Part III-B: Probability Theory and Mathematical Statistics

Lecture by 李漫漫

Note by THF

2024 年 10 月 18 日

目录

Lecture 6

Corollary. X 是连续性随机变量，密度函数为 $f_X(x)$ ，随机变量 $Y=g(X)$，且 $\exists D, P\{Y\in D\}=1$，$g(x)$ 存在反函数 $h(y)$ 且严格单调可导，则：

$$f_Y(y)=\begin{cases}|h'(y)|\,f_X(h(y)), y\in D\\ 0, \text{Others}\end{cases}.$$

Notation. 指数分布 $X\sim \Gamma(1,\lambda)$ 的数学期望 $E(X)=\dfrac{1}{\lambda}$

1 多维随机变量函数及其分布

在实际问题中，试验结果有时需要使用两个或两个以上的随机变量 (random value, r.v.) 来描述

Example. 天气预报：温度、湿度、风力、降水等

1.1 二维随机变量及其分布

Definition. 设 Ω 为随机试验的样本空间，则

$$\forall\omega\in\Omega \xrightarrow{\text{某种变化}} \exists(X(\omega), Y(\omega))\in\mathbb{R}^2.$$

或：

$$\{X\le x, Y\le y\}=\{\omega|X(\omega)\le x, Y(\omega)\le y, \omega\in\Omega\}\in\mathscr{F}.$$

称 (X,Y) 为概率空间 $(\Omega,\mathscr{F},P)$ 上的二维随机变量

Notation. $\{X\le x, Y\le y\}=\{X\le x\}\cap\{Y\le y\}$

性质：

1. $F(x,y)\in[0,1]$
2. 关于每个变量单调不减，即固定 x ，对 $\forall y_1<y_2$，

$$F(x,y_1)\le F(x,y_2).$$

3. 对每个变量右连续，即：

$$F(x_0, y_0) = F\left(x_0 + 0^+, y_0\right) = F\left(x_0, y_0 + 0^+\right).$$

4. 对 $\forall a < b, c < d$ ，有：

$$F(b, d) - F(b, c) - F(a, d) + F(a, c) \geq 0.$$

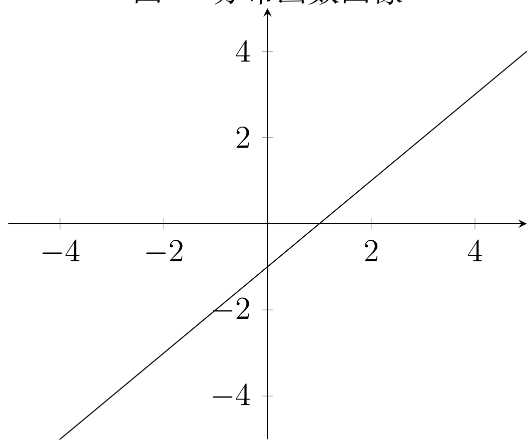
即：在任意地方框一个矩形，内部区域的概率必须大于等于 0

Example. 性质 4 例题：设

$$F(x, y) = \begin{cases} 0, x + y < 1 \\ 1, x + y \geq 1 \end{cases}.$$

讨论 $F(x, y)$ 能否成为二维随机变量的分布函数

图 1: 分布函数图像



Notation.

$$P\{X > a, Y > c\} \neq 1 - F(a, c).$$

1.2 二维随机变量的边缘分布函数

边缘分布：降一维

$$\begin{aligned} F_X(x) &= P\{X \le x\} \\ &= P\{X \le x, Y < +\infty\} \\ &= F(x, +\infty) \end{aligned}$$

.

Example. 设随机变量 (X, Y) 的联合分布函数为

$$F(x,y) = A\left(B + \arctan\frac{x}{2}\right)\left(C + \arctan\frac{y}{2}\right), x, y \in (-\infty, +\infty).$$

求 A, B, C

解：$\arctan x$ 的性质：

$$\lim_{x \to \pm\infty} \arctan x = \pm\frac{\pi}{2}.$$

则

$$F(+\infty,+\infty) = A\left(B + \frac{\pi}{2}\right)\left(C + \frac{\pi}{2}\right) = 1.$$
$$F(-\infty,+\infty) = A\left(B - \frac{\pi}{2}\right)\left(C + \frac{\pi}{2}\right) = 0.$$
$$F(-\infty,-\infty) = A\left(B - \frac{\pi}{2}\right)\left(C - \frac{\pi}{2}\right) = 1.$$

