

随机变量及其分布

一、条件概率

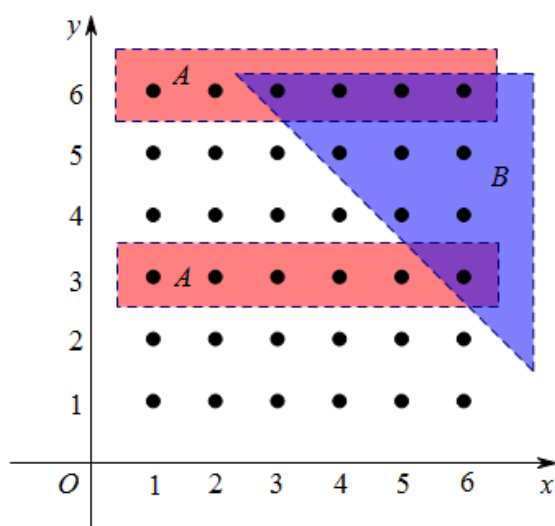
1. 条件概率

在很多实际问题中，需要考虑一个事件在“某事件已发生”这个附加条件下的概率。

我们来看下面的问题：

抛掷红、蓝两颗骰子，设事件 A = “蓝色骰子的点数为3或6”，事件 B = “两颗骰子的点数之和大于8”。

我们用 x 代表抛掷红骰子所得到的点数，用 y 代表抛掷蓝骰子所得到的点数，则这个试验的基本事件空间为 $S = \{(x, y) | x \in \mathbf{N}^*, y \in \mathbf{N}^*, 1 \leq x \leq 6, 1 \leq y \leq 6\}$ 。作出下图：



容易看出，基本事件空间的元素与上图中的点一一对应。所以抛掷红、蓝两颗骰子这一试验的基本事件总数为36。事件 B 包含所包含的基本事件对应图中三角虚线所包围的十个点，所以，事件 B 发生的概率 $P(B) = \frac{10}{36} = \frac{5}{18}$ 。

当一只蓝色骰子的点数为3或6时，事件 B 发生的概率是多少呢？也就是说，要求事件 B 在“事件 A ”已发生这个附加条件下的概率是多少。

事件 A 已发生的所有可能的结果对应图中长条虚线所包围的12个点，其中重叠部分的5个点的“点数之和大于8”。所以事件 B 在“事件 A 已发生”条件下的概率是 $\frac{5}{12}$ 。

从这个例子中看到，事件 B 在“事件 A 已发生”这个附加条件下的概率与没有这个附加条件的概率是不同的。

条件概率定义

对于任何两个事件 A 和 B ，在已知事件 A 发生的条件下，事件 B 发生的概率叫做条件概率，用符号“ $P(B|A)$ ”来表示。

条件概率公式

上图中重叠部分的5个点对应的事件为“事件 A 发生并且事件 B 也发生”，我们把由事件 A 和 B 同时发生所构成的事件 D ，称为事件 A 与 B 的交（或积），记作 $D = A \cap B$ （或 $D = AB$ ）。

容易得到上面的例子中 $P(A) = \frac{1}{3}$ ， $P(A \cap B) = \frac{5}{36}$ ，而：

$$P(B|A) = \frac{5}{12} = \frac{\frac{5}{36}}{\frac{1}{3}} = \frac{P(A \cap B)}{P(A)}.$$

为此，我们有如下的条件概率公式： $P(B|A) = \frac{P(A \cap B)}{P(A)}$ ， $P(A) > 0$ 。

条件概率的求法：

(1) 利用定义，分别求出 $P(A)$ 和 $P(A \cap B)$ ，得 $P(B|A) = \frac{P(A \cap B)}{P(A)}$ 。

(2) 借助古典概型概率公式，先求事件 A 包含的基本事件数，即 $n(A)$ 再求事件 $n(A \cap B)$ ，得 $P(B|A) = \frac{n(A \cap B)}{n(A)}$ 。

例1

- 袋中有10个大小相同但编号不同的球，6个红球和4个白球，无放回地依次摸出2个球，在第一次摸出红球的条件下，第二次也摸到红球的概率为（ ）。
A. $\frac{3}{5}$ B. $\frac{2}{5}$ C. $\frac{1}{10}$ D. $\frac{5}{9}$
- 从1, 2, 3, 4, 5中任取2个不同的数，事件 A = “取到的2个数之和为偶数”，事件 B = “取到的2个数均为偶数”，则 $P(B|A)$ = （ ）。
A. $\frac{1}{8}$ B. $\frac{1}{4}$ C. $\frac{2}{5}$ D. $\frac{1}{2}$

