# 概率论与数理统计

## 第二章

杨青
yangq@ustc.edu.cn

# 第二章： 随机变量及其分布

大纲：

- 随机变量概念
- 离散型随机变量
  - 定义
  - 常用离散型分布
- 连续型随机变量
  - 定义
  - 常用连续型分布
- 随机变量函数的分布

# 1. 随机变量的概念

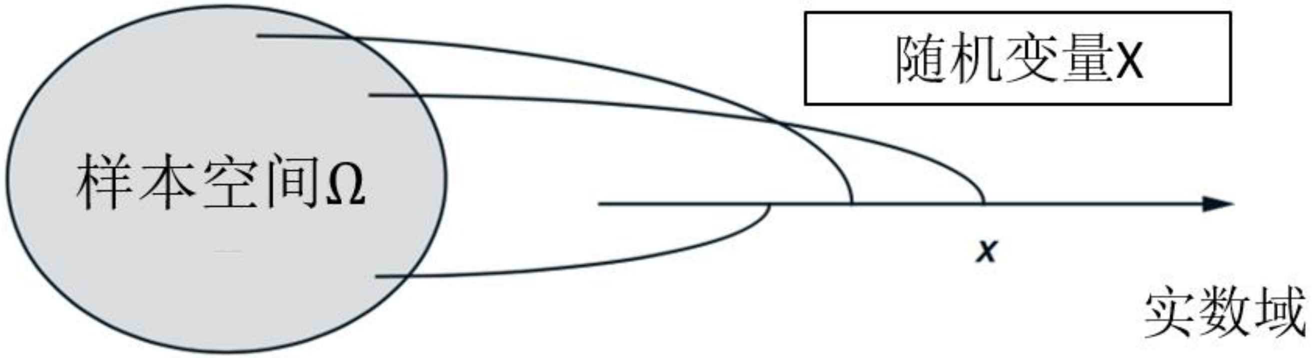
- 引例：

| |
|---|
| 掷一枚硬币出现正面或反面. |
| 产品被分为正品或废品. |

- 这两个样本空间描述不同, 但是样本空间的的概率机制相同.
- 上面两例中的结果均可用一个取值 0,1 的随机变量来描述.

- 如何统一来研究样本空间中随机事件发生的概率机制? 从数学的角度看, 最好把它们都放在同一个空间(一维直线)来比较和研究.
  - 映射：样本空间到直线 R --- 随机变量

# 随机变量——样本空间的简单化




- 随机变量是把随机试验的结果，也就是样本空间，与一组实数联系起来. 这样的处理简化了原来的概率结构.

- 通常用大写英文字母$X, Y, Z, W$等表示随机变量. 小写字母$x, y, a, b$表示实数取值.

**定义 2.1 随机变量**

随机变量 $X$ 是一个映射, $X: \Omega \to \mathbb{R}$, 对每个 (可测集) $A \subset \mathbb{R}$, $\{X \in A\}$ 是一个 (可定义概率的) 随机事件, 且

$$P(X \in A) = P(\{\omega \in \Omega : X(\omega) \in A\}). \tag{2.1}$$

♣

- 样本空间中随机事件发生的概率机制的研究化为对随机变量的研究
- 事件 → 随机变量的取值范围
- 事件发生的概率 → 随机变量取这些值的概率

随机变量的分布
(distribution)

# 例子

- 记$X$表示掷一颗骰子出现的点数，则$X$的可能取值为1，2，3，4，5，6。
  - 事件$\mathrm{A}=\{$点数小于等于3$\}$
  - $\mathrm{A}=\{X\leq 3\}$

事件转化为随机变量的取值范围

- 记$Y$表示一天内达到之心城的顾客数，则$Y$的可能取值为0，1，2，…。
  - 事件$B=\{$至少来了1000位顾客$\}$
  - $B=\{Y\geq 1000\}$

- 记$T$表示某品牌手机的使用寿命，则$T$的可能取值充满了区间$[0,+\infty]$。
  - 事件$\mathrm{C}=\{$使用寿命在365到1000天之间$\}$
  - $C=\{365\leq T\leq 1000\}$

常见的随机变量可以分为两大类：

- 只取有限个或可数个值的随机变量称为离散型随机变量；
- 取连续的值且密度存在的随机变量称为连续型随机变量.

