概率定义与解释

续本达

清华大学 工程物理系

2023-09-20 清华

复习

随机事件的相关概念

- 样本点、样本空间
- 随机事件、基本事件
- 必然事件、不可能事件
- 样本空间的分割

重要出发点

- 定义清楚样本空间

随机事件与集合运算

记号	概率论	集合论
S	样本空间, 必然事件	全集
$\emptyset$	不可能事件	空集
e	样本点	元素
$A \subset B$	A 发生必然导致 B 发生	A 是 B 的子集
$AB = \emptyset$	A 与 B 互不相容	A 与 B 无相同元素
$A \cup B$	A 与 B 至少有一发生	A 与 B 的并集
AB	A 与 B 同时发生	A 与 B 的交集
$A - B$	A 发生且 B 不发生	A 与 B 的差集
$\overline{A}$	A 不发生、对立事件	A 的余集

概率

事件域

设 S 为样本空间， $\mathcal{F}$ 为 S 的某些子集组成的集合类。如果 $\mathcal{F}$ 满足下列三个条件，则称其为 S 的一个事件域：

1. $S \in \mathcal{F}$,
2. $A \in \mathcal{F} \rightarrow \overline{A} \in \mathcal{F}$,
3. $A_n \in \mathcal{F}, n = 1, 2, \cdots \rightarrow \bigcup_{n=1}^{K} A_n \in \mathcal{F}$.

公理化定义

- 1933 年柯尔莫哥洛夫（Kolmogorov）基于集合论，给出了概率的公理化定义。
- 设 S 为样本空间，$\mathcal{F}$ 是由 S 的某些子集组成的一个事件域。如果对任意事件 $A \in \mathcal{F}$，定义在 $\mathcal{F}$ 上的一个实值函数 $P(A)$ 满足：

  非负性公理 $A \in \mathcal{F} \to P(A) \geq 0,$
  正则性公理 $P(S) = 1,$
  可加性公理 若 $A_1, A_2, \cdots, A_n, \cdots$ 互斥，则 $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$

  则称 $P(A)$ 为事件 A 的概率。三元素 $(S, \mathcal{F}, P)$ 称为概率空间。
- “正则性公理”也叫“规范性公理”，
  “可加性公理”全称是“可列可加性公理”。

公理化定义解读

- 概率的公理化定义描述了概率的本质：概率是事件（集合）的函数。

$$P : \mathcal{F} \mapsto [0, 1] \in \mathbb{R}$$

- 若在事件域 $\mathcal{F}$ 上给出一个函数，只要这个函数满足柯氏公理，就可以称该函数为概率。
- 可以证明历史上出现过的频率定义、几何定义、古典定义、主观定义等，均满足柯氏公理。
  更准确地说，这些定义是概率的不同解释。

性质

- $P(A) \geq 0, P(\emptyset) = 0, P(S) = 1$
- 有限可加性：设 $A_1, A_2, \cdots, A_n$ 两两互斥，则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- $P(\bar{A}) = 1 - P(A) \implies P(A) \leq 1$
- $A \subset B \implies P(B - A) = P(B) - P(A) \implies P(A) \leq P(B)$
- 加法公式，任意两个事件 A, B, C,

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

性质

- $P(A) \geq 0, P(\emptyset) = 0, P(S) = 1$
- 有限可加性：设 $A_1, A_2, \cdots, A_n$ 两两互斥，则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- $P(\bar{A}) = 1 - P(A) \implies P(A) \leq 1$
- $A \subset B \implies P(B - A) = P(B) - P(A) \implies P(A) \leq P(B)$
- 加法公式，任意两个事件 A, B, C,

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

性质

- $P(A) \geq 0, P(\emptyset) = 0, P(S) = 1$
- 有限可加性：设 $A_1, A_2, \cdots, A_n$ 两两互斥，则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- $P(\bar{A}) = 1 - P(A) \implies P(A) \leq 1$
- $A \subset B \implies P(B - A) = P(B) - P(A) \implies P(A) \leq P(B)$
- 加法公式，任意两个事件 A, B, C,

