达尔文的理论,现在被称为新达尔文主义。到费希尔的这本书第 2 版(1958 年)出版时,由放射性引起的突变的存在已被充分证实,由 DNA 复制失败引起的突变已经变得非常可信,基因重组(早在 1886 年就由魏斯曼提出)被认为是提供自然选择赖以生存的个体变异的另一种机制,但其起源令达尔文感到困惑。因此,费希尔以小字添加了许多新的段落,指出这种新的理解及其含义。达尔

文如果活着，会多么愿意看到这些对他的问题的漂亮回答！费希尔对科学真正、永久的贡献是在这样的作品中，而不是在他的统计学中。费希尔的统计学在 20 世纪 20 年代是一种进步，但自从杰弗里斯 1939 年的工作以来一直是一种阻碍的力量。

Fisher, R. A. (1933), 'Probability, likelihood and quantity of information in the logic of uncertain inference', *Proc. Roy. Soc.* **146**, 1-8.

一篇试图攻击杰弗里斯的著名作品，我们在第 16 章中有所讨论。

Fisher, R. A. (1934), 'Two new properties of mathematical likelihood', *Proc. Roy. Soc. London A* **144**, 285-307.

Fisher, R. A. (1950), *Contributions to Mathematical Statistics*, W. A. Shewhart, ed., J. Wiley & Sons, Inc., New York.

他的早期著名论文的合集。

Fisher, R. A. (1956), *Statistical Methods and Scientific Inference*, Oliver & Boyd, London; second revised edition, Hafner Publishing Co., New York (1959).

费希尔关于统计学的最后一本书，他试图总结他对不确定推断的逻辑性质的看法。人们可以看出这相对于他的早期作品有相当大的观点转变——他甚至偶尔承认自己以前错了。他现在更认同先验信息的作用，说可识别子集应该予以考虑，并且先验无知对于置信估计的有效性至关重要。他对哥德尔定理的简洁解释展示了他一贯的直觉洞察力，但书中也有一些明显的记忆失误和数值错误。每位认真学习统计学的学生都应该至少放慢速度仔细地阅读这本书两遍，因为其中的思考非常深刻，阅读一遍无法完全理解费希尔的意思。此外，费希尔谈到了我们在目前的工作中没有讨论的几个专门化主题。

Fisher, R. A. (1962), 'Some examples of Bayes' method of the experimental determination of probability a priori', *J. Roy. Stat. Soc. B* **24**, 118-124.

Fisher, R. A. (1974), *Collected Papers of R. A. Fisher*, J. H. Bennett, ed., University of Adelaide.

Fowler, R. H. (1929), *Statistical Mechanics; The Theory of the Properties of Matter in Equilibrium*, Cambridge University Press.

Fraser, D. A. S. (1966), 'On fiducial inference', *Ann. Math. Stat.* **32**, 661-676.

Fraser, D. A. S. (1980), 'Comments on a paper by B. Hill', in J. M. Bernardo *et al.*, eds., *Bayesian Statistics*, University Press, Valencia, Spain, pp. 56-58.

声称找到一个对似然原理的反例。但它与第 15 章中贝叶斯方法解决的翻滚四面体问题相同，人们在 1980 年还不知道该问题的正确解。

Galilei, Galileo (1638), *Dialogues Concerning Two New Sciences*; English translation by Henry Crew & Alfonso de Salvio, MacMillan Company, London (1914).

中译本 伽利略. 关于两门新科学的对话. 武际可, 译. 北京: 北京大学出版社, 2020.

平装本由 Dover Publications, Inc., New York 重印, 日期不详 (约 1960 年).

Galton, F. (1886), 'Family likeness in stature', *Proc. Roy. Soc. Lond.* **40**, 42-73.

Galton, F. (1908), *Memories of My Life*, Methuen, London.

更多的传记性和技术细节在皮尔逊的著作 (Pearson, 1914-1930) 中.

Gardner, M. (1957), *Fads and Fallacies in the Name of Science*, Dover Publications, Inc., New York.

[中译本] 马丁·加德纳. 西方伪科学种种. 贝金, 译. 长正, 校. 北京: 知识出版社, 1984.

似乎是德摩根的著作 (de Morgan, 1872) 的一本 20 世纪的续作, 更多地关注科学而非数学领域的造假者. 在这里, 我们遇到了真诚但悲惨的被误导者, 以及故意欺诈以从轻信者中赚取不当钱财的人.

Gardner, M. (1981), *Science - Good, Bad, and Bogus*, Paperbound edition (1989), Prometheus Books, Buffalo NY; paperbound edn (1989).

加德纳的著作 (Gardner, 1957) 的续作, 其中有每个人都应该注意的发人深省的内容, 记录着最近几种趋势的细节: 神创论者利用电视面向数百万人攻击达尔文理论, 同时严重歪曲达尔文理论; 特异功能倡导者闯入科学会议, 试图援引量子理论来支持特异功能, 尽管他不了解什么是量子理论; 宣传奇才将知识 (人工智能、混沌、灾难理论、分形) 的每一次微小进步都变成了革命性的宗教崇拜; 专业的灾难贩卖者, 通过从人类的每一项活动中营造越来越荒谬的危险来达到个人宣传的目的; 最可怕的是, 新闻媒体对这一切随时给予支持并进行免费的宣传. 今天, 我们的电视广播中充斥着虚假的科学和中世纪的迷信, 企图贬低和歪曲真正的科学. 在引言中, 加德纳表达了他对网络管理人员以其有利可图为由拒绝纠正这一点的愤慨. 那么, 这种持续、蓄意滥用言论自由以谋取利益的行为什么时候会成为对公共社会的明显与现实的威胁? 另见罗思曼 (Rothman, 1989) 和胡贝尔 (Huber, 1992) 的著作.

Gauss, K. F. (1809), *Theoria Motus Corporum Celestium*, Perthes, Hamburg; English translation, *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections*, Dover Publications, Inc., New York (1963).

Gauss, K. F. (1823), *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae*, Dieterich, Göttingen; *Supplementum* (1826).

Geisser, S. & Cornfield, J. (1963), 'Posterior distribution for multivariate normal parameters', *J. Roy. Stat. Soc.* **B25**, 368-376.

正确处理了一个后来沦为边缘化悖论的问题, 如第 15 章所述, 在杰恩斯的著作 (Jaynes, 1983, 第 337~339 页和第 374 页) 中有更全面讨论.

Geisser, S. (1980), 'The contributions of Sir Harold Jeffreys to Bayesian inference', in Zellner, A., ed., *Bayesian Analysis in Econometrics and Statistics*, North-Holland Publishing Co., Amsterdam, pp. 13-20.

Gell-Mann, M. (1992) 'Nature conformable to Herself', *Bull. Santa Fe Inst.* **7**, 7-10.

关于数学与物理关系的一些评论. 这位诺贝尔物理学奖获得者和我们一样, 对“布尔巴基主义瘟疫”终于消失感到高兴, 希望数学和理论物理学可以再次成为互助伙伴而不是敌手.

Gibbs, J. W. (1875), 'On the equilibrium of heterogeneous substances'; reprinted in *The Scientific Papers of J. Willard Gibbs*, Vol. I, Longmans, Green & Co. (1906), and Dover Publications, Inc., New York (1961).

Gibbs, J. W. (1902), *Elementary Principles in Statistical Mechanics*, Yale University Press,

New Haven, Connecticut.

[中译本] 约西亚·威拉德·吉布斯. 统计力学的基本原理. 毛俊雯, 译. 汪秉宏, 审校. 合肥: 中国科学技术大学出版社, 2016.

重印于 *The Collected Works of J. Willard Gibbs*, Vol. 2, 由 Longmans, Green & Co. (1928) 和 Dover Publications, Inc., New York (1960) 出版.

Glymour, C. (1985), 'Independence assumptions and Bayesian updating', *Artificial Intell.* **25**, 25-99.

Gnedenko, B. V. (1962), *The Theory of Probability*, Chelsea Publ. Co., New York, pp. 40-41.

[中译本] Б. В. 格涅坚科. 概率论教程. 丁寿田, 译. 北京: 高等教育出版社, 1956.

Gödel, K. (1931), 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I', *Monatshefte für Math. & Phys.*, **38**, 173-198.

英译版: 'On formally undecidable propositions of Principia Mathematica and related systems', Basic Books, Inc., New York (1962); 重印于 Dover Publications, Inc., New York (1992).

Goldman, S. (1953), *Information Theory*, Prentice-Hall, Inc., New York.

即使这部著作在该领域有奇怪的名声, 我们仍想为其添加一个友好的说明. 作者重叙了维纳和香农的工作, 而且为初学者着想, 比维纳和香农更清楚地解释了它们. 它的奇怪名声是两个不幸的偶然的結果: (1) 第 1 章标题中的一个拼写错误没有引起作者和出版者的注意, 但为 20 世纪 50 年代流传的数十个冷笑话提供了素材; (2) 在第 295 页有一张吉布斯的照片, 标题是“吉布斯 (1839—1903), 他的遍历假设是信息论中基本思想的前驱”, 但是由于吉布斯从未提到遍历性, 这是更多笑话的来源. 然而, 作者只是因为相信维纳的工作 (Wiener, 1948) 的真实性而受到谴责.

Goldstein, H. (1980), *Classical Mechanics*, Addison-Wesley, Reading, MA.

[中译本] H. 戈德斯坦. 经典力学. 陈为恂, 译. 汤家镛, 校. 北京: 科学出版社, 1981, 1986.

Good, I. J. (1950), *Probability and the Weighing of Evidence*, C. Griffin & Co., London.

一部重要性与篇幅不成比例的著作. 它仍然是每位做科学推断的学生的必读之作, 并且可以在一个晚上读完.

Good, I. J. (1962), *The Scientist Speculates*, Heinemann Basic Books, New York.

Good, I. J. (1967), 'The white shoe is a red herring', *Brit. J. Phil. Sci.* **17**, 322.