**定义 2.2　离散型随机变量和分布律**

如果随机变量 $X$ 只取有限多个或可数多个值, 那么称 $X$ 为离散型随机变量. 设 $X$ 取的一切可能值为 $x_1, x_2, \cdots, x_n, \cdots$, 则

$$P(X = x_k) = p_k,\ k = 1, 2, \cdots, n, \cdots, \tag{2.2}$$

其中

$$p_k \geqslant 0,\ k = 1, 2, \cdots, n, \cdots;\ \sum_{k=1}^{\infty} p_k = 1. \tag{2.3}$$

(2.2) 式称为离散型随机变量 $X$ 的分布律或概率质量函数 (probability mass function, 简写为 pmf).

♣

分布律也可以用表格或矩阵表达

$$\begin{array}{c|ccccc} X & x_1 & x_2 & \cdots & x_n & \cdots \\ \hline P & p_1 & p_2 & \cdots & p_n & \cdots \end{array} \quad 或 \quad \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

- 设$X$为一离散型随机变量，它以概率$\{p_1, \ldots, p_n, \ldots\}$取值$\{a_1, \ldots, a_n, \ldots\}$。令事件$A \subseteq \{a_1, \ldots, a_n, \ldots\}$，则

$$\mathrm{P}(\mathrm{A}) = \sum_{a_i \in A} p_i$$

这样知道了离散型随机变量 $X$ 的pmf，我们就能给出关于$X$的任何概率问题的回答.

# 常用离散分布

- 0-1分布(Bernoulli 分布)
- 离散的均匀分布
- 二项分布(Binomial distribution)
- 负二项分布(Pascal distribution)
- 几何分布(Geometric distribution)
- 泊松(Poisson)分布

# 0-1分布(Bernoulli分布)

**定义 2.3　0–1 分布**

若随机变量 $X$ 的分布律为

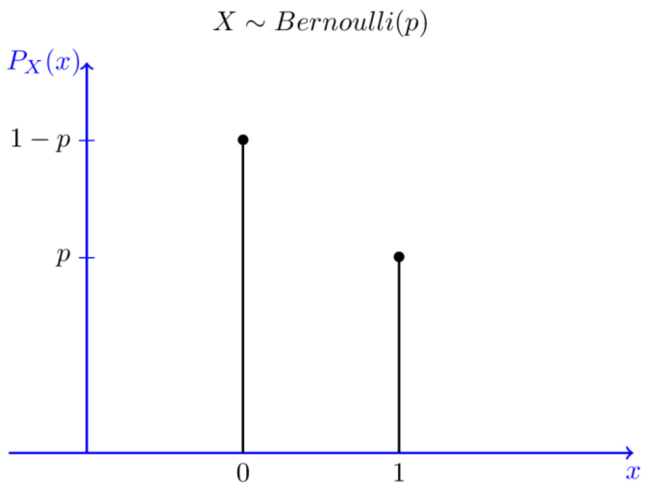
| $X$ | $0$ | $1$ |
|---|---|---|
| $P$ | $1-p$ | $p$ |

其中 $0<p<1$. 则称 $X$ 服从 $0-1$ 分布或者伯努利分布或两点分布. ♣

设一个随机试验只有两个可能结果 $A$ 和 $\bar{A}$, 则称此试验为一 Bernoulli 试验.

