

标题 title

作者 *author*

2023 年 8 月 1 日

前言

目录

前言	i
第一部分 科学的逻辑	1
第一章 合情推理	2
§1.1 回顾：命题逻辑的演绎推理	2
§1.2 合情推理的数学模型	4
1.2.1 似然，合情推理的原则	4
1.2.2 似然与概率	6
§1.3 合情推理的归纳强论证	8
1.3.1 先验与基率谬误	8
1.3.2 归纳强论证	9
1.3.3 有效论证和归纳强论证的比较	12
第二章 Markov 链与决策	22
§2.1 Markov 链	22
§2.2 Markov 奖励过程 (MRP)	26
§2.3 Markov 决策过程 (MDP)	29
§2.4 隐 Markov 模型 (HMM)	33
2.4.1 评估问题	34
2.4.2 解释问题	35
第二部分 信息与数据	37
第三章 信息论基础	43

§3.1 熵	43
3.1.1 概念的导出	43
3.1.2 概念与性质	46
3.1.3 熵与通信理论	51
§3.2 Kullback-Leibler 散度	54
3.2.1 定义	54
3.2.2 两个关于信息的不等式	56
3.2.3 在机器学习中的应用：语言生成模型	57
§3.3 附录：Shannon 定理的证明	58
§3.4 习题	59
§3.5 章末注记	61
第四章 Johnson-Lindenstrauss 引理	63
§4.1 机器学习中的数据	63
§4.2 矩法与集中不等式	64
§4.3 J-L 引理的陈述与证明	68
§4.4 J-L 引理的应用	72
§4.5 习题	73
§4.6 章末注记	73
第五章 差分隐私	74
§5.1 数据隐私问题	74
§5.2 差分隐私的定义与性质	76
§5.3 差分隐私的应用	80
5.3.1 随机反应算法	80
5.3.2 全局灵敏度与 Laplace 机制	81
5.3.3 DP 版本 Llyod 算法	83
§5.4 差分隐私与信息论	84
§5.5 习题	85
§5.6 章末注记	85
第三部分 决策与优化	86
第六章 凸分析	87

§6.1 决策与优化的基本原理	87
6.1.1 统计决策理论	87
6.1.2 优化问题	88
6.1.3 例子：网格搜索算法	91
§6.2 凸函数	93
§6.3 凸集	96
6.3.1 基本定义和性质	96
6.3.2 分离超平面定理	98
第七章 对偶理论	100
§7.1 条件极值与 Lagrange 乘子法	101
§7.2 Karush-Kuhn-Tucker 条件	104
§7.3 Lagrange 对偶	107
7.3.1 Lagrange 定理	107
7.3.2 弱对偶定理，强对偶定理	111
§7.4 应用：支持向量机 (SVM)	115
第八章 不动点理论	118
§8.1 Banach 不动点定理	118
§8.2 Brouwer 不动点定理	121
§8.3 不动点的一般视角	124
第四部分 逻辑与博弈	125
第九章 动态博弈	126
§9.1 输赢博弈	126
§9.2 随机博弈 (Markov 博弈)	131
第十章 静态博弈	137
§10.1 正则形式博弈	137
10.1.1 生成对抗网络	138
10.1.2 混合策略	140
§10.2 不完全信息博弈 (Bayes 博弈)	141

第五部分 认知逻辑	146
第十一章 模态逻辑基础	147
§11.1 模态逻辑的起源	147
11.1.1 三段论	147
11.1.2 非经典逻辑	148
§11.2 模态语言	149
§11.3 Kripke 语义与框架语义	152
§11.4 模态可定义性	157
第十二章 认知逻辑与共同知识	165
§12.1 “泥泞的孩童”谜题	165
§12.2 认知逻辑的基本模型与性质	170
§12.3 对不一致达成一致	180
§12.4 Rubinstein 电子邮件博弈	186

第一部分

科学的逻辑

第一章 合情推理

§1.1 回顾：命题逻辑的演绎推理

命题逻辑的命题公式由如下定义递归形成

- 命题变元 p, q, r, \dots 是命题公式.
- 如果 ϕ 和 ψ 是命题公式, 那么 $(\neg\phi)$, $(\phi \vee \psi)$, $(\phi \wedge \psi)$, $(\phi \leftrightarrow \psi)$ 和 $(\phi \rightarrow \psi)$ 都是命题公式.

例如, $(p \vee (q \rightarrow r))$ 是命题公式, 但是 $p \vee \vee q$ 不是.