重印、转载于古德的文章集 (Good, 1983), 其中指出了亨佩尔悖论中的错误.

Good, I. J. (1980), 'The contributions of Jeffreys to Bayesian statistics', in A. Zellner, ed., *Bayesian Analysis in Econometrics and Statistics*, North-Holland Pub. Co., Amsterdam.

Good, I. J. (1983), *Good Thinking*, University of Minnesota Press.

重印了古德的 23 篇文章, 涵盖了多年完成的许多主题, 以及其他作品的参考列表. 古德大约写过 2000 篇这样的短文章, 从 20 世纪 40 年开始在统计和哲学文献中都可以找到. 该领域的工作者普遍认为, 现代统计学中的每一个想法都可以在这些文章中的一篇或多篇中找到源头, 但是它们的数量之多让我们无法找到或引用它们. 而且大多数文章只有一两页长, 在一小时内匆匆写完, 再也没有进一步发展. 因此, 许多年来, 无论人们在贝叶斯统计

中做了什么,都只是默认将首先发明权让给古德,而没有尝试搜索相关文献.找到相关文献可能需要几天时间.最后,这本书提供了这些文章的前 1517 篇中的大部分的列表(大概是按照它们的写作顺序,而不是出版顺序排列),并且带有一个长索引,因此现在可以对他到 1983 年为止的作品给予适当引用.请务必阅读第 15 章,他在那里指出了卡尔·波普尔工作中具体、定量的错误,并证明了波普尔拒绝的贝叶斯方法实际上能纠正这些错误.

Gould, S. J. (1989), *Wonderful Life: The Burgess Shale and the Nature of History*, W. W. Norton & Co., New York.

中译本 斯蒂芬·杰·古尔德. 奇妙的生命: 六亿年地球生命演化的秘密. 郑浩, 译. 海口: 海南出版社, 2019.

加拿大落基山脉中的一个小地区拥有完全合适的地质变迁历史,因此软体动物几乎完美地保存了下来.因此,我们现在知道寒武纪早期存在的生命种类比想象的要多得多.这对我们的进化观念有着深远的影响.古尔德近乎狂热地坚持认为“进化”并不等同于“进步”.当然,任何熟悉物理和化学原理的人都会同意,以一个方向进行的过程也能以相反方向进行.尽管如此,在我们看来,至少 99% 的观察到的进化实际上是朝着进步的方向发展的(更有能力、适应性更强的生物).我们还认为,用符合当前基本知识和贝叶斯推断原理的术语正确表述的达尔文理论,恰恰预测了这一点.

Graunt, J. (1662), *Natural and Political Observations Made Upon the Bills of Mortality*, Roycroft, London.

重印于 Newman, J. R., ed. (1956), *The World of Mathematics*, Vol. 3, pp. 1420–1435, Simon & Schuster, New York. 首次意识到可以从出生和死亡记录中推断出的有用事实,是社会学推断的开始,区别于单纯的统计数据收集.这项工作有时被归功于威廉·佩蒂,详情参见格林伍德的文章 (Greenwood, 1942).

Gray, C. G. & Gubbins, K. E. (1984), *Theory of Molecular Fluids*, Oxford University Press.

Greenwood, Major (1942), 'Medical statistics from Graunt to Farr', *Biometrika*, **32**, 203–225.

三部分工作的第 2 部分. 对约翰·格兰特 (1620—1674)、威廉·佩蒂 (1623—1687) 和埃德蒙·哈雷 (1656—1742) 第 1 张死亡率表的冗长但混乱的描述. 佩蒂 (格兰特的朋友,也是那些躁动不安但没有受到良好训练的人之一,几乎对所有东西都简短地涉猎,但从未真正掌握任何东西) 试图在哈雷之前许多年对爱尔兰进行人口普查,但没有足够仔细地推断得出有意义的结果. 格林伍德对佩蒂是否是格兰特工作的真正作者感到很困惑,他显然不知道佩蒂是在格兰特死后编辑了其死亡率表的第 5 版. 哈雷参考了佩蒂的版本,并看到了应如何纠正. 奥古斯塔斯·德摩根 (Augustus de Morgan, 1872, Vol. I, pp. 113–115) 很久以前就用有趣而讽刺的语言对所有这些进行了解释.

Grosser, M. (1979), *The Discovery of Neptune*, Dover Publications, Inc., New York.

Haldane, J. B. S. (1932), *Proc. Camb. Phil. Soc.* **28**, 58.

提倡和使用非正常先验. 哈罗德·杰弗里斯 (Harold Jeffreys, 1939) 承认这是他自己一些想法的来源.

Halley, E. (1693), 'An estimate of the degrees of mortality of mankind', *Phil. Trans. Roy. Soc.* **17**, 596–610, 654–656.

- 重印于 Newman, J. R., ed. (1956), *The World of Mathematics*, Simon & Schuster, New York, Vol. 3, pp. 1436-1447. 第一张死亡率表, 基于 1687~1691 年布雷斯劳的出生和死亡记录. 另见格林伍德的文章 (Greenwood, 1942).
- Hamilton, A. G. (1988), *Logic for Mathematicians*, revised 2nd edn, Cambridge University Press.
- 中译本 A. G. 哈密尔顿. 数学家的逻辑. 骆如枫, 陈慕昌, 茹季札, 黄万徽, 译. 沈百英, 校. 北京: 商务印书馆, 1989.
- Hansel, C. E. M. (1980), *ESP and Parapsychology - A Critical Re-evaluation*, Prometheus Books, Buffalo, NY, Chap. 12.
- Hardy, G. H. (1911), 'Theorems connected with MacLearin's test for the convergence of series', *Proc. Lond. Math. Soc.* **9**, (2), 126-144.
- Haar, A. (1933), 'Der Massbegriff in der Theorie der Kontinuierlichen Gruppen', *Ann. Math. Stat.* **34**, 147-169.
- Hartigan, J. (1964), 'Invariant prior distributons', *Ann. Math. Stat.* **35**, 836-845.
- Hedges, L. V. & Olkin, I. (1985), *Statistical Methods for Meta-Analysis*, Academic Press, Inc., Orlando, FL.
- Helmholtz, H. von (1868), 'Ueber discontinuirliche flussigkeitsbewegungen', *Monatsberichte d. Konigl. akademie der wissenschaften zu Berlin*.
- Hempel, C. G. (1967), *Brit. J. Phil. Sci.* **18**, 239-240.
- 对古德 (Good, 1967) 的回复.
- Herschel, J. (1850), *Edinburgh Rev.* **92**, 14.
- 在麦克斯韦 (Maxwell, 1860) 之前给出二维“麦克斯韦速度分布”.
- Hill, B. M. (1980), 'On some statistical paradoxes and nonconglomerability', in Bernardo, J. M. et al., eds., *Bayesian Statistics*, University Press, Valencia.
- Holland, J. D. (1962), 'The Reverend Thomas Bayes F.R.S. (1702-1761)', *J. Roy. Stat. Soc. (A)*, **125**, 451-461.
- Howson, C. & Urbach, P. (1989), *Scientific Reasoning: The Bayesian Approach*, Open Court Publishing Co., La Salle, Illinois.
- 一部奇怪的过时著作, 它若在 60 年前出版可能有用. 主要是对过去哲学家开始的错误观点的重述, 没有提供对它们的新见解, 并且忽略了使得它们过时的科学家、工程师和经济学家的现代发展部分. 很少有贝叶斯统计的材料能超越哈罗德·杰弗里斯 50 年前的理解水平, 除去应用它所需的数学. 他们坚持使用前杰弗里斯的记号表示法, 无法在概率符号中表示先验信息, 不注意冗余参数, 也没有解决任何问题.
- Howson, C. & Urbach, P. (1991), 'Bayesian reasoning in science', *Nature* **350**, 371-374.
- 豪森和乌尔巴赫的著作 (Howson and Urbach, 1989) 的广告, 也有同样的缺点. 由于他们使用 60 年前就存在的形式阐述贝叶斯原理, 因此安东尼·爱德华兹 (*Nature* **352**, 386-387) 用 60 年前他的老师费希尔给出的标准反驳来回应是恰当的. 但是对于那些在当今真正的科学问题中积极实际使用贝叶斯方法的人来说, 这种交流似乎是在对两种不同

的本轮系统进行争论。

Hoyt, W. G. (1980), *Planets X and Pluto*, University of Arizona Press, Tucson.

Huber, P. J. (1981), *Robust Statistics*, J. Wiley & Sons, Inc., New York.

Huber, P. J. (1992), *Galileo's Revenge: Junk Science in the Courtroom*, Basic Books, Inc., NY.

记录了伪装成科学家的江湖骗子和疯子正在造成的破坏性影响。他们受雇提供“专家”证词,声称存在各种实际并不存在的奇怪因果关系,以支持各种诉讼。这使得消费者和企业浪费数十亿美元的费用。亲因果偏见和反因果偏见现象在本书第5章、第16章和第17章中有讨论。目前似乎没有什么有效的方法来对付这种行为,正如加德纳(Gardner, 1981)指出的那样,新闻媒体总是倾向于报道能引起轰动的新闻,从而对江湖骗子给予支持和鼓励,同时拒绝负责任的科学家举行听证会来陈述事实。看起来,什么是有效及无效的科学推断问题必须很快走出学术界,成为立法问题——这方面的前景比目前的滥用更可怕。

Hume, D. (1739), *A Treatise of Human Nature*, London; as revised by P. H. Nidditch, Clarendon Press, Oxford (1978).

[中译本] 休谟. 人性论. 关文运, 译. 北京: 商务印书馆, 2016, 2020.

Hume, D. (1777), *An Inquiry Concerning Human Understanding*, Clarendon Press, Oxford.

Jaynes, E. T. (1956), 'Probability theory in science and engineering', no. 4 in *Colloquium Lectures in Pure and Applied Science*, Socony-Mobil Oil Co., USA.

Jaynes, E. T. (1957a), 'Information theory and statistical mechanics', *Phys. Rev.* **106**, 620-630; **108**, pp. 171-190.

