AI 中的数学

第三、四讲

方聪，概率统计部分参考章复熹和张原老师课件

2024 年秋季

1. 随机变量
2. 离散随机变量

例 1.2. 盒中有 5 个球, 其中有 2 个白球, 3 个黑球.
从中任取 3 个球, 将其中所含的白球的数目记为 X.

- 建模: 将球编号, $1 \sim 3$ 表示黑球, 4,5 表示白球.
- $\omega = (i, j, k)$, 其中 $1 \leqslant i < j < k \leqslant 5. \Omega = C_5^3 = 10.$
- 事件是 Ω 的集合，如果只关注第一次结果，所有事件对应的每个集合中可以包括 $2, 3$ 次的所有情况

例 1.2. 盒中有 5 个球, 其中有 2 个白球, 3 个黑球. 从中任取 3 个球, 将其中所含的白球的数目记为 X.

- 建模: 将球编号, $1 \sim 3$ 表示黑球, 4,5 表示白球.
- $\omega = (i, j, k)$, 其中 $1 \leqslant i < j < k \leqslant 5. \Omega = C_5^3 = 10.$
- 事件是 Ω 的集合, 如果只关注第一次结果, 所有事件对应的每个集合中可以包括 $2, 3$ 次的所有情况
- 满足 $X = 0$ 的 ω 有 $C_2^0 C_3^3 = 1$ 个;
  满足 $X = 1$ 的 ω 有 $C_2^1 C_3^2 = 6$ 个;
  满足 $X = 2$ 的 ω 有 $C_2^2 C_3^1 = 3$ 个.
- 事件: $\{X = 1\} = \{\omega : X(\omega) = 1\}$,
  $\{X \leqslant 1\} = \{\omega : X(\omega) \leqslant 1\}$.
- 将 $P(\{X = 1\})$ 简记为 $P(X = 1)$. 例如,

$$P(X = 1) = \frac{6}{10}, P(X \leqslant 1) = \frac{7}{10}.$$

例 1.6. 某公共汽车站每隔 10 min 会有一两某路公交车到达. 某乘客随机在任意时刻到达车站.

- 候车时间 X (单位: min) 为随机变量.
- $0 \leqslant X \leqslant 10$.
- 几何概型 (参阅 1.8): 例,

$$P(X \leqslant 3) = \frac{3}{10}, \quad P(2 \leqslant X \leqslant 6) = \frac{4}{10}.$$

- 定义 2.1. X 是离散型随机变量指: X 取有限个值 $x_1, \cdots, x_n$, 或可列个值 $x_1, x_2, \cdots$ .X 的概率分布 (列) 指:

$$p_k = P(X = x_k), \quad k = 1, \cdots, n \text{ 或} k = 1, 2, \cdots .$$

- 概率分布表:

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p	p_1	p_2	$\cdots$	p_k	$\cdots$

- 非负: $p_k \geqslant 0, \forall k$; 规范: $\sum_{k=1}^{n} p_k = 1$ 或 $\sum_{k=1}^{\infty} p_k = 1$.
- p_k, $k = 1, 2, 3 \ldots$, 是 X 的概率分布, 也称为概率函数或者概率分布律.

1. 两点分布 (伯努利分布), $X \sim B(1,p)$ (参数 $0 \leqslant p \leqslant 1$):

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

- 模型：投币，

$$投到 H \text{ 则} X=1; 投到 T \text{ 则} X=0.$$

- 示性函数 1_A : 事件 A 发生则取 1; A 不发生则取 0 .
- 例 2.1. 100 件产品中有 3 件次品. 从中任取一件.

$$A = \text{“取到合格品”}, X = 1_A, p = 0.97.$$

2. 二项分布, $X \sim B(n,p)$(参数 $n \geqslant 1, 0 \leqslant p \leqslant 1$) :

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

- 模型: 独立投币 n 次, 正面的总次数.
- 定理 2.1. 分布列的最大值点 k_0 如下:

若 $(n+1)p \notin \mathbb{Z}$, 则 $k_0 = [(n+1)p]$;
若 $(n+1)p \in \mathbb{Z}$, 则 $k_0 = (n+1)p$ 或 $(n+1)p - 1$.