联立解出 A, B, C

1.3 联合分布律

二维离散随机变量的联合分布函数

$$F(x,y) = \sum_{x_i \le x} \sum_{y_j \le y} p_{ij}.$$

如何求 p_{ij} ：

1. 古典概型

2. 乘法公式：

$$p_{ij} = P\{X = x_i\} P\{Y = y_i | X = x_i\} = P\{X = x_i, Y = y_i\}.$$

表 1: 联合分布律

X \ Y	b_1	b_2	$\dots$	b_j	$p_{a\cdot}$
a_1	p_{11}	p_{12}	$\dots$	p_{1j}	$\sum_{n=1}^{j} p_{nj}$
a_2	p_{21}	p_{22}	$\dots$	p_{2j}	$\sum_{n=2}^{j} p_{nj}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
a_i	p_{i1}	p_{i2}	$\dots$	p_{ij}	$\sum_{n=i}^{j} p_{nj}$
$p_{b\cdot}$	$\sum_{m=1}^{i} p_{im}$	$\sum_{m=2}^{i} p_{im}$	$\dots$	$\sum_{m=1}^{i} p_{im}$	1

1.4 二维连续性随机变量及其概率特性

Definition. 若

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,\mathrm{d}v\mathrm{d}u.$$

则称 $f(x,y)$ 为二维随机变量 (X,Y) 的联合密度函数，称 (X,Y) 为二维连续型随机变量

Notation. 联合密度与联合分布函数的性质:

1. $f(x,y)\geq 0$
2. $$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y\mathrm{d}x=1=F(-\infty,+\infty).$$
3. 对每个边缘连续，在 $f(x,y)$ 的连续点处:
$$\frac{\partial^2 F}{\partial x\partial y}=f(x,y).$$

从而有: $P(x<X\leq x+\Delta x, y<Y\leq y+\Delta y)\approx f(x,y)\,\Delta x\Delta y$

Lecture 7

两个随机变量的独立性

证明. 当 A, B 独立时：$P(AB) = P(A)P(B)$

$$\forall i, j : P\{X = x_i, Y = y_j\} = P\{X = x_i\} P(Y = y_j).$$

□

Notation. 离散随机变量独立的情况下：$P_{ij} = P_{i\cdot} \cdot P_{\cdot j}$

如何证明 X, Y 独立：

$$F(x, y) = P\{X \le x, Y \le y\} = P\{X \le x\} P\{Y \le y\} = F_X(X) \cdot F_Y(y).$$

由联合分布律得边缘分布律：

$$F_X(x) = \lim_{y \to +\infty} F(x, y) = \lim_{y_0 \to +\infty} P\{X \le x, Y \le y_0\}.$$

Notation. 连续性随机变量的区间 D 概率：

$$F(x, y) = P\{X \le x, Y \le y\} = \iint_{u=x, v=y} f(u, v)\,\mathrm{d}x\mathrm{d}y.$$

概率函数

Notation. 二维均匀分布：

$$f(x, y) = \begin{cases} \dfrac{1}{S(D)}, & (x, y) \in D \\ 0, & (x, y) \notin D \end{cases}.$$

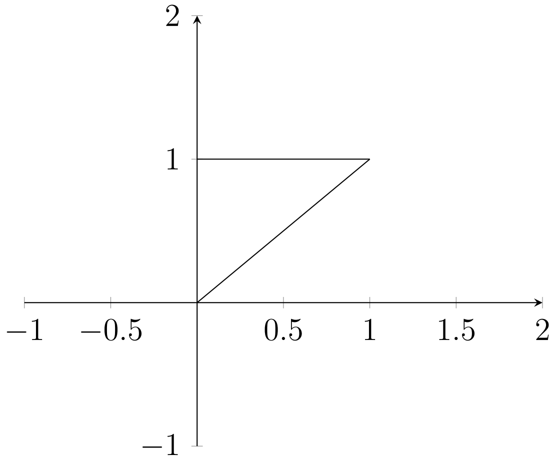
Example.

$$f(x, y) = \begin{cases} Axy, & x \in (0, 1), y \in (0, 1) \\ 0, & \text{Others} \end{cases}.$$