- $0-1$ 分布随机变量 $X$ 的分布函数也可以写为

$$P(X=x)=p^x(1-p)^{1-x}, x=0 \text{ 或 } 1.$$




概率函数图

- 一般在试验中仅考虑事件 $A$ 是否发生时，引入示性函数, 为$0-1$分布的随机变量.

$$I_A = \begin{cases} 1, & A \text{ 发生}, \\ 0, & A \text{ 不发生}. \end{cases}$$

# 离散均匀分布

**定义 2.4　离散均匀分布**

若随机变量 $X$ 的分布律为

$$P(X = x_k) = \frac{1}{n},\ k = 1, 2, \cdots, n, \tag{2.5}$$

其中 $x_k, k = 1, 2, \cdots, n$ 为 $n$ 个不同的实数. 则称随机变量 $X$ 服从离散均匀分布, 其中常用的 $x_k = k, k = 1, 2, \cdots, n$.

- 例：古典概型是离散均匀分布.

# 二项分布(Binomial distribution)

- 设某事件 $A$ 在一次试验中发生的概率为 $p$. 现把试验独立地重复$n$次. 以 $X$ 记 $A$ 在这 $n$ 次试验中发生的次数，则 $X$ 取值 0, 1, ..., $n$，此$X$ 服从二项分布.

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

是概率质量函数.　　$\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=$ ?　(二项式定理/展开)