练1

- 篮子里装有2个红球，3个白球和4个黑球。某人从篮子中随机取出两个球，记事件 A = “取出的两个球颜色不同”，事件 B = “取出一个红球，一个白球”，则 $P(B|A)$ = （ ）。
A. $\frac{1}{6}$ B. $\frac{3}{13}$ C. $\frac{5}{9}$ D. $\frac{2}{3}$
- 某地气象台预计，7月1日该地区下雨的概率为 $\frac{4}{15}$ ，刮风的概率为 $\frac{2}{15}$ ，既刮风又下雨的概率为 $\frac{1}{10}$ ，设 A 表示下雨， B 表示刮风，则 $P(A|B)$ = （ ）。

- A. $\frac{1}{2}$ B. $\frac{3}{4}$ C. $\frac{2}{5}$ D. $\frac{3}{8}$

5. 在5道题中有3道理科题和2道文科题，如果不放回地依次抽取2道题，则在第一次抽到理科题的条件下，第二次抽到理科题的概率为（ ）.

- A. $\frac{1}{4}$ B. $\frac{1}{3}$ C. $\frac{1}{2}$ D. $\frac{2}{3}$

6. 某种电路开关闭合后会出现红灯或绿灯闪烁，已知开关第一次闭合后出现红灯的概率为 $\frac{1}{2}$ ，两次闭合后都出现红灯的概率为 $\frac{1}{5}$ ，则在第一次闭合后出现红灯的条件下第二次闭合后出现红灯的概率为（ ）

- A. $\frac{1}{10}$ B. $\frac{1}{5}$ C. $\frac{2}{5}$ D. $\frac{1}{2}$

7. 某地区空气质量监测资料表明，一天的空气质量为优良的概率是0.8，连续两天为优良的概率是0.6，已知某天的空气质量为优良，则随后一天的空气质量为优良的概率是（ ）.

- A. 0.48 B. 0.12 C. 0.75 D. 0.32

2. 事件的独立性

如果事件A是否发生对事件B发生的概率没有影响，即 $P(B|A) = P(B)$ ，这时，我们称事件A、B**相互独立**，并把这两个事件叫做**相互独立事件** .

对于两个独立事件A和B，有如下两个性质：

① A 与 $\bar{B}$ ， $\bar{A}$ 与 B ， $\bar{A}$ 与 $\bar{B}$ 也相互独立；

② $P(A \cap B) = P(A) \times P(B)$.

【补充说明】

“互斥事件”和“相互独立事件”是两个不同的概念，前者表示两个事件不可能同时发生，后者指一个事件是否发生对另一个事件发生的概率没有影响 .

拓展

对于n个事件 $A_1, A_2, \dots, A_n$ ，如果其中任一个事件发生的概率不受其他事件是否发生的影响，则称这n个事件相互独立 .

同样也有： $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \times P(A_2) \times \dots \times P(A_n)$.

例2

8. 抛掷一枚均匀的骰子2次，在下列事件中，与事件“第一次得到6点”不相互独立的是（ ）.

- A. “第二次得到6点” B. “第二次的点数不超过3点”
C. “第二次的点数是奇数” D. “两次得到的点数和是12”

2. 离散型随机变量

如果随机变量 X 的所有可能的取值都能一一列举出来，则称 X 为**离散型随机变量**。

上面所举出的都是离散型随机变量的例子。接下来我们也主要研究离散型随机变量。

【补充说明】

(1) 随机变量是将随机试验的结果数量化的结果，随即变量的每一个取值要与某一个或某几个随机试验的结果相对应；

有些随机试验的记录虽然不具有数量性质，不过我们可以通过对应的手段将试验结果数量化。如抛掷一枚硬币这种随机试验，可能出现两种结果——正面向上和反面向上，虽然这两种结果不具备数量性质但我们可以用 $X = 0$ 表示正面向上， $X = 1$ 表示反面向上，这样就建立起所有可能的试验结果与实数的一个对应关系。

