

第三章 随机变量及其分布（上）

第一节 随机变量和分布函数

在随机现象中，有很大一部分问题与数值发生关系。

例如，

- (1)在产品检验问题中出现的废品数；
- (2)在车间供电问题中某一时刻正在工作的车床数；
- (3)测量的误差；
- (4) 灯泡的寿命等

都与数值有关。

因此，在随机试验中，我们的观测对象常常是一个或若干随机取值的变量。

有些初看起来与数值无关的随机现象，也常常能用数值来描述。

例如，在掷一枚硬币问题中，每次出现的结果为正面（记为H）或反面（记为T），与数值没有关系，但是我们可以用下面方法使它与数值联系起来，当出现正面时对应数“1”，而出现反面时对应数“0”，

即相当于引入一个定义在样本空间 $\Omega = \{H, T\}$ 上的变量 $X(\omega)$ ，其中

$$X(\omega) = \begin{cases} 1, \omega = H \\ 0, \omega = T \end{cases}$$

由于试验结果的出现是随机的，因而 $X(\omega)$ 的取值也是随机的。

通过以上的分析，我们可以看到：

(1) 一类试验的结果，自然地对应着一个实数。试验的结果就是样本点 ω ，即 ω 是一个数，此时我们就让 $X(\omega) = \omega$ 。

(2)另一类试验的结果需要人为地建立试验结果与数值的关系。

由此可见，无论是那一种情况，都是试验结果（即样本点 ω ）和实数 $X(\omega)$ 之间的一个对应关系。

实际上， $X(\omega)$ 是样本点 ω 的函数，此函数定义在样本空间 Ω 上，取值为实数。

由此我们引入随机变量的定义。

定义3-1 设随机试验的样本空间为 Ω 。
若对于每一个 $\omega \in \Omega$ ，有一个实数 $X(\omega)$ 与之对应，这样就得到一个定义在 Ω 上的实值单值函数 $X(\omega)$ ，则称 $X(\omega)$ 为随机变量（**random variable**），简记为 X 。

随机变量通常用字母 X, Y, X_k 或 ξ, η ，
等表示。

为什么在变量前加上随机两字呢？

主要是由于在一次试验中，若出现了样本点 ω ，则 X 就取 $X(\omega)$ ；

又由于在一次试验中究竟出现哪一个样本点带有随机性，因而 $X(\omega)$ 的取值也就带有随机性。

我们加上随机两字以区别高等数学中的函数。

注：随机变量 $X(\omega)$ 不是微积分研究的函数

现在的问题是前二章所讨论的“随机事件”和现在讨论的“随机变量”有什么关系呢？

设随机试验的样本空间为 Ω 。若 A 为 Ω 的某个随机事件，则定义

$$X = X(\omega) = \begin{cases} 1, & A \text{发生} (\omega \in A) \\ 0, & A \text{不发生} (\omega \notin A) \end{cases}$$

这样就有

$$A = \{\omega | X(\omega) = 1\}$$

通常将 $\{\omega|X(\omega)=1\}$ 记为 $\{X=1\}$,
即

$$A = \{X = 1\}$$

从而可得出：任意事件都可以用一个随机变量表示。

相应地，求事件 A 的概率 $P(A)$ 等价于求随机变量 X 等于1的概率，即

$$P(A) = P(\{X = 1\}) \triangleq P(X = 1)$$

这样一来，事件的研究就可纳入随机变量的研究之中，所有关于事件的性质和定理都可搬过来用。

从现在起，我们的研究对象主要集中在随机变量及它的分布上。

在随机试验中，我们的观测对象常常是一个随机取值的变量 x （即随机变量 X ），观测的目的是要了解它取各可能值或在各个范围内取值的概率。

这些概率描述了 x 取值的规律，因而掌握所有这些概率就成为我们研究的中心任务。

为了对随机变量有直观的了解，
先举一个例子

杨勇制作

34页例3-1

例1 有五件产品，其中两件是次品（用 a_1, a_2 表示），三件是正品（用 b_1, b_2, b_3 表示）。从中任意取出两件，此时随机试验的样本空间为

$$\Omega = \left\{ (a_1, a_2), (a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), \right. \\ \left. (a_2, b_2), (a_2, b_3), (b_1, b_2), (b_1, b_3), (b_2, b_3) \right\}$$

我们将 Ω 中的样本点依次记为 $\omega_1, \omega_2, \dots, \omega_{10}$

考虑抽取的两件产品中次品的个数 X ,
显然 X 是定义在 Ω 上的一个随机变量,
即 $X = X(\omega)$ 。可具体表示为

$$X(\omega) = \begin{cases} 0, & \text{当 } \omega = \omega_8, \omega_9, \omega_{10} \text{ 时} \\ 1, & \text{当 } \omega = \omega_2, \dots, \omega_7 \text{ 时} \\ 2, & \text{当 } \omega = \omega_1 \text{ 时} \end{cases}$$