1. 求 A

$$\iint_{x \in [0, y], y \in [0, 1]} Axy\mathrm{d}x\mathrm{d}y = 1.$$

图 2: f(x,y)



$A = 8$

2. $P\left\{X+Y \geq 1\right\}$

$$
\begin{aligned}
P\left\{X+Y \geq 1\right\} &= \iint_{x+y \geq 1} f\left(x, y\right) \mathrm{d}x\mathrm{d}y \\
&= \iint_{x+y \geq 1} 8xy\mathrm{d}x\mathrm{d}y \\
&= \int_{0.5}^{1} \mathrm{d}y \int_{1-y}^{y} 8xy\mathrm{d}x \\
&= \frac{5}{6}
\end{aligned}
$$

3. X, Y 的分布函数

3.1 $x \in (-\infty, 0), y \in (-\infty, 0)$

3.2 $x \in [0, 1), y \in [0, x)$

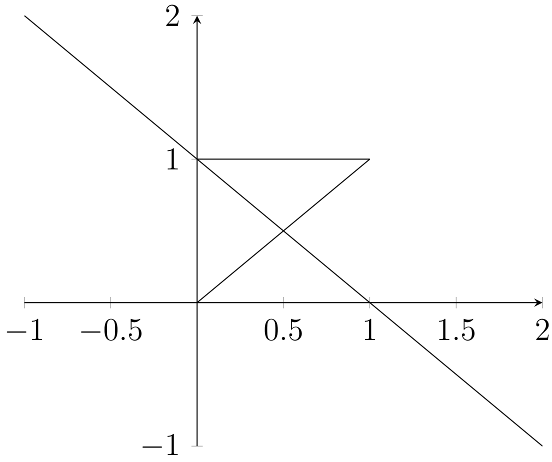
3.3 $x \in [0, 1), y \in [x, 1)$

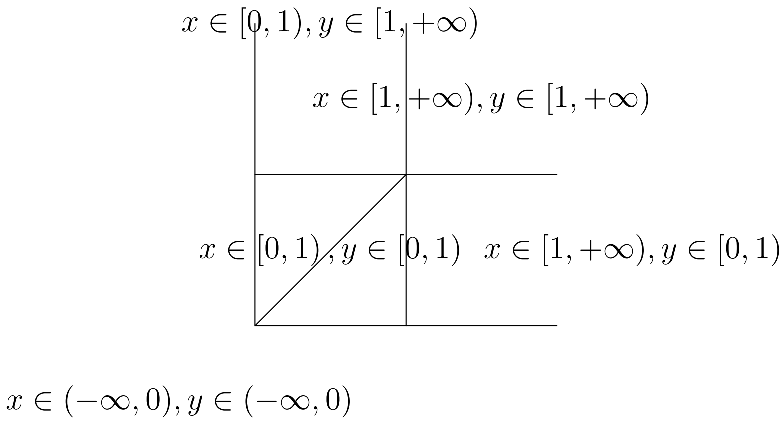
3.4 $x \in [0, 1), y \in [1, +\infty)$

3.5 $x \in [1, +\infty), y \in [0, 1)$

3.6 $x \in [1, +\infty), y \in [1, +\infty)$

图 3: P(x+y>= 1)





分段对 Axy 积分:

$$F(x, y) = \iint_{x \le u, y \le v} f(u, v)\, \mathrm{d}x\mathrm{d}y$$

1.5 多维随机变量及分布

Definition. 二维推广至多维:

$$F(x_1, x_2, \cdots, x_n) = P\{X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\}.$$

称 F 为 n 维随机变量的联合分布函数

Definition. 多维联合分布律:

$$P\left\{X_1=a_{1k_1},X_2=a_{2k_2},\cdots,X_n=a_{nk_n}\right\}.$$

联合分布函数和联合分布律的关系:

$$F\left(x_1,x_2,\cdots,x_n\right).$$

$$=\sum_{a_{1k_1}\le x_1}\sum_{a_{2k_2}\le x_2}\cdots\sum_{a_{nk_n}\le x_n}P\left\{X_1=a_{1k_1},X_2=a_{2k_2},\cdots,X_n=a_{nk_n}\right\}.$$

Notation. 二项分布推广多项分布:

$A_1,A_2,\cdots,A_r$ 是 E 的完备事件组,$P\left(A_i\right)=p_i,i=1,2,\cdots,r$,对 E 进行 n 次独立重复试验,X_i 表示 A_i 发生的次数,则:

$$P\left\{X_1=k_1,X_2=k_2,\cdots,X_r=k_r\right\}=\frac{n!}{k_1!k_2!\cdots k_r!}\prod_{i=1}^{r}p_i^{k_i}.$$

其中 $k_i\ge 0,\sum_{i=1}^{r}k_i=n$,当 $n=2$ 时为二项分布

1.5.1 多维随机变量的独立性

Definition.

$$F\left(x_1,x_2,\cdots,x_n\right)=\prod_{i=1}^{n}F_{X_i}\left(x_i\right).$$

称随机变量相互独立

等价于:

$$f\left(x_1,x_2,\cdots,x_n\right)=\prod_{i=1}^{n}f_{X_i}\left(x_i\right).$$

Lecture 8

10.15

1.5.2 条件分布

Notation. 离散型条件分布:

$$P\left\{X=a_i|Y=b_j\right\}=\frac{P\left\{X=a_i,Y=b_j\right\}}{P\left\{Y=b_j\right\}}.$$

$$P\left\{Y=b_j\right\}=\sum_{i=1}^{+\infty} P\left\{Y=b_j, X=a_i\right\}.$$

性质:

1. $P\left\{X=a_i \mid Y=b_j\right\} \geq 0$
2. $\sum\limits_{i=1}^{+\infty} P=1$

Notation. 连续型条件分布:

$$P\{X=a \mid Y=b\}=0.$$

(无穷多个点)

通过微元法:

$$\begin{aligned}
P\{X \leq x \mid Y=y\} &= \lim_{\varepsilon \rightarrow 0^{+}} P\{X \leq x \mid y-\varepsilon<Y \leq y\} \\
&= \lim_{\varepsilon \rightarrow 0^{+}} \frac{P\{X \leq x, Y \in(y-\varepsilon, y]\}}{P\{Y \in(y-\varepsilon, y]\}} \\
&= \lim_{\varepsilon \rightarrow 0^{+}} \frac{F(x, y)-F(x, y-\varepsilon)}{F_Y(y)-F_Y(y-\varepsilon)} \\
&= \frac{\dfrac{\partial F(x, y)}{\partial y}}{\dfrac{\mathrm{d} F_Y(y)}{\mathrm{d} y}} \\
&= \frac{\displaystyle\int_{-\infty}^{x} f(u, y) \,\mathrm{d} u}{f_Y(y)} \\
&= \int_{-\infty}^{x} \frac{f(u, y)}{f_Y(y)} \mathrm{d} u
\end{aligned}$$

.

1.6 二维随机变量函数的分布

$$\begin{cases} aX+bY+c \\ \max\{X, Y\} \\ \min\{X, Y\} \end{cases}.$$

Lecture 8

重点公式：3.4.5 ~ 3.4.8

假设随机变量 Z=aX+bY+c，有分布函数 $(X,Y) \to F(x,y)$

$$
\begin{aligned}
F_Z(z) &= P\{Z \le z\} \\
&= P\{aX + bY + c \le z\} \\
&= \begin{cases} \displaystyle\sum_{ax_i+by_j+c \le z}\sum P\{X = x_i, Y = y_j\}, & (X,Y)\text{ 离散} \\ \displaystyle\iint_{ax+by+c \le z} P\{X = x, Y = y\}\,\mathrm{d}x\mathrm{d}y, & (X,Y)\text{ 连续} \end{cases}
\end{aligned}
$$

.

Notation. 二项分布可加性: 有两个相互独立的试验 $X \sim B(m,p), Y \sim B(n,p)$ ，相当于一个试验 $Z \sim B(m+n,p)$

泊松分布可加性：相互独立的两个随机变量 $X \sim P(\lambda_1), Y \sim P(\lambda_2)$ ，相当于一个分布 $Z \sim P(\lambda_1 + \lambda_2)$

Notation. 极值公式:

两个随机变量（连续）X, Y 相互独立，求 $Z_1 = \max(X,Y), Z_2 = \min(X,Y)$ 的分布函数和密度函数

1. $F_{Z_1}(z)$： 最大的不超过 z 等价于每一个都不超过 z

$$
\begin{aligned}
F_{Z_1}(z) &= P\{X \le z, Y \le z\} \\
&= P\{X \le z\} \cdot P\{Y \le z\} \\
&= F_X(z) \cdot F_Y(z)
\end{aligned}
$$

.