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

性质

- $P(A) \geq 0, P(\emptyset) = 0, P(S) = 1$
- 有限可加性：设 $A_1, A_2, \cdots, A_n$ 两两互斥，则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- $P(\bar{A}) = 1 - P(A) \implies P(A) \leq 1$
- $A \subset B \implies P(B - A) = P(B) - P(A) \implies P(A) \leq P(B)$
- 加法公式，任意两个事件 A, B, C,

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

性质

- $P(A) \geq 0, P(\emptyset) = 0, P(S) = 1$
- 有限可加性：设 $A_1, A_2, \cdots, A_n$ 两两互斥，则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- $P(\bar{A}) = 1 - P(A) \implies P(A) \leq 1$
- $A \subset B \implies P(B - A) = P(B) - P(A) \implies P(A) \leq P(B)$
- 加法公式，任意两个事件 A, B, C,

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

经典定义都是公理定义的解释

直观定义 事件 A 发生的可能性的大小。

频率定义 在大量重复试验下，事件 A 出现的频率会稳定在某个值，将其定义为事件 A 发生的概率。

古典概型 等可能基本事件。

几何概型 测度思想的萌芽。

主观定义 人们曾经对“贝叶斯概率”的误解。

频率

频率的稳定性

- 设在 n 次试验中，事件 A 发生了 m 次，称

$$f_n(A) = \frac{m}{n}$$

为事件 A 发生的频率。

- 稳定性：趋于常数 $P(A)$

$$\lim_{n\to\infty} f_n(A) = P(A)$$

频率解释

非负性 $0 \le f_n(A) \le 1$

正则性 $f_n(S) = 1$

可加性 若 A, B 互斥，则 $f_n(A \cup B) = f_n(A) + f_n(B)$ 可推广到有限个两两互斥事件的和事件

公理定义

$$\lim_{n\to\infty} f_n(A) = P(A)$$

定义的 $P(A)$ 满足柯氏公理

古典概型

定义

- 设随机试验 E 具有下列特点：
  1. 基本事件的个数有限
  2. 每个基本事件等可能性发生

  则称 E 为古典（等可能）概型
- 古典概型中概率：

  n S 中包含基本事件的个数
  k 组成 A 的基本事件的个数

  则 $P(A) = \frac{k}{n}$
- 古典概型的概率满足柯氏公理。
- 古典概型在概率论中占有相当重要的地位，它具有简单、直观的特点，且应用广泛。

古典概型的计算

- 正确判断试验为古典概型，即试验必须有两个特点：
  - 试验的所有可能结果只有有限个；
  - 每个结果发生的可能性相同。
- 恰当选取样本空间 S，并计算 S 中含样本点的个数。
- 计算事件 A 的基本事件数 k。
- 由等可能概型计算公式得到 $P(A) = \frac{k}{n}$。

例：样本空间的重要性

抛一枚硬币三次 $\iff$ 抛三枚硬币一次，判断下面两个样本空间是否符合古典概型。

$$S_1 = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$$

$$S_2 = \{\mathrm{3H}, \mathrm{2H1T}, \mathrm{1H2T}, \mathrm{3T}\}$$

例：圆桌座位

圆桌座位

$n>2$ 个人围圆桌而坐，求 A 和 B 两人相邻而坐的概率。

考虑 A 先坐到某个位置，则 B 有 $n-1$ 个位置可供选择，其中与 A 相邻而坐只有两种可能性。
所以

$$P=\frac{2}{n-1}$$

例：盲盒

盲盒

盲盒中有 a 件正品，b 件次品，按放回与不放回两种方式从中取 m 件，求其中恰有 k 件正品的概率。$m \le a+b, k \le a, k \le m$

- 放回

$$P = \binom{m}{k}\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{m-k}$$

二项分布

- 不放回

$$P = \frac{\binom{a}{k}\binom{b}{m-k}}{\binom{a+b}{m}}$$

超几何分布

例：盲盒

盲盒

盲盒中有 a 件正品，b 件次品，按放回与不放回两种方式从中取 m 件，求其中恰有 k 件正品的概率。$m \leq a+b, k \leq a, k \leq m$