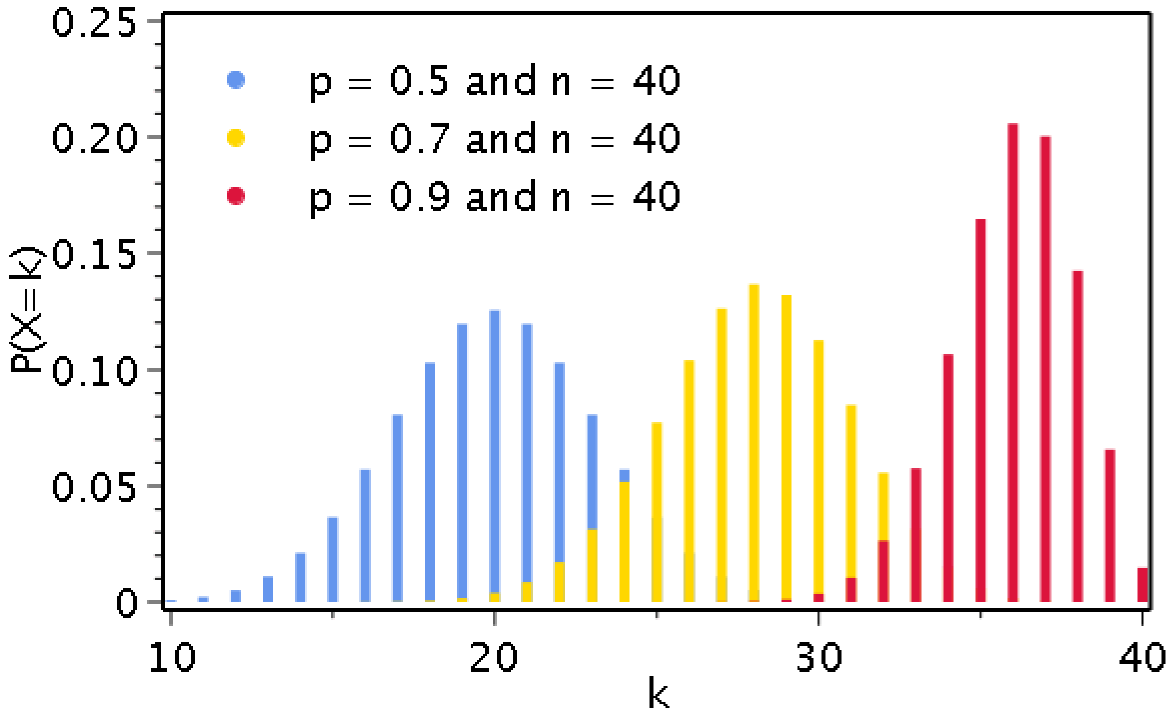
**定义 2.5　二项分布**

设离散型随机变量 $X$ 所有可能取值为 $\{0,1,\cdots,n\}$, $0<p<1$, 如果其分布律为

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,\cdots,n. \tag{2.7}$$

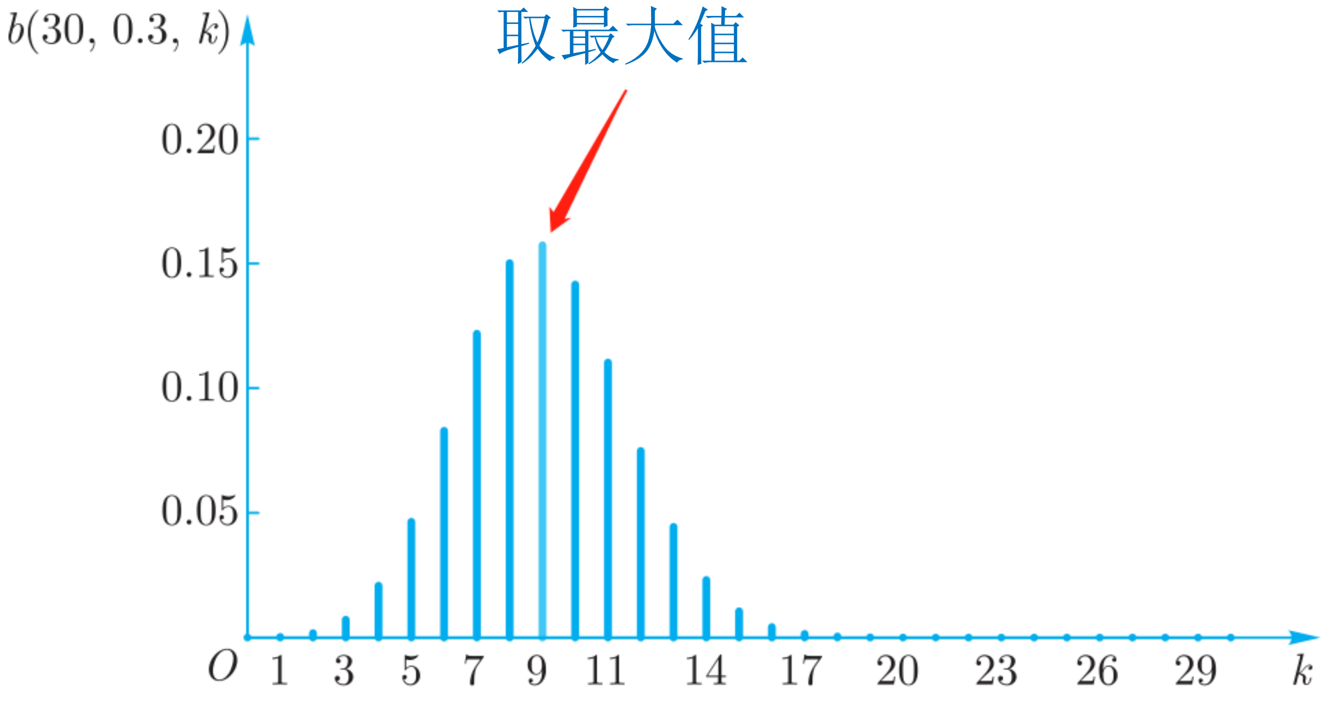
那么称 $X$ 服从二项分布, 常记为 $X\sim B(n,p)$, 而 $P(X=k)$ 常记为 $b(n,p,k)$. ♣

- 二项分布描述了n次重复独立的Bernoulli试验中成功次数的分布。
- 一个随机变量服从二项分布有以下两个条件：
  - 各次试验的条件是稳定的，这保证了事件 $A$ 发生的概率 $p$ 在各次试验中保持不变.
  - 各次试验之间相互独立.