重印于杰恩斯的文章集 (Jaynes, 1983).

Jaynes, E. T. (1957b), 'How does the brain do plausible reasoning?', Stanford University Microwave Laboratory Report 421.

重印于埃里克森和史密斯的著作 (Erickson & Smith, 1988, 第1卷, 第1~23页).

Jaynes, E. T. (1963a), 'New engineering applications of information theory', in Bogdanoff, J. L. & Kozin, F., eds., *Engineering Uses of Random Function Theory and Probability*, J. Wiley & Sons, Inc., NY, pp. 163-203.

Jaynes, E. T. (1963b), 'Information theory and statistical mechanics', in K. W. Ford, ed., *Statistical Physics*, W. A. Benjamin, Inc., New York, pp. 181-218.

重印于杰恩斯的文章集 (Jaynes, 1983).

Jaynes, E. T. (1963c), 'Comments on an article by Ulric Neisser', *Science* **140**, 216.

就人机交互交换意见.

Jaynes, E. T. (1965), 'Gibbs vs. Boltzmann entropies', *Am. J. Phys.* **33**, 391.

重印于杰恩斯的文章集 (Jaynes, 1983).

Jaynes, E. T. (1967), 'Foundations of probability theory and statistical mechanics', in Bunge, M. (ed.), *Delaware Seminar in Foundations of Physics*, Springer-Verlag, Berlin.

重印于杰恩斯的文章集 (Jaynes, 1983).

- Jaynes, E. T. (1968), 'Prior probabilities', *IEEE Trans. Systems Science and Cybernetics*, **SSC-4**, 227-241.
重印于 Tummala, V. M. Rao & Henshaw, R. C., eds., *Concepts and Applications of Modern Decision Models*, Michigan State University Business Studies Series (1976), 以及杰恩斯的文章集 (Jaynes, 1983)。
- Jaynes, E. T. (1976), 'Confidence intervals vs Bayesian intervals', in W. L. Harper & C. A. Hooker, eds., *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, vol. II, Reidel Publishing Co., Dordrecht, Holland, pp. 175-257.
重印于杰恩斯的文章集 (Jaynes, 1983)。
- Jaynes, E. T. (1978) 'Where do we stand on maximum entropy?', in Levine, R. D. and Tribus, M., eds., *The Maximum Entropy Formalism*, M.I.T. Press, Cambridge MA, pp. 15-118.
重印于杰恩斯的文章集 (Jaynes, 1983)。
- Jaynes, E. T. (1980), 'Marginalization and prior probabilities', in Zellner, A., ed., *Bayesian Analysis in Econometrics and Statistics*, North-Holland Publishing Co., Amsterdam.
重印于杰恩斯的文章集 (Jaynes, 1983)。
- Jaynes, E. T. (1983), in Rosenkrantz, R. D., ed., *Papers on Probability, Statistics, and Statistical Physics*, D. Reidel Publishing Co., Dordrecht, Holland; second paperbound edition, Kluwer Academic Publishers (1989).
1957~1980 年间 13 篇文章的重印。
- Jaynes, E. T. (1984), 'The intuitive inadequacy of classical statistics', *Epistemologia, Special Issue on Probability, Statistics, and Inductive Logic VII*, 43-73.
带讨论。
- Jaynes, E. T. (1985), 'Highly informative priors', in Bernardo, *et al.*, eds., *Bayesian Statistics 2*, Elsevier Science Publishers, North-Holland, pp. 329-360.
带讨论。先是历史综述, 然后是一个解决的示例 (计量经济学中的季节性调整问题), 显示先验信息可以如何影响我们的最终结论, 甚至是以一种正统统计理论语言无法描述的方式, 因为它不承认后验分布函数的相关性这一概念。
- Jaynes, E. T. (1986a), 'Bayesian methods: general background', in Justice, J. H., ed., *Maximum Entropy and Bayesian Methods in Geophysical Inverse Problems*, Cambridge University Press.
面向初学者的非技术性通用入门教程, 旨在解释术语和观点, 并提醒常见的误解和沟通难点。
- Jaynes, E. T. (1986b), 'Monkeys, kangaroos, and N ', in Justice, J. H., ed., *Maximum Entropy and Bayesian Methods in Geophysical Inverse Problems*, Cambridge University Press.
使用狄利克雷先验对图像重建中的更深层假设空间做初步探索。
- Jaynes, E. T. (1986c), 'Predictive statistical mechanics', in Moore, G. T. and Scully, M. O., eds., *Frontiers of Nonequilibrium Statistical Physics*, Plenum Press, NY, pp. 33-55.

Jaynes, E. T. (1987), 'Bayesian spectrum and chirp analysis', in Smith, R. C. & Erickson, G. J., eds., *Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems*, D. Reidel, Dordrecht-Holland, pp. 1-37.

对图基 (Tukey, 1984; 见参考文献) 的答复, 由布雷特索斯特 (Bretthorst, 1988) 进一步阐述。

Jaynes, E. T. (1989), 'Clearing up mysteries – the original goal', in Skilling, J., ed., *Maximum Entropy and Bayesian Methods*, Kluwer Publishing Co., Holland, pp. 1-27.

包含我们认为是贝叶斯定理在动力学理论中的首次应用, 在贝尔定理中隐藏假设的首次识别, 以及热力学第二定律在生物学中的首次定量应用。

Jaynes, E. T. (1990a), 'Probability in quantum theory', in Zurek, W. H., ed. *Complexity, Entropy and the Physics of Information*, Addison-Wesley Pub. Co., Redwood City, CA, pp. 381-404.

使用作为逻辑的概率论使得量子理论的意义显得非常不同, 并暗示量子力学概念上的困难可能在未来得到解决。

Jaynes, E. T. (1990b), 'Probability theory as logic', in Fougère, P., ed., *Proceedings of the Ninth Annual Workshop on Maximum Entropy and Bayesian Methods*, Kluwer Publishers, Holland.

通过一个不平凡例子表明, 条件概率不需要表达任何波普尔类型的因果作用——这一事实与杰恩斯 (Jaynes, 1989) 讨论的贝尔定理中的隐藏假设高度相关。

Jaynes, E. T. (1992), 'Commentary on two articles by C. A. Los', *Computers & Math. Appl.* **24**, 267-273.

这位经济学家令人惊讶地不仅谴责我们的贝叶斯分析, 而且谴责几乎所有在数据分析中做过的有用的事情, 仿佛回到高斯时代。

Jefferson, T. (1781), *Notes on Virginia*; reprinted in Koch, A. & Peden, W. (eds.), *The Life and Selected Writings of Thomas Jefferson*, The Modern Library, New York (1944).

中译本 托马斯·杰斐逊. 杰斐逊选集. 朱曾汶, 译. 北京: 商务印书馆, 2017. 第 VI 章: 弗吉尼亚笔记。

1972 年由 Random House, Inc 重印。

Jeffrey, R. C. (1983), *The Logic of Decision*, 2nd edn, University of Chicago Press.

尝试以特别的方式修改贝叶斯定理. 正如第 5 章讨论的, 这必然违反了我们的必备条件之一。

Jeffreys, H. (1931), *Scientific Inference*, Cambridge University Press; later editions, 1937, 1957, 1973.

中译本 哈罗德·杰弗里. 科学推断. 龚凤乾, 译. 厦门: 厦门大学出版社, 2011.

请务必阅读引言部分, 其中包含与伽利略的对话, 展示了归纳法如何在科学中实际应用。

Jeffreys, H. (1932), 'On the theory of errors and least squares', *Proc. Roy. Soc.* **138**, 48-55.

表达对尺度参数完全无知的 $d\sigma/\sigma$ 先验的一个漂亮推导, 受到费希尔 (Fisher, 1933) 的猛烈攻击. 我们在第 7 章和第 16 章中对此有讨论。

Jeffreys, H. (1939), *Theory of Probability*, Clarendon Press, Oxford; later editions, 1948, 1961, 1967, 1988.

中译本 哈罗德·杰弗里. 概率论. 龚凤乾, 译. 厦门: 厦门大学出版社, 2014.

在我们的前言中对此有感谢和赞誉.

Jeffreys, H. & Jeffreys, Lady Bertha Swirles (1946), *Methods of Mathematical Physics*, Cambridge University Press.

Jevons, W. S. (1874), *The Principles of Science: A Treatise on Logic and Scientific Method*, 2 vols., Macmillan, London.

重印于 Dover Publications, Inc., NY (1958). 杰文斯是德摩根的学生, 也阐述了拉普拉斯的观点. 因此, 杰文斯和德摩根都受到了维恩和其他人的攻击, 萨贝尔 (Zabell, 1989) 称之为“杰文斯大争论”.

Johnson, R. W. (1985), ‘Independence and Bayesian updating methods’, U. S. Naval Research Laboratory Memorandum Report 5689, November 1985.

Johnson, W. E. (1924), *Logic, Part III: The Logical Foundations of Science*, Cambridge University Press; reprinted by Dover Publications, Inc., NY (1964).

Johnson, W. E. (1932), ‘Probability, the deduction and induction problem’, *Mind* **44**, 409–413.

Justice, J. H. (1986), ed., *Maximum Entropy and Bayesian Methods in Geophysical Inverse Problems*, Cambridge University Press.

1984 年 8 月在卡尔加里举行的第四届年度“最大熵研讨会”的论文集.

Kac, M. (1956), ‘Some stochastic problems in physics and mathematics’, Field Research Laboratory. Magnolia Petroleum Co., Dallas, Colloquium Lectures in Pure and Applied Science no. 2.

Kadane, J. B., Schervish, M. J. & Seidenfeld, T. (1986), ‘Statistical implications of finitely additive probability’, in Goel, P. K. & Zellner, A., eds., *Bayesian Inference and Decision Techniques*, Elsevier Science Publishers, Amsterdam.

KSS 的工作在第 15 章中有详细讨论.

