杨柳　刘陶唐　编著

智能规划理论和方法研究

ZHINENG GUIHUA LILUN HE FANGFA YANJIU

冶金工业出版社
www.cnmip.com.cn

智能规划理论和方法研究

杨 柳 刘陶唐 编著

北 京
冶 金 工 业 出 版 社
2017

内 容 提 要

智能规划是人工智能的一个重要研究领域。本书主要介绍了与经典智能规划和不确定性规划相关的理论基础，主要包括概率论基础、布尔表达式及其描述、经典规划求解技术、智能规划语言、马尔科夫决策过程、代数决策图、贝叶斯网络等，旨在让感兴趣的研究者对这一研究领域的相关内容有一个完整的了解。

本书可作为计算机专业学生和研究人员及工程技术人员的参考用书。

本书由黑龙江省教育厅项目（项目编号：12541835）资助出版。

图书在版编目(CIP)数据

智能规划理论和方法研究/杨柳，刘陶唐编著．—北京：冶金工业出版社，2017.7

ISBN 978-7-5024-7525-3

Ⅰ.①智…　Ⅱ.①杨…　②刘…　Ⅲ.①人工智能—研究　Ⅳ.①TP18

中国版本图书馆 CIP 数据核字(2017)第 105366 号

出 版 人　谭学余

地　　址　北京市东城区嵩祝院北巷 39 号　邮编　100009　电话　(010)64027926

网　　址　www.cnmip.com.cn　电子信箱　yjcbs@cnmip.com.cn

责任编辑　曾　媛　美术编辑　彭子赫　版式设计　孙跃红

责任校对　李　娜　责任印制　李玉山

ISBN 978-7-5024-7525-3

冶金工业出版社出版发行；各地新华书店经销；固安华明印业有限公司印刷

2017 年 7 月第 1 版，2017 年 7 月第 1 次印刷

169mm×239mm；6.75 印张；131 千字；100 页

39.00 元

冶金工业出版社　投稿电话　(010)64027932　投稿信箱　tougao@cnmip.com.cn

冶金工业出版社营销中心　电话　(010)64044283　传真　(010)64027893

冶金书店　地址　北京市东四西大街 46 号(100010)　电话　(010)65289081(兼传真)

冶金工业出版社天猫旗舰店　yjgycbs.tmall.com

（本书如有印装质量问题，本社营销中心负责退换）

前　言

智能规划是用人工智能理论与技术自动或半自动地生成一组动作序列（或称一个“计划”），用以实现期望的目标。智能规划是智能系统理论与应用研究的重要分支。

不确定性推理是人工智能研究的重要课题之一，从20世纪60~70年代以来，人们提出了多种方法，如概率方法、非单调逻辑、模糊逻辑等。在这些方法中，概率方法是最自然也是最早被尝试的方法之一，因为概率论本身是关于随机现象和不确定性的数学理论。

不确定性通过对经典规划问题假设的适当放宽形成了非经典问题，主要研究系统初始状态不确定和动作效果不确定对求解的影响。对动作执行效果的不确定性，决策理论的应用是一种自然的选择。采用马尔科夫决策过程是一种主要的表达和求解方法。

贝叶斯网络是为了处理人工智能研究中的不确定性（uncertainty）问题而发展起来的，贝叶斯网络是将概率统计应用于复杂领域进行不确定性推理和数据分析的工具，是一种系统地描述随机变量之间关系的工具。建立贝叶斯网络的目的主要是进行概率推理，用概率论处理不确定性的主要优点是保证推理结果的正确性。

布尔代数是逻辑系统设计的重要基石，人工智能规划与调度等领域的许多问题都可归结为逻辑函数及其序列的操作和运算，布尔函数的表达及操作对合理有效解决这些领域中的问题起着重要的作用。二元判定图（BDD）、有序的二元判定图（OBDD）、代数决策图（ADD）是布尔函数表述和操作中最为有效的技术，可以有效减缓甚至避免问题处理过程中的状态组合复杂性。

规划问题必须用形式化语言来进行描述，才能被规划系统读入进

行求解。因此，规划语言的描述能力和自身的完善性对智能规划研究领域是至关重要的。

本书对经典规划和非经典规划中的相关理论和方法进行了介绍分析。全书共 8 章：第 1 章介绍了概率论的基础知识；第 2 章为布尔表达式及相关描述；第 3 章介绍了经典规划求解技术；第 4 章对智能规划语言进行了总结；第 5 章介绍了马尔科夫决策过程；第 6 章介绍了代数决策图和相关操作；第 7 章对贝叶斯网络进行了简要介绍；第 8 章介绍了最新的 multi-agent 规划系统国际比赛情况。

