概率论

——科学逻辑

E. T. Jaynes 著

明夷船 译

1　可信推理

2 量化规则

概率论，不过是生活约化而成的计算。

拉普拉斯，1819

问题已被形式化了，这是我们所作的公设在数学上的必然结果。这些公设可粗略描述如下：

(I) 命题的可信性由实数表示；
(II) 定性符合常识；
(III) 一致性。

本章仅基于这些公设推导与推断相关的量化规则。所得结果，曾经有着一段漫长、复杂又令人难以置信的历史。这段历史充满着广义科学方法论方面的教训（见某些章尾部的评注部分）。

2.1 乘法规则

我们首先为逻辑乘 AB 的可信性分别与 A 与 B 的可信性之间的关系找出一个一致性的规则。我们在意的其实是 $AB|C$。由于推理过程本身有些微妙，这需要我们从一些不同的视角去审视。

先考虑将 AB 为真的裁决打碎为对 A 与 B 所作的更为基本的裁决。机器人可以

(1) 裁决 B 为真； $(B|C)$
(2) 认可 B 为真，裁决 A 为真。 $(A|BC)$

或者，等同地，

(1) 裁决 A 为真； $(A|C)$
(2) 认可 A 为真，裁决 B 为真。 $(B|AC)$

上面，在每种情况中，我们给出了每一步骤相应的可信性。

现在，解释一下第 1 个过程。为了命题 AB 为真，B 必须为真。因而，必须考虑 $B|C$ 的可信性。此外，若 B 为真，则需要进一步保证 A 为真，因此还需要考虑 $A|BC$ 的可信性。但是若 B 为假，则无需裁决 A，亦即无需裁决 $A|\overline{B}C$，便可知晓 AB 为假；若机器人先推理 B，那么，A 的可信性仅在 B 为真时才值得考虑。因此，机器人在有了 $B|C$ 与 $A|BC$ 的情况下，就不需要 $A|C$ 了，因为 $A|C$ 并未给它能带来更多的信息。

类似地，$A|B$ 与 $B|A$ 是不需要的；无论 A 与 B 在缺乏 C 的情况下多么可信，都与机器人在知道 C 为真的情况下所作的裁决无关。例如，若机器人知晓地球是圆的，那么在裁决与现在的宇宙学相关的问题时，它就可以忽略那些在它尚不知地球是圆的之时需要考虑的观点。

由于逻辑乘法运算遵循交换律，$AB = BA$，因此上文中的 A 与 B 毫无疑问，可以互换；亦即 $A|C$ 与 $B|AC$ 的知识也能用于确定 $AB|C = BA|C$。对于 AB，机器人必定能从这两个过程中得到相同的可信性，这是我们的一致性条件之一，即公设 (IIIa)。

可以用更明确的形式来描述。$(AB|C)$ 可以是 $B|C$ 与 $A|BC$ 的某个函数：

$$(AB|C) = F[(B|C), (A|BC)] \tag{2.1}$$

现在，如果上述推理尚不完全清晰，那么就审视一下其他的替代形式。例如，可以假设

$$(AB|C) = F[(A|C), (B|C)] \tag{2.2}$$

是容许的形式。但是，很容易揭示，这种形式的关系无法满足公设 (II) 的那些定性条件。给定 C，命题 A 或许非常可信，命题 B 或许可信，但是 AB 可能依然非常可信或非常不可信。

例如，下一个遇到的人，蓝眼睛，这相当可信，黑头发，这也相当可信；这两样生理特征同时出现于此人身上，也没什么不合理之处。不过，左眼睛是蓝色的，这相当可信，右眼睛是褐色的，也相当可信，但是若它们同时为真，那就极为不可信。如果使用这种形式的公式，那么就没法考虑这些情况。用这种形式的函数，我们的机器人无法像人类那样去作推理，它甚至连定性的推理都做不到。

还有一些其他的可能的形式。尝试所有的可能形式的方法——“穷举证明”——可像下面这样进行。先引入一些实数

$$u = (AB|C), \quad v = (A|C), \quad w = (B|AC), \quad x = (B|C), \quad y = (A|BC). \tag{2.3}$$

如果 u 被表示成 u、v、x、y 中两个或更多个实数的函数，那么有 11 种可能的形式。可以写出每一种可能，将它置于各种极端条件下，如同褐色眼睛与蓝色眼睛那样（抽象描述：A 蕴含了 B 为假[1]）。其他极端条件有，$A = B$，$A = C$，$C \Rightarrow \overline{A}$，等。Tribus（1969）作了冗长乏味的分析，发现只有两种可能的形式，在某种极端的情况下，遵循着定性符合常识这一公设，它们分别是 $u = F(x, y)$ 与 $u = F(w, v)$，这正是之前我们推理出来的那两种函数形式。

