4. 数字特征

4.1 期望

离散型 r.v. 的数学期望　　设 X 分布列为

X	x_1	x_2	$\cdots$
P	p_1	p_2	$\cdots$

若 $\sum_{i=1}^{+\infty} |x_i| p_i < +\infty$，则称 $\sum_{i=1}^{+\infty} x_i p_i$ 为 X 的(数学)期望(或均值)，

记为
$$EX = \sum_{i=1}^{+\infty} x_i p_i .$$

注：$\sum |x_i| p_i < +\infty \Rightarrow \sum x_i p_i$ 与求和顺序无关. 下面为与求和顺序有关的级数：

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots = \ln 2,$$
$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\cdots = \frac{3}{2}\ln 2.$$

例：设 $X \sim B(n,p)$. 证 $EX = np$.　(直观?)

证：
$$EX = \sum_{k=0}^{n} k\, P(X=k)$$
$$= \sum_{k=0}^{n} k\, C_n^k\, p^k (1-p)^{n-k}$$
$$= \sum_{k=1}^{n} np C_{n-1}^{k-1}\, p^{k-1} (1-p)^{n-1-(k-1)}$$
$$\overset{(l=k-1)}{=} np \sum_{l=0}^{n-1} C_{n-1}^{l}\, p^{l} (1-p)^{n-1-l} = np. \quad \square$$

$$k\, C_n^k = \frac{n!}{(k-1)!(n-k)!} = n\, C_{n-1}^{k-1}$$

例　设 $X \sim P(\lambda)$. 证 $EX = \lambda$　($\because$ λ=平均顾客数)

证：
$$EX = \sum_{k=0}^{+\infty} k\, P(X=k)$$
$$= \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda. \quad \square$$

例. 设 X 服从几何分布,

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

证 $EX=\frac{1}{p}$. （直观？）

证: $EX=\sum_{k=1}^{+\infty} k\,P(X=k)=\sum_{k=1}^{+\infty} k(1-p)^{k-1}p$

$|x|<1$ 时, $\sum_{k=1}^{+\infty} kx^{k-1}=\sum_{k=1}^{+\infty}(x^k)' \doteq \left(\sum_{k=1}^{+\infty}x^k\right)'$

$$=\left(\frac{x}{1-x}\right)'=\frac{1}{(1-x)^2}$$

$\therefore\ EX=\frac{1}{p}$. $\square$

例. 设 X 服从超几何分布, $M<N$, $n<N$.

$$P(X=k)=\frac{C_M^k\, C_{N-M}^{n-k}}{C_N^n}, \quad 0\leq k\leq n.$$

则 $EX=\frac{Mn}{N}$.

证. $EX=\sum_{k=0}^{n} k\,\frac{C_M^k\, C_{N-M}^{n-k}}{C_N^n}$

$$\sum_{k_1+k_2=n} C_{n_1}^{k_1}C_{n_2}^{k_2}=C_{n_1+n_2}^{n}$$

$$=\sum_{k=1}^{n}\frac{M\,C_{M-1}^{k-1}\,C_{N-M}^{n-k}}{C_N^n}=M\,\frac{C_{N-1}^{n-1}}{C_N^n}=\frac{Mn}{N} \quad \square$$

例: (Coupons) 有 n 类卡片; 你每天买一张, 买到的卡片随机. 令 T 表示集齐 n 类卡片所需时间. 求 ET.

解: 令 T_j 表示从集齐 j 类卡片到集齐 $j+1$ 类卡片所需时间. 则 $T=\sum_{j=0}^{n-1}T_j$ 且 $T_j\sim Geo\left(1-\frac{j}{n}\right)$.

$\therefore\ E[T_j]=\frac{n}{n-j}$ $\quad E[T]=\sum_{j=0}^{n-1}E[T_j]=\sum_{j=0}^{n-1}\frac{n}{n-j}=n\sum_{j=1}^{n}\frac{1}{j}$ $\square$

例 罐子中有 n 个球，标号从 1 到 n. 随机取出 k 个. 求取出的 k 个球标号和的期望.

解：记 $S=$ 取出的 k 个球和的期望.

$$A_i = \{\text{球 } i \text{ 被取出}\}$$

则 $S = \sum_{i=1}^{n} i \mathbb{1}_{A_i}$

$$\Rightarrow ES = \sum_{i=1}^{n} i\, p(A_i) = \frac{k}{n} \sum_{i=1}^{n} i = \frac{k(n+1)}{2} \qquad \square.$$

例. Pepy put a simple version of the following problem to Newton in 1693:

Sam 掷 $6n$ 次骰子，需要至少 n 个 6；Isaac 掷 $6(n+1)$ 次，需要至少 $n+1$ 个 6. 谁更可能如愿？

解：令 $X \sim B(6n, \frac{1}{6})$，$Y \sim B(6, \frac{1}{6})$，$X$ 与 Y 独立，

$Z = X+Y$. 则 $Z \sim B(6(n+1), \frac{1}{6})$. 即比较 $p(Z \geq n+1)$ 与 $p(X \geq n)$

$$p(Z \geq n+1) - p(X \geq n)$$

$$= \sum_{k=0}^{6} p(X \geq n+1-k)\, p(Y=k) - p(X \geq n)$$

$$\because\ p(X \geq n+1-k) - p(X \geq n)$$

$$= \begin{cases} \sum_{r=2}^{k} p(X = n+1-r) \leq (k-1)p(X=n), & k \geq 2 \\ -p(X=n), & k=0 \end{cases}$$

$$\leq (k-1)\, p(X=n)$$

$\therefore \quad p(Z \geq n+1) - p(X \geq n)$

$$\leq \sum_{k=0}^{6} (k-1)\, p(Y=k)\, p(X=n)$$

$$= (EY-1)\, p(X=n) = 0$$

$\therefore$ Sam 更可能如愿. $\square$

Example* (The probabilistic method)　$G=(V,E)$ 有限图

$W \subset V$, $e \in E$, 定义

$$I_W(e) = \begin{cases} 1, & 若 e 连接 W 与 W^c \\ 0, & 其它. \end{cases}$$

令 $N_W = \sum_{e \in E} I_W(e)$. 证明存在 $W \subseteq V$ s.t.

$$N_W \geq \frac{1}{2}|E|.$$

证: $\forall\, v \in V$, 令 $v \in W$ 以 $\frac{1}{2}$ 概率. 则

$$E[I_W(e)] = \frac{1}{2}, \quad \forall\, e \in E$$

$$\Rightarrow E[N_W] = \frac{1}{2}|E|$$

$\therefore\ \exists\, W$ s.t. $N_W \geq \frac{1}{2}|E|$. $\square$

连续型 r.v. 的数学期望　设 X 概率密度为 $f(x)$. 若 $\int_{-\infty}^{+\infty} |x| f(x)\, dx < +\infty$,

则定义期望(或均值)

$$EX = \int_{-\infty}^{+\infty} x f(x)\, dx.$$

直观: 划分 $\mathbb{R}$ 为 $x_0 < x_1 < \cdots < x_n$.

$$P(X \in (x_i, x_{i+1})) \approx f(x_i)\Delta x_i, \quad \Delta x_i = x_{i+1} - x_i$$

在 (x_i, x_{i+1}) 上, $X \approx x_i$.

$$\therefore EX \approx \sum x_i f(x_i) \Delta x_i \longrightarrow \int x f(x)\, dx.$$

例 若 $X \sim U(a,b)$，则 $EX = \frac{a+b}{2}$.

解. $EX = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{a+b}{2}$. ▯

例. 若 $X \sim E(\lambda)$，则 $EX = \frac{1}{\lambda}$.

解. $EX = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = \lambda^{-1}$. ▯

例. 若 $X \sim N(\mu, \sigma^2)$，则 $EX = \mu$.

解:
$$EX = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$\overset{(y=\frac{x-\mu}{\sigma})}{=} \int_{-\infty}^{+\infty} (\mu + y\sigma) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

$$= \mu$$ ▯

一般地，若 $f(x+\mu) = f(\mu - x)$，则 $EX = \mu$.

$\because \int x f(x) dx = \int x f(x+\mu) dx + \mu \int f(x+\mu) dx$

且 $x f(x+\mu) = x f(\mu - x) = -(-x f(\mu - x))$

即 $x f(x+\mu)$ 为奇函数

$\therefore EX = \mu$.

例. 若 $f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$，$-\infty < x < +\infty$，则称 X 服从柯西分布. 易证 柯西分布不存在期望.

期望的性质

1) $EC = C$.

2) $E(X+Y) = EX + EY$. $E(aX) = aEX$. a 常数.

一般地，$E\left(\sum_i k_i X_i\right) = \sum_i k_i EX_i$.

3) 若 $X_1, \cdots, X_n$ 相互独立，则

$$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} EX_i.$$

4) 若 $X \geq 0$，则 $EX \geq 0$. 若 $X_1 \geq X_2$，则 $EX_1 \geq EX_2$.

5) $|EX| \leq E|X|$.

6) 柯西-施瓦茨 (Cauchy-Schwarz) 不等式：

$$E(XY)^2 \leq E(X^2)E(Y^2).$$

$g(t) = E(tX-Y)^2 \geq 0, \ \forall t.$

证明见板书.

<u>例</u>. $X \sim B(n,p)$. 求 $P(X=$偶数$)$.

<u>解</u>

X_i	-1	1
σ	p	$1-p$

$X_i, \ i \geq 1$，独立

$$X = \#\{1 \leq i \leq n: \ X_i = -1\} \sim B(n,p)$$

令 $Y = \prod_{i=1}^{n} X_i$

则 $Y = \begin{cases} 1, & X \text{ 偶} \\ -1, & X \text{ 奇} \end{cases}$

$\therefore EY = P_{偶} - P_{奇} = \prod_{i=1}^{n} EX_i = (1-2p)^n.$

又 $P_{偶} + P_{奇} = 1$

$\therefore P_{偶} = \dfrac{1+(1-2p)^n}{2}$ $\square$.