呈现两头小中间大的上凸形状

- 若 LED 灯的寿命超过 50 000 h, 则称为一级品, 已知某厂以往的 LED 灯一级品率为 0.3. 现从仓库某一大批产品中随机抽取了 30 支检验, 问其中恰有 $k$ 支一级品 LED 灯的概率有多大?




$b(30, 0.3, k)$ 随 $k$ 的变化

# 负二项分布(Pascal distribution)

- 将伯努利试验一直独立地重复下去, 以$X_r$表示第 $r$ 次试验成功发生时的试验次数, $p = 1 - q$ 为成功的概率, 那么$X_r$的分布律为负二项分布, 也称为帕斯卡分布.

$$\begin{aligned} p_k = P(X_r = k) &= P(\{\text{前}k-1\text{次恰有}r-1\text{次成功且第}k\text{次成功}\}) \\ &= P(\{\text{前}k-1\text{次恰有}r-1\text{次成功}\})P(\{\text{第}k\text{次成功}\}) \\ &= \binom{k-1}{r-1} p^{r-1} q^{k-r} \cdot p \\ &= \binom{k-1}{r-1} p^r q^{k-r}, \quad k = r, r+1, \cdots . \end{aligned}$$

$$\sum_{k=r}^{\infty} p_k = \sum_{k=r}^{\infty} \binom{k-1}{r-1} p^r q^{k-r} = p^r \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} q^k = p^r (1-q)^{-r} = 1$$

**定义 2.6　负二项分布**

设随机变量 $X_r$ 取正整数值, 其分布律为

$$P(X_r = k) = \binom{k-1}{r-1} p^r q^{k-r}, \quad k = r, r+1, \cdots . \tag{2.9}$$

其中 $r$ 为正整数, $0 < p < 1$, 则称 $X$ 服从参数为 $r, p$ 的负二项分布或者帕斯卡分布. 记为 $X \sim NB(r,p)$. $P(X_r = k)$ 则记为 $nb(r,p,k)$. ♣