现在，运用第 1 章讨论的量化要求。给定的先验信息的任何变化 $C \to C'$，以及 B 变得更可信，A 没有变化，

$$B|C' > B|C \tag{2.4}$$

$$A|BC' = A|BC \tag{2.5}$$

根据常识，AB 只会变得更可信，而不是更不可新：

$$AB|C' \geq AB|C \tag{2.6}$$

当且仅当 $A|BC$ 为不可能时，相等关系成立。同样，给定先验信息 C''，以及

$$B|C'' = B|C \tag{2.7}$$

[1] 译注：$A \Rightarrow \overline{B}$

$$A|BC'' > A|BC \tag{2.8}$$

就会有

$$AB|C'' \geq AB|C \tag{2.9}$$

当且仅当 $B|C$ 为不可能时，相等关系成立（即使此时未去定义 $A|BC$，$AB|C''$ 依然不可能为真）。此外，函数 $F(x,y)$ 必须是连续的，不然，(2.1) 右侧的某个可信性的微小递增可能会导致 $AB|C$ 的大幅递增。

总之，$F(x,y)$ 必须是 x 与 y 的连续的单调递增函数。若假设它可微（并非必须如此；见 (2.13) 的讨论），便有

$$F_1(x,y) \equiv \frac{\partial F}{\partial x} \geq 0 \tag{2.10.a}$$

仅当 x 表示不可能时，上式中的相等关系成立；还有，

$$F_2(x,y) \equiv \frac{\partial F}{\partial y} \geq 0 \tag{2.10.b}$$

后文会继续使用这些符号，无论 F 是什么样的函数，F_i 表示 F 的第 i 个参数对应的微分。

接下来，我们动用公设 (IIIa)，“结构上”的一致性。假设想获得 $(ABC|D)$ 的可信性，即三个命题同时为真的可信性，由于布尔运算遵循结合律 $ABC = (AB)C = A(BC)$，因此有两种方式来做此事。若规则具有一致性，一定能从这两种顺序的运算中获得相同的结果。首先，可以将 BC 视为单一的命题，利用 (2.1)：

$$(ABC|D) = F[(BC|D),(A|BCD)] \tag{2.11}$$

然后，对可信性 $(BC|D)$ 应用 (2.1)，可得：

$$(ABC|D) = F\{F[(C|D),(B|CD)],(A|BCD)\} \tag{2.12.a}$$

不过，也可以在一开始将 AB 视为单一的命题，这样就可以从另一种顺序推出不同的结果：

$$(ABC|D) = F[(C|D),(AB|CD)] = F\{(C|D),F[(B|CD),(A|BCD)]\} \tag{2.12.b}$$

若我们寻找的这个规则是要用于表示推理的一致性方法，那么 (2.12.a) 与 (2.12.b) 必定恒等。在这种情况下，我们的机器人所作的一致性推理，其必要条件可表示为一个函数方程，

$$F[F(x,y),z] = F[x,F(y,z)] \tag{2.13}$$

在数学里，这个方程历史悠久，源起于 N. H. Abel (1826) 的研究。Aczél (1966) 在他的讲述函数方程的巨著中，贴切地成这个方程为“关联方程”（Associativity Equation），并列举了 98 份讨论或使用这个方程的参考文献。Aczél 在不假设函数可微的条件下，求出了通解 (2.27)，见下文。不过，他的书中（Aczél, 1987），用了 11 页的篇幅证明此解的存在。在此，我们给出 R. T. Cox (1961) 在假设函数可微的前提下给出的更短的证明；也可参考附录 B 中的讨论。

显然，(2.13) 有一个平凡解，$F(x,y) =$ 常数。但是这个解违背了单调性要求 (2.10)，并且毫无用处可言。除非 (2.13) 有一个非平凡解，不然，这条路就走不通了；因此，必须寻求最为广义的非平凡解。先定义一些缩写

$$u \equiv F(x,y), \quad v \equiv F(y,z) \tag{2.14}$$

现在，依然认为 (x,y,z) 是互不依赖的变量，待求解的函数方程可写为

$$F(x,v) = F(u,z) \tag{2.15}$$

分别对 x 与 y 求微分，按照 (2.10) 的记法，可得

$$\begin{aligned} F_1(x,v) &= F_1(u,z)F_1(x,y) \\ F_2(x,v)F_1(y,z) &= F_1(u,z)F_2(x,y) \end{aligned} \tag{2.16}$$

从这两个方程中消除 $F_1(u,z)$，可得

$$G(x,v)F_1(y,z) = G(x,y) \tag{2.17}$$

其中，$G(x,y) \equiv F_2(x,y)/F_1(x,y)$。显然，(2.16) 的左部一定是不依赖 z。可将 (2.17) 等价地写为

$$G(x,v)F_2(y,z) = G(x,y)G(y,z) \tag{2.18}$$

用 U 与 V 分别表示 (2.17) 与 (2.18) 的左部，可以得出 $\partial V/\partial y = \partial U/\partial z$。因而 $G(x,y)G(y,z)$ 必定不依赖 y。具有这一性质的最具一般性的函数 $G(x,y)$ 为

$$G(x,y) = r\frac{H(x)}{H(y)} \tag{2.19}$$

其中，r 为常数，$H(x)$ 为任意函数。基于 F 的单调性，可确定 $G > 0$，因此需要 $r > 0$，至于 $H(x)$ 的正负则无关大体。应用 (2.19)，则 (2.17) 与 (2.18) 变成：