其中 $\omega_2 = (a_1, b_1)$, $X(\omega_2) = 1$

一般地，若出现了样本点 $\omega \in \Omega$ ，则 X 就取对应的值 $X(\omega)$ 。

由于在每次试验中究竟出现哪一个样本点带有随机性，所以， X 的取值也就带有随机性。

X 的所有可能取值为 0, 1, 2，
令 $A_i =$ “抽取的两件产品中有 i 件次品”，

$$i = 0, 1, 2$$

则

$$A_0 = \{\omega_8, \omega_9, \omega_{10}\} = \{\omega \mid X(\omega) = 0\} = \{X = 0\}$$

$$A_1 = \{\omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7\} = \{\omega \mid X(\omega) = 1\} = \{X = 1\}$$

$$A_2 = \{\omega_1\} = \{\omega \mid X(\omega) = 2\} = \{X = 2\}$$

由于 Ω 为古典概型，因而

$$P(X = 0) = P(A_0) = \frac{3}{10}, \quad P(X = 1) = P(A_1) = \frac{6}{10},$$

$$P(X = 2) = P(A_2) = \frac{1}{10}$$

这就是随机变量 X 取值的规律，通常记为

X	0	1	2
p	$\frac{3}{10}$	$\frac{3}{5}$	$\frac{1}{10}$

称它为 X 的概率分布。

所有其他的概率都可以此概率分布中得到，例如

$$P(X = 4) = P(\emptyset) = 0,$$

$$P(0.5 < X \leq 2) = P(X = 1) + P(X = 2) = \frac{7}{10},$$

其中 $\{X = 1\}$ 和 $\{X = 2\}$ 互不相容。

由以上讨论可知，仅知道随机变量的取值是不够的，我们还要了解随机变量取值的规律（概率）。为此我们引入随机变量分布函数的定义。

定义3-2 设 X 是一个随机变量, x 是任意实数, 概率 $P(X \leq x)$ 是 x 的函数, 则称此函数为随机变量 X 的分布函数 (distribution function), 记为 $F(x)$ 。
即

$$F(x) = P(X \leq x) = P(\{\omega | X(\omega) \leq x\})$$

注: 分布函数 $F(x)$ 是微积分研究的函数。

分布函数 $F(x)$ 具有下列四个基本性质:

$$(1) 0 \leq F(x) \leq 1, x \in \mathbb{R}$$

(2) 若 $x_1 < x_2$, 则 $F(x_1) \leq F(x_2)$, 即单调不减

$$(3) F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$$

$$(4) \lim_{x \rightarrow x_0^+} F(x) = F(x_0), \forall x_0 \in \mathbb{R}, \text{ 即右连续}$$

证明: (1) 由于 $F(x)$ 是事件 $\{X \leq x\}$ 的概率,

显然结论成立

$$(2) \text{ 由于 } \{X \leq x_2\} = \{X \leq x_1\} + \{x_1 < X \leq x_2\}$$

则由可加性 $P(X \leq x_2) = P(X \leq x_1) + P(x_1 < X \leq x_2)$

即 $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$

又由概率的非负性知,

$$F(x_1) \leq F(x_2)$$

(3)做直观的说明

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} P(X \leq x) = P(\emptyset) = 0$$

$$F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} P(X \leq x) = P(\Omega) = 1$$

(4)做直观的说明

$$x > x_0, P(x_0 < X \leq x) = F(x) - F(x_0)$$

$$\begin{aligned}\lim_{x \rightarrow x_0^+} (F(x) - F(x_0)) &= \lim_{x \rightarrow x_0^+} P(x_0 < X \leq x) \\ &= P(x_0 < X \leq x_0) = P(\emptyset) = 0\end{aligned}$$

所以, $F(x)$ 右连续

$$\text{注: } x < x_0, P(x < X \leq x_0) = F(x_0) - F(x)$$

$$\begin{aligned}\lim_{x \rightarrow x_0^-} |F(x) - F(x_0)| &= \lim_{x \rightarrow x_0^-} P(x < X \leq x_0) \\ &= P(X = x_0)\end{aligned}$$

分布函数一定具有以上四个基本性质。
反之，可以证明：若某一个函数 $F(x)$
具有以上四个基本性质，则 $F(x)$ 一
定可作为某一个随机变量的分布函数。
由此可知，这四个基本性质完全刻画
了分布函数的本质特性。

在性质（1）中，我们证明了一个公式：

当 $x_1 < x_2$ 时，有

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$$

再利用概率的运算，就可得到其他事件的概率。例如，

$$P(X < a) = \lim_{x \rightarrow a^-} P(X \leq x) = \lim_{x \rightarrow a^-} F(x) = F(a^-)$$

$$P(X = a) = P(X \leq a) - P(X < a) = F(a) - F(a^-)$$

$$P(X > a) = 1 - P(X \leq a) = 1 - F(a)$$

$$P(a < X < b) = P(X < b) - P(X \leq a) = F(b^-) - F(a)$$

$$P(a \leq X < b) = P(a < X < b) + P(X = a) = F(b^-) - F(a^-)$$

$$P(a \leq X \leq b) = P(X = a) + P(a < X \leq b) = F(b) - F(a^-)$$

由以上讨论可知，我们只要知道一个随机变量的分布函数，就能知道该随机变量取值于任何一个区间，乃至取任一实数值的概率。

所以，随机变量的分布函数是全面地描述了随机变量的情况。

只要知道分布函数，也就把握住了随机变量取值的统计规律。由此可领会到分布函数对于随机变量的重要意义。

顺便指出，有些书中将分布函数定义为

$$F(x) = P(X < x)$$

这与我们的定义无本质区别。在此定义下，上述四个基本性质中（1），

（2），和（3）同样成立，性质（4）则由“右连续”变为“左连续”。

虽然分布函数有些不同，但用此分布函数求出的概率和用前一个定义求的概率是一样的。

随机变量的分类:

- (1) 离散型随机变量: 随机变量仅取数轴上的有限个或可列个孤立点;
- (2) 连续型随机变量: 随机变量的可能取值充满数轴上的一个或若干区间;
- (3) 奇异型随机变量: 既不是离散型随机变量, 也不是连续型随机变量。在理论上很有价值而实际问题中很少有应用。本课程不作讨论。

分布函数可以描述一切随机变量，具体到离散性和连续性随机变量还可以用更方便的方法来描述。

第二节 离散型随机变量及其分布

一.离散型随机变量和概率分布

定义3-3 若随机变量 X 所有可能取值为 $x_1, x_2, \dots, x_n, \dots$ (可以是有限个, 也可以是可列个), 则称随机变量 X 为离散型随机变量 (discrete random variable)。

注: $x_1, x_2, \dots, x_n, \dots$ 是两两互不相等的

例如，从一批产中抽取 n 件，抽到的次品个数 X 只能取有限个可能值 $0, 1, \dots, n$

对于一个离散型随机变量，要掌握它的统计规律，首先要了解它的所有可能取值，除此之外，更主要的是要了解它取各可能值的概率。

定义3-4 设离散型随机变量 X 的所有可能取值为 $x_1, x_2, \dots, x_n, \dots$, 且 p_k 是 X 取 x_k 的概率, $k = 1, 2, \dots, n, \dots$, 即

$$P(X = x_k) = p_k \quad , \quad k = 1, 2, \dots, n, \dots$$

则称上式为离散型随机变量 X 的概率分布 (probability distribution) 或分布律 (distribution law)

通常还可以形象地表示为

X	x_1	x_2	\cdots	x_n	\cdots
P	p_1	p_2	\cdots	p_n	\cdots

或

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

对于离散型随机变量 X ， X 落在任何区间或等于任何一点的概率都可以用概率分布表示，由此可知，**概率分布完整地描述了离散型随机变量的统计规律。**

概率分布具有下列两个基本性质：

$$(1) p_k \geq 0, k = 1, 2, \cdots, n, \cdots;$$

$$(2) \sum_{k=1}^{\infty} p_k = 1。$$

证明：（1）由概率的非负性知，显然成立。

（2）由于 $x_1, x_2, \dots, x_n, \dots$ 是 X 所有可能取值，因此

$$\{X = x_1\} + \{X = x_2\} + \dots + \{X = x_n\} + \dots = \Omega$$

再由概率的可列可加性知，

$$P(X = x_1) + P(X = x_2) + \dots + P(X = x_n) + \dots = 1$$

即

$$\sum_{k=1}^{\infty} p_k = 1$$

概率分布一定具有以上两个基本性质。
反之，若一串 $p_k (k = 1, 2, \dots, n, \dots)$ 具有以上两个性质，则 $p_k (k = 1, 2, \dots, n, \dots)$ 一定可作为某一离散型随机变量的概率分布。

I 是一个区间

$$P(X \in I) = \sum_{i: x_i \in I} p_i$$

37页例3-3

例2 袋中有一个白球和二个黑球，每次从中任取一个球，直到取得白球为止，求取球次数的概率分布。

假定：

- (1) 每次取出的黑球不放回；
- (2) 每次取出的黑球仍放回。

解：用 X 表示取球次数。

- (1) X 所有可能取值为 1, 2, 3 。

令 A_k = “第 k 次抽到白球” , $k = 1, 2, 3$, 则

$$P(X = 1) = P(A_1) = \frac{1}{3}$$

$$P(X = 2) = P(\bar{A}_1 A_2)$$

$$= P(\bar{A}_1)P(A_2|\bar{A}_1)$$

$$= \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$$

$$P(X = 3) = P(\bar{A}_1 \bar{A}_2 A_3)$$

$$= P(\bar{A}_1)P(\bar{A}_2|\bar{A}_1)P(A_3|\bar{A}_1\bar{A}_2)$$

$$= \frac{2}{3} \times \frac{1}{2} \times 1 = \frac{1}{3}$$

则

$$P(X = k) = \frac{1}{3}, k = 1, 2, 3$$

或者表示为

X	1	2	3
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

(2) X 所有可能取值为 $1, 2, \dots$ 。

令 $A_k =$ “第 k 次抽到白球”， $k = 1, 2, \dots$ ，则

$$P(X = i) = P(\bar{A}_1 \cdots \bar{A}_{i-1} A_i)$$

$$= P(\bar{A}_1) \cdots P(\bar{A}_{i-1}) P(A_i)$$

$$= \left(\frac{2}{3}\right)^{i-1} \cdot \left(\frac{1}{3}\right), i = 1, 2, \cdots$$

