概率论札记

阚成章 张远航 编辑

根据巩馥洲老师《概率论》课程整理而成

2017 年 3 月 20 日

目录

本书所用记号 **1**

第 1 章 引言 **2**

1.1 研究对象 2

1.2 组合分析与计数基本原理 2

1.3 组合恒等式 2

1.4 方程的整数解个数 2

第 2 章 概率论公理 **3**

2.1 从人工智能谈起 3

2.2 概率性事件与 Knight 不确定性 3

2.3 测度论与 Kolmogorov 概率论公理体系 3

2.4 古典概型 3

2.5 样本空间与维数论 3

2.6 概率论与现代数学体系：遍历论、代数几何与理论物理 3

2.7 贝叶斯概率公理体系 3

2.7.1 概率：确信程度的度量 3

2.7.2 可信度推理及其基本公设（原理） 3

2.7.3 可信度推理的定量化 4

参考文献 **6**

术语 **7**

本书所用记号

To be completed

第 1 章 引言

1.1 研究对象

1.2 组合分析与计数基本原理

1.3 组合恒等式

1.4 方程的整数解个数

第 2 章　概率论公理

2.1　从人工智能谈起

2.2　概率性事件与 Knight 不确定性

2.3　测度论与 Kolmogorov 概率论公理体系

2.4　古典概型

2.5　样本空间与维数论

2.6　概率论与现代数学体系：遍历论、代数几何与理论物理

2.7　贝叶斯概率公理体系

2.7.1　概率：确信程度的度量

- 贝叶斯公理体系 (1) $\sim$ (3)
- 演绎推理、布尔代数、数理逻辑与多项式代数系统
- 合情推理作为布尔逻辑的推广

2.7.2　可信度推理及其基本公设（原理）

- 作为布尔值 0/1 的推广，命题的**可信度**（plausibility）用实数（一般是区间 $[0,1]$ 上的值）来表示。此处可引申对比参看**模糊数学**（fuzzy mathematics）体系。
- **可信度推理**（plausible reasoning）无处不在。
- 之所以用“公设”，措辞解释：“公理”之间须有独立性，而“公设”不严格独立。

(1) 条件概率公式。Keynes 在完整阐述其经济学理论之前，曾著作《概率论》(A Treatise on Probability)，引入 $A \mid B$ 来表示给定命题 B 正确时，命题 A 的可信度，或 B 事件发生的情况下，A 发生的可能性。当然，在不引起混淆的情况下，我们也会以 $A \mid B$ 来指称相对应的事件。

现若有 $A \mid B > A \mid C$, 则对任意的 D：

(1)-1 以事件 A/D 为研究对象，$D \mid AB = D \mid AC$；

理解：前提可比较即可。在结果已经出现的时候，可信度增益已无意义。

(1)-2 $AD \mid B > AD \mid C$；

(1)-3 $A^{-1} \mid B < A^{-1} \mid C$。

(2) 可信性推理规则应符合常识，即可信性推理为**常识推理**（commonsense reasoning）。

(3) 可信度推理要有 **相容性**（consistency）。

(3)-1 若结论是由两个以上方式得到的，则对任何方式来说，其结论导致的可信度是相同的。简单来讲，即殊途同归。此处我们可以联想到前面提到的流程图方法，不同的流程描述同样的事件，其结果应当相同。

(3)-2 可信度的确定应该基于掌握的所有事实，不可遗漏。

(3)-3 对相同的命题，若基于相同的事实，其可信度相同。即表面上陈述方式不同的命题可信度可能相同，命题整体的可信度只取决于命题词汇所描述的事实本身。

2.7.3 可信度推理的定量化

乘法规则

乘法规则（product rule）确定了 $AB \mid C$ 满足的规则。让我们反思一下我们如何考虑 AB 的可信度：

第一，不失一般性，我们可以先考虑命题“若 AB 正确，则 B 正确”的可信度，即 AB 的可信度如何与 B 自己的可信度联系起来。进一步，AB 的可信度取决于“若 B 正确，则 A 也正确”的可信度。也即，$AB \mid C$ 的可信度取决于 $B \mid C$ 和 $A \mid BC$ 的可信度。

第二，显然的事实是，B 错误时，AB 一定错误。也即，当我们确定 $B \mid C$ 和 $A \mid BC$ 的可信度时，$AB \mid C$ 的可信度与 $AB \mid C$ 和 $A \mid C$ 的可信度无关。

