概率论

胡太忠
thu@ustc.edu.cn

安徽 合肥

2023 年 9 月

第 5 章　数字特征和特征函数

（18 课时）

- 数字特征
- 求数学期望的基本技巧和方法
- 特征函数的性质
- 特征函数的应用

5.1.1 数学期望的定义

▶ 定义 5.1.1 (一般 rv 期望的定义) 设 $X \sim F$, 若

$$\int_{-\infty}^{\infty} |x| \, \mathrm{d}F(x) < \infty,$$

则 X 的 (数学) 期望存在有限, 定义为

$$\mathrm{E}\, X = \int_{-\infty}^{\infty} x \, \mathrm{d}F(x).$$

※ 数学期望 (Expectation) 与均值 (Mean)

※ Lebesgue-Stieltjes 积分与 Riemann 积分

5.1.1 数学期望的定义

- **定义 5.1.1** (一般 rv 期望的定义) 设 $X \sim F$, 若

$$\int_{-\infty}^{\infty} |x| \,\mathrm{d}F(x) < \infty,$$

则 X 的 (数学) 期望存在有限, 定义为

$$\mathrm{E}\, X = \int_{-\infty}^{\infty} x \,\mathrm{d}F(x).$$

※ 数学期望 (Expectation) 与均值 (Mean)

※ Lebesgue-Stieltjes 积分与 Riemann 积分

- **定义 5.1.2** (离散 rv 期望的定义) 设 X 取值于 $\{a_1, a_2, \ldots, a_n\}$, $n \le \infty$, X 的 pmf 为 $\{p_i\}$. 若
$$\sum_{i=1}^{n} |a_i| p_i < \infty, \tag{1}$$
则 X 的 (数学) 期望存在有限, 定义为
$$\mathrm{E}\, X = \sum_{i=1}^{n} a_i p_i.$$

若 $\sum_{i=1}^{n} a_i p_i$ 条件收敛, 则称 X 的期望不存在.

※ 命题 设 $\sum_{n=1}^{\infty} b_n$ 收敛, 但 $\sum_{n=1}^{\infty} |b_n| = +\infty$, 则对 $\forall\ r \in \Re$, 必存在 $\{b_n\}$ 的重新排列 $\{b_n^*\}$, 使得
$$\sum_{n=1}^{\infty} b_n^* = r.$$

- **定义 5.1.2** (离散 rv 期望的定义) 设 X 取值于 $\{a_1, a_2, \ldots, a_n\}$, $n \le \infty$, X 的 pmf 为 $\{p_i\}$. 若

$$\sum_{i=1}^{n} |a_i| p_i < \infty, \tag{1}$$

则 X 的 (数学) 期望存在有限, 定义为

$$\mathrm{E}\, X = \sum_{i=1}^{n} a_i p_i.$$

若 $\sum_{i=1}^{n} a_i p_i$ 条件收敛, 则称 X 的期望不存在.

※ **命题** 设 $\sum_{n=1}^{\infty} b_n$ 收敛, 但 $\sum_{n=1}^{\infty} |b_n| = +\infty$, 则对 $\forall\ r \in \Re$, 必存在 $\{b_n\}$ 的重新排列 $\{b_n^*\}$, 使得

$$\sum_{n=1}^{\infty} b_n^* = r.$$

- 【例 5.1.2】 一个人连续抛掷一枚均匀的硬币，直到抛出反面为止. 如果此前他抛出了 n 次正面，则发给他 2^n 元奖金. 以 X 表示他所得到的奖金数目，试求 X 的数学期望.

解：容易看出，X 的 pmf 为

$$\mathrm{P}\left(X=2^{n-1}\right)=\frac{1}{2^n},\quad n\geq 1.$$

由于

$$\sum_{n=1}^{\infty} a_n p_n=\sum_{n=1}^{\infty}\frac{2^{n-1}}{2^n}=+\infty,$$

所以 X 的数学期望不是有限的. 但我们说其数学期望是存在的，$\mathrm{E}X=+\infty$. ∎

- 【例 5.1.2】一个人连续抛掷一枚均匀的硬币，直到抛出反面为止。如果此前他抛出了 n 次正面，则发给他 2^n 元奖金。以 X 表示他所得到的奖金数目，试求 X 的数学期望。

解：容易看出，X 的 pmf 为

$$\mathrm{P}\left(X = 2^{n-1}\right) = \frac{1}{2^n}, \quad n \geq 1.$$

由于

$$\sum_{n=1}^{\infty} a_n p_n = \sum_{n=1}^{\infty} \frac{2^{n-1}}{2^n} = +\infty,$$

所以 X 的数学期望不是有限的。但我们说其数学期望是存在的，$\mathrm{E}\, X = +\infty$. ∎

- 定义 5.1.3 (连续型 rv 期望的定义) 设 X 的 pdf 为 f. 若

$$\int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x < \infty, \tag{2}$$

则 X 的 (数学) 期望存在有限, 定义为

$$\mathrm{E}\,X = \int_{-\infty}^{\infty} x\, f(x) dx.$$

若上积分条件收敛, (2) 不成立, 则称 X 的期望不存在.

- 新的看法 (离散 rv 期望)

一个新观点去看待离散 rv X 的期望 $\mathrm{E}\,X$；设 X 取值于 $\{a_i, i\in I\}$，$p_i=\mathrm{P}\,(X=a_i)$，记

$$A_i=\{w\in\Omega : X(w)=a_i\},$$

则由 $\mathrm{E}\,X$ 的定义得

$$\mathrm{E}\,X=\sum_{i\in I}a_ip_i=\sum_{i\in I}a_i\mathrm{P}\,(A_i)=\sum_{i\in I}\int_{A_i}X\,\mathrm{d}P=\int_\Omega X\,\mathrm{d}P.$$

※ 概率论中数学期望的定义的标准路线：

$$\text{阶梯 rv}\ \underbrace{\longrightarrow}_{\text{单调收敛定理}}\ \text{非负 rv}\ \longrightarrow\ \text{一般 rv}$$

▶ 新的看法 (离散 rv 期望)

一个新观点去看待离散 rv X 的期望 $\mathrm{E}\,X$；设 X 取值于 $\{a_i, i \in I\}$，$p_i = \mathrm{P}\,(X = a_i)$，记

$$A_i = \{w \in \Omega : X(w) = a_i\},$$

则由 $\mathrm{E}\,X$ 的定义得

$$\mathrm{E}\,X = \sum_{i \in I} a_i p_i = \sum_{i \in I} a_i \mathrm{P}\,(A_i) = \sum_{i \in I} \int_{A_i} X \,\mathrm{d}P = \int_{\Omega} X \,\mathrm{d}P.$$

※ 概率论中数学期望的定义的标准路线：

$$\text{阶梯 rv} \underbrace{\longrightarrow}_{\text{单调收敛定理}} \text{非负 rv} \longrightarrow \text{一般 rv}$$

- $\mathrm{E}\,X$ 的标准定义

Step 1 阶梯 rv X: $\mathrm{E}\,X$ $\surd$

Step 2 非负 rv X: 存在 rv 序列 $\{X_n\}$, 满足 $0 \le X_n \uparrow X$, 其中

$$X_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I\left(\frac{k-1}{2^n} \le X < \frac{k}{2^n}\right) + n\, I(X \ge n), \quad n \ge 1.$$

由单调收敛定理得

$$\mathrm{E}\,X = \lim_{n\to\infty} \mathrm{E}\,X_n.$$

Step 3 一般 rv X: $X = X_+ - X_-$, $\mathrm{E}\,X_+$ 和 $\mathrm{E}\,X_-$ 可定义.

- 若 $\mathrm{E}\,X_+ < \infty$, $\mathrm{E}\,X_- < \infty$, 则称 $\mathrm{E}\,X$ 存在有限, 且
$$\mathrm{E}\,X = \mathrm{E}\,X_+ - \mathrm{E}\,X_-;$$
- 若 $\mathrm{E}\,X_+ < \infty$ 或 $\mathrm{E}\,X_- < \infty$, 则称 $\mathrm{E}\,X = \mathrm{E}\,X_+ - \mathrm{E}\,X_-$ 存在, 但未必有限.

- $\mathrm{E}\,X$ 的标准定义

Step 1 阶梯 rv X: $\mathrm{E}\,X$ $\surd$

Step 2 非负 rv X: 存在 rv 序列 $\{X_n\}$, 满足 $0 \le X_n \uparrow X$, 其中

$$X_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I\left(\frac{k-1}{2^n} \le X < \frac{k}{2^n}\right) + n\, I(X \ge n), \quad n \ge 1.$$

由单调收敛定理得

$$\mathrm{E}\,X = \lim_{n\to\infty} \mathrm{E}\,X_n.$$

Step 3 一般 rv X: $X = X_+ - X_-$, $\mathrm{E}\,X_+$ 和 $\mathrm{E}\,X_-$ 可定义.

- 若 $\mathrm{E}\,X_+ < \infty$, $\mathrm{E}\,X_- < \infty$, 则称 $\mathrm{E}\,X$ 存在有限, 且

$$\mathrm{E}\,X = \mathrm{E}\,X_+ - \mathrm{E}\,X_-;$$

- 若 $\mathrm{E}\,X_+ < \infty$ 或 $\mathrm{E}\,X_- < \infty$, 则称 $\mathrm{E}\,X = \mathrm{E}\,X_+ - \mathrm{E}\,X_-$ 存在, 但未必有限.

- $\mathrm{E}\,X$ 的标准定义

Step 1 阶梯 rv X: $\mathrm{E}\,X$ $\surd$

Step 2 非负 rv X: 存在 rv 序列 $\{X_n\}$, 满足 $0 \le X_n \uparrow X$, 其中

$$X_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I\left(\frac{k-1}{2^n} \le X < \frac{k}{2^n}\right) + n\, I(X \ge n), \quad n \ge 1.$$

由单调收敛定理得

$$\mathrm{E}\,X = \lim_{n\to\infty} \mathrm{E}\,X_n.$$

Step 3 一般 rv X: $X = X_+ - X_-$, $\mathrm{E}\,X_+$ 和 $\mathrm{E}\,X_-$ 可定义.

- 若 $\mathrm{E}\,X_+ < \infty$, $\mathrm{E}\,X_- < \infty$, 则称 $\mathrm{E}\,X$ 存在有限, 且
$$\mathrm{E}\,X = \mathrm{E}\,X_+ - \mathrm{E}\,X_-;$$
- 若 $\mathrm{E}\,X_+ < \infty$ 或 $\mathrm{E}\,X_- < \infty$, 则称 $\mathrm{E}\,X = \mathrm{E}\,X_+ - \mathrm{E}\,X_-$ 存在, 但未必有限.

- 定理 5.1.1 (积分变换定理)

  设 X 为 $(\Omega, \mathscr{F}, \mathrm{P})$ 上的 rv，$X \sim F$, 则对任意 Borel 可测函数 g, 有

$$\int_{\Re} g(x)\,\mathrm{d}F(x) = \int_{\Omega} g(X)\,\mathrm{d}P,$$

  其中假设一端积分存在 (则另一端自然也存在).

※

$$\mathrm{E}\,[g(X)] = \int_{\Re} g(x)\,\mathrm{d}F(x).$$

※ 定理 5.1.1 求 $\mathrm{E}\,[g(X)]$ 的便捷之处.

※ $\mathrm{E}\,[g(X_1, X_2, \ldots, X_n)]$ 的类似表达：设 $(X_1, \ldots, X_n) \sim F$, 则

$$\mathrm{E}\,[g(X_1, \ldots, X_n)] = \int_{\Re^n} g(x_1, \ldots, x_n)\,\mathrm{d}F(x_1, \ldots, x_n).$$

- 定理 5.1.1 (积分变换定理)

  设 X 为 $(\Omega, \mathscr{F}, \mathrm{P})$ 上的 rv, $X \sim F$, 则对任意 Borel 可测函数 g, 有

$$\int_{\Re} g(x)\,\mathrm{d}F(x) = \int_{\Omega} g(X)\,\mathrm{d}P,$$

  其中假设一端积分存在 (则另一端自然也存在).

※

$$\mathrm{E}\,[g(X)] = \int_{\Re} g(x)\,\mathrm{d}F(x).$$

※ 定理 5.1.1 求 $\mathrm{E}\,[g(X)]$ 的便捷之处.

※ $\mathrm{E}\,[g(X_1, X_2, \ldots, X_n)]$ 的类似表达: 设 $(X_1, \ldots, X_n) \sim F$, 则

$$\mathrm{E}\,[g(X_1, \ldots, X_n)] = \int_{\Re^n} g(x_1, \ldots, x_n)\,\mathrm{d}F(x_1, \ldots, x_n).$$

- 定理 5.1.1 (积分变换定理)

  设 X 为 $(\Omega, \mathscr{F}, \mathrm{P})$ 上的 rv，$X \sim F$, 则对任意 Borel 可测函数 g, 有

$$\int_{\Re} g(x)\,\mathrm{d}F(x) = \int_{\Omega} g(X)\,\mathrm{d}P,$$

  其中假设一端积分存在 (则另一端自然也存在).

※

$$\mathrm{E}\,[g(X)] = \int_{\Re} g(x)\,\mathrm{d}F(x).$$

※ 定理 5.1.1 求 $\mathrm{E}\,[g(X)]$ 的便捷之处.

※ $\mathrm{E}\,[g(X_1, X_2, \ldots, X_n)]$ 的类似表达：设 $(X_1, \ldots, X_n) \sim F$, 则

$$\mathrm{E}\,[g(X_1, \ldots, X_n)] = \int_{\Re^n} g(x_1, \ldots, x_n)\,\mathrm{d}F(x_1, \ldots, x_n).$$

- 定理 5.1.1 (积分变换定理)

  设 X 为 $(\Omega, \mathscr{F}, \mathrm{P})$ 上的 rv, $X \sim F$, 则对任意 Borel 可测函数 g, 有

$$\int_{\Re} g(x)\,\mathrm{d}F(x) = \int_{\Omega} g(X)\,\mathrm{d}P,$$

  其中假设一端积分存在 (则另一端自然也存在).