其中由于采用有放回抽球，所以 A_1, A_2, \cdots, A_i 相互独立。即 X 的概率分布为

$$P(X = i) = \frac{1}{3} \left(\frac{2}{3}\right)^{i-1}, \quad i = 1, 2, \cdots$$

现在我们来讨论概率分布和分布函数如何相互表示。

已知概率分布求分布函数：

设 $P(X = x_i) = p_i$, $i = 1, 2, \dots, n, \dots$

且 $x_1 < x_2 < \dots < x_n < \dots$, 则

$$F(x) = P(X \leq x) = \sum_{i: x_i \leq x} p_i$$

$$= \left\{ \begin{array}{ll} 0 & , x < x_1 \\ p_1 & , x_1 \leq x < x_2 \\ p_1 + p_2 & , \quad x_2 \leq x < x_3 \\ \cdots & , \quad \cdots \\ p_1 + p_2 + \cdots + p_{n-1}, x_{n-1} \leq x < x_n \\ \cdots & , \quad \cdots \end{array} \right.$$

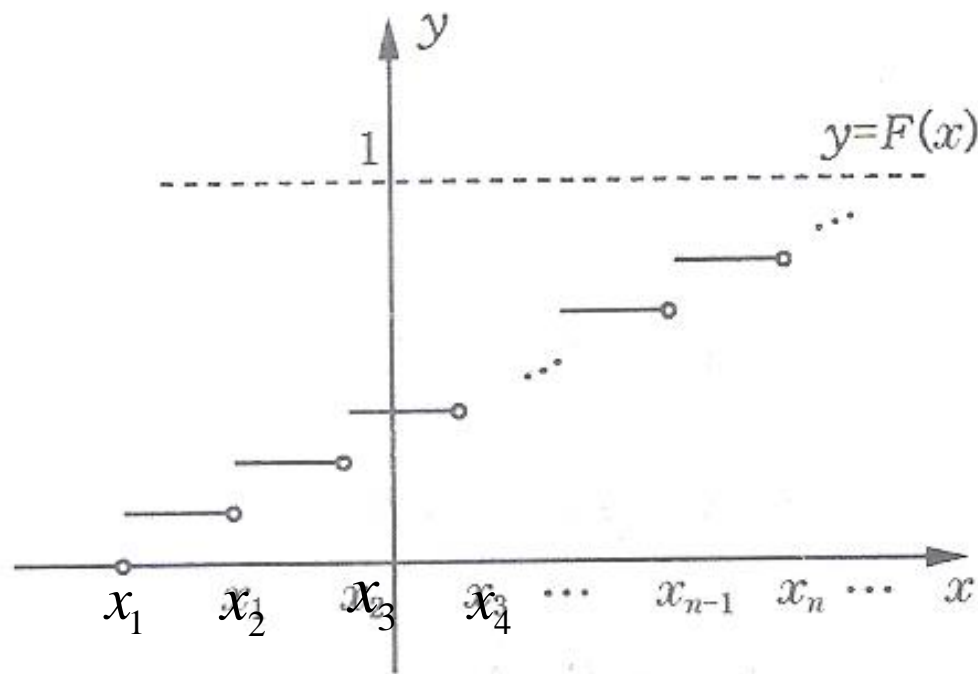


图 3—1 离散型随机变量的分布函数

离散性随机变量分布函数图形的重要特征：
阶梯型

已知分布函数求概率分布：

设分布函数为 $F(x)$ ， 则

$$\begin{cases} p_i = F(x_i) - F(x_{i-1}) \\ p_1 = F(x_1) \end{cases}, \quad i = 2, 3, \dots, n, \dots$$

