第一章复习:

- ω, Ω, $\mathcal{A}$, $P : A \mapsto P(A)$ 三条.
- 条件概率: $P(B|A) := P(AB)/P(A)$
- 乘法公式: $P(AB) = P(A)P(B|A)$.
- 全概公式: $P(A) = \sum_i P(B_i)P(A|B_i)$.
- 逆概公式: $P(B_i|A) = P(B_i)P(A|B_i)/\sum_j P(B_j)P(A|B_j)$.
- 独立性: $P(A_{i_1} \cdots A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$, $\forall \ldots$.

第二章 随机变量与概率分布

- 第四次

  §2.1 随机变量的概念

  §2.2 离散型随机变量

- 第五次

  §2.3 连续型随机变量

  §2.4 随机变量的严格定义和分布函数

  §2.5 随机变量的函数

- 第六次

  §2.6 随机变量的数学期望

  §2.7 随机变量的方差及其他数字特征

定义 (随机变量, random variable)

随机变量指样本空间上的函数, $X : \Omega \to \mathbb{R}$, $\omega \mapsto X(\omega)$.

概率是权分配方案, P 是事件(子集)的函数, $P(A)$;

随机变量是观测值, X 是样本(点)的函数, $X(\omega)$.

例1.2 盒中有5个球, 2白3黑. 任取3个, 取到的白球数为X.

- X 是随机变量.

  $\omega = \{i, j, k\}$. $X(\omega) = |\{i, j, k\} \cap \{1, 2\}|$.
- $\{X = 1\} = \{\omega : X(\omega) = 1\}$, $\{X \leqslant 1\}$ 是事件.
- $P(\{X = 1\})$ 简记为$P(X = 1)$.

模型参数 vs 随机变量

例: N 个产品中有M 个次品. 抽出n 个, 次品数为X.
参数: N, M, n. 随机变量: X.

习题一, 11 在未名湖中捕鱼80条, 标记后放回, 再捕鱼100条, 其中4条有标记. 猜湖中的总鱼数N.

- 随机试验: 捕获(80条)、再捕获(100条).
- N: 试验前确定, 参数;
  X =有标记的鱼数, 试验后确定, 随机变量.
- 事件"$\{X = 4\}$"发生. $P_N(X = 4) = \dfrac{C_{80}^4 C_{N-80}^{96}}{C_N^{100}} =: p_N$.
- 求p_N 的最大值点. §7.1 最大似然估计.
- **注:** $P(N = 2000|X = 4)$ 无意义. $N = 2000$ 不是事件!

§2.2 离散型随机变量

定义 (离散型随机变量)

若X只可能取可数个值x_i, $i \geqslant 1$, 则称X 是离散型随机变量.

称$p_i := P(X = x_i)$, $i \geqslant 1$ 为X 的(概率) 分布(列).

- 分布列: $p_i \geqslant 0$; $\sum_i p_i = 1$.

定义 (离散型随机变量(推广))

若存在可数个值x_i, $i \geqslant 1$, 使得$p_i := P(X = x_i) > 0$, $i \geqslant 1$,

且$\sum_i p_i = 1$, 则称X 是离散型随机变量.

称$p_i := P(X = x_i)$, $i \geqslant 1$ 为X 的(概率)分布(列).

1. 两点分布

伯努利分布(Bernoulli distribution):

$$P(X=1)=p,\quad P(X=0)=1-p,\quad 0<p<1.$$

记为$X\sim B(1,p)$. Jacob Bernoulli (1655-1675 瑞士数学家)

- (1) 掷硬币, 正面出现的概率为p. $X(H)=1$, $X(T)=0$.

  (2) 掷色子, $\Omega=\{1,2,3,4,5,6\}$, $Y:1,2,3\mapsto 0;\ 4,5,6\mapsto 1$.
- 样本空间、函数映射关系不重要. 重要的是随机变量的分布.
- 示性函数(index function), $1_A(\omega)=\begin{cases}1, & 若\omega\in A;\\ 0, & 否则.\end{cases}$

  $A=\{X=1\}$; $X=1_A\sim B(1,P(A))$.
- 一般情形:

$$P(X=a)=1-P(X=b)=p.$$

例如: $X=a1_A+b1_{A^c}$, $A=\{X=a\}$.