(2) 并不是所有的随机变量的取值都能一一列出，也就是说不是所有的随机变量都是离散型随机变量。如：电灯泡的寿命 X 的可能取值是任意一个非负实数，无法一一列出。

正如上面的例子，如果随机变量可以取得某一区间内的任意一个值，则称这样的随机变量为连续型随机变量。

3. 离散型随机变量的分布列

将离散型随机变量 X 所有可能的取值 x_i 与该取值对应的概率 $p_i (i = 1, 2, \dots, n)$ 列表表示：

X	x_1	x_2	$\dots$	x_i	$\dots$	x_n
P	p_1	p_2	$\dots$	p_i	$\dots$	p_n

我们称这个表为离散型随机变量 X 的**概率分布**，或称为离散型随机变量 X 的**分布列**。

通过我们之前对概率的学习，我们知道：

$$(1) p_i \geq 0, i = 1, 2, 3, \dots, n;$$

$$(2) p_1 + p_2 + \dots + p_n = 1.$$

离散型随机变量的分布列的两个性质是检验一个分布列正确与否的标准；

还可以利用性质(2)间接求出某一个随机变量对应的概率（前提是你对其他随机变量的概率的正确性足够自信）。

例3

14. 引入随机变量后，下列说法正确的有：_____（填写出所有正确的序号）。

- ①随机事件个数与随机变量一一对应；
- ②随机变量与自然数一一对应；

③随机变量的取值是实数 .

练3

15. 下列随机变量中，不是离散随机变量的是 ()
- A. 从10只编号的球 (0号到9号) 中任取一只，被取出的球的号码
- B. 抛掷两个骰子，所得的最大点数
- C. $[0, 10]$ 区间内任一实数与它四舍五入取整后的整数的差值
- D. 一电信局在未来某日内接到的电话呼叫次数

例4

16. 已知随机变量 X 的分布列为

X	1	2	3	4
P	0.2	$0.4 - a$	$0.5 - a$	a

则实数 a 等于 _____ .

17. 随机变量 X 的分布列为 $P(X = k) = \frac{c}{k(k+1)} (k = 1, 2, 3, 4)$ ，其中 c 为常数，则

$$P\left(\frac{1}{2} < X < \frac{5}{2}\right) = \text{_____} .$$

练4

18. 若离散型随机变量 ξ 的概率分布如表所示，则 a 的值为 () .

ξ	-1	1
P	$4a - 1$	$3a^2 + a$

- A. $\frac{1}{3}$ B. -2 C. $\frac{1}{3}$ 或 -2 D. $\frac{1}{2}$

19. 设离散型随机变量 x 的概率分布列为 $p(x = k) = \frac{m}{k+1} (k = 0, 1, 2, 3)$ ，则 $m = ()$.

- A. $\frac{8}{25}$ B. $\frac{12}{25}$ C. $\frac{16}{25}$ D. $\frac{4}{5}$

20. 设随机变量 X 的概率分布列如下图，则 $P(|X - 2| = 1) = \text{_____}$.

X	1	2	3	4
P	$\frac{1}{6}$	$\frac{1}{4}$	m	$\frac{1}{3}$

21. 一袋中装有6个同样大小的黑球，编号分别为1, 2, 3, 4, 5, 6，现从中随机取出3个球，以 X 表示取出球的最大号码，求 X 的分布列 .

三、两点分布

如果随机变量 X 的分布列为

X	1	0
P	p	$1 - p$

其中 $0 < p < 1$ ，则称离散型随机变量 X 服从参数为 p 的**两点分布**，并且称 $P = p(X = 1)$ 为成功概率。两点分布又称**0-1分布**，由于只有两个可能结果的随机试验叫做**伯努利试验**，所以这种分布又称为**伯努利分布**。

例5

22. 下列随机变量中服从两点分布的是（ ）。

- A. 射击一次命中目标的次数
- B. 抛掷三枚骰子，所得的点数之和
- C. 抛掷一枚骰子，所得的点数
- D. 6张卡片上分别标有号码1, 2, 3, 4, 5, 6，从中任取3张，三张卡片中最大的号码

23. 在射击试验中，令 $X = \begin{cases} 1, & \text{射中} \\ 0, & \text{未射中} \end{cases}$ ，如果射中的概率是0.9，则随机变量的分布列为 _____。

练5

24. 设随机变量 X 服从两点分布，若 $P(X = 1) - P(X = 0) = 0.2$ ，则成功概率 $P(X = 1) = ()$ 。

- A. 0.2
- B. 0.4
- C. 0.6
- D. 0.8

25. 袋内有10个白球，5个红球，从中摸出两球，记 $X = \begin{cases} 0, & \text{两球全红,} \\ 1, & \text{两球非全红,} \end{cases}$ ，求随机变量 X 的分布列。

四、二项分布

独立重复试验

如果每次试验，只考虑有两个可能的结果 A 及 $\bar{A}$ ，并且事件 A 发生的概率相同。在相同的条件下，重复地做 n 次试验，各次试验的结果相互独立，那么一般就称它们为 **n 次独立重复试验**。