38页例3-4

例3 设随机变量 X 的概率分布为

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

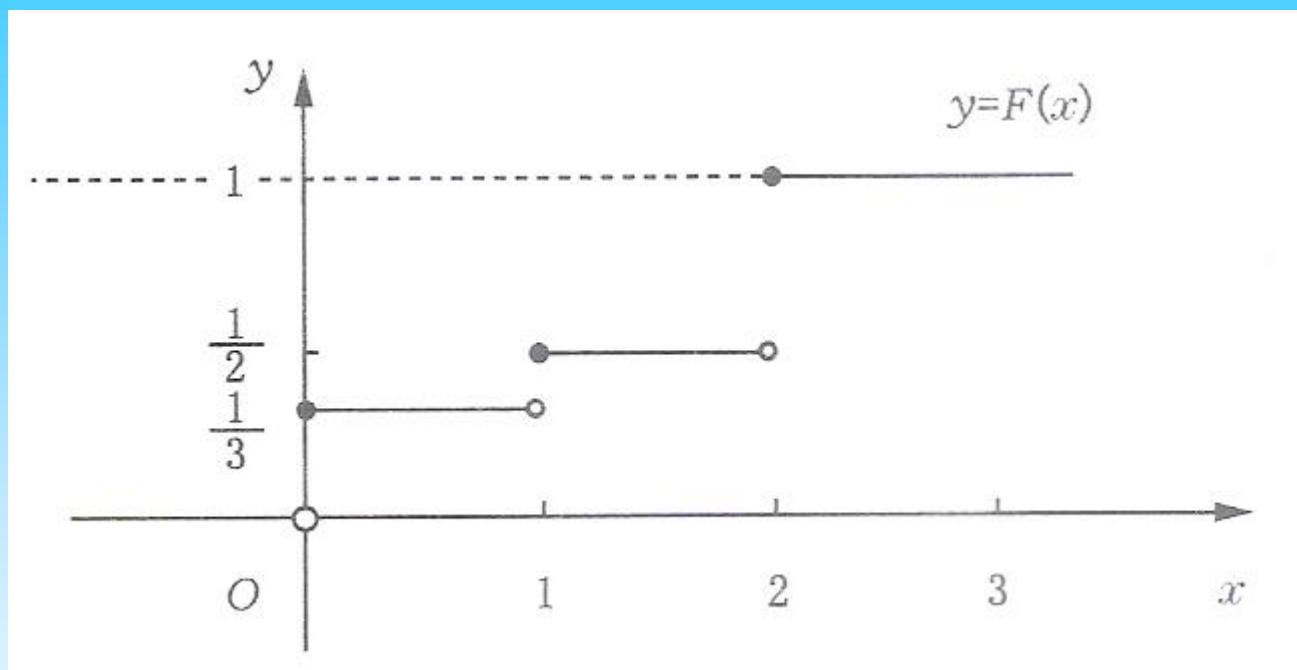
求 (1) X 的分布函数;

(2) $P\left(X < \frac{1}{2}\right)$ 和 $P\left(0 \leq X \leq \frac{3}{2}\right)$ 。

解：（1）

$$F(x) = P(X \leq x) = \sum_{i: x_i \leq x} p_i$$
$$= \begin{cases} 0, & x < 0 \\ \frac{1}{3}, & 0 \leq x < 1 \\ \frac{1}{3} + \frac{1}{6}, & 1 \leq x < 2 \\ \frac{1}{3} + \frac{1}{6} + \frac{1}{2}, & x \geq 2 \end{cases}$$

$$= \begin{cases} 0 & , x < 0 \\ \frac{1}{3}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$



注：若已知分布函数 $F(x)$ ，则

$$p_1 = F(0) = \frac{1}{3},$$

$$p_2 = F(1) - F(0) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6},$$

$$p_3 = F(2) - F(1) = 1 - \frac{1}{2} = \frac{1}{2}$$

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

$$(2) \quad P\left(X < \frac{1}{2}\right) = P(X = 0) = \frac{1}{3},$$

$$\text{或} \quad P\left(X < \frac{1}{2}\right) = F\left(\frac{1}{2}^{-}\right) = \frac{1}{3};$$

$$P\left(0 \leq X \leq \frac{3}{2}\right) = P(X = 0) + P(X = 1) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

或

$$P\left(0 \leq X \leq \frac{3}{2}\right) = F\left(\frac{3}{2}\right) - F(0^-) = \frac{1}{2} - 0 = \frac{1}{2}$$

二. 常用离散型随机变量的分布

1. 0—1分布

定义3-5 设随机变量的概率分布为

$$P(X = k) = p^k (1-p)^{1-k}, \quad k = 0, 1$$

或写为

X	0	1
P	$1-p$	p

其中 $0 < p < 1$ ，则称随机变量 X 服从 0—1 分布（0—1 distribution）或两点分布（two-point distribution）。

设随机试验中事件 A 发生的概率为 $p(0 < p < 1)$ ， 令

$$X = \begin{cases} 1, & \text{当} A \text{发生时} \\ 0, & \text{当} A \text{不发生时} \end{cases}$$

则 X 服从 0—1 分布。

由此可知，任何一个只有两种可能结果的随机试验，都可以用一个服从0—1分布的随机变量来描述。

例如，从一批产中随机地抽取一个产品进行观测，并设

$$X = \begin{cases} 1, & \text{产品为次品} \\ 0, & \text{产品为正品} \end{cases}$$

且 $P(X = 1) = p, 0 < p < 1$ ，则 X 就表示这一次观测中次品的个数。显然 X 服从0—1分布，其中 p 就是次品率。

2. 二项分布

定义3-6 若随机变量 X 的所有可能取值为 $0, 1, \dots, n$ ，且它的概率分布为

$$p_k = P(X = k) = C_n^k p^k q^{n-k}, k = 0, 1, \dots, n,$$

其中 $q = 1 - p$ ， $0 < p < 1$ ，则称随机变量 X 服从二项分布(binomial distribution)，记作 $X \sim B(n, p)$