$\neg, \vee, \wedge, \leftrightarrow, \rightarrow$ 被称为连接词, \neg 被称为否定, \vee 被称为析取, \wedge 被称为合取, \leftrightarrow 被称为双向蕴含, \rightarrow 被称为蕴含. 在不产生混淆的时候也会省略括号. $A \wedge B$ 也常写作 AB , $\neg A$ 也常写作 \bar{A} .

命题逻辑最重要的问题是: 什么样的公式是真的? 给定一个公式集 Γ , 对一个公式 ϕ , 我们有两种真的概念:

- 语义: $\Gamma \models \phi$.
- 语型: $\Gamma \vdash \phi$.

首先看语义. 每一个命题变元可以赋值真假: \top (真) 或 \perp (假). 对于一般公式, 可以利用真值表递归地定义公式的真假赋值. 例如, $p \rightarrow q$ 的真值表为

p	q	$p \rightarrow q$
\top	\top	\top
\top	\perp	\perp
\perp	\top	\top
\perp	\perp	\top

符号 $\Gamma \models \phi$ 意味着对任意赋值，只要 Γ 全为真， ϕ 就为真。

然后再看语型。推导规则（即语法）描述了从一些公式出发如何得到另外一些公式。语法推导的形式是

$$\frac{\phi \quad \phi \rightarrow \psi}{\psi}$$

横线上方的称为前提，横线下方的称为结论。上面的推导法则被称为肯定前件（MP），是三段论的基础。

语法推导可以引入新的连接词，例如

$$\frac{\phi \quad \psi}{\phi \wedge \psi}$$

也可以消除连接词，例如

$$\frac{\phi \wedge \psi}{\phi}$$

在给定一组推导法则下，我们就可以根据规则来推演命题，这就是**演绎推理**。

例如，将归谬法（RAA）加入推导法则：将 $\neg\phi$ 作为前提，得到了结论 \perp ，那么可以推出 ϕ 才是结论。写作

$$\frac{\begin{array}{c} [\neg\phi] \\ \vdots \\ \perp \end{array}}{\phi}$$

方括号表示假设 $[\neg\phi]$ 是前提，省略号表示推导的中间步骤。我们就可以得到反证法。

记号 $\Gamma \vdash \phi$ 的意思是以 Γ 作为前提，依据推导法则，可以得到 ϕ 作为结论。从前提出发，按照法则，得出结论的过程，即称谓演绎推理。命题逻辑的主要定理是

定理 1.1 (完备性定理) 对任意公式集 Γ 和任意公式 ϕ ，

$$\Gamma \models \phi \iff \Gamma \vdash \phi.$$

推论：检验一个演绎推理的正确性可以用真值表完成。

如果 $\models \phi$ ，那么称 ϕ 为重言式。重言式即是不需要加任何假设也一定成立的公式，表明了这一推理逻辑所包含的“正确的废话”。如果 $\psi \leftrightarrow \phi$ 是重言式，我们就说 ψ, ϕ 是等值的。等值的公式在演绎推理中起着相同的作用。例如： $p \rightarrow q$ 与 $\neg q \rightarrow \neg p$ 是等值的，所以证明一个定理也可以去证明它的逆否命题。

§1.2 合情推理的数学模型

现在我们从演绎推理过渡到合情推理。生活中的推理并不总是演绎推理。走在北大校园里，看到有人突然倒地不起，一个合理的第一反应是，他突发疾病了。在常识的意义下，这是一种合理的推理。我们来仔细分析一下这个推理的过程。

演绎推理包含两个强三段论：

$$\begin{array}{cc} A \rightarrow B & A \rightarrow B \\ \hline A \text{ 是真的} & B \text{ 是假的} \\ \hline B \text{ 是真的} & A \text{ 是假的} \end{array}$$

具体到上面的例子， A 是“突发疾病”， B 是“突然倒地不起”。因此演绎推理应该是：因为突发疾病，所以突然倒地不起。

然而，我们实际上作出的推理形如：因为突然倒地不起，所以更可能突发疾病。这是一个弱三段论：，数学上写作

$$\begin{array}{c} A \rightarrow B \\ B \text{ 是真的} \\ \hline A \text{ 变得更合理} \end{array}$$

虽然我们不能得知 A 的真假，但是根据已有的事实 B ，我们可以得到 A 的合理性。这样没有严格真假的推理模式就是合情推理。

再次考虑同样的问题，走在北大校园里，看到有人突然倒地不起，你的第一反应是？

- 选项 1：他突发疾病了。
- 选项 2：他接到情报有一枚来自俄罗斯的导弹意外偏航，马上要落在北京，他正在躲避。

虽然选项 1 和选项 2 都是可能的，但是他们的合理程度并不同。这里产生了一个自然的问题：如何刻画不同推理的合理程度？为此，我们引入似然的概念。

1.2.1 似然，合情推理的原则

用符号 $A|B$ 表示当 B 已知（或假设）为真时，命题 A 合理程度的非负实数度量。如果 $A|B$ 是一个度量，那么任意非负函数 $f(A|B)$ 也会是一个度量，因此 $A|B$ 还不是一个唯一的数学概念。因此，我们现在称 $A|B$ 为以 B 为假设时命题 A 的似然。似然不能定义在任意两个命题之间，只有有合理的关联的命题才能定义似然。