【补充说明】

(1) 独立重复试验必须满足下列两个特征：

①每次试验的条件都完全相同，相关事件的概率保持不变；

②各次试验的结果互不影响（独立）；

（2）独立重复试验的每次试验只有两个可能的结果 A 与 $\bar{A}$ ，且它们彼此对立，服从两点分布．

一般地，事件 A 在 n 次试验中发生 k 次，共有 C_n^k 种情形，由试验的独立性知 A 在 k 次试验中发生而在其余 $n - k$ 次试验中不发生的概率是 $p^k(1 - p)^{n-k}$ ，所以由概率加法公式知，如果在一次试验中事件 A 发生的概率是 p ，那么在 n 次独立重复试验中，事件 A 恰好发生 k 次的概率为：

$$P_n(k) = C_n^k p^k (1 - p)^{n-k} (k = 0, 1, 2, \dots, n).$$

二项分布

若将事件 A 发生的次数设为 X ，事件 A 发生的概率为 p ，事件 A 不发生的概率为 $q = 1 - p$ ，那么在 n 次独立重复试验中，事件 A 恰好发生 k 次的概率是 $P(X = k) = C_n^k p^k q^{n-k}$ ，其中 $k = 0, 1, 2, \dots, n$ ．于是得到 X 的分布列：

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 p^1 q^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

由于表中的第二行恰好是二项展开式

$(q + p)^n = C_n^0 p^0 q^n + C_n^1 p^1 q^{n-1} + \dots + C_n^k p^k q^{n-k} + \dots + C_n^n p^n q^0$ 各对应项的值，所以称这样的散型随机变量 X 服从参数为 n, p 的二项分布，记作 $X \sim B(n, p)$ ．

例6

26. 已知随机变量 X 服从二项分布 $B(6, \frac{1}{3})$ ，则 $P(X = 2) = ()$ ．

- A. $\frac{13}{16}$ B. $\frac{4}{243}$ C. $\frac{13}{243}$ D. $\frac{80}{243}$

27. 某人射击的命中率是0.9，他射击5次恰有3次命中的概率是()．

- A. $C_5^3 \times 0.9^3 \times 0.1^2$ B. $C_5^3 \times 0.1^3 \times 0.9^2$ C. $1 - (1 - 0.9)^3$ D. 0.9^3

练6

28. 设随机变量 $\xi \sim B(2, p)$ ， $\eta \sim B(3, p)$ ，若 $P(\xi \geq 1) = \frac{5}{9}$ ，则 $P(\eta \geq 2)$ 的值为()．

- A. $\frac{20}{27}$ B. $\frac{8}{27}$ C. $\frac{7}{27}$ D. $\frac{1}{27}$

29. 乒乓球单打比赛在甲、乙两名运动员间进行，比赛采用7局4胜制（即先胜4局者获胜，比赛结束），假设两人在每一局比赛中获胜的可能性相同，那么甲以4比2获胜的概率为()．

- A. $\frac{5}{64}$ B. $\frac{15}{64}$ C. $\frac{5}{32}$ D. $\frac{5}{16}$

30. 一袋中有5个白球，3个红球，现从袋中往外取球，每次任取一个记下颜色后放回，直到红球出现10次时停止，设停止时共取了 ξ 次球，则 $P(\xi = 12)$ 等于()．

A. $C_{12}^{10} \left(\frac{3}{8}\right)^{10} \left(\frac{5}{8}\right)^2$ B. $C_{11}^9 \left(\frac{3}{8}\right)^{10} \left(\frac{5}{8}\right)^2$ C. $C_{11}^9 \left(\frac{5}{8}\right)^9 \left(\frac{3}{8}\right)^2$ D. $C_{11}^9 \left(\frac{3}{8}\right)^9 \left(\frac{5}{8}\right)^2$

31. 已知每次试验的成功概率为 $\frac{2}{3}$ ，重复进行试验直至第5次才能3次成功的概率为（ ）。

A. $\frac{8}{243}$ B. $\frac{80}{243}$ C. $\frac{8}{81}$ D. $\frac{16}{81}$

32. 某次知识竞赛规则如下：在主办方预设的固定顺序的5个问题中，选手若能正确回答出三个问题，即停止答题，晋级下一轮；否则不能晋级。假设某选手正确回答每个问题的概率都是 $\frac{2}{3}$ ，且每个问题回答的正确与否都相互独立。

(1) 设“该选手连续答对三道题晋级下一轮”为事件A，求事件A发生的概率。

(2) 记该选手在本轮中答对问题的个数为随机变量 ξ ，求随机变量 ξ 的分布列和期望。

五、超几何分布

【例】某校组织一次认识大自然夏令营活动，有10名同学参加，其中有6名男生，4名女生，为了活动需要，要从这10名同学中随机抽取3名同学去采集自然标本，那么其中恰有一名女生的概率有多大？