$$F_1(y,z) = \frac{H(v)}{H(y)} \tag{2.20}$$

$$F_2(y,z) = r\frac{H(v)}{H(z)} \tag{2.21}$$

关系 $dv = dF(y,z) = F_1 dy + F_2 dz$ 可化为

$$\frac{dv}{H(v)} = \frac{dy}{H(y)} + r\frac{dz}{H(z)} \tag{2.22}$$

或者化为积分形式

$$w[F(y,z)] = w(v) = w(y)w^r(z) \tag{2.23}$$

其中

$$w(x) = \exp\left\{\int^{x} \frac{dx}{H(x)}\right\} \tag{2.24}$$

积分符号无下界，意味着 w 会有一个任意的倍增因子。但是，将 (2.15) 代入函数 $w(\cdot)$ 并应用 (2.23)，可得 $w(x)w^r(v) = w(u)w^r(z)$；再次应用 (2.23)，函数方程便可化为

$$w(x)w^r(y)[w(z)]^{r^2} = w(x)w^r(y)w^r(z) \tag{2.25}$$

若 $r = 1$，便可获得一个非平凡解，最终结果可表示为以下两种形式：

$$w[F(x, y)] = w(x)w(y) \tag{2.26}$$

或

$$F(x, y) = w^{-1}[w(x)w(y)] \tag{2.27}$$

因此，逻辑乘法所遵守的结合律与交换律必须体现为以下的函数形式

$$w(AB|C) = w(A|BC)w(B|C) = w(B|AC)w(A|C) \tag{2.28}$$

我们将这种形式称为**乘法规则**。由 (2.24) 的构造可知，$w(x)$ 必定是个正的连续的单调函数，至于它是递增的还是递减的，这有赖于 $H(x)$ 的符号；目前，它另有深意。

结果得到了 (2.28)，某种意义上，它是公设 (IIIa) 所述一致性的必要条件。不过，对于任意多个联结的命题，(2.28) 显然也能保证这种一致性的存在。例如，用 (2.12) 的办法可以将 $(ABCDEFG|H)$ 展开为数目繁多的不同形式；但是，只要 (2.28) 成立，这些形式的结果必定相同。

定性要符合常识，这一要求对 $w(x)$ 有着更为严格的限定。例如，根据 (2.28) 所给出的形式，假设在给定 C 的情况下 A 为真，那么在由 C 的知识所营造的“逻辑环境”里，从当且仅当一个为真时另一个也必定为真这一意义来说，命题 AB 与 B 并无区别。基于第 1 节讨论的最基本的公理，同样为真的命题一定具备相同的可信性[2]：

$$AB|C = B|C \tag{2.29}$$

还有

$$A|BC = A|C \tag{2.30}$$

由于给定 C 的时候 A 是确信的（亦即 C 蕴含 A），那么给定任何其他不与 C 矛盾的 B，A 依然是确信的。在这种情况下，(2.28) 约化为

$$w(B|C) = w(A|C)w(B|C) \tag{2.31}$$

对于机器人而言，无论 B 有多么可信或不可信，该式必定成立。因此，函数 $w(x)$ 必定具有以下性质

---

[2] 见 1.5 节。

$$w(A|C)=1 \text{ 表示确信性} \tag{2.32}$$

现在，在给定 C 的情况下，假设 A 不可信，那么在给定 C 时，命题 AB 也不可信：

$$AB|C = A|C \tag{2.33}$$

若给定 C 的情况下，A 不可信（亦即 C 蕴含 $\overline{A}$），那么给定任何不与 C 矛盾的更充分的信息 B，A 依然不可信：

$$A|BC = A|C \tag{2.34}$$

在这种情况下，(2.28) 可化为

$$w(A|C) = w(A|C)w(B|C) \tag{2.35}$$

无论 B 有多么可信，这个方程必定成立。$w(A|C)$ 只有两个可能的值能够满足这个条件，要么为 0，要么为 $+\infty$（排除了 $-\infty$，不然基于连续性，$w(B|C)$ 必须为负值；于是 (2.35) 自相矛盾）。

综上所述，定性要符合常识，这决定了 $w(x)$ 必须是一个正值的连续单调函数。它可能递增，也可能递减。若它递增，取值范围必须从 0 到 1，前者表不可信，后者表确信。若它递减，取值范围必须从 $+\infty$ 到 1，前者表不可信，后者表确信。至于它在这两种区间里具体如何变化，我们所给出的公设则没有对其进行限定。

然而，这两种可能的表示在内涵上并不相同。给定任意函数 $w_1(x)$，令它符合上述标准并且用 $+\infty$ 表示不可信，便可以定义一个新函数 $w_2(x)=\dfrac{1}{w_1(x)}$，这个函数同样可被接受，它是以 0 来表示不可信。因而，作为一种约定，若我们采纳 $0 \le w(x) \le 1$，不失一般性；也就是说，就内涵而言，我们所作的公设里面的所有的一致性皆被包含于这种形式。（读者不妨检验一下，未尝不可选择相反的那种约定，从这一点发展出一套完整的理论，也包括它的全部应用，结果是一样的，只不过方程在形式上有些另类，但内涵是相同的。）