我们容易验证二项满足：

$$(1) \quad p_k \geq 0, k = 0, 1, \dots, n$$

$$(2) \quad \sum_{k=0}^n p_k = \sum_{k=0}^n C_n^k p^k q^{n-k} = (p+q)^n = 1$$

由于 $p_k = C_n^k p^k q^{n-k}$ 恰好是二项式 $(q+p)^n$ 展开式中的第 $k+1$ 项，所以二项分布由此得名。

我们还常常记

$$b(k; n, p) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \dots, n$$

我们回想第二章贝努利概型

求在 n 次独立试验中事件 A 发生 k 次的概率。

$B_k = “n$ 次独立试验中事件 A 发生 k 次”

$$P(B_k) = C_n^k p^k q^{n-k}, k = 0, 1, \dots, n$$

请问怎么设随机变量 X 用来表示 B_k ?

$X = “n$ 次独立试验中事件 A 发生的次数”

则

$$B_k = \{X = k\}$$

$$P(X = k) = P(B_k) = C_n^k p^k q^{n-k}, k = 0, 1, \dots, n$$

由以上讨论知，二项分布的背景是 **n 重贝努利概型**，定义

X = “ n 次独立试验中事件A 发生的次数”

则 X 服从二项分布，即

$$X \sim B(n, p)$$

特别地，二项分布 $B(n, p)$ 中当 $n = 1$ 时的特殊分布为**0—1分布**。

二项分布是离散型分布中最重要的分布之一，它概括了许多实际问题，很有实用价值。

40页例3-5

例4 已知一公司生产某零件的次品率为0.01，并设各零件是否为次品是相互独立的。该公司将每10个零件装成一盒出售，并承诺若发现某盒内次品多于1个则可退款，问售出的各盒零件将被退回公司的概率为多少？

解：设 X 表示某盒零件中次品的个数，又各零件是否为次品是相互独立的，则 X 显然服从二项分布，即

$$X \sim B(10, 0.01)$$

$$P(X = k) = C_{10}^k 0.01^k 0.99^{10-k}, k = 0, 1, \dots, 10$$

又因为“售出的各盒零件被退回公司”等价于事件 $\{X > 1\}$ ，所以售出的各盒零件被退回公司的概率为

$$P(X > 1) = \sum_{k=2}^{10} P(X = k)$$

$$= 1 - \sum_{k=0}^1 P(X = k)$$

$$= 1 - C_{10}^0 0.01^0 \cdot 0.99^{10} - C_{10}^1 0.01 \cdot 0.99^9$$

$$\approx 0.00427$$

41页例3-6

例5 某一大学的校网球队与该校某系网球队举行对抗赛。一般地，校队的实力比系队为强，每个校队队员获胜的概率为**0.55**。现有校和系双方代表商讨对抗赛的比赛方式，提出了三种方案供选择：

(1) 双方各出**3**人；

(2) 双方各出**5**人；

(3) 双方各出**7**人；

三种方案中均以比赛中获胜人数多的一方为胜。问对系队而言，哪一种方案较为有利？

解：

设系队获胜队员人数为 x ，系队队员获胜概率为 **0.45**，一般可以认为队员间比赛的胜负结果是相互独立的，则 x 服从二项分布。因此，系队获胜的概率分别为：

$$(3) P(X \geq 4) = \sum_{i=4}^7 C_7^i 0.45^i 0.55^{7-i}$$
$$\approx 0.39171$$

由此可知，第一种方案对系队最有利，也即对校队最不利。

性质：若 $X \sim B(n, p)$, 则存在 X_1, X_2, \dots, X_n 独立，且 X_1, X_2, \dots, X_n 都服从参数为 p 的 0-1 分布，使得

$$X = \sum_{k=1}^n X_k$$

此性质反过来也成立。

证明：二项分布是以 n 重贝努利概型为背景

X = “ n 次独立试验中事件 A 发生的次数”

令

$$X_k = \begin{cases} 1, & \text{第}k\text{次试验事件}A\text{发生} \\ 0, & \text{第}k\text{次试验事件}A\text{没发生} \end{cases}, k = 1, 2, \dots, n$$

X_1, X_2, \dots, X_n 独立,

且都服从参数为 p 的0-1分布

请问

$$\sum_{k=1}^n X_k = X_1 + X_2 + \cdots + X_n$$

表示什么意思？

$\sum_{k=1}^n X_k$ 表示 n 次独立试验中事件 A 发生的次数

$$X = \sum_{k=1}^n X_k$$

3. 普阿松分布

定义3-7 若随机变量 X 的取值为非负整数，且其概率分布为

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \dots,$$

其中 $\lambda > 0$ 是某个常数，则称随机变量 X 服从普阿松分布（Poisson distribution），记作

$$X \sim P(\lambda)$$

我们容易验证 p_k 满足:

$$(1) \quad p_k \geq 0, k = 0, 1, 2, \dots;$$

$$(2) \quad \sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

普阿松分布分布是概率论中最重要的概率分布之一。

由于许多实际问题中的随机变量都可以用普阿松分布来描述，从而使得普阿松分布对于概率论的应用来说，有着很重要的作用；而概率论理论的研究又表明普阿松分布在理论上也有其特殊重要的地位。