【分析】

记X为所选取的3名同学中女生的个数。根据组合数知识，从10名同学中任意选出3名同学，共有 $C_{10}^3 = 120$ 种不同的选法。由分步乘法计数原理得到恰有1名女生的基本事件为 $C_4^1 C_6^2 = 60$ ，因此其中恰有1名女生的概率为：

$$P(X=1) = \frac{1}{2}.$$

进一步我们还可得如下分布列：

X	0	1	2	3
P	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

可见结果全部是女生的概率非常的小，这一现象有其独特的使用价值，后面我们会提到。

从上面这个实例中我们可以抽取出“超几何分布”的模型。

一般地，设有总数为N件的两类物品，其中第一类物品有M件，从所有物品中任取n件($n \leq N$)，这n件中所含第一类物品的件数X是一个离散型随机变量，它取值为m时的概率为

$$P(X=m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \quad (0 \leq m \leq l, l \text{ 为 } n \text{ 和 } M \text{ 中较小的一个}).$$

我们称离散型随机变量X的这种形式的概率分布为超几何分布，也称X服从参数为N, M, n的超几何分布。

【补充说明】

- (1) 超几何分布的概率计算公式可由分步乘法计数原理推导得出；
 (2) 当抽取的物品的件数不大于于总体中第一类物品的件数（即 $n \leq M$ ）时， k 的最大值为 n ；
 当抽取的物品的件数大于总体中第一类物品的件数（即 $n > M$ ）时， k 的最大值为 M 。

例7

33. 盒子中有大小相同的3只白球，1只黑球，若从中随机地摸出两只球，两只球颜色不同的概率是 _____。
34. 设袋中有80个红球，20个白球，若从袋中任取10个球，则其中恰有6个红球的概率为（ ）。
- A. $\frac{C_{80}^4 \cdot C_{10}^6}{C_{100}^{10}}$ B. $\frac{C_{80}^6 \cdot C_{10}^4}{C_{100}^{10}}$ C. $\frac{C_{80}^4 \cdot C_{20}^6}{C_{100}^{10}}$ D. $\frac{C_{80}^6 \cdot C_{20}^4}{C_{100}^{10}}$
35. 一个盒子里装有相同大小的10个黑球，12个红球，4个白球，从中任取2个，其中白球的个数记为 X ，则下列概率等于 $\frac{C_{22}^1 C_4^1 + C_{22}^2}{C_{26}^2}$ 的是（ ）。
- A. $P(0 < X \leq 2)$ B. $P(X \leq 1)$ C. $P(X = 1)$ D. $P(X = 2)$

练7

36. 某学习小组有4名男生和3名女生，若从中随机选出2名同学代表该小组参加知识竞赛，则选出的2名同学中恰好1名男生1名女生的概率为 _____。
37. 有20件产品，其中15件合格品，5件次品。现从中任意选取10件产品，用 X 表示这10件产品中次品的件数，下列概率中等于 $\frac{C_5^3 C_{15}^7}{C_{20}^{10}}$ 的是（ ）。
- A. $P(X \leq 3)$ B. $P(X = 3)$ C. $P(X = 7)$ D. $P(X \leq 7)$
38. 已知一个袋中装有大小相同的黑球、白球和红球，共有10个球，从袋中任意取出1个球，得到黑球的概率是 $\frac{2}{5}$ ，则从中任意取出2个球，得到的都是黑球的概率是 _____。
39. 盒中共有9个球，其中有4个红球，3个黄球和2个绿球，这些球除颜色外完全相同，从盒中一次随机抽出4个球，其中红球、黄球、绿球的个数分别为 x_1, x_2, x_3 ，随机变量 X 表示 x_1, x_2, x_3 的最大数，求 X 的分布列。
40. 某学校为了了解在校同学们对学校某一政策的看法，学校进行了调查，同时选三个班，同学们的看法情况如下：

对学校的看法	非常好，奠定了我一生成长的起点	很好，我的中学很快乐很充实
A班人数比例	$\frac{1}{2}$	$\frac{1}{2}$

B 班人数比例	$\frac{2}{3}$	$\frac{1}{3}$
C 班人数比例	$\frac{3}{4}$	$\frac{1}{4}$

- (1) 从这三个班中各选一个同学，求恰好有2人认为学校“非常好”的概率（用比例作为相应概率）。
- (2) 若在 B 班按所持态度分层抽样，抽取9人，在这9人中任意选取3人，认为学校“非常好”的人数记为 ξ ，求 ξ 的分布列。