Kahneman, D. & Tversky, A. (1972), ‘Subjective probability: a judgment of representativeness’, *Cog. Psychol.* **3**, 430–454.

另见特韦尔斯基和卡奈曼的文章 (Tversky & Kahneman, 1981).

Kemphorne, O. & Folks, L. (1971), *Probability, Statistics and Data Analysis*, Iowa State University Press.

Kendall, M. G. (1963), ‘Ronald Aylmer Fisher, 1890–1962’, *Biometrika* **50**, 1–15.

重印于皮尔逊和肯德尔的著作 (Pearson & Kendall, 1970). 与之前的引用文献一样, 这里告诉我们更多的是关于作者而不是相关主题的信息.

Kendall, M. G. & Moran, P. A. P. (1963), *Geometrical Probability*, Griffin, London.

很多有用的数学材料, 所有这些材料都很容易适配贝叶斯主义.

Kendall, M. G. & Plackett, R. L. (1977), *Studies in the History of Statistics and Probability*, 2 vols, Griffin, London.

Kendall, M. G. & Stuart, A. (1961), *The Advanced Theory of Statistics: Volume 2, Inference and Relationship*, Hafner Publishing Co., New York.

这代表了置信区间的终结, 尽管作者以“客观性”为由继续支持它, 但是他们注意到很多由此产生的荒谬之处, 以至于读者以后不再敢使用置信区间. 杰恩斯 (Jaynes, 1976) 解释了产生这种困难的根源, 并表明这些荒谬结果如何可以通过应用贝叶斯方法自动纠正.

Kendall, M. G. & Stuart, A. (1977), *The Advanced Theory of Statistics: Volume 1, Distribution Theory*, Macmillan, New York.

Kennard, E. H. (1938), *Kinetic Theory of Gases*, McGraw-Hill Book Co., NY.

Keynes, J. M. (1921), *A Treatise on Probability*, Macmillan, London; reprinted by Harper & Row, New York (1962).

首次明确解释逻辑独立性和因果独立性之间的区别. 直到今天依然很重要, 因为它在历史上曾是考克斯工作的灵感来源. 有关凯恩斯的有趣评论见博雷尔的文章 (Borel, 1924).

Keynes, J. M. (1936), *Allgemeine theorie der beschuiftigung, des Zinses und des Geldes*, von Duncker & Humblot, Munchen, Leipzig.

Khinchin, A. I. (1957), *Mathematical Foundations of Information Theory*, Dover Publications, Inc., New York.

一位数学家试图“严格化”香农的工作, 但是我们认为不需要这个. 无论如何, 当人们从一开始就试图直接在无限集合上工作时, 由此产生的定理不涉及现实世界中的任何内容. 辛钦可能足够小心地避免实际错误, 从而产生了在他想象世界中有效的定理, 但是我们在第 15 章中注意到其他试图以这种方式进行数学计算的人造成的一些恐怖结果.

Kline, M. (1980), *Mathematics: The Loss of Certainty*, Oxford University Press.

中译本 莫里斯·克莱因. 数学简史: 确定性的消失. 李宏魁, 译. 北京: 中信出版集团, 2019.

相当完整的数学史, 记录了数百个有趣的轶事, 但是表达的观点与菲利克斯 (Félix, 1960) 的布尔巴基主义观点截然不同.

Kolmogorov, A. N. (1933), *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Math. (2), Berlin; English translation, *Foundations of the Theory of Probability*, Chelsea Publishing Co., New York (1950).

中译本 柯尔莫哥洛夫. 概率论基本概念. 丁寿田, 译. 北京: 商务印书馆, 1952.

在附录 A 中有描述.

Koopman, B. O. (1936), 'On distributions admitting a sufficient statistic', *Trans. Am. Math. Soc.* **39**, 399–509.

证明 NASC 存在充分统计量的条件是分布具有指数形式, 后来被认为与最大熵自动生成的相同. 与皮特曼 (Pitman, 1936) 同时发现了这一结论.

Kronecker, L. (1901), *Vorlesungen über Zahlentheorie*, Teubner, Leipzig; republished by Springer-Verlag, 1978.

- Kullback, S. & Leibler, R. A. (1951), 'On information and sufficiency', *Ann. Math. Stat.* **22**, 79-86.
- Kurtz, P. (1985), *A Skeptic's Handbook of Parapsychology*, Prometheus Books, Buffalo, NY.
里面的几章有相关材料, 特别见贝蒂·马克威克写的第 11 章.
- Kyburg, H. E. & Smokler, H. E. (1981), *Studies in Subjective Probability*, 2nd edn, J. Wiley & Sons, Inc., New York.
- Lancaster, H. O. (1969), *The Chi-squared Distribution*, J. Wiley & Sons, Inc., New York.
- Lane, D. A. (1980), 'Fisher, Jeffreys, and the nature of probability', in Fienberg, S., *et al.*, eds., *R. A. Fisher, an Appreciation*, Springer-Verlag, New York, pp. 148-160.
- Laplace, P. S. (1774), 'Mémoire sur la probabilité des causes par les évènements', *Mem. Acad. Roy. Sci.* **6**, 621-656.
重印于拉普拉斯的著作 (Laplace, 1878-1912, 第 8 卷, 第 27~65 页). 由施蒂格勒 (Stigler, 1986b) 翻译成英文.
- Laplace, P. S. (1781), 'Memoire sur les probabilités', *Mem. Acad. Roy. Sci. (Paris)*.
重印于拉普拉斯的著作 (Laplace, 1878-1912, 第 9 卷, 第 384~485 页). "高斯" 分布特性的早期阐述, 并且认为它是如此重要, 以至于建议制成表格.
- Laplace, P. S. (1810), 'Mémoire sur les approximations des formules qui sont fonctions de très grands nombres et sur leur application aux probabilités,' *Mem. Acad. Sci. Paris*, 1809, pp. 353-415, 559-565.
重印于拉普拉斯的著作 (Laplace, 1878-1912, 第 12 卷, 第 301~353 页). 高斯分布起源、性质及使用的大量概略性描述.
- Laplace, P. S. (1812), *Théorie Analytique des Probabilités*, 2 vols., Courcier Imprimeur, Paris; 3rd edition with supplements, 1820.
重印于拉普拉斯的著作 (Laplace, 1878-1912, 第 7 卷). 这部罕见但非常重要的作品的重印版本可以通过以下来源获得: Editions Culture et Civilisation, 115 Ave. Gabriel Lebron, 1160 Brussels, Belgium.
- Laplace, P. S. (1814, 1819), *Essai Philosophique sur les Probabilités*, Courcier Imprimeur, Paris.
重印于 *Oeuvres Complètes de Laplace*, Vol. 7, Gauthier-Villars, Paris (1886). 可从以下来源获得: Editions Culture et Civilisation, 115 Ave. Gabriel Lebron, 1160 Brussels, Belgium. 英文翻译: F. W. Truscott & F. L. Emory, Dover Publications, Inc., New York (1951). 请注意, 这一 "翻译" 只不过是字面上的转录, 在许多方面扭曲了拉普拉斯的意思. 在接受该译本中的任何解释性声明之前, 一定记得参阅法语原版.
- Laplace, P. S. (1878-1912), *Oeuvres Complètes*, 14 vols., Gauthier-Villars, Paris.
- Lee, Y. W. (1960), *Statistical Theory of Communication*, J. Wiley & Sons, New York.
源自维纳的可用但是有折扣的教学性著作 (Wiener, 1949). 有大量解释得很清楚的示例, 但没有任何诸如维纳原作中使用的佩利-维纳分解或维纳测度上的函数积分等数学技巧. 极大地扩充始于维纳 (Wiener, 1948) 的关于吉布斯的传说. 评论见 E. T. Jaynes, *Am.*

J. Phys. **29**, 276 (1961).

Lehmann, E. L. (1959), *Testing Statistical Hypotheses*, Wiley, New York; 2nd edn, 1986.

Lighthill, M. J. (1957), *Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press.

所有不相信德尔塔函数的人的必读读物. 见戴森评论 (Dyson, 1958). 莱特希尔和戴森是哈代在剑桥大学著名的“纯数学”课程的同班同学, 当时傅里叶分析主要专注于收敛理论, 如蒂奇马什的著作 (Titchmarsh, 1937) 中所述的那样. 现在, 我们在附录 B 中对术语“函数”进行了重新定义, 所有这些都变得几乎无关紧要. 戴森表示, 莱特希尔“将哈代的著作毁于一旦, 但是哈代会比其他任何人都更喜欢它”.

Lindley, D. V. (1957), ‘A statistical paradox’, *Biometrika* **44**, 187–192.

提到索尔和贝特曼 (Soal & Bateman, 1954) 的超心理学实验.

Little, J. F. & Rubin, D. B. (1987), *Statistical Analysis with Missing Data*, J. Wiley & Sons, New York.

中译本 利特尔, 鲁宾. 缺失数据统计分析. 孙山泽, 译. 北京: 中国统计出版社, 2004.

缺失数据会对正统统计方法造成严重的破坏性影响, 因为这会改变样本空间, 从而不仅会改变估计量的抽样分布, 甚至还会改变其分析形式. 对于这种情况, 人们必须回到起点. 但是无论抽样分布变化如何复杂, 似然函数的变化都是非常简单的. 贝叶斯方法可以毫不费力地适应缺失数据的情形. 在所有情况下, 我们只是将我们拥有的所有数据包含在似然函数中, 贝叶斯定理会自动返回该数据集的新的最优估计量.

Luce, R. D. & Raiffa, H. (1989), *Games and Decisions*, Dover Publications, NY.

Lukasiewicz, J. (1957), *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*, 2nd edn, Clarendon Press, Oxford; reprinted 1972.

中译本 卢卡西维茨. 亚里士多德的三段论. 李真, 李先焜, 译. 北京: 商务印书馆, 2017.

Lusted, L. (1968), *Introduction to Medical Decision Making*, C. C. Thomas, Springfield, IL.

