§2.1 随机变量

我们在第一章中实际上已经接触过随机变量。

例 投掷一枚骰子，用X表示得到的点数。

例 检测一件产品可能出现的两个结果，也可以用一个随机变量来描述

$$X(\omega)=\begin{cases}1, & \text{取到次品}\\ 0, & \text{取到正品}\end{cases}$$

随机变量的描述性定义：

根据试验结果取值的变量

随机变量一般用大写字母 $X, Y, Z, \ldots$ 表示。

有些书也用小写希腊字母 ξ, η, ζ 表示.

随机变量的数学定义：

定义 设S是试验E的样本空间, 定义在S上的函数称为随机变量。随机变量可能取的值称为可能值。

例1 在一袋中装有编号为1, 2, 3 的3 个球.在袋中任取一个球, 放回以后再取一只球。记录它们的编号。试验的样本空间

$$S=\{e\}=\{(i,j)\,|\,i,j=1,2,3\}$$

对于这样的样本空间, 我们很容易可以建立一个它到实数集的映射X, 例如对每一个样本点$e=(i,j)$, 让$X(e)=i+j$, 则$X:S\to R$ 是随机变量

- 引入随机变量后，可用随机变量的等式或不等式表达随机事件，例如

$(X > 100)$—— 表示 “某天9:00 ~ 10:00 接到电话次数超过100次” 这一事件

- 随机变量的函数一般也是随机变量

随机变量分类
- 离散型
- 非离散型

其中一种重要的类型为 连续型 随机变量.

§2.2 离散型随机变量及其分布律

一、定义

离散型随机变量----只有有限个或可列个可能值的随机变量。

我们想知道什么？

X取什么值？

但是，这不可能做到。于是我们退一步，希望知道取一个可能值的概率是多少。

设离散型随机变量X所有可能值为x_k （$k=1,2,\cdots$），X取各个可能值的概率，即事件$\{X=x_k\}$的概率，为

$$P\{X=x_k\}=p_k, k=1,2,\cdots$$

称这组等式为X的概率分布或者分布律。如果分布律不能写成通式就列成表：

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

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

或者

X	x_1	x_2	…	x_k	…
P	p_1	p_2	…	p_k	…

如果可以用通式表示就用通式，不要列表。

注意：

1、如果随机变量只有k 个可能值，就把上面的表格中的最后一列省略号删除。

2、如果有无穷多个可能值，不能写成：

X	x_1	x_2	$\cdots$	x_n
p_k	p_1	p_2	$\cdots$	p_n

必须写成

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

概率分布的性质：　$1^0\quad p_k \geq 0,\quad k=1,2\cdots,\qquad (2.2)$

$$2^0\quad \sum_{k=1}^{\infty} p_k = 1 \qquad (2.3)$$

如果上述两条成立，则前面的表一定是某个离散型随机变量X的分布律。

例. 设X的概率分布为

$$P\{X=i\} = a\left(\frac{2}{3}\right)^i,\qquad i=1, 2, 3;$$

求上述式子中的系数 a 。

解. 由概率分布的性质得$1=\sum_{i=1}^{3} a\left(\frac{2}{3}\right)^i = a\frac{38}{27}$，所以 $a=\frac{27}{38}$.

例1　设一汽车在开往目的地的道路上需经过四组信号灯,每组信号灯以1/2 的概率允许或禁止汽车通过. 以X 表示汽车首次停下时，它已通过的信号灯的组数（设各组信号灯的工作是相互独立的）. 求X 的分布律.

解：则X的可能值是0，1，2，3，4，

“X＝k”表示“连续遇到了k次允许，第k+1次遇到禁止或到达目的地”。故

$$P(X=k)=\begin{cases} 0.5^k(1-0.5) & k=0,1,2,3 \\ 0.5^4 & k=4 \end{cases}$$

求离散型随机变量分布的步骤：

1、找出它的所有可能值。

2、分别计算它取每个可能值的概率。

3、如果这些概率可以写成通式，则该通式可以作为答案。否则就必须列出概率分布的表格。

注意：在计算$P(X=a)$的时候，首先要想清楚事件“$X=a$”的确切含义。

错误的写法

把$P(X=x_k)$简写为$P(X)$，$P(x_k)$或$p(X)$

例如：把P(X=2) 写成 P(2)，p_2 或者 p(2)。

当然，我们可能会记 p_2 为 P(X=2)，但不是直接用p_2

二、常见分布

(一) (0-1)分布，也叫两点分布：

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

或：$P\{X=0\}=(1-p)$，$P\{X=1\}=p \quad (0<p<1)$

(二) 二项分布、伯努利试验

设试验E只有两个可能结果：A 及$\overline{A}$，

则称E 为伯努利($Bernoulli$) 试验.