- 放回

$$P = \binom{m}{k}\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{m-k}$$

二项分布

- 不放回

$$P = \frac{\binom{a}{k}\binom{b}{m-k}}{\binom{a+b}{m}}$$

超几何分布

例：生日悖论

生日悖论

教室里有 n 个人，求至少有两人生日相同 (设为事件 A) 的概率

- $\overline{A}$ 为 n 个人的生日均不相同

$$P(\overline{A}) = \frac{P_{365}^{n}}{365^{n}}$$

$$P(A) = 1 - P(\overline{A})$$

```r
prod <- 1
n <- 14
for (i in (365 - n):365) {
    prod <- prod * (i/365)
}
print(1-prod)
```

```
0.252901319763686
```

例：生日悖论

生日悖论

教室里有 n 个人，求至少有两人生日相同 (设为事件 A) 的概率

- $\overline{A}$ 为 n 个人的生日均不相同

$$P(\overline{A}) = \frac{P_{365}^{n}}{365^{n}}$$

$$P(A) = 1 - P(\overline{A})$$

```
prod <- 1
n <- 14
for (i in (365 - n):365) {
    prod <- prod * (i/365)
}
print(1-prod)
```

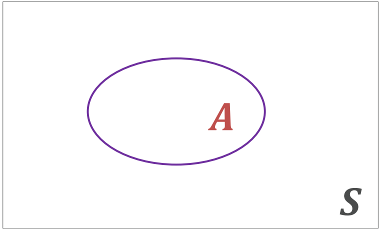
```
0.252901319763686
```

几何概型

定义

当随机试验的样本空间是某个区域，并且任意一点落在度量（长度、面积、体积）相同的子区域是等可能的，则事件 A 的概率可定义为

$$P(A) = \frac{m(A)}{m(S)}$$

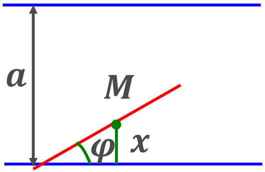


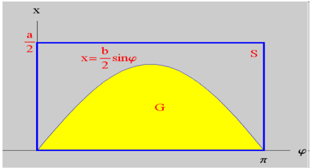
例：投针试验

1777 年，法国科学家蒲丰 (Buffon) 提出了投针试验问题。平面上画有等距离为 $a(>0)$ 的一些平行直线，现向此平面任意投掷一根长为 $b(<a)$ 的针，试求针与任一平行直线相交的概率。

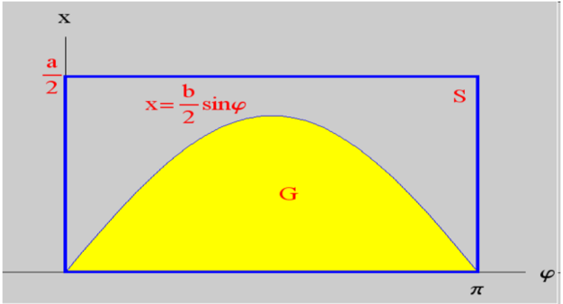
- 以 x 表示针投到平面上时，针的中点 M 到最近一条平行线的距离，ϕ 表示该针与该平行直线夹角，那么针落在平面上的位置可由 (x, ϕ) 完全确定。
- 投针实验的所有可能结果与矩形区域 $S = \{(x, \phi) \mid x \in [0, \frac{a}{2}], \phi \in [0, \pi]\}$ 中的所有点一一对应。
- 所关心的事件 A “针与任一平行直线相交” 发生的充分必要条件为

$$x \in \left[0, \frac{b}{2}\sin\phi\right], \phi \in [0, \pi]$$





例：投针试验



$$P(A) = \frac{m(G)}{m(S)} = \frac{\int_0^\pi \frac{b}{2}\sin\phi \mathrm{d}\phi}{\frac{a\pi}{2}} = \frac{2b}{a\pi}$$