※

$$\mathrm{E}[g(X)] = \int_{\Re} g(x)\,\mathrm{d}F(x).$$

※ 定理 5.1.1 求 $\mathrm{E}[g(X)]$ 的便捷之处.

※ $\mathrm{E}[g(X_1, X_2, \ldots, X_n)]$ 的类似表达: 设 $(X_1, \ldots, X_n) \sim F$, 则

$$\mathrm{E}[g(X_1, \ldots, X_n)] = \int_{\Re^n} g(x_1, \ldots, x_n)\,\mathrm{d}F(x_1, \ldots, x_n).$$

- 【例 5.1.1】 设 X 服从柯西分布, 其 pdf 为

$$f(x)=\frac{1}{\pi(1+x^2)},\quad x\in\Re,$$

则

$$\mathrm{E}\,X_-=\int_{-\infty}^{0}-xf(x)\,\mathrm{d}x=\infty,\quad \mathrm{E}\,X_+=\int_{0}^{\infty}xf(x)\,\mathrm{d}x=\infty.$$

于是 X 的数学期望不存在.

在计算机上独立重复产生 10^7 个和 X 同分布的随机变量的观测值, 利用前 n 个观测值计算的平均数 $\overline{X}_n$ 如下:

n	10	10^2	10^3	10^4	10^5	10^6	10^7
$\overline{X}_n$	-0.103	0.3884	-11.79	-3.628	-0.810	0.2442	-11.28

可见, 当 $n\to\infty$, 上述的 $\overline{X}_n$ 并不收敛.

- 【例 5.1.1】 设 X 服从柯西分布, 其 pdf 为

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \Re,$$

则

$$\mathrm{E}\, X_- = \int_{-\infty}^{0} -xf(x)\,\mathrm{d}x = \infty, \quad \mathrm{E}\, X_+ = \int_{0}^{\infty} xf(x)\,\mathrm{d}x = \infty.$$

于是 X 的数学期望不存在.

在计算机上独立重复产生 10^7 个和 X 同分布的随机变量的观测值, 利用前 n 个观测值计算的平均数 $\overline{X}_n$ 如下:

n	10	10^2	10^3	10^4	10^5	10^6	10^7
$\overline{X}_n$	-0.103	0.3884	-11.79	-3.628	-0.810	0.2442	-11.28

可见, 当 $n \to \infty$, 上述的 $\overline{X}_n$ 并不收敛.

- 【例 5.1.1 (续)】 设 X 服从柯西分布，则 $Y=|X|$ 有 pdf

$$g(x)=\frac{2}{\pi(1+x^2)},\quad x\in\Re_+,$$

则

$$\mathrm{E}\,Y=\int_0^\infty xg(x)\,\mathrm{d}x=\infty.$$

在计算机上独立重复产生 10^7 个和 Y 同分布的 rv 的观测值，利用前 n 个观测值计算的平均数 $\overline{y}_n$ 如下：

n	10	10^2	10^3	10^4	10^5	10^6	10^7
$\overline{Y}_n$	2.413	2.365	9.588	7.443	15.672	11.441	21.991

可以看出，当 $n\to\infty$ 时，上述的 $\overline{Y}_n$ 有越来越大的趋势.

Example (1)

- 若 $X \sim B(n, p)$, $0 < p < 1$, 则 $\mathrm{E}\, X = np$.
- 若 $X \sim \mathrm{Poi}(\lambda)$, $\lambda > 0$, 则 $\mathrm{E}\, X = \lambda$.
- 若 $X \sim \mathrm{Geo}(p)$, $0 < p < 1$, 则

$$\mathrm{E}\, X = \frac{1-p}{p}.$$

- 若 $X \sim \mathrm{NB}(\alpha, p)$, $\alpha > 0$, $0 < p < 1$, 则

$$\mathrm{E}\, X = \frac{\alpha(1-p)}{p}.$$

Example (2)

- 若 $X \sim U(a,b)$, $a<b$, 则 $\mathrm{E}\,X=(a+b)/2$.
- 若 $X \sim \mathrm{Exp}(\lambda)$, $\lambda>0$, 则
$$\mathrm{E}\,X=\frac{1}{\lambda}.$$
- 若 $X \sim \Gamma(\alpha,\beta)$, $\alpha>0$, $\beta>0$, 则
$$\mathrm{E}\,X=\frac{\alpha}{\beta}.$$
- 若 $X \sim N(\mu,\sigma^2)$, $\mu\in\Re$, $\sigma>0$, 则 $\mathrm{E}\,X=\mu$.

Example (3)

设 X 的 pdf f 关于点 μ 对称, 即

$$f(x+\mu)=f(\mu-x), \quad \forall\, x\in\Re,$$

则 $\mathrm{E}\,X=\mu$ (假定 $\mathrm{E}\,X$ 存在有限).

证: 令 $g(x)=xf(\mu+x)$, 则 g 为奇函数, 于是

$$\begin{aligned}
\mathrm{E}\,X &= \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x \\
&= \mu+\int_{-\infty}^{\infty}(x-u)f((x-\mu)+\mu)\,\mathrm{d}x \\
&= \mu+\int_{-\infty}^{\infty} tg(t)\,\mathrm{d}t \\
&= \mu.
\end{aligned}$$

Example (3)

设 X 的 pdf f 关于点 μ 对称, 即

$$f(x+\mu) = f(\mu - x), \quad \forall\, x \in \Re,$$

则 $\mathrm{E}\,X = \mu$ (假定 $\mathrm{E}\,X$ 存在有限).

证: 令 $g(x) = xf(\mu + x)$, 则 g 为奇函数, 于是

$$\begin{aligned}
\mathrm{E}\,X &= \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x \\
&= \mu + \int_{-\infty}^{\infty} (x-u)f((x-\mu)+\mu)\,\mathrm{d}x \\
&= \mu + \int_{-\infty}^{\infty} tg(t)\,\mathrm{d}t \\
&= \mu.
\end{aligned}$$

- 【例 5.1.6】 一盒中有编号 $1, 2, \ldots, n$ 的小球, 从中不放回任意取出 m $(m < n)$ 个, X 表示这 m 个小球号码之和，求 $\mathrm{E}\,X$.

解法一：记 $\Lambda = \{1, 2, \ldots, n\}$, 样本空间为

$$\Omega = \{A : A \subset \Lambda, |A| = m\},$$

A 也表示事件 “取出的球号码恰好为 A 中的元”. 古典概型：

$$\mathrm{P}\,(A) = 1 \Big/ \binom{n}{m}, \quad A \in \Omega.$$

于是,

$$\mathrm{E}\,X = \frac{1}{\binom{n}{m}} \sum_{1 \le a_1 < a_2 < \cdots < a_m \le n} \sum_{k=1}^{m} a_k = \frac{\binom{n-1}{m-1}}{\binom{n}{m}} \sum_{k=1}^{n} k = \frac{m}{2}(n+1),$$

其中每个 $j \in \Lambda$ 在第一个求和号中出现 $\binom{n-1}{m-1}$ 次. ∎

- 【例 5.1.6】 一盒中有编号 $1, 2, \ldots, n$ 的小球, 从中不放回任意取出 m $(m < n)$ 个, X 表示这 m 个小球号码之和, 求 $\mathrm{E}\, X$.

解法一: 记 $\Lambda = \{1, 2, \ldots, n\}$, 样本空间为

$$\Omega = \{A : A \subset \Lambda, |A| = m\},$$

A 也表示事件 “取出的球号码恰好为 A 中的元”. 古典概型:

$$\mathrm{P}(A) = 1 \Big/ \binom{n}{m}, \qquad A \in \Omega.$$

于是,

$$\mathrm{E}\, X = \frac{1}{\binom{n}{m}} \sum_{1 \le a_1 < a_2 < \cdots < a_m \le n} \sum_{k=1}^{m} a_k = \frac{\binom{n-1}{m-1}}{\binom{n}{m}} \sum_{k=1}^{n} k = \frac{m}{2}(n+1),$$

其中每个 $j \in \Lambda$ 在第一个求和号中出现 $\binom{n-1}{m-1}$ 次. ∎

- 【例 5.1.a】 现有 n 把外形相同的钥匙，其中只有一把能打开锁，用它们逐一去试开锁，每一次开锁是相互独立的，每把钥匙试开锁一次后拿走，记 X 表示打开锁时已经试开过锁的次数．求 $\mathrm{E}X$.

解法一： 直观容易看出 X 服从 $\{1,2,\ldots,n\}$ 上的均匀分布，于是，

$$\mathrm{E}X=\frac{1}{2}(n+1).$$

另解：对任意 $k\in\{1,2,\ldots,n\}$,

$$\mathrm{P}\,(X=k)=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\cdots\cdot\frac{1}{n-k+1}=\frac{1}{n}.$$

因此，$\mathrm{E}X=(n+1)/2$. ∎

- 【例 5.1.a】 现有 n 把外形相同的钥匙，其中只有一把能打开锁，用它们逐一去试开锁，每一次开锁是相互独立的，每把钥匙试开锁一次后拿走，记 X 表示打开锁时已经试开过锁的次数. 求 $\mathrm{E}X$.

解法一： 直观容易看出 X 服从 $\{1,2,\ldots,n\}$ 上的均匀分布，于是，

$$\mathrm{E}X = \frac{1}{2}(n+1).$$

另解：对任意 $k \in \{1,2,\ldots,n\}$,

$$\mathrm{P}(X=k) = \frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\cdots\cdot\frac{1}{n-k+1} = \frac{1}{n}.$$

因此，$\mathrm{E}X = (n+1)/2$. ∎

5.1.2 数学期望的性质

- **定理 5.1.2** 假设以下出现的数学期望存在有限.
  - 线性性:

$$\mathrm{E}[aX] = a\mathrm{E}X,\ \forall\, a \in \Re; \quad \mathrm{E}[X+Y] = \mathrm{E}X + \mathrm{E}Y.$$

  - 非负性: 若 $X \geq 0$, 则 $\mathrm{E}X \geq 0$.
  - 单调性: 若 $X \geq Y$, 则 $\mathrm{E}X \geq \mathrm{E}Y$.

- **定理 5.1.4** 设 X, Y 相互独立, 若 $\mathrm{E}X, \mathrm{E}Y$ 存在有限, 则 $\mathrm{E}[XY]$ 存在, 且

$$\mathrm{E}[XY] = \mathrm{E}X \cdot \mathrm{E}Y.$$

证: 设 $(X, Y) \sim F$, $X \sim F_1$, $Y \sim F_2$, 则

$$\mathrm{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,\mathrm{d}F(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,\mathrm{d}F_1(x)\,\mathrm{d}F_2(y).$$

5.1.2 数学期望的性质

- **定理 5.1.2** 假设以下出现的数学期望存在有限.
  - 线性性:
  $$\mathrm{E}[aX] = a\mathrm{E}\,X,\ \forall\ a \in \Re; \quad \mathrm{E}[X+Y] = \mathrm{E}\,X + \mathrm{E}\,Y.$$
  - 非负性: 若 $X \geq 0$, 则 $\mathrm{E}\,X \geq 0$.
  - 单调性: 若 $X \geq Y$, 则 $\mathrm{E}\,X \geq \mathrm{E}\,Y$.
- **定理 5.1.4** 设 X, Y 相互独立，若 $\mathrm{E}\,X, \mathrm{E}\,Y$ 存在有限，则 $\mathrm{E}[XY]$ 存在，且
$$\mathrm{E}[XY] = \mathrm{E}\,X \cdot \mathrm{E}\,Y.$$

证: 设 $(X, Y) \sim F$, $X \sim F_1$, $Y \sim F_2$, 则

$$\mathrm{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,\mathrm{d}F(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,\mathrm{d}F_1(x)\,\mathrm{d}F_2(y).$$

5.1.2 数学期望的性质

- **定理 5.1.2** 假设以下出现的数学期望存在有限.
  - 线性性:
$$\mathrm{E}[aX] = a\mathrm{E}X,\ \forall\ a \in \Re;\quad \mathrm{E}[X+Y] = \mathrm{E}X + \mathrm{E}Y.$$
  - 非负性: 若 $X \geq 0$, 则 $\mathrm{E}X \geq 0$.
  - 单调性: 若 $X \geq Y$, 则 $\mathrm{E}X \geq \mathrm{E}Y$.
- **定理 5.1.4** 设 X, Y 相互独立, 若 $\mathrm{E}X, \mathrm{E}Y$ 存在有限, 则 $\mathrm{E}[XY]$ 存在, 且
$$\mathrm{E}[XY] = \mathrm{E}X \cdot \mathrm{E}Y.$$

证: 设 $(X, Y) \sim F$, $X \sim F_1$, $Y \sim F_2$, 则
$$\mathrm{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,\mathrm{d}F(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,\mathrm{d}F_1(x)\,\mathrm{d}F_2(y).$$

- 求数学期望的方法一

  随机变量分解法：若

$$Y = X_1 + X_2 + \cdots + X_n,$$

则

$$\mathrm{E}\, Y = \mathrm{E}\, X_1 + \mathrm{E}\, X_2 + \cdots + \mathrm{E}\, X_n.$$

※ 优点：

- 不需要去考虑 $X_1, X_2, \ldots, X_n$ 之间的相依关系；
- $X_1, X_2, \ldots, X_n$ 通常为简单的 Bernoulli rv.

※ 技巧：按照“时间”的先后顺序进行分解 (特别当 Y 是与时间有关的计数变量时).

- 求数学期望的方法一

随机变量分解法: 若

$$Y = X_1 + X_2 + \cdots + X_n,$$

则

$$\mathrm{E}\, Y = \mathrm{E}\, X_1 + \mathrm{E}\, X_2 + \cdots + \mathrm{E}\, X_n.$$

※ 优点:

- 不需要去考虑 $X_1, X_2, \ldots, X_n$ 之间的相依关系;
- $X_1, X_2, \ldots, X_n$ 通常为简单的 Bernoulli rv.