由于作者水平所限，书中疏漏和不妥之处，恳请读者批评指正。

作　者

2017 年 1 月

目　　录

1 概率论基础

概率论是研究随机现象数量规律的数学分支。随机现象是相对于决定性现象而言的。在一定条件下必然发生某一结果的现象称为决定性现象。事件的概率是衡量该事件发生的可能性的量度。虽然在一次随机试验中某个事件的发生是带有偶然性的，但那些可在相同条件下大量重复的随机试验却往往呈现出明显的数量规律。下面简要介绍概率论中的基本概念。

（1）事件空间。基本事件空间（space of elementary event）概率论的基本概念之一，指同一问题中的所有基本事件组成的集合。在统计学中习惯称为样本空间，它的元素称为样本点，即基本事件。事件空间所含元素个数为有限或可数时，分别称为有限基本事件空间或离散基本事件空间。

（2）随机变量。随机变量的定义是，如果对于试验的样本空间 Ω 中的每一个样本点 ω，变量 X 都有一个确定的实数值与之对应，则变量 X 是样本点 ω 的实函数。这样的变量称为随机变量，记作：$X = X(\omega)$。随机变量通常用大写英文字母 X，Y，Z，… 表示。随机变量首先是变量，取值是实数。它的取值随着试验结果的变化而变化。

用随机变量刻画随机事件：

随机变量 X 取得某一数值 x，记作：$X=x$，这是一个随机事件。随机变量 X 取得不大于实数 x 的值，记作：$X\leqslant x$，也是一个随机事件。以下都是随机事件：$a\leqslant X<b$，$a<X<b$，$a<X\leqslant b$，$a\leqslant X\leqslant b$。一般来说，关于随机变量的等式或不等式都是随机事件。

随机变量的类型如下：

1）离散型随机变量。仅可能取得有限个或可数无穷多个数值。例如，孵化一枚种蛋可能结果只有两种，即“孵出小鸡”与“未孵出小鸡”。若用变量 x 表示试验的两种结果，则可令 $X=0$ 表示“未孵出小鸡”，$X=1$ 表示“孵出小鸡”。

2）连续型随机变量。可以取得某一区间内的任何数值。例如，测定某品种猪初生重，表示测定结果的变量 X 所取的值为一个特定范围 (a, b)，如 0.5~1.5kg，X 值可以是这个范围内的任何实数。

（3）概率分布。概率分布是概率论的基本概念之一，用以表述随机变量取值的概率规律。为了使用的方便，根据随机变量所属类型的不同，概率分布取不同的表现形式，可以为描述随机变量值 X_i 及这些值对应概率 $P(X = X_i)$ 的表格、

公式或图形。

（4）先验概率。事件发生前的预判概率。可以是基于历史数据的统计，可以由背景常识得出，也可以是人的主观观点给出。一般都是单独事件概率，如 $P(X)$ ，$P(Y)$ 。

（5）联合分布和边缘分布。设 E 是一个随机试验，它的样本空间是 Ω，设 x_1，x_2，…，x_n 是定义在 Ω 上的 n 个随机变量，由它们构成的一个向量（x_1，x_2，…，x_n）称为 n 维随机向量或 n 维随机变量。

特别地，当 $n=2$ 时，即（x_1，x_2），称为二维随机向量或二维随机变量。

【例 1-1】 炮弹的弹着点的位置（X，Y）就是一个二维随机变量。

【例 1-2】 考察某一地区学前儿童的发育情况，则儿童的身高 H 和体重 W 就构成二维随机变量（H，W）。

联合分布描述了多个随机变量的概率分布，是对单一随机变量的自然拓展。联合分布的多个随机变量都定义在同一个样本空间中。

对多个随机变量 X_1，X_2，…，X_n，可以用联合概率分布 P（X_1，X_2，…，X_n），简称联合概率分布来描述各变量所有可能的状态组合的概率。

它是一个定义在所有变量状态空间的笛卡尔乘积之上的函数，其所有函数值之和为1，即 $\sum\limits_{X_1,\ \cdots,\ X_n} P(X_1,\ \cdots,\ X_n)=1$。联合分布常被表示为一张表，如果所有变量都只有两个状态，则联合分布表中共有 2^n 个项，刻画了变量之间的各种关系。