意义

$$P(A) = \frac{2b}{a\pi}$$

根据频率的稳定性，当投针试验次数 n 很大时，算出针与平行直线相交的次数 m，则频率值 $\frac{m}{n}$ 即可作为的近似值代入上式，那么

$$\frac{m}{n} \approx \frac{2b}{a\pi} \implies \pi \approx \frac{2bn}{am}$$

利用上式可计算圆周率 π 的近似值。挑战祖冲之?

自行投针

历史上一些学者的计算结果

学者位争相试验，使用随机投针的方法逼近圆周率。

试验者	时间	针长	投掷次数	相交次数	π的近似值
Wolf	1850	0.8	5000	2532	3.1596
Smith	1855	0.6	3204	1218	3.1554
De Morgan	1860	1.0	600	382	3.137
Fox	1884	0.75	1030	489	3.1595
Lazzerini	1901	0.83	3408	1808	3.1415929
Reina	1925	0.5419	2520	859	3.1795

清小华，2023，针长 $=1$，投掷次数 $=100$，相交次数 $=60$，$\pi \approx 3.3333$

历史上一些学者的计算结果

学者位争相试验，使用随机投针的方法逼近圆周率。

试验者	时间	针长	投掷次数	相交次数	π的近似值
Wolf	1850	0.8	5000	2532	3.1596
Smith	1855	0.6	3204	1218	3.1554
De Morgan	1860	1.0	600	382	3.137
Fox	1884	0.75	1030	489	3.1595
Lazzerini	1901	0.83	3408	1808	3.1415929
Reina	1925	0.5419	2520	859	3.1795

清小华，2023，针长 $= 1$，投掷次数 $= 100$，相交次数 $= 60$，$\pi \approx 3.3333$

推广

- 如果投其他形状的物体到画满平行直线的平面上，结果会如何?
- 例如投一个小三角形（三个边长 $l_1, l_2, l_3 < a$）

$$P(A) = \frac{l_1 + l_2 + l_3}{a\pi}$$

推广

- 如果投其他形状的物体到画满平行直线的平面上，结果会如何?
- 例如投一个小三角形（三个边长 $l_1, l_2, l_3 < a$）

$$P(A) = \frac{l_1 + l_2 + l_3}{a\pi}$$

贝叶斯概率

定义

- ~~主观~~ 又称为贝叶斯概率，主要用于处理不能重复或不能大量重复的随机现象。
- 贝叶斯概率的样本空间中的样本点为一系列“假设（hypotheses）”，即要么为“真”要么为“否”的表述。
  贝叶斯学派将“假设”为真的概率定义为个人“信心程度”的度量：$P(A)$ 是 A 发生（为真）的信心程度。
- 样本空间 S 中的基本“假设”必须是互斥的。任何情况下必须只存在一个基本假设为“真”。
- 贝叶斯概率满足柯氏公理。