合情推理的原则实际表明了似然之间的关系。给定相同的假设，似然应该保持命题逻辑中的等值命题。例如：

- $(A \vee B)|C = (B \vee A)|C.$
- $(AB)C|D = A(BC)|D.$
- $\neg\neg A|B = A|B.$

考虑从 C 出发推出 AB 为真. 一种可能的推理方法是先推断出 A 为真, 然后在 A 为真的前提下, 推断出 B 为真. 所以似然 $AB|C$ 应该取决于这两段推理的似然, 即 $A|C$ 和 $B|AC$, 因此存在函数 F 满足

$$AB|C = F(A|C, B|AC).$$

然而, 这并不能写成 $AB|C = F(B|C, A|C)$, 考虑 A 是指小明的左眼是蓝色的, B 是指小明的右眼是棕色的. $A|C$ 和 $B|C$ 都可以很合理, 但 $AB|C$ 则几乎不可能发生.

我们定义的 F 具有如下性质:

定理 1.2 如果 F 是连续函数, 那么 F 充分必要地满足

$$cf(F(p, q)) = f(p)f(q),$$

这里, c 是某个常数, f 是某个函数. 因此,

$$cf(AB|C) = f(A|C)f(B|AC).$$

我们之前说过, 如果 $A|B$ 是一个合理性度量, 那么 $f(A|B)$ 也是, 因此, 遵循惯例, 不妨就考虑似然 $A|B$ 本身, 并且取 $c = 1$. 因此

$$AB|C = (A|C)(B|AC).$$

合情推理的与规则表述如下:

规则 1.1 (与规则) 对命题 A, B, C , 成立

$$AB|C = (A|C)(B|AC).$$

第二个考虑的是似然 $A|B$ 和 $\neg A|B$ 的关系. 两个似然应该存在某种函数关系, 即存在函数 S 使得

$$\neg A|B = S(A|B).$$

实际上, S 具有如下性质:

定理 1.3

$$S(x) = (1 - x^m)^{1/m},$$

这里的 m 是一个正常数.

由上面的定理, 我们可以得到以下性质

推论 1.1 $(A|B)^m + (\neg A|B)^m = 1$.

证明 $(\neg A|B) = S(A|B) = (1 - (A|B)^m)^{1/m}$. □

注意到 $(A|B)^m$ 也是一个满足与规则的似然, 所以我们不妨就取似然为 $(A|B)^m$, 于是有

规则 1.2 (否定规则) 对命题 A, B , 成立

$$(A|B) + (\neg A|B) = 1.$$

连接词 $\{\wedge, \neg\}$ 组成了演绎逻辑的完备集: 所有真值函数都可以用这个集合的词来表示. 回忆, 我们要求合情推理保持演绎逻辑的等值公式. 与规则对应 \wedge , 否定规则对应 \neg . 因此合情推理的两个规则是完备的: 任何命题都可以通过这两个公式计算出似然.

例 1.1 例如, 计算 $A \vee B|C$,

$$\begin{aligned} A \vee B|C &= 1 - (\overline{AB}|C) \\ &= 1 - (\overline{A}|C)(\overline{B}|\overline{AC}) \\ &= 1 - (\overline{A}|C)(1 - (B|\overline{AC})) \\ &= (A|C) + (\overline{AB}|C) \\ &= (A|C) + (B|C)(\overline{A}|BC) \\ &= (A|C) + (B|C)(1 - (A|BC)) \\ &= (A|C) + (B|C) - (AB|C). \end{aligned}$$

1.2.2 似然与概率

我们已经知道, 似然是一个合理性度量, 在本节中, 我们将似然与概率联系起来: 他们实际上是等价的.