这本书中包含许多对重要医学问题有用的贝叶斯解, 带有计算机源代码. 李·卢斯特德 (1923—1994) 是多年前我在康奈尔大学物理专业的同学. 然后, 我们的经历出奇地相似: 卢斯特德在哈佛无线电研究实验室从事微波雷达对抗研究, 我在阿纳卡斯蒂亚的海军研究实验室从事雷达目标识别工作; 二战后, 卢斯特德进入哈佛医学院攻读医学博士学位, 我在普林斯顿大学研究生院攻读理论物理学博士学位. 我们都主要对这些领域中使用的推理过程感兴趣. 然后我们都独立地发现了贝叶斯分析, 认为它是解决我们问题的方法, 并且把我们的余生都奉献给了它. 基本上在同一时间, 阿诺德·泽尔纳的经历也类似, 从物理学转向了经济学. 因此, 现代贝叶斯对三个完全不同领域的影响都来自物理学家, 他们的年龄和兴趣几乎相同. 一位社会学家抱怨: “上帝把简单问题给了物理学家.” 虽然我们中的一些人愿意承认这一点, 但我们只是想加一句: “……上帝如此安排的原因是使物理学家找到的解也有助于解决其他人的问题.”

McClave, J. T. & Benson, P. G. (1988), *Statistics for Business and Economics*, 4th edn, Dellen Publ. Co., San Francisco.

中译本 詹姆斯·麦克拉夫, 乔治·本森, 特里·辛西奇. 商务与经济统计学 (第 12 版). 易

丹辉, 李扬, 译. 北京: 中国人民大学出版社, 2015.

Machol, R. E., Ladany, S. P. & Morrison, D. G. (eds.) (1976), *Management Science in Sports*, Vol. 4, TIMS Studies in the Management Sciences, North-Holland, Amsterdam.

概率论的奇特应用, 得到了更奇特的结论. 同样, 莫里斯 (Morris, 1977) 对网球比赛进行了分析. 他将 1 分的“重要性”定义为如果球员赢得该分会赢得比赛的概率, 减去如果输了该分同样会赢得比赛的概率. 比较能力相当的球员, 他发现最重要的 1 分重要性值是 30~40. 然后他得出结论说, 通过在最重的分上更努力一点儿, 球员可以大大提高他的胜利前景. 批评这里的逻辑对于读者是一个很好的练习.

Maxwell, J. C. (1860), 'Illustration of the dynamical theory of gases. Part I. On the motion and collision of perfectly elastic spheres', *Phil. Mag.* 56.

Middleton, D. (1960), *An Introduction to Statistical Communication Theory*, McGraw-Hill Book Co., New York.

一本包含大量数学材料的巨著 (1140 页). 书名具有误导性, 因为该材料实际上适用于一般的统计推断. 不幸的是, 大部分工作完成得有点儿早, 所以结果是抽样理论和奈曼-皮尔逊决策规则, 现在由于沃尔德决策理论和贝叶斯方法的进步已经过时了. 尽管如此, 其中的数学问题——例如求解奇异积分方程的方法——与一个人的推断哲学无关, 因此它包含有很多适用于当前问题的有用材料. 人们应该浏览这本书, 并记录下其中的可用内容.

Middleton, D. & Van Meter, D. (1955), 'Detection and extraction of signals in noise from the point of view of statistical decision theory', *J. Soc. Ind. Appl. Math.* 3, (4), 192.

Middleton, D. & Van Meter, D. (1956), 'Detection and extraction of signals in noise from the point of view of statistical decision theory', *J. Soc. Ind. Appl. Math.* 4, (2), 86.

Molina, E. C. (1963), *Two Papers by Bayes with Commentaries*, Hafner Publishing Co., New York.

包含关于拉普拉斯和布尔关系的深入历史评论, 指出那些引用布尔以支持他们对拉普拉斯的攻击的人可能误解了布尔的意图.

Monod, Jacques (1970), *Le Hazard et la Nécessité*, Seuil, Paris.

Morris, C. (1977), 'The most important points in tennis', in Ladany, S. P. & Machol, R. E., eds., *Studies in Management Science and Systems*, vol. 5 (North-Holland, Amsterdam).

Mosteller, F. (1965), *Fifty Challenging Problems in Probability with Solutions*, Addison-Wesley, Reading, MA.

Newcomb, S. (1881), 'Note on the frequency of use of the different digits in natural numbers', *Am. J. Math.* 4, 39-40.

Neyman, J. (1950), *First Course in Probability and Statistics*, Henry Holt & Co., New York.

Neyman, J. (1952), *Lectures and Conferences on Mathematical Statistics and Probability*, Graduate School, US Dept of Agriculture.

包含贝叶斯区间与置信区间估计的令人难以置信的比较内容. 一个很好的作业是找出其中推理的错误.

Northrop, E. P. (1944), *Riddles in Mathematics; a Book of Paradoxes*, van Nostrand, New York, pp. 181–183.

Pearson, E. S. (1967), 'Some reflections on continuity in the development of mathematical statistics 1890–94', *Biometrika* **54**, 341–355.

Pearson E. S. & Clopper C. J. (1934), 'The use of confidence in fiducial limits illustrated in the case of the binomial', *Biometrika* **26**, 404–413.

Pearson, E. S. & Kendall, M. G. (1970), *Studies in the History of Statistics and Probability*, Hafner Publishing Co., Darien, Conn.

Pearson, K. (1914–1930), *The Life, Letters and Labours of Francis Galton*, 3 vols., Cambridge University Press.

弗朗西斯·高尔顿继承了一笔不多的财产, 1911 年去世时, 他在伦敦大学学院捐赠设立了优生学教授席位。卡尔·皮尔逊是它的第一获得者。这使他能够不再对工程师和物理学家讲授应用数学, 而专注于生物学和统计学。

Pearson, K. (1920), 'Notes on the history of correlation', *Biometrika* **13**, 25–45.

重印于皮尔逊和肯德尔的著作 (Pearson & Kendall, 1970)。

Pearson, K. (1970), 'Walter Frank Raphael Weldon 1860–1906', in Pearson, E. S. & Kendall, M. G., *Studies in the History of Statistics and Probability*, London.

Penrose, O. (1979), 'Foundations of statistical mechanics', *Rep. Prog. Phys.* **42**, 1937–2006.

发表在“进展报告”中, 但是实际上没有报告任何进展。

Pitman, E. J. G. (1936), 'Sufficient statistics and intrinsic accuracy', *Proc. Camb. Phil. Soc.* **32**, 567–579.

几乎与库普曼 (Koopman, 1936) 同时证明 NASC 的充分性, 现在称为皮特曼-库普曼定理。

Poincaré, H. (1899), 'L'Oeuvre mathématique de Weierstraß', *Acta. Math.* **22**, 1–18.

包含对克罗内克和魏尔斯特拉斯作品之间关系的权威描述, 指出它们之间的差异更多在于品味而不是实质。可与贝尔的著作 (Bell, 1937) 做对比, 后者似乎将他们描述为死对头。这在附录 B 中有讨论。

Poincaré, H. (1904), *Science et Hypothèse*; English translation, Dover Publications, Inc., NY (1952).

中译本 昂利·彭加勒. 科学与假设. 李醒民, 译. 北京: 商务印书馆, 2021.

庞加莱有一种天赋, 就是一句话能比大多数作家一页说得更多。这本书中充满了可引用的评论, 到今天仍与其写作时一样真实和重要。

Poincaré, H. (1909), *Science et Méthode*; English translation, Dover Publications, Inc., New York (1952).

中译本 昂利·彭加勒. 科学与方法. 李醒民, 译. 北京: 商务印书馆, 2010.

就像克莱因的著作 (Kline, 1980) 一样, 这是对数学和逻辑学当代工作的强烈控诉, 布尔巴基主义者从未因此原谅过他。然而, 在知识和判断力上, 庞加莱远远胜过他的现代批评家, 因为他与现实世界有更好的联系。

Poincaré, H. (1912), *Calcul des Probabilités*, 2nd edn, Gauthier-Villars, Paris.

包含通过群不变性原则分配概率分布的第一个示例。

Pólya, G. (1920), 'Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem', *Math. Zeit.* 8, 171–181.

重印于波利亚的著作 (Pólya, 1984, 第 IV 卷)。“中心极限定理”一词首次出现在印刷品中。他实际上并没有证明该定理 (他将其归功于拉普拉斯), 而是指出了可用于该定理各种证明的、关于一系列单调函数一致收敛的一个定理, 但是我们在第 7 章中的证明更简单。

Pólya, G. (1945), *How to Solve It*, Princeton University Press. Second paperbound edition by Doubleday Anchor Books (1957).

[中译本] G. 波利亚. 怎样解题: 数学思维的新方法. 涂泓, 冯承天, 译. 上海: 上海科技教育出版社, 2018.

Pólya, G. (1954), *Mathematics and Plausible Reasoning*, 2 vols, Princeton University Press.

[中译本] G. 波利亚. 数学与猜想 (第一卷) (第二卷). 李心灿, 王日爽, 李志尧, 译. 杨禄荣, 张理京, 校. 北京: 科学出版社, 2001.

Pólya, G. (1984), *Collected Papers*, 4 vols., ed. G-C. Rota, MIT Press, Cambridge, MA.

第 IV 卷包含关于概率论和组合学的论文, 几篇关于合情推理的短文章, 以及波利亚的 248 篇论文的参考列表. 波利亚一直声称他的主要兴趣在于解决特定问题的心理过程, 而不是推广. 尽管如此, 他的一些结果通过其他人的推广开创了新的数学分支. 本书在比我们前言中提到的更多方面受到波利亚的影响: 我们大部分阐述的目的不是为了自身原因阐述一般性, 而是为了学习如何解决特定问题——尽管是通过一般方法。

Pólya, G. (1987), *The Pólya Picture Album: Encounters of a Mathematician*, G. L. Alexanderson, ed., Birkhäuser, Boston.