将一个伯努利试验E 独立重复地做n 次, 则这一串重复的独立试验称为n 重伯努利试验.

设$P(A)=p$，考虑事件"A在n重伯努利试验中发生k次"。记为X为"A在n重伯努利试验中发生的次数"。则这个事件可以表示为"$X=k$"。

由于这里的“k次”可以是任意的k次。有$\binom{n}{k}$种情况

每一种情况的概率是 $p^k(1-p)^{n-k}$ 。所以,

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k},\quad k=0,1,2,\dots,n$$

这个分布称为二项分布。记为$X\sim b(n,\ p)$

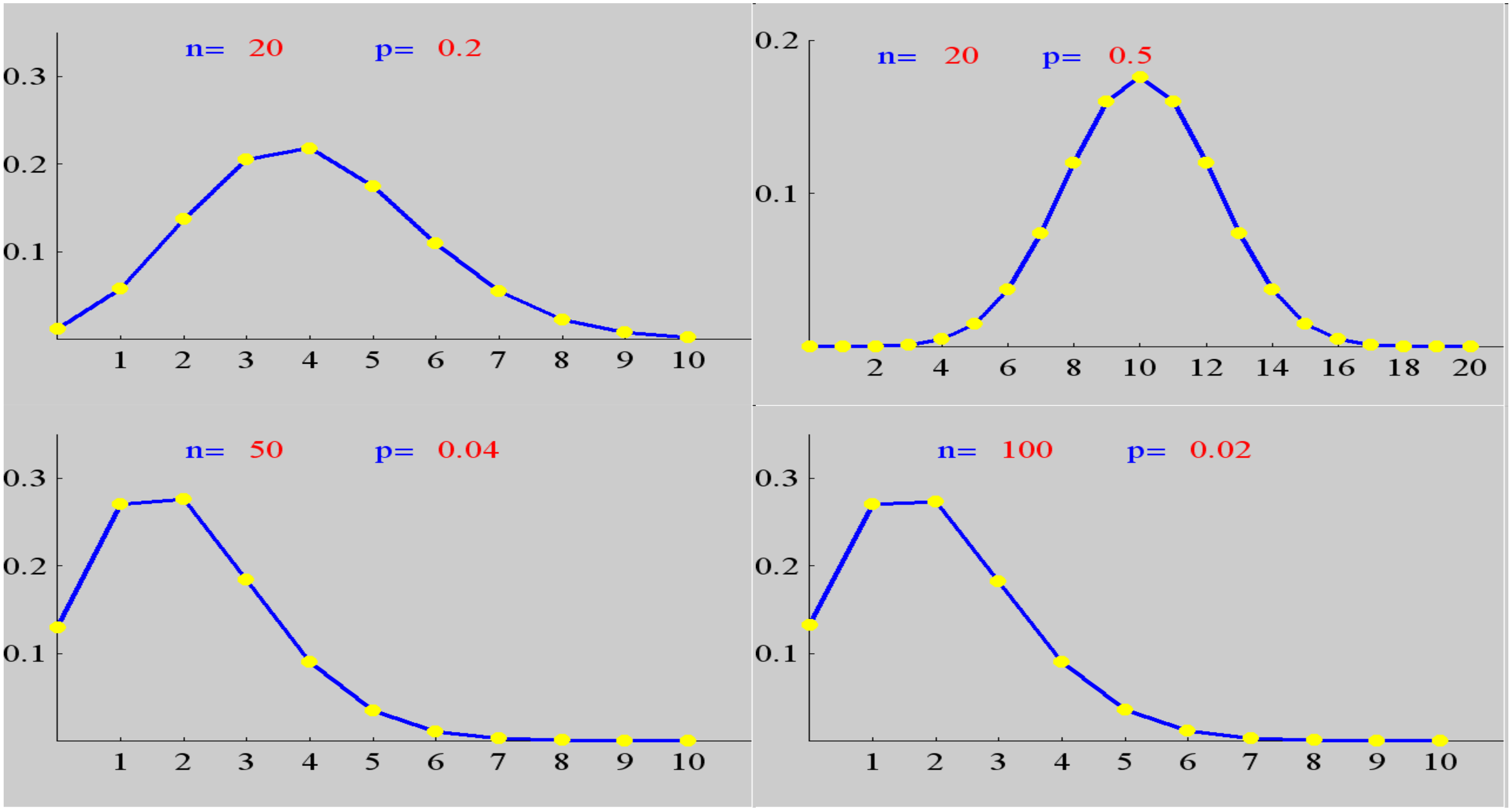
或$X\sim B(n,p)$

1、（0－1）分布是二项分布的特例。

2、由二项式定理知

$$\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=[p+(1-p)]^n=1$$

二项分布的图形



例2　按规定,某种型号电子元件的使用寿命超过1500小时的为一级品. 已知某一大批产品的一级品率为0.2, 现在从中随机地抽查20只. 问20只元件中恰有k只$(k=0,1,\cdots,20)$为一级品的概率是多少?

解：以X记20只元件中一级品的只数。则

$$P\{X=k\}\approx\binom{20}{k}0.2^k(1-0.2)^{20-k},k=0,1,2,\ldots,20$$

即X～b (20, 0.2)

例题3： 自己看

例4 设有80台同类型设备, 各台工作是相互独立的, 发生故障的概率都是 0.01, 且同一台设备的故障能由一个人处理. 考虑两种配备维修工人的方法, 其一是由4个人维护, 每人负责20台; 其二是由3人共同维护80台. 试比较这两种方法在设备发生故障时不能及时维修的概率的大小.

解:按第一种方法: 以X_i 记"第i 人维护的20 台中同一时刻发生故障的台数", 则知80 台中发生故障而不能及时维修的概率为

$$\begin{aligned}
&P(\{X_1>1\}\cup\{X_2>1\}\cup\{X_3>1\}\cup\{X_4>1\}) \\
&\quad =1-P(\{X_1\le 1\}\cap\{X_2\le 1\}\cap\{X_3\le 1\}\cap\{X_4\le 1\}) \\
&\quad =1-(P\{X_1\le 1\})^4