首先介绍 *Kolmogorov* 的概率论。这一概率论研究事件空间上的概率测度。其研究对象的总体被称为样本空间，记为 Ω 。我们只能观测到某种可观测的特性 P ，而不能直接观测样本点，即我们只能观察事件，或者说集合

$$\{\omega \in \Omega : P(\omega)\}.$$

我们可以观测的所有事件的集合称为事件域，记为 \mathcal{F} 。

事件域 \mathcal{F} 中的事件之间互相有关联。我们自然可以观测到 Ω ，因此 $\Omega \in \mathcal{F}$ 。如果我们观测到事件 A ，那么我们也可以通过没有观测到 A 来判断观测到了 $\Omega \setminus A$ 。因此， $A \in \mathcal{F} \implies \Omega \setminus A \in \mathcal{F}$ 。如果我们观测到了 A 或者 B ，我们其实也观测到了 $A \cup B$ ，即 $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$ 。

事实上，我们可以形式上定义概率：

定义 1.1 (概率) 概率 \Pr 是一个事件域 \mathcal{F} 到实数的映射，并且满足：

- 规范性： $\Pr(\Omega) = 1$ 。
- 非负性： $\forall A \in \mathcal{F} \Pr(A) \in [0, 1]$ 。
- 可列可加性：对 A_1, A_2, \dots 满足 $A_i \cap A_j = \emptyset, i \neq j$ ，有

$$\Pr\left(\bigcup_i A_i\right) = \sum_i \Pr(A_i).$$

事件是命题的集合论描述。具体来说，有如下对应

事件	命题
Ω	\top
\emptyset	\perp
$A \cap B$	$A \wedge B$
$A \cup B$	$A \vee B$
$A \subseteq B$	$A \rightarrow B$
$A = B$	$A \leftrightarrow B$

回忆条件概率的定义，我们可以得到链式法则：

$$\Pr(AB|C) = \frac{\Pr(ABC)}{\Pr(C)} = \frac{\Pr(B|AC) \Pr(AC)}{\Pr(C)} = \Pr(B|AC) \Pr(A|C).$$

回忆补事件公式： $\Pr(A|C) + \Pr(\bar{A}|C) = \Pr(\Omega|C) = 1$ 。这两个公式恰好对应了合情推理的与规则以及否定规则！

公理化概率论	合情推理
事件	命题
(条件) 概率	似然
链式法则	与规则
补事件公式	否定规则

表 1.1: 概率论与合情推理的对应

这并不是巧合. 可以证明, 合情推理与事件域的公理化概率具有一一对应的关系, 见表 1.1.

从这个意义上说, 概率是似然唯一的数学模型! 从此, 我们将似然 $A|B$ 定义为概率 $\Pr(A|B)$.

§1.3 合情推理的归纳强论证

1.3.1 先验与基率谬误

在前面, 我们有意模糊了条件概率和概率在合情推理中的区别. 然而, 这样的区别是非常重要的. 在合情推理中, 非条件的概率被称为先验概率或者基率, 它表示了对这个命题合理程度的一种无条件的信念. 对应地, 条件概率就是后验概率或似然, 它表示了对合情推理合理程度的一种信念. 先验概率和后验概率有若干相互转化的公式.

回忆, 条件概率有全概率公式:

定理 1.4 (全概率公式) 设 A_i 是彼此互斥的事件, $\cup_i A_i = \Omega$, 那么

$$\Pr(B) = \sum_i \Pr(B|A_i) \Pr(A_i).$$

全概率公式表明了如何使用似然建立起不同先验概率之间的联系.

回忆, 条件概率还有 Bayes 定理:

定理 1.5 (Bayes 定理)

$$\Pr(A|B) = \Pr(B|A) \frac{\Pr(A)}{\Pr(B)}.$$

Bayes 定理表明了两个不同的后验概率如何基于先验概率相互转化. 在合情推理中, 这表明了前提推结果的强三段论和结果推前提合理性的弱三段论之间的关系.

下面我们看一个合情推理的例子。

一辆出租车在夜间发生了一起肇事逃逸事故。这座城市有两家出租车公司，绿色和蓝色。这个城市历史上肇事逃逸车辆 85% 是绿色的，15% 是蓝色的。一名目击者指认出租车是蓝色的，这里的指认并不一定正确。考虑一种理想的假设，法庭知道这位证人 80% 的概率能正确识别颜色，20% 的概率会把颜色识别错。问：事故车辆是那种颜色的可能性更大？忽略先验概率会产生答案是蓝车的结论。基率谬误指因为忽略先验概率（即基率）而产生的错误判断。

下面我们考虑先验概率再次做计算。记 B 为肇事逃逸的出租车为蓝色， G 为肇事逃逸的出租车为绿色， R 为目击者指认出租车是蓝色。先验概率为 $\Pr(B) = 0.15$ ， $\Pr(G) = 0.85$ 。似然为 $\Pr(R|B) = 0.8$ ， $\Pr(R|G) = 0.2$ 。利用全概率公式， $\Pr(R) = \Pr(B)\Pr(R|B) + \Pr(G)\Pr(R|G) = 0.29$ 。利用 Bayes 定理， $\Pr(B|R) = \Pr(R|B)\Pr(B)/\Pr(R) \approx 0.41$ 。然而， $\Pr(G|R) = \Pr(R|G)\Pr(G)/\Pr(R) \approx 0.59$ 。因此我们更应该倾向于认为肇事逃逸的出租车是绿色的！