※ 技巧: 按照"时间"的先后顺序进行分解 (特别当 Y 是与时间有关的计数变量时).

随机变量分解的方法

- 【例 5.1.4】 一盒中有 a 个白球和 b 个黑球, 从盒子中无放回地取出 n 个球, 其中 $n \le a+b$, 记 X 为取出的球中白球个数, 求 $\mathrm{E}\,X$.

解: 设 n 个球是依序从盒子中取出, 定义

$$I_k = \begin{cases} 1, & \text{取出的第 } k \text{ 个球为白色}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\, I_k = \frac{a}{a+b}$$

于是

$$\mathrm{E}\,X = n\,\mathrm{E}\,I_1 = \frac{na}{a+b} \quad \blacksquare$$

※ 有放回情形如何?

随机变量分解的方法

- 【例 5.1.4】 一盒中有 a 个白球和 b 个黑球，从盒子中无放回地取出 n 个球，其中 $n \le a+b$，记 X 为取出的球中白球个数，求 $\mathrm{E}\,X$.

解：设 n 个球是依序从盒子中取出，定义

$$I_k = \begin{cases} 1, & \text{取出的第 } k \text{ 个球为白色}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\,I_k = \frac{a}{a+b}$$

于是

$$\mathrm{E}\,X = n\,\mathrm{E}\,I_1 = \frac{na}{a+b} \quad \blacksquare$$

※ 有放回情形如何？

随机变量分解的方法

- 【例 5.1.4】 一盒中有 a 个白球和 b 个黑球，从盒子中无放回地取出 n 个球，其中 $n \le a+b$，记 X 为取出的球中白球个数，求 $\mathrm{E}\,X$.

解：设 n 个球是依序从盒子中取出，定义

$$I_k = \begin{cases} 1, & \text{取出的第 } k \text{ 个球为白色}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\, I_k = \frac{a}{a+b}$$

于是

$$\mathrm{E}\,X = n\,\mathrm{E}\,I_1 = \frac{na}{a+b} \quad \blacksquare$$

※ 有放回情形如何？

随机变量分解的方法

- 【例 5.1.6 (续)】 一盒中有编号 $1, 2, \ldots, n$ 的小球, 从中不放回任意取出 m ($m < n$) 个, X 表示这 m 个小球的号码之和，求 $\mathrm{E}\,X$.

解法二: 对 $\forall\, j = 1, \ldots, m$, 定义

$$Y_j = \text{第 } j \text{ 次取出的小球上的号码},$$

则

$$X = \sum_{i=1}^{m} Y_i.$$

利用 Y_j 服从 $\{1, 2, \ldots, n\}$ 上的离散均匀分布，于是 $\mathrm{E}\,Y_j = (n+1)/2$, $\forall j$. 因此,

$$\mathrm{E}\,X = \sum_{i=1}^{m} \mathrm{E}\,Y_i = \frac{m(n+1)}{2}. \quad \blacksquare$$

随机变量分解的方法

- 【例 5.1.6 (续)】 一盒中有编号 $1, 2, \ldots, n$ 的小球, 从中不放回任意取出 m $(m < n)$ 个, X 表示这 m 个小球的号码之和, 求 $\mathrm{E}\,X$.

解法二: 对 $\forall\, j = 1, \ldots, m$, 定义

$$Y_j = \text{第 } j \text{ 次取出的小球上的号码},$$

则

$$X = \sum_{i=1}^{m} Y_i.$$

利用 Y_j 服从 $\{1, 2, \ldots, n\}$ 上的离散均匀分布, 于是 $\mathrm{E}\,Y_j = (n+1)/2$, $\forall j$. 因此,

$$\mathrm{E}\,X = \sum_{i=1}^{m} \mathrm{E}\,Y_i = \frac{m(n+1)}{2}. \quad \blacksquare$$

- 【例 5.1.a (续)】 现有 n 把外形相同的钥匙，其中只有一把能打开锁，现逐一去试开锁，每一次开锁是相互独立的，每把钥匙试开锁一次后拿走，记 X 表示打开锁时已经试开过锁的次数. 求 $\mathrm{E}\,X$.

解法二： 记 A_i 表示第 i 次取出的钥匙恰好能打开锁, 并定义

$$X_i = \begin{cases} 1, & \text{前 } i-1 \text{ 次摸取的钥匙未能打开锁;} \\ 0, & \text{否则,} \end{cases} \qquad i = 1, \ldots, n.$$

于是, $X = \sum_{i=1}^{n} X_i$. 显然, $X_1 = 1$, $\mathrm{E}\,X_1 = 1$,

$$\begin{aligned} \mathrm{E}\,X_i &= \mathrm{P}\,(A_1^c \cdots A_{i-1}^c) = \mathrm{P}\,(A_{i-1}^c | A_{i-2}^c \cdots A_1^c) \cdots \mathrm{P}\,(A_2^c | A_1^c) \mathrm{P}\,(A_1^c) \\ &= \frac{n-i+1}{n-i+2} \cdots \frac{n-2}{n-1} \cdot \frac{n-1}{n} = \frac{n-i+1}{n} \end{aligned}$$

因此,

$$\mathrm{E}\,X = 1 + \sum_{i=2}^{n} \mathrm{E}\,X_i = \sum_{i=1}^{n} \frac{n-i+1}{n} = \frac{n+1}{2}.$$

- 【例 5.1.a (续)】 现有 n 把外形相同的钥匙，其中只有一把能打开锁，现逐一去试开锁，每一次开锁是相互独立的，每把钥匙试开锁一次后拿走，记 X 表示打开锁时已经试开过锁的次数. 求 $\mathrm{E}\,X$.

解法二： 记 A_i 表示第 i 次取出的钥匙恰好能打开锁, 并定义

$$X_i = \begin{cases} 1, & \text{前 } i-1 \text{ 次摸取的钥匙未能打开锁;} \\ 0, & \text{否则,} \end{cases} \qquad i = 1, \ldots, n.$$

于是, $X = \sum_{i=1}^n X_i$. 显然, $X_1 = 1$, $\mathrm{E}\,X_1 = 1$,

$$\begin{aligned} \mathrm{E}\,X_i &= \mathrm{P}\,(A_1^c \cdots A_{i-1}^c) = \mathrm{P}\,(A_{i-1}^c | A_{i-2}^c \cdots A_1^c) \cdots \mathrm{P}\,(A_2^c | A_1^c) \mathrm{P}\,(A_1^c) \\ &= \frac{n-i+1}{n-i+2} \cdots \frac{n-2}{n-1} \cdot \frac{n-1}{n} = \frac{n-i+1}{n} \end{aligned}$$

因此,

$$\mathrm{E}\,X = 1 + \sum_{i=2}^n \mathrm{E}\,X_i = \sum_{i=1}^n \frac{n-i+1}{n} = \frac{n+1}{2}.$$

随机变量分解的方法

- 【例 5.1.b】 在一次聚会上，n 个人将自己的帽子混在一起，会后每人随机取帽子，记 X 为能正确取到自己帽子的人数. 求 $\mathrm{E}\,X$.

解：首先，将 n 人编号，然后依序取帽. 定义

$$I_k = \begin{cases} 1, & \text{第 } k \text{ 人取到自己的帽子}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\, I_k = \frac{1}{n}, \quad \forall k,$$

于是

$$\mathrm{E}\, X = 1. \quad \blacksquare$$

随机变量分解的方法

- 【例 5.1.b】 在一次聚会上，n 个人将自己的帽子混在一起，会后每人随机取帽子，记 X 为能正确取到自己帽子的人数. 求 $\mathrm{E}\,X$.

解：首先，将 n 人编号，然后依序取帽. 定义

$$I_k = \begin{cases} 1, & \text{第 } k \text{ 人取到自己的帽子}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\,I_k = \frac{1}{n}, \quad \forall\, k,$$

于是

$$\mathrm{E}\,X = 1. \quad \blacksquare$$

随机变量分解的方法

- 【例 5.1.c】 一辆客车载有 20 人从机场开往市区, 沿途有 10 个站可以下车. 如果到一战有人下车, 则停车; 否则不停车. 假设每个旅客在每站下车是等可能的, 记 X 为该车停靠的站的个数, 求 $\mathrm{E}\,X$.

解: 定义

$$I_k = \begin{cases} 1, & \text{第 } k \text{ 站有人下车}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{10} I_k$. 利用

$$\mathrm{E}\,I_k = 1 - \mathrm{P}\,(I_k = 0) = 1 - \left(\frac{9}{10}\right)^{20},$$

于是

$$\mathrm{E}\,X = 10\mathrm{E}\,I_1 = 10\left[1 - \left(\frac{9}{10}\right)^{20}\right] \approx 8.787. \quad \blacksquare$$

随机变量分解的方法

- 【例 5.1.c】 一辆客车载有 20 人从机场开往市区，沿途有 10 个站可以下车. 如果到一战有人下车，则停车；否则不停车. 假设每个旅客在每站下车是等可能的，记 X 为该车停靠的站的个数，求 $\mathrm{E}\,X$.

解：定义

$$I_k = \begin{cases} 1, & \text{第 } k \text{ 站有人下车}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{10} I_k$. 利用

$$\mathrm{E}\,I_k = 1 - \mathrm{P}\,(I_k = 0) = 1 - \left(\frac{9}{10}\right)^{20},$$

于是

$$\mathrm{E}\,X = 10\mathrm{E}\,I_1 = 10\left[1 - \left(\frac{9}{10}\right)^{20}\right] \approx 8.787. \quad \blacksquare$$

游程

设有 m 个 0 和 n 个 1 随机排成一行，把其中相连排列的 0 或者 1 都称为一个游程. 例如,

$$0\ 0\ 0\ \underbrace{1\ 1\ 1\ 1\ 1}_{5}\ 0\ 1\ \underbrace{0\ 0\ 0\ 0}_{4}\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1,$$

由 0 构成的游程有 7 个:

$$000,\ 0,\ 0000,\ 0, 00,\ 00,\ 00,$$

最长的游程为 4. 由 1 构成的游程有 7 个:

$$11111,\ 1,\ 11,\ 111,\ 1,\ 11,\ 1,$$

最大的 1-游程为 5.

随机变量分解的方法

- 【例 3.1.6】 设有 m 个 0 和 n 个 1 随机排成一行，记 Z 为该序列中游程的个数，求 $\mathrm{E}\,Z$.

解：设 X 和 Y 分别表示该序列 “$a_1a_2\cdots a_{m+n}$” 中 1-游程和 0-游程的个数，定义 $A_1=\{a_1=1\}$,

$$A_k=\{a_{k-1}=0,a_k=1\},\quad k=2,\ldots,m+n,$$

$$X_j=I_{A_j},\quad j=1,\ldots,m+n,$$

则 $X=\sum_{i=1}^{m+n}X_k$. 利用

$$\mathrm{E}\,X_1=\frac{n}{m+n},\quad \mathrm{E}\,X_k=\frac{mn}{(m+n)(m+n-1)},\quad k\geq 2,$$

于是

$$\mathrm{E}\,X=\sum_{k=1}^{m+n}\mathrm{E}\,X_k=\frac{n+mn}{m+n}.\quad \mathrm{E}\,Y=\frac{m+mn}{m+n}\ (\text{类似}),\quad \cdots$$

随机变量分解的方法

- 【例 3.1.6】 设有 m 个 0 和 n 个 1 随机排成一行，记 Z 为该序列中游程的个数，求 $\mathrm{E}\,Z$.

解：设 X 和 Y 分别表示该序列 "$a_1a_2\cdots a_{m+n}$" 中 1-游程和 0-游程的个数，定义 $A_1=\{a_1=1\}$,

$$A_k=\{a_{k-1}=0, a_k=1\},\quad k=2,\ldots,m+n,$$

$$X_j=I_{A_j},\quad j=1,\ldots,m+n,$$

则 $X=\sum_{i=1}^{m+n}X_k$. 利用

$$\mathrm{E}\,X_1=\frac{n}{m+n},\quad \mathrm{E}\,X_k=\frac{mn}{(m+n)(m+n-1)},\quad k\geq 2,$$

于是

$$\mathrm{E}\,X=\sum_{k=1}^{m+n}\mathrm{E}\,X_k=\frac{n+mn}{m+n}.\quad \mathrm{E}\,Y=\frac{m+mn}{m+n}\ (类似),\quad \cdots$$

随机变量分解的方法

- 【例 3.1.7】 设有 m 个 0 和 n 个 1 随机排成一行, 记 W 为该序列中 1 首次出现的时刻, 求 $\mathrm{E}\,W$.

解: 将 m 个 0 依次编号为 $0_1, 0_2, \ldots, 0_m$, 定义

$$B_k = \{0_k \text{ 出现在所有 } n \text{ 个 } 1 \text{ 之前}\}, \quad k = 1, \ldots, m,$$

则

$$W = 1 + \sum_{k=1}^{m} I_{B_k}.$$

利用

$$\mathrm{P}(B_k) = \frac{1}{n+1}, \quad k \geq 2,$$

于是

$$\mathrm{E}\,X = 1 + \sum_{k=1}^{m} \mathrm{E}\left[I_{B_k}\right] = 1 + \frac{m}{n+1}. \quad ∎$$

随机变量分解的方法

- 【例 3.1.7】 设有 m 个 0 和 n 个 1 随机排成一行, 记 W 为该序列中 1 首次出现的时刻, 求 $\mathrm{E}\,W$.