设二维离散型随机变量（X，Y）的所有可能取的值为（x_i，y_i），i，$j=1$，2，…，记作 $P\{X=x_i,\ Y=y_i\}=P_{ij}$，$i$，$j=1$，2，…，称此为二维离散型随机变量（$X$，$Y$）的联合分布律，或随机变量 X 和 Y 的联合分布律。注意：$(X=x_i,\ Y=y_i)=(X=x_i)\cap(Y=y_i)$，$i=1$，2，…。

二维随机变量（X，Y）的联合分布列也可表示为表 1-1。

表 1-1 二维随机变量的联合分布

Y \ X	x_1	x_2	…	x_i	…
y_1	P_{11}	P_{21}	…	P_{i1}	…
y_2	P_{12}	P_{22}	…	P_{i2}	…
⋮	⋮	⋮	⋮	⋮	⋮
y_i	P_{1j}	P_{2j}	…	P_{ij}	…
⋮	⋮	⋮	⋮	⋮	⋮

联合分布的性质如下：

1) $P_{ij}=P(x_i,\ y_j)\geqslant 0$;

2) $\sum\limits_i\sum\limits_j P_{ij}=\sum\limits_i\sum\limits_j P(x_i,\ y_j)=1$;

3) $P(X=x_i)=\sum\limits_{j=1}^{\infty}p_{ij}=p_{i\cdot}$, $P(Y=y_i)=\sum\limits_{i=1}^{\infty}p_{ij}=p_{\cdot j}$。

分别称 $p_{i\cdot}$ $(i=1,\ 2,\ \cdots)$、$p_{\cdot j}(j=1,\ 2,\ \cdots)$ 为 $(X,\ Y)$ 关于 X 和关于 Y 的边缘分布列，见表 1-2。

表 1-2　联合分布及边缘分布列

Y \ X	x_1	…	x_i	…	$P_{\cdot j}$
y_1	p_{11}	…	p_{i1}	…	$p_{\cdot 1}$
⋮	⋮	⋮	⋮	⋮	⋮
y_i	p_{1j}	…	p_{ij}	…	$p_{\cdot j}$
⋮	⋮	⋮	⋮	⋮	⋮
$P_{i\cdot}$	$p_{1\cdot}$	…	$p_{i\cdot}$	…	1

联合分布可以唯一地确定边缘分布，边缘分布却不能唯一确定联合分布。

【例 1-3】　考虑市场上所有出租房屋，从中随机抽取一间，考查其月租 (R) 和类型 (T) 这两个随机变量，月租分为 4 等：{low(低于 1000 元)，medium(1000 ~ 2500 元)，upper medium(2500 ~ 4000 元)，high(高于 4000 元)}。类型有 3 种：{public(公屋)，private(私家屋)，others(其他)}。联合分布 $P(R,\ T)$ 见表 1-3。

表 1-3　联合分布 P $(R,\ T)$

R \ T	public	private	others
low	0.17	0.01	0.02
medium	0.44	0.03	0.01
upper medium	0.09	0.07	0.01
high	0	0.14	0.01

从表中可知，随机抽到中价公屋的可能性最大，为 44%。

边缘概率分布：在例 1-3 中，由于有了联合分布 $P(R,\ T)$，所以可以回答这样的问题：随机抽取一间出租房屋为公屋的概率 $P(T=\text{public})$ 是多少?

根据概率的有限可加性：

$$P(T=\text{public})=P(T=\text{public},\ R=\text{low})+P(T=\text{public},\ R=\text{medium})+$$
$$P(T=\text{public},\ R=\text{upper medium})+$$

$$P(T=\text{public}, R=\text{high})$$
$$=0.7$$

同样地，可以计算 $P(T=\text{private})$，$P(T=\text{others})$：

$$\begin{aligned}P(T=\text{private}) = &P(T=\text{private}, R=\text{low}) + \\ &P(T=\text{private}, R=\text{medium}) + \\ &P(T=\text{private}, R=\text{upper medium}) + \\ &P(T=\text{private}, R=\text{high}) \\ = &0.25\end{aligned}$$

$$\begin{aligned}P(T=\text{others}) = &P(T=\text{others}, R=\text{low}) + \\ &P(T=\text{others}, R=\text{medium}) + \\ &P(T=\text{others}, R=\text{upper medium}) + \\ &P(T=\text{others}, R=\text{high}) \\ = &0.05\end{aligned}$$