1.3.2 归纳强论证

在基率谬误的例子中，我们看到合情推理必须要完整地考虑先验的影响。另一方面，我们看到最终做出决策的方式是最大似然，即似然更高的那个命题更有可能是对的。这说明合情推理中有很不同于演绎推理的论证方式。

我们首先指出合情推理包含了命题逻辑中的两个强三段论：

$$\begin{array}{cc} A \rightarrow B & A \rightarrow B \\ \frac{A \text{ 是真的}}{B \text{ 是真的}} & \frac{B \text{ 是假的}}{A \text{ 是假的}} \end{array}$$

我们以第一个三段论为例，记 $C \equiv A \rightarrow B$ 。由链式法则， $\Pr(B|AC) = \Pr(AB|C)/\Pr(A|C)$ 。 $A \rightarrow B$ 意味着作为事件 $A \subseteq B$ ，即 $AB = A$ ， $\Pr(AB|C) = \Pr(A|C)$ 。代入上式得到 $\Pr(B|AC) = 1$ ，这就是说，当 A 为真时， B 也为真。

除了演绎逻辑中的强三段论，合情推理还包含了弱三段论的定量形式：

$$\begin{array}{c} A \rightarrow B \\ \frac{B \text{ 是真的}}{A \text{ 变得更合理}} \end{array}$$

$\Pr(A|C)$ 是 A 的似然，而 $\Pr(A|BC)$ 是假设 B 为真时， A 的似然。由链式法则， $\Pr(A|BC) = \Pr(A|C)\Pr(B|AC)/\Pr(B|C)$ 。因为 $\Pr(B|AC) = 1$ 且 $\Pr(B|C) \leq 1$ ，所以 $\Pr(A|BC) \geq \Pr(A|C)$ 。也就是当 B 为真时， A 的合理程度会变大。

仿照三段论，现在我们将演绎推理中的若干概念推广到合情推理中。

回忆记号 $X \vdash Y$ 表示从前提为 X 出发可以跟据推导规则得到结论 Y 。此时，我们说从 X 到 Y 的过程是一个有效论证，或者说 Y 是 X 的逻辑结论。它与以下三个表述等价：

- $X \models Y$.
- $X \rightarrow Y$ 是重言式.
- $X \wedge \neg Y$ 是矛盾式.

等价性证明由蕴含的推导法则（或演绎定理）和完备性定理得出。那么，如何在合情推理中定义类似的概念呢？为此，我们引入随机真值表。

回忆，事件是命题的集合论描述。合情推理中，每个事件被赋予一个概率（似然），对应的命题也会被赋予同样的概率。于是对应于演绎推理中的语义真值表，合情推理中有随机真值表，例如

Pr	A	B	$A \vee B$
0.4	⊤	⊤	⊤
0.2	⊤	⊥	⊤
0.25	⊥	⊤	⊤
0.15	⊥	⊥	⊥

在合情推理中，我们也有和有效论证对应的归纳强论证。考虑如下推理：

$$\frac{X}{Y}$$

利用随机真值表，我们可以尝试定义归纳强论证为 $X \wedge \neg Y$ 的不太可能为真（或 $\neg X \vee Y$ 很可能为真）。然而我们会看到，仅仅用随机真值表得到的概念是不符合合情推理的直觉的。我们通过两个例子来引入归纳强论证的限制条件。

例 1.2 (奇怪的例子一) 记 X 为一个北京大学的同学今年 2000 岁, Y 为一个北京大学的同学今年 2000 岁, 并且有三条胳膊。直观来讲, 如上 $X \vdash Y$ 不是归纳强论证。但是, 其等效的公式 $\neg X \vee Y$, 成立的概率足够大。这个悖论却为这个应该不成立的结论给了一个归纳强论证。所以, 从这个角度来看, 判断是否为归纳强论证不能只关注结论成立的概率。

从这个例子出发, 我们得到限制条件一: 证据支持, 它的定义如下。假设 X 和 Y 是公式, $t \in [0.5, 1]$ 。如果 $\Pr(Y|X) > t$, 我们称: X 支持 Y 的证据强度大于 t 。证据支持是比最大似然准则的更进一步的要求。

然而, 证据支持并不能解决所有问题。考虑下面的例子。

例 1.3 (奇怪的例子二) 记 X 为小明是一位北京大学的学生, Y 为小明不会飞. 表面看来, $\Pr(Y|X) = 1$, 但我们知道 $X \models Y$ 并不应该是归纳强论证. 问题出在了 $\Pr(Y)$ 本身就等于 1, 所以 $\Pr(Y|X) = 1$ 并没有什么实际意义.