42页例3-7

例6 一家商店销售某一型号的彩色电视机。由该商店过去的销售记录知道，该型号彩色电视机每月的销售数可以用参数 $\lambda=10$ 的普阿松分布来描述。为了以 **95%**以上的把握保证不脱销，问商店在月底至少应该进该型号彩色电视机多少台？

解：设 X 表示商店每月销售该型号彩色电视机的台数，且月底的进货为 t 台，则当 $X \leq t$ 时就不会脱销，
因而按题意要求为

$$P(X \leq t) \geq 0.95,$$

又因为 $X \sim P(10)$ ，所以上式也可以表示为

$$\sum_{k=0}^t \frac{10^k}{k!} e^{-10} \geq 0.95,$$

由查表和计算得

$$\sum_{k=0}^{14} \frac{10^k}{k!} e^{-10} \approx 0.9166 < 0.95,$$

$$\sum_{k=0}^{15} \frac{10^k}{k!} e^{-10} \approx 0.9513 > 0.95,$$

所以，满足条件的最小 t 为**15**。

由此可知，这家商店只要在月底最少进货**15**台该型号彩色电视机（假定上个月无存货，否则应扣除），就可以有**95%**以上的把握保证这种电视机下个月不会脱销。

从以上的例子我们可以看出，这家商店采用了科学管理后，可以使每月的进货计划有的放矢，不会因为进货太多而使资金积压，影响该商店的经济效益；同样，也不会因为进货太少而完不成当月的营业额，以至影响广大居民的生活需要。

在二项分布中，要计算 $C_n^k p^k (1-p)^{n-k}$ ，当 n 较大时，计算量是很大的。但是，当 n 充分大， p 充分小， np 大小适中时，我们有如下定理，可以简化计算。

普阿松逼近定理：

若 n 充分大， p 充分小，而 $np = \lambda$ 大小适中，则有

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, \dots, n$$

在实际计算中，当 $n \geq 100, p \leq 0.1, np \leq 10$ 时，我们就可以使用以上近似公式计算，当然，当 n 越大， p 越小， np 大小适中时，近似公式计算就越精确。

在 n 重贝努利试验中，我们把在每次试验中出现的概率 p 很小的事件称为稀有事件。

因此，由以上定理可知，在实际应用中，通常很多重贝努利试验中稀有事件出现的次数可以近似用普阿松分布来描述。

例如，不幸事件、意外事故、故障、非常见病和自然灾害等都是稀有事件，所以这类事件在大量重复试验中出现的次数均近似地服从普阿松分布。这是普阿松分布会有如此广泛应用的重要原因。

43页例3-9

例7 (寿命保险问题)在保险公司里有2500个同一年龄和同社会阶层的人参加了人寿保险。在一年里每个人死亡的概率为0.002, 每个参加保险的人在元旦付12元保险费, 而在死亡时家属可以从保险公司领取2000元。问

(1)“保险公司亏本” (记为 A)的概率是多少?

(2)“保险公司获利不少于10000元, 20000元” (分别记为 B_1 和 B_2)的概率是多少?

解：显然，我们可以把考察“参加保险的人在一年中是否死亡”看作一次随机试验，因为有**2500**个人参加保险，所以我们可以把该问题看作为具有死亡概率 $p = 0.002$ 的**2500**重贝努利试验。

设 X 表示一年中死亡的人数，这样在一年的元旦保险公司收入为 **$2500 \times 12 = 30000$** (元)，而保险公司在这一年中应付出的为 $2000X$ (元)。

(1) “保险公司亏本” (不计利息)等价于 $2000X > 30000$, 即 $X > 15$ (人), 所以

$$P(A) = P(X > 15)$$

$$= \sum_{k=16}^{2500} C_{2500}^k 0.002^k 0.998^{2500-k}$$

$$\approx \sum_{k=16}^{2500} \frac{5^k}{k!} e^{-5} = 1 - \sum_{k=0}^{15} \frac{5^k}{k!} e^{-5} \approx 0.000069$$

(2)“保险公司获利不少于10000元” 等价于 $30000 - 2000X \geq 10000$, 即 $X \leq 10$, 所以 “保险公司获利不少于10000元” 的概率为

$$P(B_1) = P(X \leq 10)$$

$$= \sum_{k=0}^{10} C_{2500}^k 0.002^k 0.998^{2500-k}$$

$$\approx \sum_{k=0}^{10} \frac{5^k}{k!} e^{-5} \approx 0.9863$$

“保险公司获利不少于20000元”等价于 $30000 - 2000X \geq 20000$ ，即 $X \leq 5$ ，所以“保险公司获利不少于20000元”的概率为

$$P(B_2) = P(X \leq 5)$$

$$= \sum_{k=0}^5 C_{2500}^k 0.002^k 0.998^{2500-k}$$

$$\approx \sum_{k=0}^5 \frac{5^k}{k!} e^{-5} \approx 0.616$$

从以上的计算结果可以看出，在一年中
保险公司亏本的概率是非常小的，即在
10万年中约有7年是亏本的，而保险公
司获利不少于10000元和20000元的概
率分别在98%和61%以上。