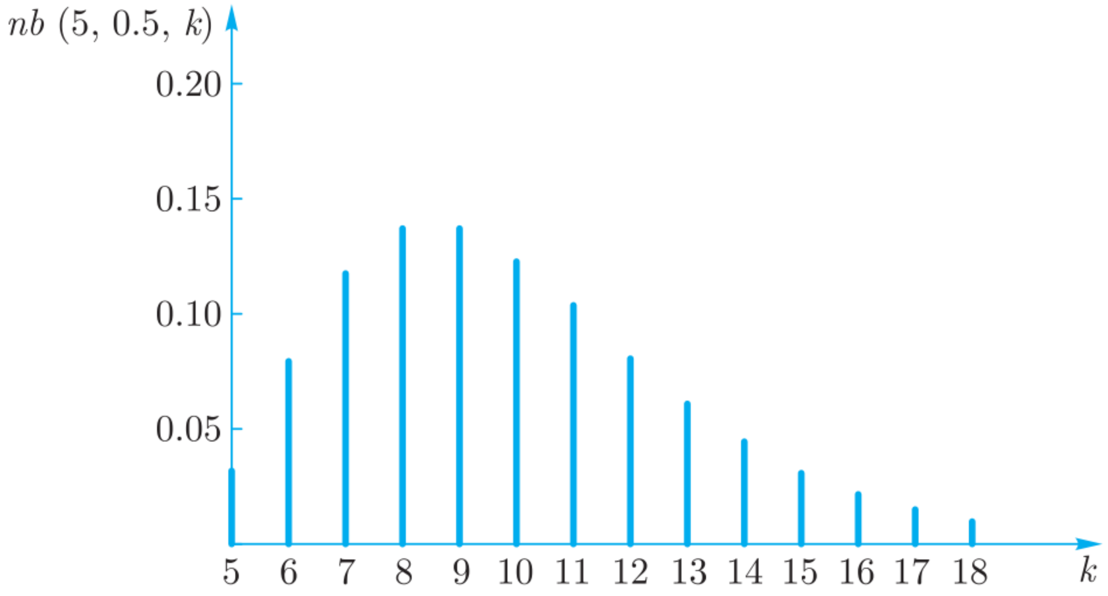



图 2.3　$nb(5, 0.5, k)$ 随 $k$ 的变化

# 例子

- (巴拿赫 (Banach) 火柴问题) 某人口袋里放有两盒火柴, 每盒装有火柴 $n$ 根. 他每次随机取出一盒, 并从中拿出一根火柴使用. 试求他取出一盒, 发现已空, 而此时另一盒中尚余 $r$ 根火柴的概率.

Hint: 记A={甲盒已空, 而乙盒中尚余 $r$ 根火柴}, 求2P(A).
将每取出甲盒一次视为取得一次成功， p=0.5.
$X$ =取得第 $n+1$ 次成功时的取盒次数

- 在独立重复的伯努利试验序列中，成功的概率为$p$，求事件 $E$ ={$n$ 次成功发生在 $m$ 次失败之前}的概率 .

Hint:　记 $F_k$={第 $n$ 次成功发生在第 $k$ 次试验}，则

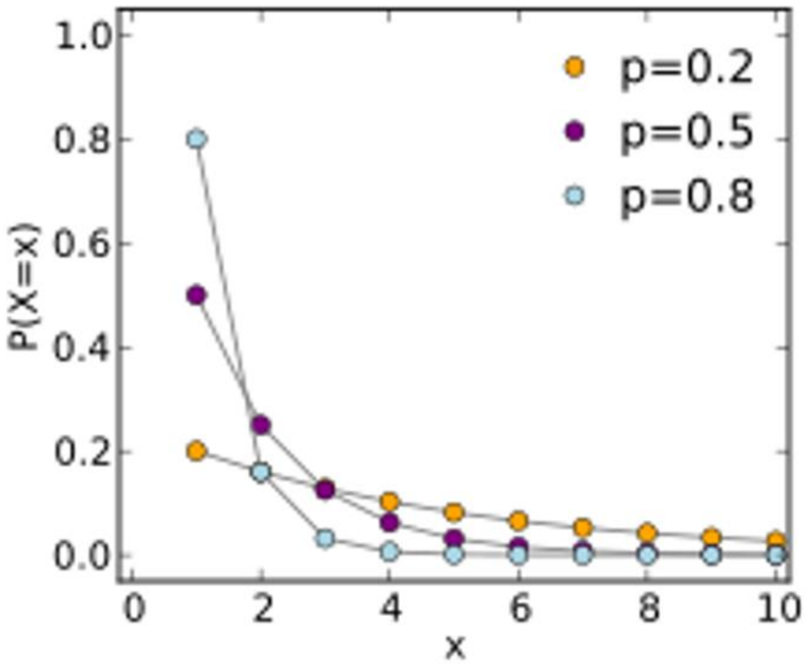
$$E = \bigcup_{k=n}^{n+m-1} F_k$$

# 几何分布(Geometric distribution)

- 在负二项分布中, 若 $r = 1$, 则$X_1$表示首次成功时的试验次数, 其分布常常称为几何分布, 记为$X_1 \sim Ge(p)$.

- 由负二项分布的分布可知几何分布的分布律为

$$P(X_1 = k) = q^{k-1}p, k = 1, 2, \cdots .$$



# 例子

一个人要开门, 他共有 $n$ 把钥匙。其中仅有一把可以打开门。现随机地有放回的从中选取一把开门，若不成功再放回去重新随机选取一把开门，问这人在第 $S$ 次才首次试开成功的概率。

**定理 2.1**

以所有正整数为取值集合的随机变量 $X$ 服从几何分布 $Ge(p)$，当且仅当对任何正整数 $m$ 和 $n$，都有

$$P(X > m + n \mid X > m) = P(X > n). \tag{2.10}$$

这个性质称为几何分布的无记忆性 (memoryless property).