从这个例子出发, 我们得到限制条件二: 正相关性, 它的定义如下. 我们称 X 与 Y 正相关, 如果 $\Pr(Y|X) > \Pr(Y)$. 等价地, 示性函数 $I(X)$ 和 $I(Y)$ 相关系数大于 0. 类似地, 如果 $\Pr(Y|X) < \Pr(Y)$ (或 $I(X)$ 和 $I(Y)$ 的相关系数小于 0), 那么 X 和 Y 负相关. 如果 $\Pr(Y|X) = \Pr(Y)$ (或 $I(X)$ 和 $I(Y)$ 的相关系数等于 0), 那么 X 和 Y 不相关. 在归纳强论证中, 我们要求 X 和 Y 正相关.

加上以上两个限制条件, 我们可以得到归纳强论证的严格定义.

定义 1.2 (归纳强论证) 如果 $X \models Y$ 满足以下三个条件, 我们称之为归纳强论证:

- X 证据支持 Y : $\Pr(Y|X) > 0.5$.
- X 与 Y 正相关: $\Pr(Y|X) > \Pr(Y)$.
- $X \rightarrow Y$ 不是有效论证.

不将有效论证定义为归纳强论证得原因之一是: 一个论证可以在前提 X 是矛盾式时成为有效. 例如: $P \wedge \neg P \models Q$ 是有效论证. 但是, 由于 $\Pr(P \wedge \neg P) = 0$, $\Pr(Q|P \wedge \neg P)$ 是无定义的, 所以 $P \wedge \neg P$ 并不证据支持 Q , 也不和 Q 正相关.

进一步, 我们还希望能够衡量前提 X 在多大程度上确认结论 Y 成立, 我们可以通过如下两个量来衡量:

认可度 认可概率增量衡量事件 X 发生后给事件 Y 的发生增加了多大的概率.

$$d(X, Y) = \Pr(Y|X) - \Pr(Y).$$

认可度 认可度似然比衡量假设 Y 发生时 X 的似然会比假设 Y 没发生时 X 的似然增加多少.

该差值越大表示观测到 X 的话越应该发生了 Y . 分母归一化使得 $\ell(X, Y) \in [-1, 1]$.

$$\ell(X, Y) = \frac{\Pr(X|Y) - \Pr(X|\neg Y)}{\Pr(X|Y) + \Pr(X|\neg Y)}$$

认可度和相关性的关系可以表达如下:

命题 1.1 设 $0 < \Pr(X), \Pr(Y) < 1$, 下列等价成立:

- X 和 Y 正相关 $\iff d(X, Y) > 0 \iff \ell(X, Y) > 0$.

• X 和 Y 不相关 $\iff d(X, Y) = 0 \iff \ell(X, Y) = 0$.

• X 和 Y 负相关 $\iff d(X, Y) < 0 \iff \ell(X, Y) < 0$.

认可度与有效论证的关系可以表达如下:

命题 1.2 设 $0 < \Pr(X), \Pr(Y) < 1$, 那么

$$d(X, Y) = \begin{cases} \Pr(\neg Y), & \text{如果 } X \models Y, \\ -\Pr(Y), & \text{如果 } X \models \neg Y. \end{cases}$$

$$\ell(X, Y) = \begin{cases} 1, & \text{如果 } X \models Y, \\ -1, & \text{如果 } X \models \neg Y. \end{cases}$$

1.3.3 有效论证和归纳强论证的比较

考虑一个论证 $X \Rightarrow Y$, 我们已经有三种方式评估 X 如何支持 Y :

1. $X \Rightarrow Y$ 是一个演绎推理: $X \models Y$.
2. 基于随机真值表, X 证据支持 Y : $\Pr(Y|X) > 0.5$.
3. 基于随机真值表, X 正相关于 Y : $\Pr(Y|X) > \Pr(Y)$.

其中, 1 对应有有效论证, 2 和 3 都是合情推理中的归纳强论证的必要条件. 我们将进一步讨论有效论证和归纳强论证的一些不同之处.

首先, 有效论证具有单调性: 论证的有效性随着前提的增加不会下降. 即: 对于任意 X, Y, Z , 若 $X \models Y$, 则 $X, Z \models Y$. 然而合情推理中, 单调性不再存在. 存在这样的例子: X, Y, Z 和对应的随机真值表, 使得 X 证据支持 Y , 但 $Z \wedge X$ 并不证据支持 Y . 增加新的前提反而可能降低结论发生概率.