2. $F_{Z_2}(z)$: 最小的不超过 z 不等价于每一个都不超过 z ，但最小的超过 z 等价与每一个都超过了 z

$$
\begin{aligned}
F_{Z_2}(z) &= 1 - P\{X > z, Y > z\} \\
&= 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)]
\end{aligned}
$$

.

Notation. 独立同分布：变量相互独立且分布律相同

对极值公式扩展：$(X_1, X_2, \ldots, X_n)$ 独立同分布：

$$F_{Z_1}(z) = \prod_{i=1}^{n} F_{X_i}(z).$$

由于同分布，因此 $F_{X_i}(z) = F_X(z)$

$$F_{Z_1}(z) = (F_X(z))^n.$$

$$f_{Z_1}(z) = (F_{Z_1}(z))' = n(F_X(z))^{n-1}.$$

Notation. 3.4.5:

$$\begin{aligned}
F_Z(z) &= P\{X + Y \le z\} \\
&= \iint_D f(x, y)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{z-x} f(x, y)\,\mathrm{d}y \\
&= \int_{-\infty}^{z} \mathrm{d}t \int_{-\infty}^{+\infty} f(x, t - x)\,\mathrm{d}x \\
f_Z(z) &= F_Z'(z) \\
&= \int_{-\infty}^{+\infty} f(x, z - x)\,\mathrm{d}x
\end{aligned}.$$

同理：

$$f_Z(z) = \int_{-\infty}^{+\infty} f(z - y, y)\,\mathrm{d}y.$$

称以上两个公式为卷积公式：$f_Z = f_X * f_Y$

Notation. 正态分布的可加性：

$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$ ，且 X, Y 相互独立，则：

$$X + Y = Z \sim N\left(\mu_1 + \mu_2, \sigma_1^2, \sigma_2^2\right).$$

多元：

$$Z=\sum_{i=1}^{n}a_iX_i\sim N\left(\sum_{i=1}^{n}a_i\mu_i,\sum_{i=1}^{n}a_i^2\sigma_i^2\right).$$

1.7 二元正态分布

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}$$
$$\times\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\}.$$

Lecture 9

10.17

Example. $X\sim U[-1,1]$ ，对应的密度函数：

$$f_X(x)=\begin{cases}0, & x<-1 \text{ or } x>1\\ \frac{1}{2}, & x\in[-1,1]\end{cases}.$$

同理 $Y\sim U[-1,1]$，求 $Z=|X-Y|$ 的分布函数

解：Z 的取值：[0,2]

$$\begin{aligned}F_Z(z)&=P\{Z\le z\}\\&=P\{|X-Y|\le z\}\\&=\begin{cases}0, & z<0\\ 1, & z\ge 2\\ P\{-z\le X-Y\le z\}, & z\in[0,2)\end{cases}.\end{aligned}$$

其中

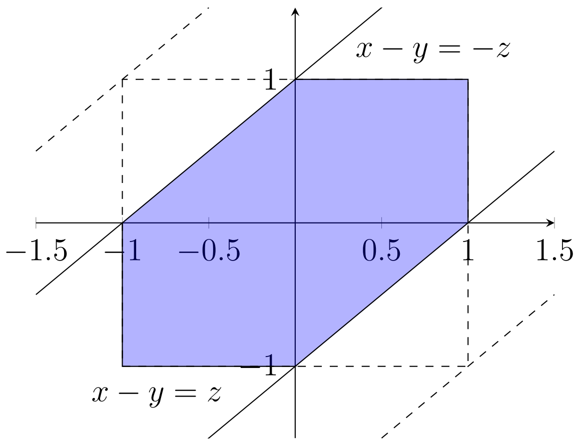
$$P\{z-z\le X-Y\le z\}=\iint_{-z\le x-y\le z}f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

通过 f_X和f_Y 求联合密度函数：X, Y 独立，即 $f(x,y) = f_X(x) f_Y(y)$

$$f(x,y) = \begin{cases} 0, & x,y < -1 \text{ or } x,y > 1 \\ \frac{1}{4}, & x,y \in [-1,1] \end{cases}.$$

$$P\{-z \le X - Y \le z\} = \iint\limits_{-z \le x-y \le z, x,y \in [-1,1]} \frac{1}{4} \mathrm{d}x \mathrm{d}y.$$