有组合数公式：

$$\frac{p_n(k+1)}{p_n(k)} = \frac{n-k}{k+1} \cdot \frac{p}{1-p}$$

由于 $\frac{n-k}{k+1} \cdot \frac{p}{1-p} > 1$ 等价于 $k < (n+1)p - 1$，于是有：

(a) 当 $k < (n+1)p - 1$ 时，$p_n(k+1) > p_n(k)$

(b) 当 $k > (n+1)p - 1$ 时，$p_n(k+1) < p_n(k)$

(c) 当 $k = (n+1)p - 1$ 时，$p_n(k+1) = p_n(k)$

3. 泊松分布, $X \sim \mathcal{P}(\lambda)$(参数: $\lambda > 0$) :

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

- 模型: 例 2.3. 研究放射性物质在 8 分钟内放射出的粒子数 X.
- 计算频次

X 近似服从 $B(n,p)$, 故

$$P(X=k)=C_n^k p^k(1-p)^{n-k}\approx\frac{n!}{k!(n-k)!}p^k(1-p)^n$$
$$\approx\frac{(np)^k}{k!}\left(1-\frac{\lambda}{n}\right)^n=\frac{\lambda^k}{k!}e^{-\lambda}$$

- 上式即为 §1.7 第一近似公式.
- 泊松分布的分布列最大值点 $k_0=[\lambda]$.

泊松分布的分布列最大值点 $k_0 = [\lambda]$.

证明：注意到 $p_{k+1} = \frac{\lambda}{k+1} p_k$，故由分布函数知
若 $k+1 \leqslant \lambda$，则 $p_{k+1} \geqslant p_k$
若 $k+1 \geqslant \lambda$，则 $p_{k+1} \leqslant p_k$
因此当 $k_0 = [\lambda]$ 时，分布列取最大值。

已知某商场一天来的顾客服从参数为 λ 的泊松分布，而每个来商场的顾客购物概率为 p，证明此商场一天内购物的顾客数服从参数为 λp 的泊松分布

已知某商场一天来的顾客服从参数为 λ 的泊松分布，而每个来商场的顾客购物概率为 p，证明此商场一天内购物的顾客数服从参数为 λp 的泊松分布

解：用 Y 表示商场内一天购物的顾客数，则由全概率公式知，对任意正整数 k 有

$$P(Y=k)=\sum_{i=k}^{\infty}P(X=i)P(Y=k\mid X=i)=\sum_{i=k}^{\infty}\frac{\lambda^{i}e^{-\lambda}}{i!}C_i^k p^k(1-p)^{i-k}$$

$$=\frac{(\lambda p)^k}{k!}e^{-\lambda}\sum_{i=k}^{\infty}\frac{[\lambda(1-p)]^{i-k}}{(i-k)!}=\frac{(\lambda p)^k}{k!}e^{-\lambda}e^{\lambda(1-p)}=\frac{(\lambda p)^k}{k!}e^{-\lambda p}$$

4. 超几何分布, $X \sim H(N, D, n)$ (参数 N, D, n) :

$$P(X = k) = \frac{C_D^k C_{N-D}^{n-k}}{C_N^n}, \quad k = 0, 1, \cdots, n.$$

- 模型: N 个产品, D 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.
- 定理 2.3. 给定 n. 当 $N \to \infty, \frac{D}{N} \to p$ 时,

$$\frac{C_D^k C_{N-D}^{n-k}}{C_N^n} \to C_n^k p^k (1-p)^{n-k}, \quad k \geqslant 0$$

- 该定理的直观解释是，如果一批产品的总量 N 很大，其中次品占比为 p，则从整批产品随机抽取 n 个，抽到次品的个数 k 近似服从参数为 p，n 的二项分布

证明：由于 $0<p<1$，当 N 充分大时，$n<D<N$，且 n 是固定的，易知

$$
\begin{aligned}
\frac{C_D^k C_{N-D}^{n-k}}{C_N^n} &= \frac{D!}{k!(D-k)!}\cdot\frac{(N-D)!}{(n-k)!(N-D-n+k)!}\cdot\frac{n!(N-n)!}{N!} \\
&= \frac{n!}{k!(n-k)!}\cdot\frac{D(D-1)\cdots(D-k+1)}{N^k} \\
&\cdot\frac{(N-D)(N-D-1)\cdots(N-D-n+k+1)}{N^{n-k}} \\
&\cdot\frac{N^n}{N(N-1)\cdots(N-n+1)} \\
&= C_n^k(\prod_{i=1}^{k}\frac{D-i+1}{N})(\prod_{i=1}^{n-k}\frac{N-D-i+1}{N})(\prod_{i=1}^{n}\frac{N}{N-i+1}) \\
&\to C_n^k p^k(1-p)^{n-k} \quad (N\to\infty)
\end{aligned}
$$

5. 几何分布, $X \sim G(p)$, 参数 $0 < p < 1$:

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

- 模型: 独立重复投币中, 第一次投到 H 时的投币次数.
- $P(X > n) = (1-p)^n, \forall n \geqslant 0.$
- 无记忆性: $P(X - n = k \mid X > n) = P(X = k).$

- 6. 负二项分布, $X \sim NB(r,p)$, 参数 $r \geqslant 1, 0 < p < 1$:

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

- 模型: 独立重复投币中, 第 r 次投到 H 时的投币次数.
- 7. 离散均匀分布,

$$P(X=k) = \frac{1}{N}, \quad k = 1, \cdots, N.$$

- 模型: 古典概型.