第三，在 C 正确的情况下，$A \mid B$ 与 $B \mid A$ 的可信度对 $AB \mid C$ 的可信性无贡献。

第四，由公设 (3)-3，AB 与 BA 的可信度相同。

以上的论证根据可信度推理的基本公设得出，更完备详细的版本可参看 [2, Jaynes 2003]。总结来讲，我们可以给出公式：

$$AB \mid C = F(B \mid C, A \mid BC)$$

其中 F 是二元连续递增函数。我们进一步根据可信度推理的基本公设，推导 F 的具体表达形式。（在这里，得益于著名数学家 Abel 在 1862 年的工作，我们可以假设 F 是可微的，从而简化在这里的论述，而结果不丧失丝毫的普遍性。）

注意结合律 $ABC = A(BC) = (AB)C$，我们有：

$$\begin{aligned} ABC \mid D &= F(BC \mid D, A \mid BCD), \\ ABC \mid D &= F(C \mid D, AB \mid CD), \end{aligned}$$

同时又有：

$$\begin{aligned} BC \mid D &= F(C \mid D, B \mid CD), \\ AB \mid CD &= F(B \mid CD, A \mid BCD), \end{aligned}$$

上述四个公式代入到一起得到结合性方程：

$$F(F(C \mid D, B \mid CD), A \mid BCD) = F(C \mid D, F(B \mid CD, A \mid BCD)).$$

进行美化公式的变量替换即：

$$F(F(X,Y),Z)=F(X,F(Y,Z)).$$

我们放弃该方程的常值平凡解，基于函数方程及可微性，（过程参看 Aczél 在 1966 年给出的证明）得出其非平凡解为：

$$F(X,Y)=w^{-1}(w(X)w(Y)),$$

其中 w 为实数轴 $\mathbb{R}$ 到区间 $[0,1]$ 的双射。因此我们得到了与直觉相符的乘法规则：

$$AB\mid C=A\mid BC\times B\mid C=A\mid C\times B\mid AC$$

加法规则

事实：A 与其对立事件 A' 的可信度互相确定。故有：

$$A'\mid B=S(A\mid B),$$

其中 $S:[0,1]\to[0,1]$ 为连续减函数。

$$AB\mid C=A\mid BC\times B\mid AC=A\mid BC\times S(B'\mid AC)=A\mid C\times S\left(\frac{AB'\mid C}{A\mid C}\right)=B\mid C\times\left(\frac{BA'\mid C}{B\mid C}\right).$$

令 $B=AD$，则 AB=B，取特殊情况以得方程的解，由特殊到一般……

再次结合乘法律，令 $X=A\mid B$，我们最终得到：

$$S(X)=(1-X^M)^{\frac{1}{M}}.$$

总结

定义 $P=X^M$，最终我们得到：

$$P(A'\mid B)+P(A\mid B)=1,\quad 0\le P(A\mid B)\le 1, \tag{2.1}$$

乘法规则：

$$P(AB\mid C)=P(A\mid C)P(B\mid AC)=P(B\mid C)P(A\mid BC), \tag{2.2}$$

以上两式合并得到可加性，取代第一个公式：

$$P(A+B\mid C)=P(A\mid C)+P(B\mid C)-P(AB\mid C). \tag{2.3}$$

以上的三个公式与贝叶斯体系下的三条公理不谋而合。也即，我们从最基本的可信度推理公设，得出了全套贝叶斯概率论公理体系。这些公式本质上是可信度推理的数学严格化，正如布尔逻辑是演绎推理的严格化一样。

以后我们会看到，贝叶斯公理体系与 Kolmogorov 测度论下的概率论体系是等价的。但显然，前者所揭示的逻辑意义要宽泛而深刻得多。这就决定了贝叶斯主义带来的丰富的方法论指导。

参考文献

[1] Ross, Sheldon M. *A First Course in Probability*. Pearson Higher Education, 2014.

[2] E. T. Jaynes, *Probability: The Logic of Science*. Cambridge University Press, 2003.

术语

常识推理 commonsense reasoning 4, 6

相容性 consistency 4, 6

模糊数学 fuzzy mathematics 3, 6

可信度 plausibility 3, 6

可信度推理 plausible reasoning 3–6

乘法规则 product rule 4–6

加法规则 sum rule 5, 6