4. 几何分布

定义3-8 设随机变量 X 的概率分布为

$$p_k = P(X = k) = (1 - p)^{k-1} p, k = 1, 2, \dots,$$

其中 $0 < p < 1$, 则称 X 服从几何分布 (geometric distribution), 记作

$$X \sim G(p)$$

我们容易验证几何分布满足：

$$(1) p_k \geq 0, k = 1, 2, \dots;$$

$$(2) \sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} (1-p)^{k-1} p = \frac{p}{1-(1-p)} = 1$$

在重复独立试验中，每次试验观察A发生与否，且 $P(A) = p$ 。记 X 为事件A首次发生时的试验次数，则 X 服从几何分布，即 $X \sim G(p)$ 。

设 A_i = “第 i 次试验事件 A 发生” , $i = 1, 2, \dots$

$$P(A_i) = P(A) = p, i = 1, 2, \dots$$

A_1, A_2, \dots, A_k 独立

$$\begin{aligned} \text{则} \quad P(X = k) &= P(\bar{A}_1 \cdots \bar{A}_{k-1} A_k) \\ &= P(\bar{A}_1) \cdots P(\bar{A}_{k-1}) P(A_k) \\ &= (1 - p)^{k-1} p, k = 1, 2, \dots \end{aligned}$$

44页例3-10

例8 某血库急需**AB**型血，需从献血者中获得，根据经验，每100个献血者中只能获得2名身体合格的**AB**型血的人，今对献血者一个接一个进行化验，用 X 表示在第一次找到合格的**AB**型血时，献血者已被化验的人数，求 X 的概率分布。

解： $p = \frac{2}{100} = 0.02$, $X \sim G(0.02)$

即 $P(X = k) = 0.02 \cdot 0.98^{k-1}, k = 1, 2, \dots$

性质**3-1** 设 $X \sim G(p)$, n, m 为任意两个自然数, 则

$$P(X > n + m | X > n) = P(X > m)$$

证明:

$$P(X > l) = \sum_{k=l+1}^{\infty} (1-p)^{k-1} p$$
$$= \frac{(1-p)^l p}{1-(1-p)} = (1-p)^l, l = 1, 2, \dots$$

$$P(X > n + m | X > n) = \frac{P(X > n + m, X > n)}{P(X > n)}$$

$$= \frac{P(X > n + m)}{P(X > n)}$$

$$= \frac{(1 - p)^{n+m}}{(1 - p)^n}$$

$$= (1 - p)^m$$

$$= P(X > m)$$

这个性质称为几何分布具有无记忆性。

可以证明，离散性随机变量只有几何分布具有无记忆性。

实际意义是：在例3-10中，若已化验了 n 个人，没有获得合格的AB型血，则再化验 m 个找不到合格AB型血的概率与已知的信息（即前 n 个人不是合格的AB型血）无关，即并不因为已查了 n 个人不合格，而第 $n+1$ 人， $n+2$ 人， \dots ， $n+m$ 人是合格AB型血的概率会因此而提高。

5. 超几何分布

定义3-9 设 $1 \leq M \leq N, 1 \leq n \leq N$ ，若随机变量 X 的概率分布为

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k = 0, 1, \dots, r,$$

其中 $r = \min\{n, M\}$ ，则称随机变量 X 服从超几何分布(hypergeometric distribution), 记作

$$X \sim H(M, N, n)$$

例如，有一批产品共 N 件，其中 M 件是次品。从中随机地(不放回)抽取 n 件产品进行检验。以 X 表示抽取的 n 件产品中次品的个数，则由古典概型的知识有

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k = 0, 1, \dots, r,$$

其中 $r = \min\{n, M\}$ ，也即随机变量 X 是服从超几何分布的。

超几何分布产生于 n 次不放回抽样，因此它在抽样理论中占有重要地位。

超几何分布与二项分布有着密切的联系。

事实上，超几何分布产生于不放回抽样，而二项分布产生于有放回抽样。

在实际工作中，抽样一般都采用不放回方式，因此计算时应该用超几何分布。

但是，当 N 较大时，超几何分布计算较繁琐。

若产品总数 N 很大，而抽样的次数 n 相对于 N 很小时，超几何分布可以用二项分布来近似。

即有以下定理。

定理3-2 对于任固定的 $n(\geq 1)$ ，当 N 充分大时，则有

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \approx C_n^k \left(\frac{M}{N} \right)^k \left(1 - \frac{M}{N} \right)^{n-k}, k = 0, 1, \dots, r,$$

其中 $r = \min\{n, M\}$ 。

以上这个定理在直观上还是比较容易理解的，这是因为：当产品总数 N 很大，而抽样的次数 n 相对于 N 很小时，可以认为不放回抽样与有放回抽样的差别应该是很小的，即超几何分布可以用二项分布来近似。

在实际计算中，一般当 $n \leq 0.1N$ 时，就可以运用以上的近似公式。