波利亚一生收藏了他认识的著名数学家的照片并做成了一个大相册, 他乐于将其展示给参观者. 在他死后, 这本迷人的书收录了该相册, 其中包含约 130 张照片, 并附有波利亚的评论, 以及编辑撰写的波利亚传记。

Popper, K. (1957), 'The propensity interpretation of the calculus of probability, and the quantum theory', in *Observation and Interpretation*, S. Körner, ed., Butterworth's Scientific Publications, London, pp. 65–70.

在这里, 批评了量子理论的波普尔向关注量子理论基础的科学家们总结了他的观点。

Popper, K. (1959), 'The propensity interpretation of probability', *Brit. J. Phil. Sci.* 10, pp. 25–42.

Popper, K. (1974), 'Replies to my critics', in P. A. Schilpp, ed., *The Philosophy of Karl Popper*, Open Court Publishers, La Salle.

这大概是对波普尔立场的最权威的陈述, 因为它比他最著名的作品晚了几年, 并试图直接回应对他的批评。

Popper, K. & Miller, D. W. (1983), 'A proof of the impossibility of inductive probability', *Nature*, 302, 687–688.

他们通过我们在第 5 章中检查过的过程得出这一结论,方法是断言一条概率论中不存在的基于直觉的特定原则.对科学家而言,这就像试图向一群专业的飞行员证明重于空气的飞行器是不可能飞行的.

Pratt, F. (1942), *Secret and Urgent; The Story of Codes and Ciphers*, 2nd edn, Blue Ribbon Books, Garden City, NY.

Pratt, J. W. (1961), 'Review of Testing Statistical Hypotheses' [Lehmann, 1959], *J. Am. Stat. Assoc.* **56**, pp. 163-166.

对正统假设检验理论的强有力的批评.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1986), *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press; 2nd edn, 1992.

中译本 W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery. C 语言数值算法程序大全(第二版).傅祖芸,赵梅娜,丁岩,等译.傅祖芸,校.北京:电子工业出版社,1995.

Quetelet, L. A. (1835), *Sur L'homme et le Développement de ses Facultés, ou Essai de Physique Sociale* (Bachelier, Paris).

1869 年以 *Physique Sociale* 为标题重新出版.

Quetelet, L. A. (1869), *Physique Sociale, ou Essai sur le Développement des Facultés, de L'homme*, C. Muquardt, Brussels.

Raiffa, H. A. & Schlaifer, R. S. (1961), *Applied Statistical Decision Theory*, Graduate School of Business Administration, Harvard University.

Raimi, R. A. (1976), 'The first digit problem', *Am. Math. Monthly* **83**, 521-538.

关于“本福德定律”的评论文章,其中有很多参考资料.

Rao, M. M. (1993), *Conditional Measures and Applications*, Marcel Dekker, Inc., New York.

我们在附录 A 中指出,条件概率的概念在柯尔莫哥洛夫系统中是多么陌生.

Reid, C. (1982), *Neyman - From Life*, Springer-Verlag, New York.

Rempe, G., Walter, H. & Klein, N. (1987), *Phys. Rev. Lett.* **58**, 353-356.

Rosenkrantz, R. D. (1977), *Inference, Method, and Decision: Towards a Bayesian Philosophy of Science*, D. Reidel Publishing Co., Boston.

由杰恩斯在 *J. Am. Stat. Assoc.*, Sept. 1979, pp. 740-741 进行评论.

Rothman, T. (1989), *Science à la Mode*, Princeton University Press.

说明当科学家失去客观性并随波逐流时会发生什么.我们想强调的是,他们不仅使自己变得荒谬,而且通过宣传耸人听闻但没有意义的想法来损害科学.例如,我们最终已经意识到“混沌”的潮流已经阻碍了六个不同领域的有序发展,而没有提供任何新的预测能力.因为,只要混沌存在,它就肯定会被哈密顿运动方程——这正是一个世纪以来我们在统计力学中一直使用的——预测到.混沌爱好者无法做出比现有的统计力学更好的预测,因为我们从来没有准确了解所需的初始条件.自麦克斯韦和吉布斯时代以来,人们一直认识到,如果我们对微观状态有准确的了解,那么原则上就可以预测目前还不能预测的未来的“热波动”细节.给定这些信息,如果存在混沌,它的细节也将被预测.但是在目前的统计力

学中, 由于缺乏这些信息, 我们只能预测与我们拥有的信息一致的所有可能的混沌行为的平均值, 这只是传统的热力学.

Routh, E. J. (1905), *The Elementary Part of A Treatise on the Dynamics of a System of Rigid Bodies*, Macmillan, New York.

关于整个主题的论文的第 1 部分和第 2 部分.

Rowe, D. E. & McCleary, J., eds. (1989), *The History of Modern Mathematics*, 2 vols., Academic Press, Inc., Boston.

Royall, R. M. & Cumberland, W. G. (1981), 'The finite-population linear regression estimator and estimators of its variance – an empirical study', *J. Am. Stat. Assoc.* **76**, 924–930.

随机化荒唐的另一个证明.

Ruelle, D. (1991), *Chance and Chaos*, Princeton University Press.

中译本 大卫·吕埃勒. 机遇与混沌. 刘式达, 李滇林, 梁爽, 译. 上海: 上海世纪出版集团, 上海科技教育出版社, 2005.

如何在科学中不使用概率论, 见第 4 章末尾的评注部分.

Savage, I. R. (1961), 'Probability inequalities of the Tchebyscheff type', *J. Res. Nat. Bureau Stand.* **65B**, 211–222.

一个有用的结果集合, 应该做得更易于获取.

Savage, L. J. (1954), *Foundations of Statistics*, J. Wiley & Sons, NY; 2nd edn (rev.), Dover Publications, Inc., NY (1972).

这项工作遭到了范丹齐格 (van Dantzig, 1957) 的强烈攻击.

Savage, L. J. (1961), 'The foundations of statistics reconsidered', *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley University Press.

Savage, L. J. (1962), *The Foundations of Statistical Inference, a Discussion*, Methuen, London.

Savage, L. J. (1981), *The Writings of Leonard Jimmie Savage – A Memorial Selection*, American Association of Statistics and the Institute of Mathematical Statistics.

吉米·萨维奇于 1971 年突然意外去世, 通过将他散布在许多不起眼之处且难以找到的作品集结在一起, 他的同事们完成了这项重要的工作. 关于他的一些个人回忆记录在杰恩斯的文章 (Jaynes, 1984, 1985) 中.

Schrödinger, E. (1948), *Statistical Thermodynamics*, Cambridge University Press.

中译本 E. 薛定谔. 统计热力学. 徐锡申, 译. 陈成琳, 校. 北京: 高等教育出版社, 2014.

Schuster, A. (1897), 'On lunar and solar periodicities of earthquakes', *Proc. Roy. Soc.* **61**, 455–465.

这标志着周期图的发明, 几乎可以称为正统显著性检验的起源. 舒斯特在不存在周期性的假设下仅考虑周期图的抽样分布来反驳一些关于地震周期性的主张! 如果某个频率的周期

性确实存在, 他从不考虑获得观测数据的概率是多少? 从那以后, 正统统计一直遵循着这种荒谬的流程.

Schwartz, L. (1950), *Théorie des Distributions*, 2 vols., Hermann et Cie, Paris.

Seal, H. (1967), 'The historical development of the Gauss linear model', *Biometrika* **54**, 1-24.

重印于皮尔逊和肯德尔的著作 (Pearson & Kendall, 1970).

Shannon, C. E. (1948), 'A mathematical theory of communication', *Bell Syst. Tech. J.*, **27**, 379, 623.

重印于 C. E. Shannon & W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1949). 另见斯隆和怀纳的著作 (Sloane & Wyner, 1993).

Shewhart, W. A. (1931), *Economic Control of Quality of Manufactured Products*, van Nostrand, New York.

Shore, J. E. & Johnson, R. W. (1980), 'Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy', *IEEE Trans. Information Theory* **IT-26**, 26-37.

许多不同的公理选择会导致相同的解决问题的实际算法. 两位作者提出了一个不同于最初 (Jaynes, 1957a) 的公理基础. 但是我们强调最大熵和最小交叉熵是相同的原理, 通过变量的变化可以将一个转换为另一个.

Simon, H. A. & Rescher, N. (1966), 'Cause and Counterfactual', *Phil. Sci.* **33**, 323-340.

Simon, H. A. (1977), *Models of Discovery*, D. Reidel Publ. Co., Dordrecht, Holland.

Sims, C. A. (1988), 'Bayesian skepticism on unit root econometrics', *J. Econ. Dyn. & Control* **12**, 463-474.

单位根假设 (见第 7 章和第 17 章) 与经济理论没有很好的联系, 但对这些模型的贝叶斯分析在正统显著性检验误导人的地方取得了成功.

Sivia, D. S. & Carlile, C. J. (1992), 'Molecular spectroscopy and Bayesian spectral analysis - how many lines are there?', *J. Chem. Phys.* **96**, 170-178.

通过贝叶斯计算机程序成功地将噪声数据分解为多达 9 个高斯分量.

Sivia, D. S. (1996), *Data Analysis - A Bayesian Tutorial*, Clarendon Press, Oxford.

这本书比较薄 (不到 200 页) 却是亟需的. 它以与我们相同的理论视角进行阐述, 其中包含大量关于贝叶斯数据处理的数值示例. 它为初学者提供了大量实用的建议, 每个理工科学生都可以轻松掌握, 应该被视为对本书的补充.

Sloane, N. J. A. & Wyner, A. D. (1993), *Claude Elwood Shannon: Collected Papers*, IEEE Press, Piscataway, NJ.

Smart, W. M. (1947), *John Couch Adams and the Discovery of Neptune*, Royal Astronomical Society, Occasional Notes No. 11.

Soal, S. G. & Bateman, F. (1954), *Modern Experiments in Telepathy*, Yale University Press, New Haven.