解: 将 m 个 0 依次编号为 $0_1, 0_2, \ldots, 0_m$, 定义

$$B_k = \{0_k \text{ 出现在所有 } n \text{ 个 } 1 \text{ 之前}\}, \quad k = 1, \ldots, m,$$

则

$$W = 1 + \sum_{k=1}^{m} I_{B_k}.$$

利用

$$\mathrm{P}\,(B_k) = \frac{1}{n+1}, \quad k \geq 2,$$

于是

$$\mathrm{E}\,X = 1 + \sum_{k=1}^{m} \mathrm{E}\,[I_{B_k}] = 1 + \frac{m}{n+1}. \quad \blacksquare$$

5.1.3 矩不等式

- 矩的定义
  - r 阶原点矩：$\mathrm{E}[X^r]$
  - r 阶绝对矩：$\mathrm{E}|X|^r$
  - 关于点 c 的 r 阶矩：$\mathrm{E}(X-c)^r$
  - 关于点 c 的 r 阶绝对矩：$\mathrm{E}|X-c|^r$
  - r 阶中心矩：$\mathrm{E}(X-\mathrm{E}X)^r$
  - r 阶中心绝对矩：$\mathrm{E}|X-\mathrm{E}X|^r$

※ 称 X 为 r 次可积，若 $\mathrm{E}|X|^r<\infty$.

※ L_r 空间定义为：

$$\mathcal{L}_r=\{X:\ \mathrm{E}|X|^r<\infty\},\quad r>0.$$

5.1.3 矩不等式

- 矩的定义
  - r 阶原点矩：$\mathrm{E}[X^r]$
  - r 阶绝对矩：$\mathrm{E}|X|^r$
  - 关于点 c 的 r 阶矩：$\mathrm{E}(X-c)^r$
  - 关于点 c 的 r 阶绝对矩：$\mathrm{E}|X-c|^r$
  - r 阶中心矩：$\mathrm{E}(X-\mathrm{E}\,X)^r$
  - r 阶中心绝对矩：$\mathrm{E}|X-\mathrm{E}\,X|^r$

※ 称 X 为 r 次可积，若 $\mathrm{E}|X|^r<\infty$.

※ L_r 空间定义为；

$$\mathcal{L}_r=\{X:\ \mathrm{E}|X|^r<\infty\},\quad r>0.$$

- **定义** 记 $\mu_r = \mathrm{E}\,(X - \mathrm{E}\,X)^r$, $r > 0$.
  - 偏度系数:
  $$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}$$
  - 峰度系数:
  $$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

※ 设 $X \sim N(\mu, \sigma^2)$, 则

$$\beta_1 = 0, \quad \beta_2 = 3,$$

因为 $\mu_2 = \sigma^2$, $\mu_3 = 0$, $\mu_4 = 3\sigma^4$.

- **定义** 记 $\mu_r = \mathrm{E}\,(X - \mathrm{E}\,X)^r$, $r > 0$.
  - 偏度系数：
  $$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}$$
  - 峰度系数：
  $$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

※ 设 $X \sim N(\mu, \sigma^2)$, 则

$$\beta_1 = 0, \quad \beta_2 = 3,$$

因为 $\mu_2 = \sigma^2$, $\mu_3 = 0$, $\mu_4 = 3\sigma^4$.

- 【例 5.1.8】 设 $X \sim N(0,1)$, n 为正整数，求 $\mathrm{E}\,|X|^n$.

解：对任意 $r > 0$,

$$\mathrm{E}\,|X|^r = \sqrt{\frac{2}{\pi}} \int_0^\infty (2u)^{(r-1)/2} e^{-u} du < \infty.$$

特别,

$$\mathrm{E}\,|X|^{2n-1} = \sqrt{\frac{2}{\pi}} \int_0^\infty (2u)^{n-1} e^{-u} du = \sqrt{\frac{2}{\pi}} 2^{n-1}\Gamma(n) = \sqrt{\frac{2}{\pi}} (2n-2)!!.$$

$$\mathrm{E}\,|X|^{2n} = \sqrt{\frac{2}{\pi}} \int_0^\infty (2u)^{n-1/2} e^{-u} du = \frac{1}{\sqrt{\pi}} 2^n \Gamma\left(n + \frac{1}{2}\right) = (2n-1)!!.$$

- 【例 5.1.8】 设 $X \sim N(0,1)$, n 为正整数，求 $\mathrm{E}|X|^n$.

解：对任意 $r > 0$,

$$\mathrm{E}|X|^r = \sqrt{\frac{2}{\pi}} \int_0^\infty (2u)^{(r-1)/2} e^{-u} du < \infty.$$

特别，

$$\mathrm{E}|X|^{2n-1} = \sqrt{\frac{2}{\pi}} \int_0^\infty (2u)^{n-1} e^{-u} du = \sqrt{\frac{2}{\pi}} 2^{n-1} \Gamma(n) = \sqrt{\frac{2}{\pi}} (2n-2)!!.$$

$$\mathrm{E}|X|^{2n} = \sqrt{\frac{2}{\pi}} \int_0^\infty (2u)^{n-1/2} e^{-u} du = \frac{1}{\sqrt{\pi}} 2^n \Gamma\left(n + \frac{1}{2}\right) = (2n-1)!!.$$

- **定理 5.1.5** (C_r 不等式) 设 $r>0$, $X_i \in \mathcal{L}_r$, $i=1,\ldots,n$, 则

$$\mathrm{E}\left|\sum_{i=1}^{n} X_i\right|^r \le C_r \sum_{i=1}^{n} \mathrm{E}\,|X_i|^r, \tag{3}$$

其中当 $r>1$ 时，$C_r = n^{r-1}$; 当 $r\in(0,1]$ 时，$C_r=1$.

证： 当 $r\ge 1$ 时, $f(x)=|x|^r$ 为凸函数, 所以

$$\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right|^r \le \frac{1}{n}\sum_{i=1}^{n}|X_i|^r,$$

当 $r\in(0,1)$ 时，利用

$$(1+x)^r \le 1+x \le 1+x^r, \quad x\in[0,1],$$

我们得到 $|x_1+x_2|^r \le |x_1|^r + |x_2|^r$, 进而有

$$\left|\sum_{i=1}^{n} X_i\right|^r \le \sum_{i=1}^{n}|X_i|^r. \quad \blacksquare$$

▶ **定理 5.1.5** (C_r 不等式) 设 $r>0$, $X_i \in \mathcal{L}_r$, $i=1,\ldots,n$, 则

$$\mathrm{E}\left|\sum_{i=1}^{n} X_i\right|^r \le C_r \sum_{i=1}^{n} \mathrm{E}\,|X_i|^r, \tag{3}$$

其中当 $r>1$ 时，$C_r = n^{r-1}$; 当 $r\in(0,1]$ 时，$C_r=1$.

证： 当 $r\ge 1$ 时, $f(x)=|x|^r$ 为凸函数, 所以

$$\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right|^r \le \frac{1}{n}\sum_{i=1}^{n} |X_i|^r,$$

当 $r\in(0,1)$ 时，利用

$$(1+x)^r \le 1+x \le 1+x^r, \quad x\in[0,1],$$

我们得到 $|x_1+x_2|^r \le |x_1|^r + |x_2|^r$, 进而有

$$\left|\sum_{i=1}^{n} X_i\right|^r \le \sum_{i=1}^{n} |X_i|^r. \qquad \blacksquare$$

- 定理 5.1.6 (Jensen 不等式) 设 X 为一个 rv，$\mathrm{E}|X| < \infty$, g 为一个 Borel 可测的凸 (convex) 函数，则

$$\mathrm{E}[g(X)] \geq g(\mathrm{E}X).$$

- 定理 5.1.7 设 $0 < s < r$, $X \in \mathcal{L}_r$, 则 $X \in \mathcal{L}_s$ 且

$$\left(\mathrm{E}|X|^s\right)^{1/s} \leq \left(\mathrm{E}|X|^r\right)^{1/r}.$$

证：利用 Jensen 不等式，取 $g(x) = |x|^{r/s}$. ∎

- 定理 5.1.6 (Jensen 不等式) 设 X 为一个 rv，$\mathrm{E}\,|X| < \infty$, g 为一个 Borel 可测的凸 (convex) 函数，则

$$\mathrm{E}\,[g(X)] \geq g(\mathrm{E}\,X).$$

- 定理 5.1.7 设 $0 < s < r$, $X \in \mathcal{L}_r$, 则 $X \in \mathcal{L}_s$ 且

$$\left(\mathrm{E}\,|X|^s\right)^{1/s} \leq \left(\mathrm{E}\,|X|^r\right)^{1/r}.$$

证：利用 Jensen 不等式，取 $g(x) = |x|^{r/s}$. ∎

- 【例 5.1.10】（熵） 设 X 的 pmf 为 $\mathrm{P}(X=k)=p_k$, $k=1,2,\ldots,n$, 其中 $\sum_{k=1}^n p_k=1$, 那么其熵定义为

$$H(X)=-\sum_{k=1}^{n} p_k \ln p_k.$$

再设 U 服从 $\{1,2,\ldots,n\}$ 上的均匀分布. 试证明 $H(X)\le H(U)$.

证: 取凸函数 $f(x)=x\ln x$, $x>0$, 利用 Jensen 不等式得

$$\frac{1}{n}\sum_{k=1}^{n} f(p_k)\ge f\left(\frac{1}{n}\sum_{k=1}^{n} p_k\right)=f\left(\frac{1}{n}\right)=\frac{1}{n}\ln\frac{1}{n}.$$

故

$$H(X)\le -\ln\frac{1}{n}=\ln n=H(U). \quad \blacksquare$$

- 【例 5.1.10】（熵） 设 X 的 pmf 为 $\mathrm{P}(X=k)=p_k$, $k=1,2,\ldots,n$, 其中 $\sum_{k=1}^n p_k=1$, 那么其熵定义为

$$H(X)=-\sum_{k=1}^{n} p_k \ln p_k.$$

再设 U 服从 $\{1,2,\ldots,n\}$ 上的均匀分布. 试证明 $H(X)\le H(U)$.

证: 取凸函数 $f(x)=x\ln x$, $x>0$, 利用 Jensen 不等式得

$$\frac{1}{n}\sum_{k=1}^{n} f(p_k)\ge f\left(\frac{1}{n}\sum_{k=1}^{n} p_k\right)=f\left(\frac{1}{n}\right)=\frac{1}{n}\ln\frac{1}{n}.$$

故

$$H(X)\le -\ln\frac{1}{n}=\ln n=H(U). \quad \blacksquare$$

- Hölder 不等式 设 $p>1$, $1/p+1/q=1$, $X\in\mathcal{L}_p$, $Y\in\mathcal{L}_q$, 则

$$\mathrm{E}\,|XY|\le\left(\mathrm{E}\,|X|^p\right)^{1/p}\left(\mathrm{E}\,|Y|^q\right)^{1/q}.$$

证：不妨设 $\mathrm{E}\,|X|^p>0$, $\mathrm{E}\,|Y|^q>0$, 仅需证明：

$$\mathrm{E}\left(\frac{|X|^p}{\mathrm{E}\,|X|^p}\right)^{1/p}\left(\frac{|Y|^q}{\mathrm{E}\,|Y|^q}\right)^{1/q}\le 1.$$

余下证明利用不等式：

$$u^\alpha v^\beta\le\alpha u+\beta v,\quad \forall\, u>0, v>0,$$

其中 $\alpha\in(0,1)$, $\beta=1-\alpha$. ∎

- **Hölder 不等式** 设 $p>1$, $1/p+1/q=1$, $X\in\mathcal{L}_p$, $Y\in\mathcal{L}_q$, 则

$$\mathrm{E}\,|XY|\le\left(\mathrm{E}\,|X|^p\right)^{1/p}\left(\mathrm{E}\,|Y|^q\right)^{1/q}.$$

证：不妨设 $\mathrm{E}\,|X|^p>0$, $\mathrm{E}\,|Y|^q>0$, 仅需证明：

$$\mathrm{E}\left(\frac{|X|^p}{\mathrm{E}\,|X|^p}\right)^{1/p}\left(\frac{|Y|^q}{\mathrm{E}\,|Y|^q}\right)^{1/q}\le 1.$$

余下证明利用不等式：

$$u^\alpha v^\beta\le\alpha u+\beta v,\quad \forall\, u>0, v>0,$$

其中 $\alpha\in(0,1)$, $\beta=1-\alpha$. ∎

- Minkowski 不等式 设 $p \geq 1$, 则

$$\left(\mathrm{E}\,|X+Y|^p\right)^{1/p} \leq \left(\mathrm{E}\,|X|^p\right)^{1/p} + \left(\mathrm{E}\,|Y|^p\right)^{1/p}.$$

证: 取 $q = p/(p-1)$, 则 $1/p + 1/q = 1$, $q(p-1) = p$. 对

$$\begin{aligned}
\mathrm{E}\,|X+Y|^p &\leq \mathrm{E}\left(|X+Y|^{p-1}|X|\right) + \mathrm{E}\left(|X+Y|^{p-1}|Y|\right) \\
&\leq \left(\mathrm{E}\,|X|^p\right)^{1/p}\left(\mathrm{E}\left[|X+Y|^{(p-1)q}\right]\right)^{1/q} \\
&\quad + \left(\mathrm{E}\,|Y|^p\right)^{1/p}\left(\mathrm{E}\left[|X+Y|^{(p-1)q}\right]\right)^{1/q} \\
&= \left[\left(\mathrm{E}\,|X|^p\right)^{1/p} + \left(\mathrm{E}\,|Y|^p\right)^{1/p}\right]\left(\mathrm{E}\left[|X+Y|^p\right]\right)^{1/q}
\end{aligned}$$

化简即得. ∎

- Minkowski 不等式 设 $p \geq 1$, 则

$$\left(\mathrm{E}\,|X+Y|^p\right)^{1/p} \leq \left(\mathrm{E}\,|X|^p\right)^{1/p} + \left(\mathrm{E}\,|Y|^p\right)^{1/p}.$$

证: 取 $q = p/(p-1)$, 则 $1/p + 1/q = 1$, $q(p-1) = p$. 对

$$\begin{aligned}
\mathrm{E}\,|X+Y|^p &\leq \mathrm{E}\left(|X+Y|^{p-1}|X|\right) + \mathrm{E}\left(|X+Y|^{p-1}|Y|\right) \\
&\leq \left(\mathrm{E}\,|X|^p\right)^{1/p}\left(\mathrm{E}\left[|X+Y|^{(p-1)q}\right]\right)^{1/q} \\
&\quad + \left(\mathrm{E}\,|Y|^p\right)^{1/p}\left(\mathrm{E}\left[|X+Y|^{(p-1)q}\right]\right)^{1/q} \\
&= \left[\left(\mathrm{E}\,|X|^p\right)^{1/p} + \left(\mathrm{E}\,|Y|^p\right)^{1/p}\right]\left(\mathrm{E}\left[|X+Y|^p\right]\right)^{1/q}
\end{aligned}$$

化简即得. ∎

5.1.4 方差

- **定义** 设 $X \in \mathcal{L}_2$, X 的方差 (Variance) 定义为

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}[(X - \mathrm{E}X)^2] \\ &= \mathrm{E}[X^2] - (\mathrm{E}X)^2.\end{aligned}$$

- **基本性质**

(1) $\mathrm{Var}(X) = 0 \Longleftrightarrow X = c$, a.s.