例 1.4 (非单调性: 例子)

Pr	X	Y	Z	$X \wedge Z$
0.1	T	T	T	T
0.2	T	T	⊥	⊥
0.2	T	⊥	T	T
0	T	⊥	⊥	⊥
0.1	⊥	T	T	⊥
0.1	⊥	T	⊥	⊥
0.1	⊥	⊥	T	⊥
0.2	⊥	⊥	⊥	⊥

$\Pr(Y|X) = (0.1 + 0.2)/(0.1 + 0.2 + 0.2 + 0) = 0.6 > 0.5$, 然而, $\Pr(Y|X \wedge Z) = (0.1)/(0.1 + 0.2) = 1/3 < 0.5$.

接下来, 我们用 Z 论证 Y , 将 X 看作某种附加的条件, 我们考虑 X 对 Y 这一论据的影响. 在演绎推理中, 若 $Z, X \models Y$ 和 $Z, \neg X \models Y$ 都满足, 则 $Z \models Y$. 如果类比到合情推理中呢? 这就涉及到确凿性原则: 如果不论条件在 X 还是 $\neg X$, Z 都是 Y 的一个“好的论据”, 那么 Z 就是 Y 的一个“好的论据”. 与之相对应的是无条件确凿性原则: 如果 $Z \wedge \neg X$ 和 $Z \wedge X$ 都是 Y 的一个“好的论据”, 那么 Z 就是 Y 的一个“好的论据”. “好的论据”可以从证据支持和正相关性两方面考虑.

在任何随机真值表中, 如果 $\Pr(Y|Z \wedge X) > 0.5$ 且 $\Pr(Y|Z \wedge \neg X) > 0.5$, 那么 $\Pr(Y|Z) > 0.5$. 因此, 从证据支持角度, 确凿性原则是成立的. 而同样的论证也说明, 从证据支持角度, 无条件确凿性原则也是成立的.

在任何随机真值表中, 如果 $\Pr(Y|Z \wedge X) > \Pr(Y)$ 且 $\Pr(Y|Z \wedge \neg X) > \Pr(Y)$, 那么 $\Pr(Y|Z) > \Pr(Y)$. 因此, 从正相关性角度, 无条件确凿性原则是成立的. 那么, 从正相关性角度, 确凿性原则成立吗?

实际上, 并不一定成立! 有反直觉的例子: 存在 X, Y, Z 和对应的随机真值表, 使得

- $\Pr(Y|Z \wedge X) > \Pr(Y|X)$.
- $\Pr(Y|Z \wedge \neg X) > \Pr(Y|\neg X)$.
- 然而, $\Pr(Y|Z) \leq \Pr(Y)$.

这样的现象和 *Simpson* 悖论有关.

举个例子, 球员甲的两分球和三分球命中率均高于球员乙, 但是球员甲的总投篮命中率却可能低于乙. 具体来说, 有一个班中一半同学来自北京大学, 另一半来自清华大学, 我们抽出一名同学 Bob, 估计 Bob 投篮命中的概率. 记 Y 为 Bob 投篮命中, 记 X 为 Bob 投出一个两分球, 则 $\neg X$ 为 Bob 投出一个三分球 (我们这里只考虑有两分和三分球), 记 Z 为 Bob 来自清华大学. $\Pr(Y)$ 表示全班学生的投篮命中率, $\Pr(Y|Z)$ 表示全班来自清华大学的学生的投篮命中率, $\Pr(Y|X)$ 表示全班学生的两分命中率, $\Pr(Y|\neg X)$ 表示全班学生的三分命中率, $\Pr(Y|Z \wedge X)$ 表示这个班来自清华大学的学生的两分命中率, $\Pr(Y|Z \wedge \neg X)$ 表示这个班来自清华大学的学生的三分命中率.

考虑这样一个投篮数据的实例:

	全班同学	来自清华大学
两分球	50/100	6/10
三分球	1/101	1/100
总命中率	51/201	7/110

$\Pr(Y) = 51/201$, $\Pr(Y|Z) = 7/110$ 为投篮命中率. $\Pr(Y|X) = 50/100 = 1/2$, $\Pr(Y|Z \wedge X) = 6/10 = 3/5$ 为两分命中率. $\Pr(Y|\neg X) = 1/101$, $\Pr(Y|Z \wedge \neg X) = 1/100$ 为三分命中率. Simpson 悖论在这一实例下的解释: 这个班里来自清华大学的学生两分命中率和三分命中率分别都比全班平均水平高, 但总体投篮命中率反倒比全班水平低.