- Stigler, S. M. (1980), 'Stigler's law of eponymy', *Trans. NY Acad. Sci.* **39**, 147–159.
- Stigler, S. M. (1983), 'Who discovered Bayes's Theorem?', *Am. Stat.* **37**, 290–296.
- Stigler, S. M. (1986a), 'John Craig and the probability of history', *JASA* **81**, 879–887.
- Stigler, S. M. (1986b), 'Translation of Laplace's 1774 memoir on Probability of causes', *Stat. Sci.* **1**, 359.
- Stigler, S. M. (1986c), *The History of Statistics*, Harvard University Press.
- 中译本 斯蒂格勒. 统计探源: 统计概念和方法的历史. 李金昌, 等译. 鲜祖德, 主审. 杭州: 浙江工商大学出版社, 2014.
- 大部头的学术著作, 值得该学科的所有学生阅读. 对我们只是简单涉及的许多主题进行了全面讨论.
- Stone, M. (1965), 'Right Haar measure for convergence in probability to quasi-posterior distributions', *Ann. Math. Stat.* **30**, 449–453.
- Stone, M. (1976), 'Strong inconsistency from uniform priors', *J. Am. Stat. Assoc.* **71**, 114–116.
- 平面上的随机游走.
- Stone, M. (1979), 'Review and analysis of some inconsistencies related to improper priors and finite additivity', in *Proceedings of the 6th International Congress of Logic, Methodology, and Philosophy of Science*, Hanover, 1979, North Holland Press.
- 第 15 章提及的翻滚四面体问题.
- Stove, D. (1982), *Popper and After: Four Modern Irrationalists*, Pergamon Press, New York.
- Székel, G. J. (1986), *Paradoxes in Probability Theory and Mathematical Statistics*, D. Reidel Publishing Co., Dordrecht, Holland.
- 在第 64 页中给出有偏硬币的误解, 对物理学没有理解. 试与本书的第 10 章进行比较.
- Takacs, L. (1958), 'On a probability problem in the theory of counters', *Ann. Math. Stat.* **29**, 1257–1263.
- 塔卡克斯有许多关于这一主题的早期论文.
- Tax, S., ed., (1960), *Evolution After Darwin*, 3 vols., University of Chicago Press.
- 第 1 卷: 生命的进化 (*The Evolution of Life*); 第 2 卷: 人类的进化 (*The Evolution of Man*); 第 3 卷: 进化中的问题 (*Issues in Evolution*). 该领域中许多研究工作者的文章和小组讨论的论文集合, 总结了达尔文最初的著作出版 100 年后的知识状态和当前的研究方向.
- Temple, G. (1955), 'Theory of generalized functions', *Proc. Roy. Soc. Lond. Ser. A* **228**, 175–190.
- Titchmarsh, E. C. (1937), *Introduction to the Theory of Fourier Integrals*, Clarendon Press, Oxford.
- 在莱特希尔的著作 (Lighthill, 1957) 出版之前, 傅里叶分析中的“最新技术”. 莱特希尔的书使得所有冗长的收敛理论变得几乎无关紧要. 然而, 这部经典著作只有一部分因此过

时了, 希尔伯特变换、埃尔米特函数和贝塞尔函数以及维纳-霍普夫积分方程的材料对于应用数学来说仍然是必不可少的。

Titchmarsh, E. C. (1939), *The Theory of Functions*, 2nd edn, Oxford University Press.

[中译本] 蒂奇马什. 函数论. 刘培杰数学工作室, 译. 哈尔滨: 哈尔滨工业大学出版社, 2014.

在第 11 章中, 读者可能会 (也许是第一次) 看到一些不可微函数的实际例子. 我们在附录 B 中对此有简要讨论.

Titterton, D. M., Smith, A. F. M. & Makov, U. E. (1985), *Statistical Analysis of Finite Mixture Distributions*, Wiley, NY.

Tribus, M. (1961), *Thermostatistics and Thermodynamics; An Introduction to Energy, Information and States of Matter, with Engineering Applications*, Van Nostrand, Princeton, NJ.

Tribus, M. (1969), *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York.

Tribus, M. & Fitts, G. (1968), 'The widget problem revisited', *IEEE Trans. SSC-4*, (3), 241-248.

Tukey, J. W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading, MA.

引入“抗性”(resistant)一词作为“稳健”(robust)的面向数据的版本.

Tversky, A. & Kahneman, D. (1981), 'The framing of decisions and the psychology of choice', *Science* **211**, 453-458.

Ulam, S. (1957), 'Marian Smoluchowski and the theory of probabilities in physics', *Am. J. Phys.* **25**, 475-481.

Uspensky, J. V. (1937), *Introduction to Mathematical Probability*, McGraw-Hill, New York, p. 251.

Valavanis, S. (1959), *Econometrics*, McGraw-Hill, New York.

现代学生会发现这本书作为正统统计教学下的计量经济学是什么样子的书面记录很有用. 对于无偏估计量的不惜一切代价的需求可能导致作者丢弃数据中的几乎所有信息, 他只是不从信息内容的方面考虑.

van Dantzig, D. (1957), 'Statistical priesthood (Savage on personal probabilities)', *Statistica Neerlandica* **2**, 1-16.

今天很难理解贝叶斯学派如何为自己的观点而抗争的年轻读者, 应该阅读这篇对吉米·萨维奇作品的攻击文章. 但是人们应该意识到, 范丹齐格在这里并不孤单. 他的观点是 20 世纪五六十年代统计学家最常表达的观点.

Venn, John (1866), *The Logic of Chance*, MacMillan & Co., London; later edns, 1876, 1888.

从库尔诺和埃利斯在反拉普拉斯事业中停下来的地方重新开始. 杰恩斯 (Jaynes, 1986b) 给出了一些细节.

Ventris, M. (1988), 'Work notes on Minoan language research and other unedited papers', Sacconi, A., ed., *Edizioni dell'Ateneo*, Rome.

- Ventris, M. & Chadwick, J. (1956), *Documents in Mycenaean Greek*, Cambridge University Press.
- Vignaux, G. A. & Robertson, B. (1996), 'Lessons from the New Evidence Scholarship', in G. R. Heidbreder, ed., *Maximum Entropy and Bayesian Methods*, Proceedings of the 13th International Workshop, Santa Barbara, California, August 1-5, 1993, pp. 391-401, Kluwer Academic Publishers, Dordrecht, Holland.
- 贝叶斯法理学领域的综述, 有很多参考资料.
- von Mises, R. (1957), *Probability, Statistics and Truth*, G. Allen & Unwin, Ltd, London.
- von Mises, R. (1964), in Geiringer, H., ed., *Mathematical Theory of Probability and Statistics*, Academic Press, New York, pp. 160-166.
- von Neumann, J. & Morgenstern, O. (1953), *Theory of Games and Economic Behavior*, 2nd edn, Princeton University Press.
- [中译本] 冯·诺伊曼, 摩根斯坦. 博弈论与经济行为. 王建华, 顾玮琳, 译. 北京: 北京大学出版社, 2018.
- Wald, A. (1947), *Sequential Analysis*, Wiley, New York.
- 评论见 G. A. Barnard, *J. Am. Stat. Assoc.* **42**, 668 (1947).
- Wald, A. (1950), *Statistical Decision Functions*, Wiley, New York.
- [中译本] A. 瓦尔特. 统计决策函数. 王福保, 译. 魏宗舒, 校阅. 上海: 上海科学技术出版社, 1960.
- 沃尔德的最后一部著作, 他认识到贝叶斯方法的基础地位, 并将自己的最优方法称为“贝叶斯策略”.
- Wason, P. C. & Johnson-Laird, P. N. (1972), *Psychology of Reasoning*, Batsford, London.
- Weaver, W. (1963), *Lady Luck, the Theory of Probability*, Doubleday Anchor Books, Inc., Garden City, NY.
- Weber, B. H., Depew, D. J. & Smith, J. D., eds. (1988), *Entropy, Information, and Evolution*, MIT Press, Cambridge, MA.
- 在 1985 年举行的一次研讨会上发表的 16 篇论文的合集. 令人震惊地说明进化论由于试图用热力学第二定律来解释而会退化到什么状态——生物学家和哲学家对于“第二定律是什么意思”意见完全不一致且感到困惑. 本书第 7 章对此有简要的讨论.
- Weyl, H. (1946), *The Classical Groups*, Princeton University Press, NJ.
- Weyl, H. (1949), *Philosophy of Mathematics and Natural Science*, Helmer, O., trans., Princeton University Press.
- [中译本] 赫尔曼·外尔. 数学与自然科学之哲学. 齐民友, 译. 上海: 上海科技教育出版社, 2007.
- Weyl, H. (1961), *The Classical Groups; Their Invariants and Representations*, Princeton University Press.
- Whittaker, E. T. & Robinson, G. (1924), *The Calculus of Observations*, Blackie & Son, London.

由于第 349 页上的假“变星”数据被布卢姆菲尔德 (Bloomfield, 1976) 使用而有名。布卢姆菲尔德继续按照这本书作者的分析方法进行分析, 并得出荒谬的结论, 这是他的频谱分析教科书的核心内容。

Whittaker, E. T. & Watson, G. N. (1927), *A Course of Modern Analysis*, 4th edn, Cambridge University Press.

Whyte, A. J. (1980), *The Planet Pluto*, Pergamon Press, NY.

Widder, D. V. (1941), *The Laplace Transform*, Princeton University Press, Princeton, NJ.

Wiener, N. (1948), *Cybernetics*, J. Wiley & Sons, Inc., New York.

中译本 诺伯特·维纳. 控制论: 或动物与机器的控制和通信的科学. 王文浩, 译. 北京: 商务印书馆, 2020.