伯努利概型

定义

假设单次试验中事件 A 发生的概率为 $p \in (0,1)$。在相同条件下重复试验 n 次，记 A 发生 k 次为事件 $A_k, k = 0, 1, 2, \cdots, n$。定义其概率为

$$P_n(A_k) = \binom{n}{k} p^k (1-p)^{n-k}$$

称为伯努利概型，也称为二项概型。

- 要求 n 次重复试验相互独立。
- 离散样本空间；
- 每个样本点不等概率。

例：动态盲盒

动态盲盒

口袋中有 $n-1$ 个黑球、1 个白球，每次从口袋中随机地摸出一球，并换入一只黑球。求第 k 次取到黑球的概率。

记 A 为“第 k 次取到黑球”，则 A 的对立事件 $\overline{A}$ 为“第 k 次取到白球”。

$$P(\overline{A}) = \left(\frac{n-1}{n}\right)^{k-1}\frac{1}{n} = \frac{(n-1)^{k-1}}{n^k}$$

$$P(A) = 1 - \frac{(n-1)^{k-1}}{n^k}$$

例：动态盲盒

动态盲盒

口袋中有 $n-1$ 个黑球、1 个白球，每次从口袋中随机地摸出一球，并换入一只黑球。求第 k 次取到黑球的概率。

记 A 为"第 k 次取到黑球"，则 A 的对立事件 $\overline{A}$ 为"第 k 次取到白球"。

$$P(\overline{A}) = \left(\frac{n-1}{n}\right)^{k-1}\frac{1}{n} = \frac{(n-1)^{k-1}}{n^k}$$

$$P(A) = 1 - \frac{(n-1)^{k-1}}{n^k}$$

例：动态盲盒二

动态盲盒二

口袋中有 2 个白球，每次从口袋中随机地摸出一球，并换入一只黑球。求第 k 次取到黑球的概率。

例：Pólya 动态盲盒

口袋中有 n 个黑球，m 个白球。任取一球记下颜色后，放回 d 个与取出颜色相同的球。再取出一球，

1. 求最初取出黑球，第二次取出也是黑球的概率；
2. 证明：任何一次取得黑球的概率都是 $\frac{n}{n+m}$，任何一次取得白球的概率都是 $\frac{m}{n+m}$；
3. 证明：第 k 次与第 l 次取出的都是黑球的概率是

$$\frac{n(n+d)}{(n+m)(n+m+d)}.$$

Pólya NB 定律

几乎所有概率论的命题，都可以用 Pólya 动态盲盒

NB = Negative-Binomial distribution 负二项分布

例：Pólya 动态盲盒

口袋中有 n 个黑球，m 个白球。任取一球记下颜色后，放回 d 个与取出颜色相同的球。再取出一球，

1. 求最初取出黑球，第二次取出也是黑球的概率；
2. 证明：任何一次取得黑球的概率都是 $\frac{n}{n+m}$，任何一次取得白球的概率都是 $\frac{m}{n+m}$；
3. 证明：第 k 次与第 l 次取出的都是黑球的概率是

$$\frac{n(n+d)}{(n+m)(n+m+d)}.$$

Pólya NB 定律

几乎所有概率论的命题，都可以用 Pólya 动态盲盒

NB = Negative-Binomial distribution 负二项分布

例：Pólya 动态盲盒

口袋中有 n 个黑球，m 个白球。任取一球记下颜色后，放回 d 个与取出颜色相同的球。再取出一球，

1. 求最初取出黑球，第二次取出也是黑球的概率；
2. 证明：任何一次取得黑球的概率都是 $\frac{n}{n+m}$，
   任何一次取得白球的概率都是 $\frac{m}{n+m}$；
3. 证明：第 k 次与第 l 次取出的都是黑球的概率是

$$\frac{n(n+d)}{(n+m)(n+m+d)}.$$

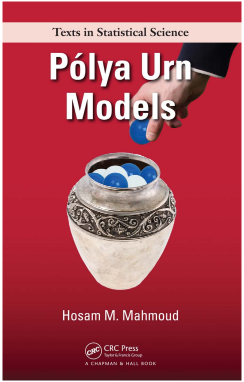
Pólya NB 定律

几乎所有概率论的命题，都可以用 Pólya 动态盲盒

NB $=$ Negative-Binomial distribution 负二项分布

前方有矿

Polya Urn Models, CRC Press 2008



学期结束之前通读全书者，总评 +15%。