(2) $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$, $\forall\, c \in \Re$.

(3) $\mathrm{Var}(X) \le \mathrm{E}[X - c]^2$, $\forall\, c \in \Re$.

※ 随机变量的标准化：设 $X \in L_2$, $\mathrm{Var}(X) > 0$, 定义

$$X^* = \frac{X - \mathrm{E}X}{\sqrt{\mathrm{Var}(X)}}.$$

5.1.4 方差

- **定义** 设 $X \in \mathcal{L}_2$, X 的方差 (Variance) 定义为

$$\begin{aligned}\mathrm{Var}\,(X) &= \mathrm{E}\,[(X - \mathrm{E}\,X)^2] \\ &= \mathrm{E}\,[X^2] - (\mathrm{E}\,X)^2.\end{aligned}$$

- **基本性质**

(1) $\mathrm{Var}(X) = 0 \Longleftrightarrow X = c$, a.s.

(2) $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$, $\forall\ c \in \Re$.

(3) $\mathrm{Var}(X) \le \mathrm{E}\,[X - c]^2$, $\forall\, c \in \Re$.

※ 随机变量的标准化：设 $X \in L_2$, $\mathrm{Var}\,(X) > 0$, 定义

$$X^* = \frac{X - \mathrm{E}\,X}{\sqrt{\mathrm{Var}(X)}}.$$

5.1.4 方差

Example (4)

- 若 $X \sim B(n,p)$, $0<p<1$, 则 $\mathrm{Var}\,(X) = np(1-p)$.
- 若 $X \sim \mathrm{Poi}(\lambda)$, $\lambda > 0$, 则 $\mathrm{Var}\,(X) = \lambda$.
- 若 $X \sim \mathrm{Geo}(p)$, $0<p<1$, 则
$$\mathrm{Var}\,(X) = \frac{1-p}{p^2}.$$
- 若 $X \sim \mathrm{NB}(\alpha, p)$, $\alpha > 0$, $0<p<1$, 则
$$\mathrm{E}\,X = \frac{\alpha(1-p)}{p^2}.$$

5.1.4 方差

Example (5)

- 若 $X \sim U(a,b)$, $a < b$, 则 $\mathrm{Var}\,(X) = (b-a)^2/12$.
- 若 $X \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$, 则
$$\mathrm{Var}\,(X) = \frac{1}{\lambda^2}.$$
- 若 $X \sim \Gamma(\alpha,\beta)$, $0 < p < 1$, 则
$$\mathrm{Var}\,(X) = \frac{\alpha}{\beta^2}.$$
- 若 $X \sim N(\mu,\sigma^2)$, $\mu \in \Re$, $\sigma > 0$, 则 $\mathrm{Var}\,(X) = \sigma^2$.

5.1.4 方差

- 【例 5.1.b】 在一次聚会上，n 个人将自己的帽子混在一起，会后每人随机取帽，记 X 为能正确取到自己帽子的人数．求 $\mathrm{Var}(X)$.

解：首先，将 n 人编号，然后依序取帽．定义

$$I_k = \begin{cases} 1, & \text{第 } k \text{ 人取到自己的帽子,} \\ 0, & \text{否则,} \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\, I_k = \frac{1}{n}, \quad \forall\, k; \qquad \mathrm{E}\,[I_i I_j] = \frac{1}{n(n-1)}, \quad \forall\, i < j.$$

于是, $\mathrm{E}\, X = 1$,

$$\mathrm{E}\,[X^2] = \sum_{i=1}^{n} \mathrm{E}\, I_i + 2 \sum_{i<j} \mathrm{E}\,[I_i I_j] = 2, \qquad \mathrm{Var}\,(X) = 1. \blacksquare$$

5.1.4 方差

- 【例 5.1.b】 在一次聚会上，n 个人将自己的帽子混在一起，会后每人随机取帽，记 X 为能正确取到自己帽子的人数．求 $\mathrm{Var}\,(X)$.

解：首先，将 n 人编号，然后依序取帽．定义

$$I_k = \begin{cases} 1, & \text{第 } k \text{ 人取到自己的帽子}, \\ 0, & \text{否则}, \end{cases}$$

则 $X = \sum_{i=1}^{n} I_k$. 利用

$$\mathrm{E}\, I_k = \frac{1}{n}, \quad \forall\, k; \qquad \mathrm{E}\,[I_i I_j] = \frac{1}{n(n-1)}, \quad \forall\, i < j.$$

于是, $\mathrm{E}\, X = 1$,

$$\mathrm{E}\,[X^2] = \sum_{i=1}^{n} \mathrm{E}\, I_i + 2 \sum_{i<j} \mathrm{E}\,[I_i I_j] = 2, \qquad \mathrm{Var}\,(X) = 1. \quad \blacksquare$$

5.1.4 方差

- 【例 5.1.12】 设 x, y, z 为正数，满足 $x+y+z=1$，证明

$$\frac{1}{x}+\frac{4}{y}+\frac{9}{z} \geq 36.$$

解: 构造随机变量 X 使其 pmf 为

$$\mathrm{P}\left(X=\frac{1}{x}\right)=x, \quad \mathrm{P}\left(X=\frac{2}{y}\right)=y, \quad \mathrm{P}\left(X=\frac{3}{z}\right)=z,$$

则

$$\mathrm{E}\,X=6, \qquad \mathrm{E}\,X^2=\frac{1}{x}+\frac{4}{y}+\frac{9}{z}.$$

利用 $\mathrm{Var}(X)=\mathrm{E}\,X^2-(\mathrm{E}\,X)^2 \geq 0$ 得证. ∎

5.1.4 方差

- 【例 5.1.12】 设 x, y, z 为正数，满足 $x+y+z=1$，证明
$$\frac{1}{x}+\frac{4}{y}+\frac{9}{z} \geq 36.$$

解: 构造随机变量 X 使其 pmf 为

$$\mathrm{P}\left(X=\frac{1}{x}\right)=x, \quad \mathrm{P}\left(X=\frac{2}{y}\right)=y, \quad \mathrm{P}\left(X=\frac{3}{z}\right)=z,$$

则

$$\mathrm{E}\,X=6, \qquad \mathrm{E}\,X^2=\frac{1}{x}+\frac{4}{y}+\frac{9}{z}.$$

利用 $\mathrm{Var}\,(X)=\mathrm{E}\,X^2-(\mathrm{E}\,X)^2 \geq 0$ 得证. ∎

5.1.4 方差

- 【例 5.1.13】 设 $0<\alpha,\beta,\gamma<\pi/2$ 且 $\sin^2\alpha+\sin^2\beta+\sin^2\gamma=1$, 证明
$$\frac{\sin^3\alpha}{\sin\beta}+\frac{\sin^3\beta}{\sin\gamma}+\frac{\sin^3\gamma}{\sin\alpha}>1.$$

解: 利用 $0<\sin\alpha\sin\beta+\sin\beta\sin\gamma+\sin\gamma\sin\alpha\le 1$, 构造 rv X 使其 pmf 为

$$\mathrm{P}\left(X=\frac{\sin\alpha}{\sin\beta}\right)=\sin\alpha\sin\beta,\quad \mathrm{P}\left(X=\frac{\sin\beta}{\sin\gamma}\right)=\sin\beta\sin\gamma,$$
$$\mathrm{P}\left(X=\frac{\sin\gamma}{\sin\alpha}\right)=\sin\gamma\sin\alpha,\quad \mathrm{P}(X=0)=1-\cdots$$

则

$$\mathrm{E}\,X=1,\qquad \mathrm{E}\,X^2=\frac{\sin^3\alpha}{\sin\beta}+\frac{\sin^3\beta}{\sin\gamma}+\frac{\sin^3\gamma}{\sin\alpha}.$$

得证. ∎

5.1.4 方差

- 【例 5.1.13】 设 $0<\alpha,\beta,\gamma<\pi/2$ 且 $\sin^2\alpha+\sin^2\beta+\sin^2\gamma=1$, 证明

$$\frac{\sin^3\alpha}{\sin\beta}+\frac{\sin^3\beta}{\sin\gamma}+\frac{\sin^3\gamma}{\sin\alpha}>1.$$

解: 利用 $0<\sin\alpha\sin\beta+\sin\beta\sin\gamma+\sin\gamma\sin\alpha\le 1$, 构造 rv X 使其 pmf 为

$$\mathrm{P}\left(X=\frac{\sin\alpha}{\sin\beta}\right)=\sin\alpha\sin\beta,\quad \mathrm{P}\left(X=\frac{\sin\beta}{\sin\gamma}\right)=\sin\beta\sin\gamma,$$

$$\mathrm{P}\left(X=\frac{\sin\gamma}{\sin\alpha}\right)=\sin\gamma\sin\alpha,\quad \mathrm{P}(X=0)=1-\cdots$$

则

$$\mathrm{E}\,X=1,\qquad \mathrm{E}\,X^2=\frac{\sin^3\alpha}{\sin\beta}+\frac{\sin^3\beta}{\sin\gamma}+\frac{\sin^3\gamma}{\sin\alpha}.$$

得证. ∎

5.1.5 中位数和 p 分位数

- 中位数: ν 称为 X 的中位数，若

$$\mathrm{P}(X \geq \nu) \geq \frac{1}{2}, \quad \mathrm{P}(X \leq \nu) \geq \frac{1}{2}.$$

- 【例 5.1.16】 设 X 服从 $\{-1, 0, 1\}$ 上的离散均匀分布, 则 X 的中位数为 0.
- 【例 5.1.17】 设 $\mathrm{P}(X=0)=\mathrm{P}(X=1)=1/2$, 则 $[0,1]$ 中任意一点都是 X 的中位数.

※ 中位数不唯一

※ 中位数与数学期望的比较

5.1.5 中位数和 p 分位数

- 中位数: ν 称为 X 的中位数，若

$$\mathrm{P}(X \geq \nu) \geq \frac{1}{2}, \quad \mathrm{P}(X \leq \nu) \geq \frac{1}{2}.$$

- 【例 5.1.16】 设 X 服从 $\{-1, 0, 1\}$ 上的离散均匀分布, 则 X 的中位数为 0.
- 【例 5.1.17】 设 $\mathrm{P}(X=0)=\mathrm{P}(X=1)=1/2$, 则 $[0,1]$ 中任意一点都是 X 的中位数.

※ 中位数不唯一

※ 中位数与数学期望的比较

5.1.5 中位数和 p 分位数

- 中位数：ν 称为 X 的中位数，若

$$\mathrm{P}(X \geq \nu) \geq \frac{1}{2}, \quad \mathrm{P}(X \leq \nu) \geq \frac{1}{2}.$$

- 【例 5.1.16】 设 X 服从 $\{-1, 0, 1\}$ 上的离散均匀分布，则 X 的中位数为 0.
- 【例 5.1.17】 设 $\mathrm{P}(X = 0) = \mathrm{P}(X = 1) = 1/2$，则 $[0, 1]$ 中任意一点都是 X 的中位数.

※ 中位数不唯一

※ 中位数与数学期望的比较

5.1.5 中位数和 p 分位数

- 中位数: ν 称为 X 的中位数，若

$$\mathrm{P}(X \geq \nu) \geq \frac{1}{2}, \quad \mathrm{P}(X \leq \nu) \geq \frac{1}{2}.$$

- 【例 5.1.16】 设 X 服从 $\{-1,0,1\}$ 上的离散均匀分布, 则 X 的中位数为 0.
- 【例 5.1.17】 设 $\mathrm{P}(X=0)=\mathrm{P}(X=1)=1/2$, 则 $[0,1]$ 中任意一点都是 X 的中位数.

※ 中位数不唯一

※ 中位数与数学期望的比较

5.1.5 中位数和 p 分位数

- p 分位数：μ_p 称为 X 的（下）p 分位数，若

$$\mathrm{P}(X \le \mu_p) \ge p, \quad \mathrm{P}(X \ge \mu_p) \ge 1-p.$$

- 上 p 分位数：μ'_p 称为 X 的上 p 分位数，若

$$\mathrm{P}(X \ge \mu'_p) \ge p, \quad \mathrm{P}(X \le \mu'_p) \ge 1-p.$$

※ 随机变量的下 p 分位数即为随机变量的上 $(1-p)$ 分位数.