尽管我们不知道他有何工作实际上使用了贝叶斯方法, 但在第 109 页, 诺伯特·维纳表明自己是一个隐藏的贝叶斯主义者. 不过无论如何, 他对现实世界的概念性理解都太幼稚了, 因此无法成功. 在第 46 页, 他得到了地球-月球系统中潮汐力的反向作用 (地球加速, 月球减速). 第 61~62 页关于吉布斯工作的陈述纯属想象, 吉布斯并没有引入或假设遍历性, 甚至根本没有提到它. 今天很清楚的是, 从奇异吸引子、混沌等的发现来看, 几乎没有真正的系统是遍历的, 遍历无论如何与统计力学无关, 因为它在实际计算中没有任何差别. 就这一点而言, 吉布斯比其他人的理解领先一个世纪. 不幸的是, 维纳关于吉布斯的陈述被戈德曼 (S. Goldman, 1953) 和李 (Y. W. Lee, 1960) 等其他作者忠实地引用, 他们反过来又被其他人引用, 从而创造了一个庞大且仍在持续增长的“传说”. 维纳没有尽心校对这部著作, 许多方程只是看起来跟正确方程差不多.

Wiener, N. (1949), *Extrapolation, Interpolation, and Smoothing of Stationary time Series*, J. Wiley & Sons, Inc., New York.

另一部粗心和晦涩的杰作, 由莱文森在附录中部分破译, 在戈德曼和李的书中更完整.

Wigner, E. P. (1931), *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Fr. Vieweg, Braunschweig.

Wigner, E. P. (1959), *Group Theory*, Academic Press, Inc., New York.

Wing, G. M. (1962), *An Introduction To Transport Theory*, John Wiley and Sons, Inc., New York.

Woodward, P. M. (1953), *Probability and Information Theory, with Applications to Radar*, McGraw-Hill, NY.

一个有趣的历史文献, 它显示了对即将发生的事情的预见能力, 但不幸的是, 它缺乏使其实际有用所需的技术细节.

Wrinch, D. M. & Jeffreys, H. (1919), *Phil. Mag.* **38**, 715-734.

这是哈罗德·杰弗里斯关于概率论的第一篇出版文章, 与对拉普拉斯继承法则的修正有关. 他一定很喜欢这个结果或者这种关联, 因为在余生中, 他在每一个可能场合下都回过头来引用了这篇论文. 多萝西·林奇是一位出生于阿根廷的数学家, 曾就读于剑桥大学, 后来在美国史密斯学院任教, 用杰弗里斯的话来说, “成了一名生物学家”. 她的照片可以在波利亚的著作 (Pólya, 1987, 第 85 页) 中找到. 后来林奇和杰弗里斯关于同一主题的两篇

论文发表于 *Phil. Mag.* **42**, 369–390 (1921); **45**, 368–374 (1923).

Zabell, S. L. (1989), 'The Rule of Succession', *Erkenntnis*, **31**, 283–321.

对该主题漫长而复杂的历史进行综述, 包含大量意想不到的细节和数量惊人的参考资料, 强烈推荐. 他试图评估归纳法过去所受批评和现状的做法相对于波普尔是一种显著进步, 但在我们看来, 他仍然未能认识到归纳法在实际科学实践中是如何使用的. 本书第 9 章对此有讨论.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, J. Wiley & Sons, Inc., New York; 2nd edn, 1987, R. E. Krieger Pub. Co., Malabar, Florida.

中译本 阿诺德·泽尔纳. 计量经济学: 贝叶斯推断引论. 张尧庭, 译. 蒋传海, 沈根祥, 校. 上海: 上海财经大学出版社, 2005.

尽管标题中有“计量经济学”这个词, 但这项工作涉及一般原理, 对所有科学家和工程师都非常有价值. 它可以被视为杰弗里斯的著作 (Jeffreys, 1939, 第 3 版) 的续作, 在杰弗里斯当时所能达到的阶段之外进行了多元问题分析. 但是符号和风格是相同的, 专注于有用的分析材料而不是数学上的无关细节. 这本书包含对线性回归先验的更高水平的理解, 这在此后 20 多年的任何教科书中都是难以找到的.

Zellner, A. (1988), 'Optimal information processing and Bayes' theorem', *Am. Stat.* **42**, 278–284.

带讨论. 指出包含最大熵和贝叶斯算法作为解的一般变分原理的可能性. 在本书第 11 章中有讨论.

参考文献

- Akhiezer, N. I. (1965), *The Classical Moment Problem*, Hafner, New York.
- Archimedes (c 220 BC), in *Works*, T. L. Heath, ed., Cambridge University Press (1897, 1912).
- 平装版由 Dover Publications, Inc., New York 重印, 未注明日期 (约 1960 年).
- Aristotle (4th century BC), *Organon*.
- 三段论的定义.
- Aristotle (4th century BC), *Physics*; translation with commentary by Apostle, H. G., Indiana University Press, Bloomington (1969).
- Ash, B. B. (1966), *Information Theory*, John Wiley, New York.
- Bacon, F. (1620), 'Novum Organum', in Spedding, J., Ellis, R. L. & Heath, D. D., eds., *The Works of Francis Bacon*, vol. 4, Longman & Co., London (1857-1858).
- Barber, N. F. & Ursell, F. (1948), 'The generation and propagation of ocean waves and swell', *Phil. Trans. Roy. Soc. Lond.*, **A240**, 527-560.
- 噪声中哨声信号的检测.
- Barlow, E. R. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart & Winston, New York.
- Barndorf-Nielsen, O. (1978), *Information and Exponential Families in Statistical Theory*, J. Wiley & Sons, New York.
- Barr, A. & Feigenbaum, E., eds. (1981), *The Handbook of Artificial Intelligence*, 3 vols., Wm. Kaufman, Inc., Los Altos, CA.
- 中译本 A. 巴尔, E. A. 费根鲍姆. 人工智能手册 (第二卷). 钟玉琢, 夏莹, 苗玉峰, 译. 石纯一, 校. 北京: 科学出版社, 1988.
- 中译本 P. R. 科恩, E. A. 费根鲍姆. 人工智能手册 (第三卷). 周少柏, 黄汛, 译. 北京: 科学出版社, 1991.
- 来自 100 多位作者的贡献. 第一卷对搜索进行了综述, 这是 AI 为数不多可用于科学推断的方面之一.
- Barron, A. R. (1986), 'Entropy and the central limit theorem', *Ann. Prob.* **14**, 336-342.
- Bartholomew, D. J. (1965), 'A comparison of some Bayesian and frequentist inference', *Biometrika*, **52**, 19-35.
- Benford, F. (1938), 'The law of anomalous numbers', *Proc. Am. Phil. Soc.* **78**, 551-572.
- 本福德可能是沃伦·韦弗在其著作 (Warren Weaver, 1963, 第 270 页) 中也神秘地提到的那个人. 但是他们不知道的是, 西蒙·纽科姆 (Simon Newcomb, 1881) 很久以前就注意到了这种现象. 更多的细节和参考资料见赖米的文章 (Raimi, 1976).

- Berkson, J. (1977), 'My encounter with neo-Bayesianism', *Int. Stat. Rev.* **45**, 1-9.
- Berkson, J. (1980), 'Minimum chi-squared, not maximum likelihood!', *Ann. Stat.* **8**, 457-487.
- Bernardo, J. M. (1977) 'Inferences about the ratio of normal means: a Bayesian approach to the Fieller-Creasy problem', in Barra, J. D., *et al.*, eds., *Recent Developments in Statistics*, North Holland Press, Amsterdam, pp. 345-350.
- Bernado, J. M. (1979a), 'Reference posterior distributions for Bayesian inference', *J. Roy. Stat. Soc. B* **41**, 113-147.

带讨论

- Bernado, J. M. (1979b), 'Expected information as expected utility', *Ann. Stat.* **7**, 686-690.
- Bernado, J. M., de Groot, M. H., Lindley, D. V. & Smith, A. F. M., eds. (1980), *Bayesian Statistics*, Proceedings of the First Valencia International Meeting on Bayesian Statistics, Valencia, May 28-June 2, 1979, University Press, Valencia, Spain.
- Bernado, J. M., de Groot, M. H. & Lindley, D. V., eds. (1985), *Bayesian Statistics 2*, Proceedings of the Second Valencia International Meeting on Bayesian Statistics, September 6-10, 1983, Elsevier Science Publishers, New York.
- Billingsley, P. (1979), *Probability and Measure*, Wiley, New York.
- 包含更多我们没有深入讨论的有关博雷尔-柯尔莫哥洛夫的内容。

Bishop, Y., Fienberg, S., & Holland, P. (1975), *Discrete Multivariate Analysis*, MIT Press, Cambridge, MA.

中译本 Yvonne M. M. Bishop, Setphen E. Fienberg, Paul W. Holland. 离散多元分析: 理论与实践. 张尧庭, 译. 史宁中, 校. 北京: 中国统计出版社, 1998.

Blanc-Lapierre, A. & Fortet, R. (1953), *Theorie des Fonctions Aleatoires*, Masson et Cie, Paris.

Boole, G. (1916), *Collected Logical Works, Vol. 1: Studies in Logic and Probability; Vol II: An Investigation of the Laws of Thought*, Open Court, Chicago.

Borel, E. (1926), *Traité du Calcul des Probabilités*, Gauthier-Villars, Paris.

球体上全等集合的豪斯多夫悖论在第 2 卷第 1 册中讨论。

Born, M. (1964), *Natural Philosophy of Cause and Chance*, Dover, New York.

中译本 M. 玻恩. 关于因果和机遇的自然哲学. 侯德彭, 译. 北京: 商务印书馆, 1964.

Boscovich, Roger J. (1770), *Voyage Astronomique et Geographique*, N. M. Tillard, Paris.

通过校正和为 0 幅度和最小的标准调整数据。

Box, G. E. P. (1982), 'An apology for ecumenism in statistics', NRC Technical Report #2408, Mathematics Research Center, University of Wisconsin, Madison.

Box, G. E. P., Leonard, T. & Wu, C-F, eds. (1983), *Scientific Inference, Data Analysis, and Robustness*, Academic Press, Inc., Orlando, FL.

1981 年 11 月在威斯康星州麦迪逊举行的会议上的论文集。

Bracewell, R. N. (1986), 'Simulating the sunspot cycle', *Nature*, **323**, 516.