画图确认积分区域：



Notation. i.i.d. ：独立同分布

Notation. 伽马分布 $\Gamma(\alpha, \beta)$ 的密度函数：

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\beta x}, & x \ge 0 \\ 0, & x < 0 \end{cases}.$$

Notation. 重点题目：3.6.2

2 数字特征

$$数字特征\begin{cases} 数学期望：E\left(X\right) \\ 方差：D\left(X\right)\ \text{ or }\operatorname{Var}\left(X\right) \\ 协方差：\operatorname{cov}\left(X,Y\right) \\ 相关系数：\rho\left(X,Y\right) \\ 矩：E\left(X\right)^k\ \text{ and }E\left(X-EX\right)^k \end{cases}.$$

Notation. $D\left(X\right)=E\left(X-EX\right)^2$

$\operatorname{cov}\left(X,Y\right)=E\left(\left(X-EX\right)\left(Y-EY\right)\right)$

$\rho\left(X,Y\right)=\dfrac{\operatorname{cov}\left(X,Y\right)}{\sqrt{DX}\sqrt{DY}}$

Notation. $|\rho_{X,Y}|\in[0,1]$

ρ 越大越线性相关，$\rho>0.8$ 时基本可以确定为线性相关

Notation. 矩 (moment) 是最一般的概念

矩分为两大类：k 阶原点矩和 k 阶中心矩

原点矩：$E\left(X\right)^k$

中心矩：$E\left(X-EX\right)^k$

$k+l$ 阶混合中心矩：$E\left(\left(X-EX\right)^k\left(Y-EY\right)^l\right)$

Example. 数学期望为一阶原点矩

方差为一阶中心矩

协方差为二阶混合中心矩

Notation. 可以写出无穷阶的中心矩等同于通过泰勒原理得出分布函数

本章重点：如何计算任意随机变量有关函数的数学期望

唯一计算公式：4.1.5 和 4.1.6

2.1 数学期望

Definition. 离散型随机变量 X 的分布律：$P\left\{X=x_i\right\}=p_i,i=1,2,\ldots$

若级数 $\sum\limits_{i=1}^{+\infty} x_i p_i$ **绝对收敛**（$\sum\limits_{i=1}^{+\infty} |x_i| \, p_i < +\infty$），则 X 的数学期望**存在**（x_i 为取值，p_i 为权重，$p_i \geq 0$）

$$E(X) = EX = \sum_{i=1}^{+\infty} x_i P\{X = x_i\} = \sum_{i=1}^{+\infty} x_i p_i.$$

Rule. 当一个随机变量的密度函数与分布律已知：$X \to f(x), P\{X = x_i\} = p_i$

即可以求关于 X 函数的数学期望（公式 4.1.5）：

$$E(g(X)) = \begin{cases} \sum\limits_{i=1}^{+\infty} g(x_i) P\{X = x_i\}, & X\text{离散} \\ \int_{-\infty}^{+\infty} g(x) f(x) \, \mathrm{d}x, & X\text{连续} \end{cases}.$$

Rule. 扩展至二阶：$(X, Y) \to P\{X = x_i, Y = y_j\}, f(x, y)$

关于 (X, Y) 的函数的数学期望（公式 4.1.6）：

$$E(g(X, Y)) = \begin{cases} \sum\limits_{i=1}^{+\infty} \sum\limits_{j=1}^{+\infty} g(x_i, y_j) P\{X = x_i, Y = y_j\}, & \text{离散} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) \, \mathrm{d}x\mathrm{d}y, & \text{连续} \end{cases}.$$

Notation. 柯西分布：

$$f(x) = \frac{1}{\pi(1 + x^2)}, x \in \mathbb{R}.$$

常见分布数学期望：

Notation. 伯努利分布 $X \sim B(n, p)$： $EX = np$

泊松分布 $X \sim P(\lambda)(\lambda > 0)$： $EX = \lambda$

柯西分布：EX 不存在（柯西分布不绝对收敛）

Notation. 柯西活了 68 岁，21 岁成名（导师拉格朗日）,27 岁当选法国科学院院士