为了简化记号，上面三式可分别缩写为：

$$P(T=\text{public}) = \sum_R P(T=\text{public}, R)$$

$$P(T=\text{private}) = \sum_R P(T=\text{private}, R)$$

$$P(T=\text{others}) = \sum_R P(T=\text{others}, R)$$

这三个式子还可以进一步合并为下式：

$$P(T) = \sum_R P(T, R)$$

相对于联合分布 $P(R, T)$，$P(T)$ 称为边缘分布。表 1-4 同时给出了联合分布 $P(R, T)$ 和边缘分布 $P(T)$，$P(R)$。

表 1-4 联合分布 $P(R, T)$ 和边缘分布 $P(T)$，$P(R)$

R \ T	public	private	others	$P(R)$
low	0.17	0.01	0.02	0.20
medium	0.44	0.03	0.01	0.48
upper medium	0.09	0.07	0.01	0.17
high	0	0.14	0.01	0.15
$P(T)$	0.70	0.25	0.05	

从联合分布 $P(X)$ 到边缘分布 $P(Y)$ 的过程称为边缘化。

(6) 联合概率。表示两个事件共同发生（数学概念上的交集）的概率。A 与 B 的联合概率表示为 $P(A \cap B)$。

（7）条件概率。条件概率就是事件 A 在另外一个事件 B 已经发生条件下的发生概率。条件概率表示为 $P(A|B)$，读作“在 B 发生的条件下 A 发生的概率”。定义为：

$$P(A|B)=\frac{P(AB)}{P(B)}$$

反过来可以用条件概率表示 A、B 的乘积概率，即有乘法公式。

若 $P(B)\neq 0$，则 $P(AB)=P(B)P(A|B)$，同样有：

若 $P(A)\neq 0$，则 $P(AB)=P(A)P(B|A)$。

【例 1-4】 在掷骰子试验中，掷出 6 的概率为 1/6。假定投掷后被告知“掷出的结果是偶数”，问此时对结果为 6 的信度是多少？

设掷出 6 为事件 A，掷出结果为偶数为事件 B，则 $P(A)=1/6$，$P(B)=1/2$，$P(A\cap B)=1/6$。所问的问题即是要计算如下条件概率：

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{1/6}{1/2}=\frac{1}{3}$$

（8）条件分布。设 X 和 Y 是两个随机变量，x 和 y 分别是它们的一个取值。考虑事件 $X=x$ 在给定 $Y=y$ 时的条件概率为 $P(X=x|Y=y)=\dfrac{(X=x,\ Y=y)}{P(Y=y)}$。

在上式中，固定 y，让 x 在 Ω_X 上变动，则得到一个在 Ω_X 上的函数，这个函数称为在给定 $Y=y$ 时变量 X 的概率分布，记为 $P(X|Y=y)$。用 $P(X|Y)$ 记作 $\{P(X|Y=y)|y\in\Omega_Y\}$，即在 Y 取不同值时 X 的条件概率分布的集合。$P(X|Y)$ 称为给定 Y 时变量 X 的条件概率分布。在上式中，让 x 和 y 在 Ω_X 和 Ω_Y 上变动，则得到一组等式。把这些等式缩写为 $P(X|Y)=\dfrac{P(X,\ Y)}{P(Y)}$。该式可视为是 $P(X|Y)$ 的直接定义。

更一般地，设 $X=\{X_1,\ \cdots,\ X_n\}$ 和 $Y=\{Y_1,\ \cdots,\ Y_m\}$ 为两个变量集合，$P(X,\ Y)$ 为 $X\cup Y$ 的联合概率分布，$P(Y)$ 为 Y 的边缘概率分布，则给定 Y 时 X 的条件概率分布定义为 $P(X|Y)=\dfrac{P(X,\ Y)}{P(Y)}$。

【例 1-5】 在例 1-3 中问：随机抽取一间私家屋，其租金为 low 的概率多大？

这即是问给定 T=private 时，R=low 的条件概率 $P(R=\text{low}|T=\text{private})$ 按定义，有

$$P(R=\text{low}|T=\text{private})=\frac{P(R=\text{low},\ T=\text{private})}{P(T=\text{private})}=\frac{0.01}{0.25}=0.04$$