2. 二项分布(binomial, $X \sim B(n,p)$)

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

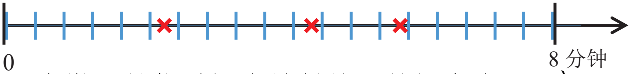
- 模型: 连续投币n 次, 正面出现的总次数; 放回抽样.
- 分布列的最大值点: $k_0 \approx np$.
  确切地(定理2.1)，最大值点在$[(n+1)p]$达到最大值;
  当$(n+1)p$为整数时有两个最大值点$(n+1)p-1, (n+1)p$.
- 由$(p+q)^n = \sum_{k=0}^n C_n^k p^k q^{n-k}$ 而得名，其中$q = 1-p$.

3. 泊松分布(Poisson, $X \sim \mathcal{P}(\lambda)$)

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

例2.3 放射性物质在8分钟内放射出的粒子数$X \sim \mathcal{P}(\lambda)$.

- 将8分钟视为一个宏观的单位时间. 以$\frac{1}{n}\times 8$分钟为一个微观单位时间, 即8分钟$=n$个微观单位时间.



- (i) 在一个微观单位时间内放射粒子的概率为$p=\frac{\lambda}{n}$,
  (ii) 不同的微观单位时间内是否放射粒子相互独立.
- X 近似服从$B(n,p)$, 故

$$\begin{aligned} P(X=k) &\approx C_n^k p^k(1-p)^{n-k}=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= \prod_{i=1}^{k-1}\left(1-\frac{i}{n}\right)\frac{(np)^k}{k!}\left(1-\frac{\lambda}{n}\right)^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}. \end{aligned}$$

- 泊松分布分布列的最大值点: $k_0 \approx \lambda$.
  确切地(定理2.2)，最大值点在$[\lambda]$达到最大值；当λ为整数时有两个最大值点$\lambda-1, \lambda$.
- Siméon-Denis Poisson (法国数学家), 1830.

4. 超几何分布(参数N, M, n)

$$P(X=k) = C_M^k C_{N-M}^{n-k} / C_N^n, \quad k = 0, 1, \cdots, n \wedge M.$$

- 模型: N个产品, M 个次品, 任取n 个, 抽到的次品数为X.
- 放回抽样vs 不放回抽样: 二项分布vs 超几何分布.
  次品率: $p = \frac{M}{N}$.
- 定理2.3: 给定n. 当$N \to \infty$, $\frac{M}{N} \to p$ 时,
  $P(X=k) \to P(Y=k)$, 其中$Y \sim B(n, p)$.

5. 几何分布

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

- 模型：第一次投到正面时的投币次数;
  打靶, 一次命中率为p, 第一次击中靶子的射击次数.
- $P(X>n)=(1-p)^n$, $\forall n \geqslant 0$.
- 无记忆性: $P(X-n=k|X>n)=P(X=k)$.
- 是唯一满足无记忆性的分布.

6. 负二项分布(negative binomial, $X \sim NB(r,p)$)

$$P(X=k) = C_{k-1}^{r-1}(1-p)^{k-r}p^r, \quad k = r, r+1, r+2, \cdots.$$

$$q := 1-p.$$

- 模型：第r次投到正面时的投币次数.
- 名字来自源负于二项展开式：
  $p^{-r} = (1-q)^{-r} = \sum_{i=0}^{\infty} C_{i+r-1}^{r-1} q^i = \sum_{k=r}^{\infty} C_{k-1}^{r-1} q^{k-r}$.
- $Y := X - r$. $P(Y=k) = P(X=k+r) = C_{k+r-1}^{r-1} q^k p^r = \frac{(r+k-1)!}{k!(r-1)!} p^r q^k = \frac{\Gamma(k+r)}{k!\Gamma(r)} p^r q^k, \quad k = 0, 1, 2, \cdots.$
  (有些书上称Y的分布为负二项分布.)

7. 离散均匀分布(自习)

8. 退化型

$$P(X=a) = 1.$$

重要的是随机变量的分布!

不可观测	Ω 抽象	ω 符号	P 概率
可观测	$\mathbb{R}$ 具体	$x = X(\omega)$ 实数	$\mu_X : B \mapsto P(X \in B)$ 分布