5.2.1 条件期望及其应用

- 条件期望的定义:

$$\mathrm{E}[Y|X=x]=\int_{-\infty}^{\infty} y\,\mathrm{d}F_{2|1}(y|x)\ (\text{假设存在})$$

- 条件期望的平滑公式 (“两步走”的思想): 设 $Y\in\mathcal{L}_1$, 则

$$\mathrm{E}\left\{\mathrm{E}[Y|X]\right\}=\mathrm{E}\,Y,$$

即

$$\mathrm{E}\,Y=\mathrm{E}[g(X)],$$

其中

$$g(x)=\mathrm{E}[Y|X=x].$$

5.2.1 条件期望及其应用

- 条件期望的定义:

$$\mathrm{E}[Y|X=x]=\int_{-\infty}^{\infty} y\,\mathrm{d}F_{2|1}(y|x)\ (\text{假设存在})$$

- 条件期望的平滑公式 (“两步走” 的思想): 设 $Y\in\mathcal{L}_1$, 则

$$\mathrm{E}\left\{\mathrm{E}[Y|X]\right\}=\mathrm{E}\,Y,$$

即

$$\mathrm{E}\,Y=\mathrm{E}[g(X)],$$

其中

$$g(x)=\mathrm{E}[Y|X=x].$$

- 求数学期望的方法二

条件期望的方法: $Y \in \mathcal{L}_1$, 则

$$\mathrm{E}\{\mathrm{E}[Y|X]\} = \mathrm{E}\,Y,$$

※ 优点:

- 体现“两步走”的思想, 给定 x, $\mathrm{E}[Y|X=x]$ 易于求;
- 导出递推公式或差分方程.

※ 技巧: 如何选择“恰当”的 X ?

- 向后的方法: 取 X 为试验刚开始某时刻的状态变量;
- 向前的方法: 取 X 为试验快结束前某时刻的状态变量.

▶ 求数学期望的方法二

条件期望的方法: $Y \in \mathcal{L}_1$, 则

$$\mathrm{E}\{\mathrm{E}[Y|X]\} = \mathrm{E}\, Y,$$

※ 优点:

- 体现“两步走”的思想, 给定 x, $\mathrm{E}[Y|X=x]$ 易于求;
- 导出递推公式或差分方程.

※ 技巧: 如何选择“恰当”的 X ?

- 向后的方法: 取 X 为试验刚开始某时刻的状态变量;
- 向前的方法: 取 X 为试验快结束前某时刻的状态变量.

- 【例 5.2.2】 设 X, Y iid $\sim B(n,p)$, 求 $\mathrm{E}[X|X+Y=m]$.

解法一: 首先, $[X|X+Y=m] \in \{0,1,\ldots,m\wedge n\}$. 对任意 $0 \le k \le m \wedge n$,

$$\mathrm{P}(X=k|X+Y=m) = \frac{\mathrm{P}(X=k)\cdot\mathrm{P}(Y=m-k)}{\mathrm{P}(X+Y=m)} = \frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}},$$

即 $[X|X+Y=m]$ 服从超几何分布. 利用随机变量的分解方法得 $\mathrm{E}[X|X+Y=m]=m/2$.

解法二: 利用 X 与 Y 的对称性得

$$\mathrm{E}[X|X+Y=m] = \mathrm{E}[Y|X+Y=m],$$

于是

$$\mathrm{E}[X|X+Y=m] = \frac{1}{2}\Big(\mathrm{E}[X|X+Y=m]+\mathrm{E}[Y|X+Y=m]\Big) = \frac{m}{2}.$$

- 【例 5.2.2】 设 X, Y iid $\sim B(n,p)$, 求 $\mathrm{E}[X|X+Y=m]$.

解法一: 首先, $[X|X+Y=m] \in \{0,1,\ldots,m\wedge n\}$. 对任意 $0\le k\le m\wedge n$,

$$\mathrm{P}(X=k|X+Y=m)=\frac{\mathrm{P}(X=k)\cdot\mathrm{P}(Y=m-k)}{\mathrm{P}(X+Y=m)}=\frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}},$$

即 $[X|X+Y=m]$ 服从超几何分布. 利用随机变量的分解方法得 $\mathrm{E}[X|X+Y=m]=m/2$.

解法二: 利用 X 与 Y 的对称性得

$$\mathrm{E}[X|X+Y=m]=\mathrm{E}[Y|X+Y=m],$$

于是

$$\mathrm{E}[X|X+Y=m]=\frac{1}{2}\Big(\mathrm{E}[X|X+Y=m]+\mathrm{E}[Y|X+Y=m]\Big)=\frac{m}{2}.$$

- 【例 5.2.2】 设 X, Y iid $\sim B(n,p)$, 求 $\mathrm{E}\,[X|X+Y=m]$.

解法一: 首先, $[X|X+Y=m] \in \{0,1,\ldots,m\wedge n\}$. 对任意 $0 \le k \le m\wedge n$,

$$\mathrm{P}\,(X=k|X+Y=m) = \frac{\mathrm{P}\,(X=k)\cdot\mathrm{P}\,(Y=m-k)}{\mathrm{P}\,(X+Y=m)} = \frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}},$$

即 $[X|X+Y=m]$ 服从超几何分布. 利用随机变量的分解方法得 $\mathrm{E}\,[X|X+Y=m] = m/2$.

解法二: 利用 X 与 Y 的对称性得

$$\mathrm{E}\,[X|X+Y=m] = \mathrm{E}\,[Y|X+Y=m],$$

于是

$$\mathrm{E}\,[X|X+Y=m] = \frac{1}{2}\Big(\mathrm{E}\,[X|X+Y=m] + \mathrm{E}\,[Y|X+Y=m]\Big) = \frac{m}{2}.$$

§5.2 条件期望与条件方差

- 【例 5.2.4】 设罐中有 a 个白球，b 个黑球，从中将球依次取出，直到取出一个白球为止，记 X 为所取出的黑球个数．求 $\mathrm{E}X$.

解法一：记 $M_{a,b} = \mathrm{E}X$, 且定义

$$Y = \begin{cases} 1, & \text{首次取出白球;} \\ 0, & \text{首次取出黑球.} \end{cases}$$

对 Y 取条件，注意到 $[X|Y=1]=0$, 得

$$\begin{aligned} M_{a,b} &= \mathrm{E}[X|Y=1]\mathrm{P}(Y=1) + \mathrm{E}[X|Y=0]\mathrm{P}(Y=0) \\ &= (1+M_{a,b-1})\frac{b}{a+b}. \end{aligned}$$

利用 $M_{a,0}=0$, 得

$$M_{a,1} = \frac{1}{a+1}(1+M_{a,0}) = \frac{1}{a+1},$$

$$M_{a,2} = \frac{2}{a+2}(1+M_{a,1}) = \frac{2}{a+1},$$

归纳递推得 $M_{a,b} = b/(a+1)$.

▶ 【例 5.2.4】 设罐中有 a 个白球，b 个黑球，从中将球依次取出，直到取出一个白球为止，记 X 为所取出的黑球个数. 求 $\mathrm{E}\,X$.

解法一: 记 $M_{a,b}=\mathrm{E}\,X$, 且定义

$$Y=\begin{cases}1, & \text{首次取出白球;}\\ 0, & \text{首次取出黑球.}\end{cases}$$

对 Y 取条件，注意到 $[X|Y=1]=0$, 得

$$\begin{aligned} M_{a,b} &= \mathrm{E}\,[X|Y=1]\,\mathrm{P}\,(Y=1)+\mathrm{E}\,[X|Y=0]\,\mathrm{P}\,(Y=0)\\ &= (1+M_{a,b-1})\,\frac{b}{a+b}. \end{aligned}$$

利用 $M_{a,0}=0$, 得

$$M_{a,1}=\frac{1}{a+1}(1+M_{a,0})=\frac{1}{a+1},$$
$$M_{a,2}=\frac{2}{a+2}(1+M_{a,1})=\frac{2}{a+1},$$

归纳递推得 $M_{a,b}=b/(a+1)$.

- 【例 5.2.4】 设罐中有 a 个白球，b 个黑球，从中将球依次取出，直到取出一个白球为止，记 X 为所取出的黑球个数. 求 $\mathrm{E}\,X$.

解法二：首先，给 b 个黑球进行编号 $1, 2, \ldots, b$, 定义

$$I_k = \begin{cases} 1, & \text{编号 } k \text{ 的黑球在所有白球之前被取出,} \\ 0, & \text{其它,} \end{cases}$$

则 $X = \sum_{k=1}^{b} I_k$. 现求 $\mathrm{E}\,I_k$. 为此，只需要考虑编号 k 的黑球和所有 a 个白球，这 $a+1$ 个球每一个都等可能最先被取出，所以

$$\mathrm{E}\,I_k = \frac{1}{a+1}, \quad \forall k.$$

因此，

$$\mathrm{E}\,X = \frac{b}{a+1}. \quad \blacksquare$$

- 【例 5.2.4】 设罐中有 a 个白球，b 个黑球，从中将球依次取出，直到取出一个白球为止，记 X 为所取出的黑球个数. 求 $\mathrm{E}\,X$.

解法二：首先，给 b 个黑球进行编号 $1, 2, \dots, b$，定义

$$I_k = \begin{cases} 1, & \text{编号 } k \text{ 的黑球在所有白球之前被取出}, \\ 0, & \text{其它}, \end{cases}$$

则 $X = \sum_{k=1}^{b} I_k$. 现求 $\mathrm{E}\,I_k$. 为此，只需要考虑编号 k 的黑球和所有 a 个白球，这 $a+1$ 个球每一个都等可能最先被取出，所以

$$\mathrm{E}\,I_k = \frac{1}{a+1}, \quad \forall k.$$

因此，

$$\mathrm{E}\,X = \frac{b}{a+1}. \quad \blacksquare$$

- 【例 5.2.5】 设 $X \sim \mathrm{Geo}(p)$, $0 < p < 1$, 求 $\mathrm{Var}\,(X)$.

解: 考虑几何分布的概率模型. 令

$$Y = \begin{cases} 1, & \text{首次试验成功发生}, \\ 0, & \text{否则}, \end{cases}$$

注意到 $[X|Y=1] = 0$, 对 Y 取条件得

$$\begin{aligned} \mathrm{E}\,X^2 &= \mathrm{E}\,[X^2|Y=0]\cdot \mathrm{P}\,(Y=0) + \mathrm{E}\,[X^2|Y=1]\cdot \mathrm{P}\,(Y=1) \\ &= \mathrm{E}\,[X^2|Y=0]q \qquad (q = 1-p) \\ &= q\cdot \mathrm{E}\,(1+X)^2 \\ &= q\cdot\left[1 + 2\frac{q}{p} + \mathrm{E}\,X^2\right] \end{aligned}$$

$\Longrightarrow$

$$\mathrm{E}\,X^2 = \frac{q+q^2}{p^2}, \quad \mathrm{Var}\,(X) = \frac{q}{p^2}. \blacksquare$$

- 【例 5.2.5】 设 $X \sim \mathrm{Geo}(p)$, $0 < p < 1$, 求 $\mathrm{Var}\,(X)$.

解: 考虑几何分布的概率模型. 令

$$Y = \begin{cases} 1, & \text{首次试验成功发生}, \\ 0, & \text{否则}, \end{cases}$$

注意到 $[X|Y=1]=0$, 对 Y 取条件得

$$\begin{aligned}
\mathrm{E}\,X^2 &= \mathrm{E}\,[X^2|Y=0]\cdot \mathrm{P}\,(Y=0) + \mathrm{E}\,[X^2|Y=1]\cdot \mathrm{P}\,(Y=1) \\
&= \mathrm{E}\,[X^2|Y=0]q \qquad (q = 1-p) \\
&= q\cdot \mathrm{E}\,(1+X)^2 \\
&= q\cdot \left[1 + 2\frac{q}{p} + \mathrm{E}\,X^2\right]
\end{aligned}$$

$\Longrightarrow$

$$\mathrm{E}\,X^2 = \frac{q+q^2}{p^2}, \quad \mathrm{Var}\,(X) = \frac{q}{p^2}. \ \blacksquare$$

- 【例 5.2.6】 连续抛掷一枚均匀的骰子，直到接连抛出两个 1 点为止. 以 X 表示所需的抛掷次数，求 $\mathrm{E}\,X$.

解法一: 以 Y 表示第一次抛掷时所掷出的点数，记 $E_k = \mathrm{E}\,[X|Y=k]$, $k=1,\ldots,6$, 则

$$E_2 = \cdots = E_6 = 1 + \mathrm{E}\,X.$$

再对第一次和第二次抛掷的结果取条件，分别得

$$\mathrm{E}\,X = \frac{1}{6}\sum_{k=1}^{6}\mathrm{E}\,[X|Y=k] = \frac{1}{6}E_1 + \frac{5}{6}E_2 = \frac{1}{6}E_1 + \frac{5}{6}(1+\mathrm{E}\,X),$$

$$E_1 = \mathrm{E}\,[X|Y=1] = \frac{2}{6} + \frac{1}{6}\sum_{k=2}^{6}(1+E_k) = 2 + \frac{5}{6}\mathrm{E}\,X.$$

求解得 $\mathrm{E}\,X = 42$. ∎

- 【例 5.2.6】 连续抛掷一枚均匀的骰子，直到接连抛出两个 1 点为止. 以 X 表示所需的抛掷次数，求 $\mathrm{E}\,X$.

解法一： 以 Y 表示第一次抛掷时所掷出的点数，记 $E_k = \mathrm{E}\,[X|Y=k]$, $k=1,\ldots,6$, 则

$$E_2 = \cdots = E_6 = 1 + \mathrm{E}\,X.$$

再对第一次和第二次抛掷的结果取条件，分别得

$$\mathrm{E}\,X = \frac{1}{6}\sum_{k=1}^{6}\mathrm{E}\,[X|Y=k] = \frac{1}{6}E_1 + \frac{5}{6}E_2 = \frac{1}{6}E_1 + \frac{5}{6}(1+\mathrm{E}\,X),$$

$$E_1 = \mathrm{E}\,[X|Y=1] = \frac{2}{6} + \frac{1}{6}\sum_{k=2}^{6}(1+E_k) = 2 + \frac{5}{6}\mathrm{E}\,X.$$

求解得 $\mathrm{E}\,X = 42$. ∎

- 【例 5.2.6】 连续抛掷一枚均匀的骰子, 直到接连抛出两个 1 点为止. 以 X 表示所需的抛掷次数, 求 $\mathrm{E}\,X$.