我们将这两个概率用全概公式展开来寻找原因:

$$\Pr(Y|Z) = \Pr(Y|Z \wedge X) \Pr(X|Z) + \Pr(Y|Z \wedge \neg X) \Pr(\neg X|Z),$$

$$\Pr(Y) = \Pr(Y|X) \Pr(X) + \Pr(Y|\neg X) \Pr(\neg X).$$

$\Pr > \Pr$. 然而, 关键是有可能发生 $\Pr(X|Z) \neq \Pr(X)$. 在上面篮球的例子中表现为清华大学的同学选择投两分球和三分球的比例和全班同学不同.

最后我们考虑合取谬误. 合取谬误是一种认知偏差. 一个经典例子是, Linda 是一位单身、外向且年龄为 31 岁的女性. 在大学期间, 她主修哲学, 十分关注种族歧视和社会公正问题, 而且曾参加过反核游行. (记为 E) 请问以下哪一件事情更可能发生?

1. Linda 是一名银行出纳员 (记为 B)
2. Linda 是一名银行出纳员, 同时她还是一名女权主义者 (记为 $B \wedge F$)

在调查实验中多数被试选择了 2. 但是, 我们可以肯定 $\Pr(B \wedge F|E) \leq \Pr(B|E)$. 如何理解这一现象?

为了理解这种谬误产生的原因, 考虑合取原则: 如果 E 是 $P \wedge Q$ 的“好论据”, 那么 E 也是 P 的“好论据”. 在演绎推理中, 因为 $E \rightarrow (P \wedge Q) \models E \rightarrow P$, 所以合取原则成立. 类似确凿性原则, 我们在合情推理中也可以从证据支持和正相关性两方面考虑. 从证据支持的角度, 合取原则成立, 这是因为, 如果 $\Pr(P \wedge Q|E) > 0.5$, 那么 $\Pr(P|E) \geq \Pr(P \wedge Q|E) > 0.5$.

然而, 从正相关性的角度, 合取原则未必成立. 也就是说, 假设 $\Pr(P \wedge Q|E) > \Pr(P \wedge Q)$, 不一定能推出 $\Pr(P|E) > \Pr(P)$. 当人们给定对 Linda 的描述 E 的时候, 很容易建立起 E 和 $B \wedge F$ 的正相关性. 然而这并不意味着 E 和 B 是正相关的! 因此发生了合取谬误. 从 Simpson 悖论和合取谬误可以看出, 只依靠正相关性进行推理很容易犯错误, 因此证据支持 (极大似然) 是归纳强论证不可缺少的要素.

第二部分

信息与数据

第三部分

决策与优化

第四部分

逻辑与博弈

第五部分

认知逻辑

参考文献

- [Bre57] Leo Breiman. The Individual Ergodic Theorem of Information Theory. *The Annals of Mathematical Statistics*, 28(3):809–811, 1957.
- [CT12] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2012.
- [Huf52] David A. Huffman. A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the IRE*, 40(9):1098–1101, September 1952.
- [Inf] Information | Etymology, origin and meaning of information by etymonline. <https://www.etymonline.com/word/information>.
- [Jay02] Edwin T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2002.
- [KL51] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- [LLG⁺19] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, October 2019.
- [McM53] Brockway McMillan. The Basic Theorems of Information Theory. *The Annals of Mathematical Statistics*, 24(2):196–219, June 1953.
- [RHW86] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, pages 318–362. MIT Press, Cambridge, MA, USA, January 1986.

- [Rob49] Robert M. Fano. *The Transmission of Information*. March 1949.
- [Sha48] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948.
- [Shi96] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer, New York, NY, 1996.
- [Tin62] Hu Kuo Ting. On the Amount of Information. *Theory of Probability & Its Applications*, 7(4):439–447, January 1962.
- [Uff22] Jos Uffink. Boltzmann’s Work in Statistical Physics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2022 edition, 2022.
- [李 10] 李贤平. 概率论基础. 高等教育出版社, 2010.

索引

MP, 3

Simpson 悖论, 13

三段论

弱 \sim , 4, 9

强 \sim , 4, 9

事件域, 7

似然, 4, 8

先验概率, 8

前提, 3

合取原则, 14

合取谬误, 14

合情推理, 4

与规则, 5

否定规则, 6

后验概率, 8

命题逻辑, 2

基率, 8

基率谬误, 9

完备性定理, 3

归纳强论证, 10, 11

最大似然, 9

有效论证, 10

概率, 7

演绎推理, 3

相关性, 11

确凿性原则, 13

无条件 \sim , 13

等值, 3

结论, 3

肯定前件, 3

认可度, 11

认可度似然比, 11

认可概率增量, 11

证据支持, 10

连接词, 2

逻辑结论, 10

重言式, 3

随机真值表, 10