给定 T 时，变量 R 的条件 $P(R|T)$ 分布，见表 1-5。

表 1-5 变量 R 的条件 $P(R|T)$ 分布

T \\ R	low	medium	upper medium	high
public	$\frac{0.17}{0.7}$	$\frac{0.44}{0.7}$	$\frac{0.09}{0.7}$	$\frac{0}{0.7}$
private	$\frac{0.01}{0.25}$	$\frac{0.03}{0.25}$	$\frac{0.07}{0.25}$	$\frac{0.14}{0.25}$
others	$\frac{0.02}{0.05}$	$\frac{0.01}{0.05}$	$\frac{0.01}{0.05}$	$\frac{0.01}{0.05}$

表中第 1 行显示的是在给定 T=public 时，R 的条件概率分布，第 2 行是在给定 T=private 时，R 的条件概率分布等。这里每行的数字之和为 1，即 $\sum_{R} P(R|T)=1$。这与例 1-3 所列的联合分布 $P(R, T)$ 不同，那个表中所有数字之和为 1，即 $\sum_{R, T} P(R|T)=1$。

【例 1-6】 设有一袋积木，每块积木有 3 个属性：颜色、材料和形状。设积木的颜色只能是红（r）或蓝（b）两种，材料只能是金属（m）或木头（w），形状只能是正方体（6）或正四面体（4）。设 C，M，S 为 3 个随机变量，分别代表从袋中随机取出一块积木的颜色、材料和形状，则 $\Omega_C=\{r, b\}$，$\Omega_M=\{m, w\}$，$\Omega_S=\{6, 4\}$。

设联合概率分布 $P(C, M, S)$ 为：

C	M	S	$P(C, M, S)$
r	m	6	0.10
r	m	4	0.10
r	w	6	0.25
r	w	4	0.05
b	m	6	0.15
b	m	4	0.10
b	w	6	0.20
b	w	4	0.05

那么，变量 C 和 M 的边缘分布 $P(C, M)$ 为：

C	M	$P(C, M)$
r	m	0.20
r	w	0.30
b	m	0.25
b	w	0.25

条件分布为 $p(C|M)$ 为:

M \ C	r	b
m	$\frac{4}{9}$	$\frac{5}{9}$
w	$\frac{6}{11}$	$\frac{5}{11}$

两个事件的乘法公式还可以推广到 n 个事件，即

$$P(A_1, A_2, \cdots, A_n) = P(A_1)P(A_2|A_1)\cdot P(A_3|A_1A_2)\cdots P(A_n|A_1A_2, \cdots, A_{n-1})$$

称为链式规则。

随机变量形式链式规则:

$$P(X_1=x_1, X_2=x_2, \cdots, X_n=x_n) = P(X_1=x_1)P(X_2=x_2|X_1=x_1)\cdots$$
$$P(X_n=x_n|X_1=x_1, X_2=x_2, \cdots, X_{n-1}=x_{n-1})$$

或简写为 $P(x_1, x_2, \cdots, x_n) = P(x_1)P(x_2|x_1)\cdots P(x_n|x_1, x_2, \cdots, x_{n-1})$。

(9) 独立性:

1) 事件的独立。事件 A、B，若其中任一事件的发生概率不受另一事件发生与否的影响，称事件 A，B 相互独立。

数学式子表示: $P(B)=P(B|A)$，由乘法公式写成

$$P(AB)=P(A)P(B|A)=P(A)P(B)$$

由此导出了事件间的相互独立:

①两个事件的独立性。定义 1 若两事件 A，B 满足

$$P(AB)=P(A)P(B)$$

则称独立，或称 A，B 相互独立。

②有限个事件的独立性。设 A，B，C 为三个事件，若满足等式

$$\begin{cases} P(AB)=P(A)P(B) \\ P(AC)=P(A)P(C) \\ P(BC)=P(B)P(C) \\ P(ABC)=P(A)P(B)P(C) \end{cases}$$

则称事件 A，B，C 相互独立。

对 n 个事件的独立性，可类似地定义:

设 A_1，A_2，$\cdots$，A_n 是 $n(n>1)$ 个事件，若对任意 $k(1<k\leqslant n)$ 个事件

$$A_{i1}, A_{i2}, \cdots, A_{ik}(1\leqslant i_1<i_2<\cdots<i_k\leqslant n)$$

均满足等式 $$P(A_{i1}A_{i2}\cdots A_{ik})=P(A_{i1})P(A_{i2})\cdots P(A_{ik})$$

则称 A_1，A_2，$\cdots$，A_n 相互独立。

2) 变量独立。两个随机变量 X 和 Y 称为相互（边缘）独立，记为 $X\perp Y$，

如果下式成立：

$$P(X, Y)=P(X)P(Y)$$

考虑变量 Y 的某个取值 y，如果 $P(Y=y)>0$，则由上式可得 $P(X)=P(X|Y=y)$，$P(X|Y=y)$ 是已知 $Y=y$ 时，变量 X 的概率分布，而 $P(X)$ 是未知 Y 的取值时 X 的概率分布。所以，变量 X 与 Y 相互独立意味着：对变量 Y 的取值的了解不会改变变量 X 的概率分布；同样，对变量 X 的取值的了解也不会改变 Y 的概率分布。

更一般地，称随机变量 $X_1, X_2, \cdots, X_n$ 相互（边缘）独立，如果

$$P(X_1, X_2, \cdots, X_n)=P(X_1)P(X_2)\cdots P(X_n)$$

【例 1-7】 设有 3 个装有黑白两色球的口袋，第 1 个口袋黑白球各半，第 2 个口袋黑白球比例为 4∶1，第 3 个则全是黑球。设随机变量 X，Y，Z 分别代表从这 3 个口袋随机抽出的球的颜色，其状态空间为 $\Omega_X=\Omega_Y=\Omega_Z=\{w, b\}$，其中 w 表示白，b 表示黑。从 3 个袋子中抽球，所得球的颜色的联合概率分布 $P(X, Y, Z)$ 如下：

X	Y	Z	$P(X, Y, Z)$
w	w	w	0
w	w	b	0.1
w	b	w	0
w	b	b	0.4
b	w	w	0
b	w	b	0.1
b	b	w	0
b	b	b	0.4

而边缘分布 $P(X)$、$P(Y)$ 和 $P(Z)$ 则分别如下：

X	w	b
$P(X)$	0.5	0.5

Y	w	b
$P(Y)$	0.2	0.8

Z	w	b
$P(Z)$	0	1

容易验证 $P(X, Y, Z)=P(X)P(Y)P(Z)$，即 X，Y，Z 相互边缘独立。

考虑 3 个随机变量 X、Y 和 Z，设 $P(Z=z)>0$，$\forall z\in\Omega_z$。我们说 X 和 Y 在给定 Z 时相互条件独立，记为 $X\perp Y|Z$。如果下式成立：

$$P(X, Y|Z)=P(X|Z)P(Y|Z)$$

设 y 和 z 分别是 Y 和 Z 的任意取值，且 $P(Y=y, Z=z)>0$，由 $P(X, Y|Z)=P(X|Z)P(Y|Z)$ 可得 $P(X|Y=y, Z=z)=P(X|Z=z)$，$P(X|Z=z)$ 是在已知 $Z=z$ 时，X 的概率分布。而 $P(X|Y=y, Z=z)$ 是在已知 $Y=y$ 以及 $Z=z$ 时，X 的概率

分布。因此 $X \perp Y|Z$ 的直观含义是：在已知 Z 的前提下，对 Y 的取值的了解不影响 X 的概率分布。注意：这并不意味着在未知 Z 的取值时，X 和 Y 相互独立。$Y=y$ 有可能含有关于 X 的信息，只是所有这样的信息也都包含于 $Z=z$ 中，所以当已知 $Z=z$ 时，进一步了解到 $Y=y$ 并不增加关于 X 的信息。当然，$X \perp Y|Z$ 也意味着，在已知 Z 的取值时，对 X 的取值的了解不影响 Y 的概率分布。

（10）贝叶斯定理。先验概率和后验概率这两个概念是相对于某组证据而言的。设 H 和 E 为两个随机变量，$H=h$ 为某一假设，$E=e$ 为一组证据。在考虑 $E=e$ 之前，对事件 $H=h$ 的概率估计 $P(H=h)$ 称为先验概率，而在考虑证据之后，对 $H=h$ 的概率估计 $P(H=h|E=e)$ 称为后验概率。贝叶斯定理描述了先验概率和后验概率之间的关系：

$$P(H=h|E=e)=\frac{P(H=h)P(E=e|H=h)}{P(E=e)}$$

这又称为贝叶斯规则，或贝叶斯公式。