解法二: 以 Y 表示第一次抛出点 1 所需要的抛掷次数, 则 $Y \sim \mathrm{Geo}^*(1/6)$, $\mathrm{E}\,Y = 6$. 注意到

$$\mathrm{E}\,[X|Y=k] = \frac{1}{6}(k+1) + \frac{5}{6}[k+1+\mathrm{E}\,X] = k+1+\frac{5}{6}\mathrm{E}\,X,$$

于是

$$EX = \mathrm{E}\,Y + 1 + \frac{5}{6}\mathrm{E}\,X$$

求解得 $\mathrm{E}\,X = 6(\mathrm{E}\,Y+1) = 42$. ▮

- 【例 5.2.6】 连续抛掷一枚均匀的骰子，直到接连抛出两个 1 点为止. 以 X 表示所需的抛掷次数，求 $\mathrm{E}\,X$.

解法二： 以 Y 表示第一次抛出点 1 所需要的抛掷次数，则 $Y \sim \mathrm{Geo}^*(1/6)$, $\mathrm{E}\,Y = 6$. 注意到

$$\mathrm{E}\,[X|Y=k] = \frac{1}{6}(k+1) + \frac{5}{6}[k+1+\mathrm{E}\,X] = k+1+\frac{5}{6}\mathrm{E}\,X,$$

于是

$$EX = \mathrm{E}\,Y + 1 + \frac{5}{6}\mathrm{E}\,X$$

求解得 $\mathrm{E}\,X = 6(\mathrm{E}\,Y+1) = 42$. ∎

- 【例 5.2.7】 一名矿工陷进一个有三扇门的矿井中. 第一扇门通到一个隧道, 走 3 小时后他可到达安全区. 第二扇门通到又一个隧道, 走 5 小时后会使他回到这矿井中. 第三扇门通到再一个隧道, 走 7 小时后也使他回到该矿井中. 假设该矿工在三扇门中是随机地选取一扇. 求 $\mathrm{E}X$.

解: 设 Y 表示该矿工所选择的门号, 则 $\mathrm{P}(Y=i)=1/3$, $i=1,2,3$. 利用两步走的思想得

$$\begin{aligned} \mathrm{E}X &= \frac{1}{3}\mathrm{E}[X|Y=1]+\frac{1}{3}\mathrm{E}[X|Y=2]+\frac{1}{3}\mathrm{E}[X|Y=3] \\ &= \frac{1}{3}\cdot 3+\frac{1}{3}(5+\mathrm{E}X)+\frac{1}{3}(7+\mathrm{E}X), \end{aligned}$$

最后求出 $\mathrm{E}X=15$. ∎

- 【例 5.2.7】 一名矿工陷进一个有三扇门的矿井中. 第一扇门通到一个隧道, 走 3 小时后他可到达安全区. 第二扇门通到又一个隧道, 走 5 小时后会使他回到这矿井中. 第三扇门通到再一个隧道, 走 7 小时后也使他回到该矿井中. 假设该矿工在三扇门中是随机地选取一扇. 求 $\mathrm{E}\,X$.

解: 设 Y 表示该矿工所选择的门号, 则 $\mathrm{P}\,(Y=i)=1/3$, $i=1,2,3$. 利用两步走的思想得

$$\begin{aligned}\mathrm{E}\,X &= \frac{1}{3}\,\mathrm{E}\,[X|Y=1]+\frac{1}{3}\,\mathrm{E}\,[X|Y=2]+\frac{1}{3}\,\mathrm{E}\,[X|Y=3]\\ &= \frac{1}{3}\cdot 3+\frac{1}{3}(5+\mathrm{E}\,X)+\frac{1}{3}(7+\mathrm{E}\,X),\end{aligned}$$

最后求出 $\mathrm{E}\,X=15$. ∎

- 【例 5.2.8】 有 n 种卡片, 每种卡片数量无限, 每次随机从 n 种卡片中选择一张, 选择到每种卡片的概率相同. 现在需要集齐 n 种卡片, 即每种卡片至少需要一张, 集齐之后则终止. 以 X 记能够集齐 n 张卡片所需的选择次数. 求 $\mathrm{E}\,X$.

解法一: 集到第 1 张卡片只需一次选择, 因为不管哪一张都行. 但是要集到第 2 张卡片, 就需要开始等待了, 必须是一种新的没有出现过的卡片, 它在每次选择中的出现概率 $p_1 = \frac{n-1}{n}$. 如果以 X_1 表示所需的等待次数, 则

$$X_1 \sim \mathrm{Geo}^*(p_2).$$

一般地, 如果已经集到 k 种不同的卡片, 那么再选到一种新的卡片, 其出现概率 $p_k = \frac{n-k}{n}$. 如果以 X_k 表示所需的等待次数, 则

$$X_k \sim \mathrm{Geo}^*(p_k), \quad k = 2, \ldots, n-1.$$

- 【例 5.2.8】 有 n 种卡片，每种卡片数量无限，每次随机从 n 种卡片中选择一张，选择到每种卡片的概率相同. 现在需要集齐 n 种卡片，即每种卡片至少需要一张，集齐之后则终止. 以 X 记能够集齐 n 张卡片所需的选择次数. 求 $\mathrm{E}\,X$.

解法一：集到第 1 张卡片只需一次选择，因为不管哪一张都行. 但是要集到第 2 张卡片，就需要开始等待了，必须是一种新的没有出现过的卡片，它在每次选择中的出现概率 $p_1 = \frac{n-1}{n}$. 如果以 X_1 表示所需的等待次数，则

$$X_1 \sim \mathrm{Geo}^*(p_2).$$

一般地，如果已经集到 k 种不同的卡片，那么再选到一种新的卡片，其出现概率 $p_k = \frac{n-k}{n}$. 如果以 X_k 表示所需的等待次数，则

$$X_k \sim \mathrm{Geo}^*(p_k), \quad k = 2, \ldots, n-1.$$

因此，集齐所有卡片所需的总次数就是

$$X = 1 + X_1 + \cdots + X_{n-1}.$$

于是由数学期望的线性性质知，

$$\mathrm{E}\, X = 1 + \mathrm{E}\, X_1 + \cdots + \mathrm{E}\, X_{n-1}. \qquad (1)$$

注意到

$$\mathrm{E}\, X_k = \frac{1}{p_k} = \frac{n}{n-k}.$$

代入 (1) 式即得

$$\mathrm{E}\, X = n\left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + 1\right).$$

解法二: 记 $M_k = \mathrm{E}\, X_k$, 我们来设法建立关于 M_k 的关系式.

只有当新抽出的卡片是没有出现过的, 那么收集到的卡片数目才会增加 1. 我们把这种卡片叫做有效的, 否则叫做无效的.

根据题意, 在已有 k 张有效卡片条件下, 再抽到一张有效卡片的概率为 $\frac{n-k}{n}$. 如果抽到一张有效卡片, 则 X_k 的值等于 1; 如果抽到的是无效卡片, 那么还要重新再抽, 所需抽取的平均次数与现在相同, 从而一共平均需要抽取 $1+M_k$ 张. 故而有

$$M_k = \frac{n-k}{n} + \frac{k}{n}(1+M_k).$$

由该式解得

$$M_k = \frac{n}{n-k}.$$

代入 ① 式, 即得

$$\mathrm{E}\, X = 1 + \sum_{k=1}^{n-1} \frac{n}{n-k} = n\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1} + \frac{1}{n}\right). \quad \blacksquare$$

- 【例 5.2.10】 设 $X_1, X_2, \ldots$ iid $\sim U(0,1)$, 记
$$Y = \min\left\{n : \sum_{i=1}^{n} X_i > 1\right\},$$
求 $\mathrm{E}\,Y$.

解：对任意 $x>0$, 记 $Y_x = \min\{n : \sum_{i=1}^{n} X_i > x\}$, 且定义 $m(x) = \mathrm{E}\,[Y_x]$. 对 X_1 取条件，注意到
$$\mathrm{E}\,[Y_x|X_1 = y] = \begin{cases} 1, & y > x; \\ 1 + m(x-y), & y \le x \end{cases}$$
得
$$\begin{aligned} m(x) &= \int_0^1 \mathrm{E}\,[Y_x|X_1 = y]dy \\ &= 1 + \int_0^x m(x-y)dy = 1 + \int_0^x m(u)du, \quad x \in (0,1). \end{aligned}$$
两边求导得 $m'(x) = m(x)$, 于是 $m(x) = ce^x, \forall x \in [0,1]$. 由 $m(0) = 1$, 得 $c = 1$. 故 $m(x) = e^x$, $\mathrm{E}\,Y = m(1) = e$. ∎

- 【例 5.2.10】 设 $X_1, X_2, \ldots$ iid $\sim U(0,1)$, 记

$$Y = \min\left\{n : \sum_{i=1}^{n} X_i > 1\right\},$$

求 $\mathrm{E}\, Y$.

解：对任意 $x > 0$, 记 $Y_x = \min\{n : \sum_{i=1}^{n} X_i > x\}$, 且定义 $m(x) = \mathrm{E}\,[Y_x]$. 对 X_1 取条件，注意到

$$\mathrm{E}\,[Y_x | X_1 = y] = \begin{cases} 1, & y > x; \\ 1 + m(x-y), & y \le x \end{cases}$$

得

$$\begin{aligned} m(x) &= \int_0^1 \mathrm{E}\,[Y_x | X_1 = y] dy \\ &= 1 + \int_0^x m(x-y) dy = 1 + \int_0^x m(u) du, \quad x \in (0,1). \end{aligned}$$

两边求导得 $m'(x) = m(x)$, 于是 $m(x) = ce^x$, $\forall\, x \in [0,1]$. 由 $m(0) = 1$, 得 $c = 1$. 故 $m(x) = e^x$, $\mathrm{E}\, Y = m(1) = e$. ∎

§5.2 条件期望与条件方差

- 【例 5.2.a】 一盒中有 n 张卡片，分别标有号码 $1, 2, \ldots, n$. 从盒中有放回地随机摸取卡片，并记录下卡片上的号码. 如果所取号码之和不小于 k $(1 \le k \le n)$, 则摸取结束. 记 X_k 为最终摸取卡片数的次数，求 $\mathrm{E}\,[X_k]$.

解：设 Y 为首次摸取的卡片号码，对 Y 取条件得

$$\begin{aligned}
\mathrm{E}\,X_k &= \sum_{j=k}^{n} \frac{1}{n} \mathrm{E}\,[X_k | Y = j] + \sum_{j=1}^{k-1} \frac{1}{n} \mathrm{E}\,[X_k | Y = j] \\
&= \frac{n-k+1}{n} + \sum_{j=1}^{k-1} \frac{1}{n} (1 + \mathrm{E}\,X_{k-j}) \\
&= 1 + \frac{1}{n} \sum_{j=1}^{k-1} \mathrm{E}\,X_{k-j} \\
&= 1 + \frac{1}{n} \sum_{j=1}^{k-1} \mathrm{E}\,X_j.
\end{aligned}$$

- 【例 5.2.a】 一盒中有 n 张卡片，分别标有号码 $1, 2, \ldots, n$. 从盒中有放回地随机摸取卡片，并记录下卡片上的号码. 如果所取号码之和不小于 k $(1 \leq k \leq n)$, 则摸取结束. 记 X_k 为最终摸取卡片数的次数，求 $\mathrm{E}[X_k]$.

解：设 Y 为首次摸取的卡片号码，对 Y 取条件得

$$
\begin{aligned}
\mathrm{E}\, X_k &= \sum_{j=k}^{n} \frac{1}{n} \mathrm{E}[X_k | Y = j] + \sum_{j=1}^{k-1} \frac{1}{n} \mathrm{E}[X_k | Y = j] \\
&= \frac{n-k+1}{n} + \sum_{j=1}^{k-1} \frac{1}{n} (1 + \mathrm{E}\, X_{k-j}) \\
&= 1 + \frac{1}{n} \sum_{j=1}^{k-1} \mathrm{E}\, X_{k-j} \\
&= 1 + \frac{1}{n} \sum_{j=1}^{k-1} \mathrm{E}\, X_j.
\end{aligned}
$$

(续)

于是,

$$\begin{aligned}
\mathrm{E}\,X_k &= 1+\frac{1}{n}\left(\sum_{j=1}^{k-2}\mathrm{E}\,X_j+1+\frac{1}{n}\sum_{j=1}^{k-2}\mathrm{E}\,X_j\right)\\
&= \frac{n+1}{n}\left[1+\frac{1}{n}\sum_{j=1}^{k-2}\mathrm{E}\,X_j\right]\\
&= \frac{n+1}{n}\mathrm{E}\,X_{k-1}\\
&= \left(\frac{n+1}{n}\right)^{k-1},
\end{aligned}$$

其中递推用到 $X_1=1$. ∎

5.2.2 通过条件概率求概率

$$\mathrm{P}(A) = \mathrm{E}\left\{\mathrm{E}[1_A|Y]\right\} = \mathrm{E}[\mathrm{P}(A|Y)].$$

- 设 Y 为离散型 rv, 取值于 $\{a_k, k=1,\dots,n\}$, 其中 $n \le +\infty$, 则

$$\mathrm{P}(A) = \sum_{k=1}^{n} \mathrm{P}(A|Y=k) \cdot \mathrm{P}(Y=a_k).$$

- 设 Y 为连续型 rv, 具有概率密度函数 $g(y)$, 则

$$\mathrm{P}(A) = \int_{-\infty}^{\infty} \mathrm{P}(A|Y=y) \cdot g(y)\,\mathrm{d}y.$$

- 【例 5.2.11】 设 $Y \sim U(0,1)$,

$$[X|Y=p] \sim B(n,p), \quad 0<p<1,$$

求 X 的分布.