\end{aligned}$$

$$\begin{aligned}
P\{X_1 \le 1\} &= P\{X_1=0\}+P\{X_1=1\} \\
&= 0.99^{20}+20\times 0.01\times 0.99^{19}=0.9831
\end{aligned}$$

$$P(\{X_1>1\}\cup\{X_2>1\}\cup\{X_3>1\}\cup\{X_4>1\})\approx 1-0.9831^4\approx 0.0659$$

按第二种方法：以Y记"这80台中同一时刻发生故障的台数"，则知80台中发生故障而不能及时维修的概率为

$$P(Y>3)=1-P(Y=0)-P(Y=1)-P(Y=2)-P(Y=3)$$
$$\approx 0.0087$$

故，采用第二种方案不仅能够减少人力，而且还可以提高工作效率。

三、泊松分布

设随机变量X所有可能取的值为$0,1,2,\cdots$，而取各个值的概率为

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,2\cdots$$

其中$\lambda>0$是常数.则称X服从参数为λ的泊松分布，记为$X\sim\pi(\lambda)$.

《高等数学》告诉我们

$$\sum_{k=0}^{+\infty}\frac{x^k}{k!}=e^x$$

所以，$$\sum_{k=0}^{+\infty}\frac{\lambda^k e^{-\lambda}}{k!}=1$$

泊松近似公式：当n很大，p很小时

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{(np)^k e^{-np}}{k!}$$

例5　计算机硬件公司制造某种特殊型号的微型芯片，次品率达0.1%，各芯片成为次品相互独立，求在1000只产品中至少有2只次品的概率。

解　以X记产品中的次品数，则$X \sim b(1000, 0.001)$,

$$P\{X \geq 2\} = 1 - P\{X \leq 1\} = 1 - \sum_{k=0}^{1} \binom{1000}{k} 0.001^k (1-0.001)^{1000-k}$$

$$\approx 1 - e^{-1} - e^{-1} \approx 0.26$$

注记

由上面的近似公式可见，如果一个随机变量近似地服从一个n很大，p很小的二项分布$B(n, p)$，则它近似地服从泊松分布。

作业：2；4；9