解: 对任意 $k=0,1,\ldots,n$,

$$\mathrm{P}(X=k)=\int_0^1 \binom{n}{k} p^k(1-p)^{n-k}dp=\frac{1}{n+1}.$$

- 【例 5.2.11】 设 $Y \sim U(0,1)$,

$$[X|Y=p] \sim B(n,p), \quad 0<p<1,$$

求 X 的分布.

解: 对任意 $k=0,1,\ldots,n$,

$$\mathrm{P}(X=k)=\int_0^1 \binom{n}{k} p^k(1-p)^{n-k}dp=\frac{1}{n+1}.$$

- 【例 5.2.14】 设 $X \sim \mathrm{Exp}(\lambda)$, $Y \sim \mathrm{Exp}(\mu)$, 且 X 与 Y 相互独立, 求 $\mathrm{P}(X < Y)$.

解: 对任意 $x > 0$, 利用 X 与 Y 相互独立性得

$$\mathrm{P}(X < Y | X = x) = \mathrm{P}(Y > x) = e^{-\mu x}.$$

于是,

$$\mathrm{P}(X < Y) = \mathrm{E}[\mathrm{P}(X < Y | X)] = \mathrm{E}[e^{-\mu X}] = \frac{\lambda}{\lambda + \mu}. \quad \blacksquare$$

- 【例 5.2.14】 设 $X \sim \mathrm{Exp}(\lambda)$, $Y \sim \mathrm{Exp}(\mu)$, 且 X 与 Y 相互独立, 求 $\mathrm{P}(X < Y)$.

解: 对任意 $x > 0$, 利用 X 与 Y 相互独立性得

$$\mathrm{P}(X < Y|X = x) = \mathrm{P}(Y > x) = e^{-\mu x}.$$

于是,

$$\mathrm{P}(X < Y) = \mathrm{E}[\mathrm{P}(X < Y|X)] = \mathrm{E}[e^{-\mu X}] = \frac{\lambda}{\lambda + \mu}. \blacksquare$$

- 【例 5.2.15】 设 $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ iid $\sim N(0,1)$, $n \geq 2$. 定义

$$Z = \frac{X_1Y_1 + X_2Y_2 + \cdots + X_nY_n}{\sqrt{Y_1^2 + Y_2^2 + \cdots + Y_n^2}},$$

证明 $Z \sim N(0,1)$.

解: 对任意 $y_1, \ldots, y_n, x \in \Re$, 对 $Y_1, \ldots, Y_n$ 取条件得

$$\mathrm{P}\left(Z \leq x | Y_1 = y_1, \ldots, Y_n = y_n\right) = \Phi(x).$$

于是,

$$\mathrm{P}\left(Z \leq x\right) = \mathrm{E}\left[\mathrm{P}\left(Z \leq x | Y_1, \ldots, Y_n\right)\right] = \Phi(x). \quad \blacksquare$$

- 【例 5.2.15】 设 $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ iid $\sim N(0,1)$, $n \geq 2$. 定义

$$Z = \frac{X_1 Y_1 + X_2 Y_2 + \cdots + X_n Y_n}{\sqrt{Y_1^2 + Y_2^2 + \cdots + Y_n^2}},$$

证明 $Z \sim N(0,1)$.

解: 对任意 $y_1, \ldots, y_n, x \in \Re$, 对 $Y_1, \ldots, Y_n$ 取条件得

$$\mathrm{P}\left(Z \leq x | Y_1 = y_1, \ldots, Y_n = y_n\right) = \Phi(x).$$

于是,

$$\mathrm{P}\left(Z \leq x\right) = \mathrm{E}\left[\mathrm{P}\left(Z \leq x | Y_1, \ldots, Y_n\right)\right] = \Phi(x). \quad \blacksquare$$

5.2.3 数学期望的一些其它应用

- 【例 5.2.17】 设 $A_1, A_2, \ldots, A_n$ 为同一概率空间上的 n 个事件，使得 $\mathrm{P}\left(\cup_{i=1}^{n} A_i\right) > 0$，证明：

$$\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) \geq \frac{\left(\sum_{i=1}^{n} \mathrm{P}(A_i)\right)^2}{\sum_{i=1}^{n} \mathrm{P}(A_i) + 2\sum_{1\leq i<j\leq n} \mathrm{P}(A_i A_j)}.$$

证：令 $X_i = I_{A_i}$, $i = 1, \ldots, n$. 由 Schwartz 不等式得到

$$\begin{aligned}\left(\sum_{i=1}^{n} \mathrm{P}(A_i)\right)^2 &= \left(\mathrm{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2 \\ &= \left\{\mathrm{E}\left[I\left(\sum_{i=1}^{n} X_i > 0\right) \cdot \sum_{i=1}^{n} X_i\right]\right\}^2 \\ &\leq \mathrm{P}\left(\sum_{i=1}^{n} X_i > 0\right) \cdot \mathrm{E}\left(\sum_{i=1}^{n} X_i\right)^2.\end{aligned}$$

……

5.2.3 数学期望的一些其它应用

- 【例 5.2.17】 设 $A_1, A_2, \ldots, A_n$ 为同一概率空间上的 n 个事件，使得 $\mathrm{P}\left(\cup_{i=1}^n A_i\right) > 0$，证明：

$$\mathrm{P}\left(\bigcup_{k=1}^n A_k\right) \geq \frac{\left(\sum_{i=1}^n \mathrm{P}(A_i)\right)^2}{\sum_{i=1}^n \mathrm{P}(A_i) + 2\sum_{1\leq i<j\leq n} \mathrm{P}(A_i A_j)}.$$

证：令 $X_i = I_{A_i}$, $i = 1, \ldots, n$. 由 Schwartz 不等式得到

$$\begin{aligned}
\left(\sum_{i=1}^n \mathrm{P}(A_i)\right)^2 &= \left(\mathrm{E}\left[\sum_{i=1}^n X_i\right]\right)^2 \\
&= \left\{\mathrm{E}\left[I\left(\sum_{i=1}^n X_i > 0\right) \cdot \sum_{i=1}^n X_i\right]\right\}^2 \\
&\leq \mathrm{P}\left(\sum_{i=1}^n X_i > 0\right) \cdot \mathrm{E}\left(\sum_{i=1}^n X_i\right)^2 .
\end{aligned}$$

······ ∎

- 【例 5.2.18】 沿着一个圆周等间距地分布着 52 棵松树，有 15 只松鼠生活在这些松树上. 证明：有某相连的 7 棵松树上至少生活着 3 只松鼠

证： 将 15 只松鼠依次编号 $1, 2, \ldots, 15$. 随机选择某相连的 7 棵松树. 定义

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 号松鼠生活在所选取的 7 棵松树之一上}, \\ 0, & \text{否则}, \end{cases}$$

于是，这 7 棵松树上的松鼠一共有 $X = \sum_{i=1}^{15} X_i$. 易知，$\mathrm{E}[X_i] = 7/52$, $i = 1, \ldots, 15$, 故

$$\mathrm{E}\,X = \sum_{i=1}^{15} \mathrm{E}[X_i] = 7 \times \frac{15}{52} = \frac{105}{52} > 2.$$

根据抽屉原理，存在某相连的 7 棵松树上至少生活着 3 只松鼠. ∎

- 【例 5.2.18】 沿着一个圆周等问距地分布着 52 棵松树，有 15 只松鼠生活在这些松树上. 证明：有某相连的 7 棵松树上至少生活着 3 只松鼠

证： 将 15 只松鼠依次编号 $1, 2, \ldots, 15$. 随机选择某相连的 7 棵松树. 定义

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 号松鼠生活在所选取的 7 棵松树之一上}, \\ 0, & \text{否则}, \end{cases}$$

于是，这 7 棵松树上的松鼠一共有 $X = \sum_{i=1}^{15} X_i$. 易知，$\mathrm{E}[X_i] = 7/52$, $i = 1, \ldots, 15$, 故

$$\mathrm{E}\, X = \sum_{i=1}^{15} \mathrm{E}[X_i] = 7 \times \frac{15}{52} = \frac{105}{52} > 2.$$

根据抽屉原理，存在某相连的 7 棵松树上至少生活着 3 只松鼠. ∎

5.2.4 条件方差及其应用

- 条件方差的定义：

$$\begin{aligned}
\mathrm{Var}(X|Y=y) &= \mathrm{E}\left\{\left(X-\mathrm{E}[X|Y=y]\right)^2|Y=y\right\} \\
&= \mathrm{E}[X^2|Y=y]-(\mathrm{E}[X|Y=y])^2
\end{aligned}$$

- 条件方差公式：

$$\mathrm{Var}(X)=\mathrm{E}\left[\mathrm{Var}(X|Y)\right]+\mathrm{Var}\Big(\mathrm{E}[X|Y]\Big).$$

- 【例 5.2.20】 设 $(0,t]$ 时段到达车站的旅客人数 $N(t)\sim \mathrm{Poi}(\lambda t)$，旅游车到达车站的时刻 Y 与乘客到达时刻相互独立，且 $Y\sim U(0,t_0)$. 求旅游车到达车站时乘客人数的期望与方差.

解：旅游车到站时乘客人数为 $N(Y)$. 对 Y 取条件，则

$$\mathrm{E}\,[N(Y)|Y=t]=\mathrm{E}\,[N(t)|Y=t]=\mathrm{E}\,N(t)=\lambda t,$$

于是, $\mathrm{E}\,[N(Y)]=\mathrm{E}\,[\lambda Y]=\lambda t_0/2$. 另一方面,

$$\begin{aligned}\mathrm{Var}(\mathrm{E}\,[N(Y)|Y]) &= \mathrm{Var}(\lambda Y)=\frac{\lambda^2 t_0^2}{12},\\ \mathrm{Var}(N(Y)|Y=t) &= \mathrm{Var}(N(t))=\lambda t.\end{aligned}$$

于是,

$$\mathrm{Var}(N(Y))=\mathrm{E}\left[\mathrm{Var}(N(Y)|Y)\right]+\mathrm{Var}\Big(\mathrm{E}\,[N(Y)|Y]\Big)=\frac{\lambda^2 t_0^2}{12}+\frac{\lambda t_0}{2}.$$

- 【例 5.2.20】 设 $(0,t]$ 时段到达车站的旅客人数 $N(t) \sim \mathrm{Poi}(\lambda t)$，旅游车到达车站的时刻 Y 与乘客到达时刻相互独立，且 $Y \sim U(0,t_0)$. 求旅游车到达车站时乘客人数的期望与方差.

解：旅游车到站时乘客人数为 $N(Y)$. 对 Y 取条件，则

$$\mathrm{E}\,[N(Y)|Y=t] = \mathrm{E}\,[N(t)|Y=t] = \mathrm{E}\,N(t) = \lambda t,$$

于是, $\mathrm{E}\,[N(Y)] = \mathrm{E}\,[\lambda Y] = \lambda t_0/2$. 另一方面,

$$\begin{aligned} \mathrm{Var}(\mathrm{E}\,[N(Y)|Y]) &= \mathrm{Var}(\lambda Y) = \frac{\lambda^2 t_0^2}{12}, \\ \mathrm{Var}(N(Y)|Y=t) &= \mathrm{Var}(N(t)) = \lambda t. \end{aligned}$$

于是,

$$\mathrm{Var}(N(Y)) = \mathrm{E}\left[\mathrm{Var}(N(Y)|Y)\right] + \mathrm{Var}\Big(\mathrm{E}\,[N(Y)|Y]\Big) = \frac{\lambda^2 t_0^2}{12} + \frac{\lambda t_0}{2}.$$

- 【例 5.2.b】 设 $X_1, X_2, \ldots$ iid, $\mathrm{E}\,X_1 = \mu$, $\mathrm{Var}(X_1) = \sigma^2$, N 为非负整值 rv, 独立于 $\{X_n, n \geq 1\}$, 满足 $\mathrm{E}\,N < \infty$, $\mathrm{Var}(N) = \tau^2 < \infty$. 记 $S = \sum_{i=1}^N X_i$, 求 $\mathrm{Var}(S)$.

解：对 N 取条件，得

$$\begin{aligned}
\mathrm{E}\,[S|N=n] &= \mathrm{E}\left[\sum_{i=1}^n X_i\right] = n\mu, \\
\mathrm{Var}(S|N=n) &= n\sigma^2
\end{aligned}$$

所以

$$\begin{aligned}
\mathrm{Var}(S) &= \mathrm{E}\left[\mathrm{Var}(S|N)\right] + \mathrm{Var}\Big(\mathrm{E}\,[S|N]\Big) \\
&= \mathrm{E}\,[N\sigma^2] + \mathrm{Var}(N\mu) \\
&= \sigma^2 \mathrm{E}\,N + \mu^2\tau^2. \quad \blacksquare
\end{aligned}$$

- 【例 5.2.b】 设 $X_1, X_2, \ldots$ iid, $\mathrm{E}\, X_1 = \mu$, $\mathsf{Var}(X_1) = \sigma^2$, N 为非负整值 rv, 独立于 $\{X_n, n \geq 1\}$, 满足 $\mathrm{E}\, N < \infty$, $\mathsf{Var}(N) = \tau^2 < \infty$. 记 $S = \sum_{i=1}^{N} X_i$, 求 $\mathsf{Var}(S)$.

解：对 N 取条件，得

$$\begin{aligned}
\mathrm{E}\,[S|N=n] &= \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = n\mu, \\
\mathsf{Var}(S|N=n) &= n\sigma^2
\end{aligned}$$

所以

$$\begin{aligned}
\mathsf{Var}(S) &= \mathrm{E}\left[\mathsf{Var}(S|N)\right] + \mathsf{Var}\Big(\mathrm{E}\,[S|N]\Big) \\
&= \mathrm{E}\,[N\sigma^2] + \mathsf{Var}(N\mu) \\
&= \sigma^2 \mathrm{E}\, N + \mu^2\tau^2. \quad \blacksquare
\end{aligned}$$

作业

第 5 章第 1 次作业

§5.1: 6, 8, 9, 14, 15, 17, 23, 24

第 5 章第二次作业

§5.2: 1–10