第二章、随机变量

§2.1 离散型随机变量

一、几个常见的分布列

- 单点分布(distribution)/退化的: $\mathrm{P}(X=c)=1$.
- 等概率分布: $\mathrm{P}(X=x_i)=1/n$, 其中$i=1,\cdots,n$.
- 超几何分布(Hypergeometric), $X\sim H(N,M,n)$:

$$\mathrm{P}(X=k)=h(N,M,n;k):=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\quad k=0,1,\cdots,n.$$

  - 例2.1.1 (最大似然估计). 湖中有N 条鱼, 捕捞100 条, 标记、放回、三天后再捕捞80 条, 发现其中6 条带标记. 试估计N.
    - 取$M=100$, $n=80$, $\{X=6\}$ 发生了.

$$\mathrm{P}(X=6)=C_{100}^{6}C_{N-100}^{74}/C_N^{80}\triangleq p(N).$$

    - 最大值点$\hat{N}=1333$. 未知参数vs 随机变量.

- 伯努利分布: $\mathrm{P}(X=1)=p$, $\mathrm{P}(X=0)=1-p$.
- 二项(Binomial)分布, $X \sim B(n,p)$:

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

- §2.1 习题7. 固定n,p. 当$N\to\infty$, $\frac{M}{N}\to p$ 时,

$$h(N,M,n;k)\to b(n,p;k), \quad \forall k\geqslant 0.$$

  - 例, $n=5$. 产品检验模型中H-T 字符串的权重:

$$\mathrm{HHTHT}: \frac{M}{N}\cdot\frac{M-1}{N-1}\cdot\frac{N-M}{N-2}\cdot\frac{M-2}{N-3}\cdot\frac{N-M-1}{N-4}\to p^3(1-p)^2.$$

- 几何分布, $X \sim G(p)$:

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

- 负二项分布, $X \sim NB(r,p)$:

$$P(X=k)=nb(r,p;k)=C_{r+k-1}^{k}(1-p)^k p^r, \quad k=0,1,2,\cdots.$$

- 分赌注问题: 先胜t 局者赢. 甲已胜n 局, 乙已胜m 局.

如何分赌注?

(1) H: 甲胜一局, 概率为p.

甲还需胜$r=t-n$ 次, 乙还需$s=t-m$ 次.

(2) 接下来, 甲第r 次胜时, 乙恰胜X 次, $X \sim NB(r,p)$.

(3) $P(\text{“甲赢”})=P(X<s)=\sum_{k=0}^{s-1} nb(r,p;k)$.

- 泊松分布, $X \sim P(\lambda)$:

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

- 例2.1.5 (放射性物质).
  - 细分空间或时间. 假设

(1) 特定小块放射粒子的概率$p_n=\dfrac{\lambda}{n}$;
不放射的概率为$1-p_n$;

(2) 不同小块是否放射粒子的事件相互独立.

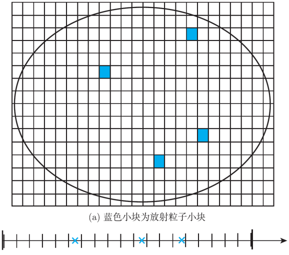


图 2.1

  - 当$n\to\infty$, $np\to\lambda$ 时, $b(n,p;k)\to p(\lambda;k)$:

$$b(n,p;k)\approx\frac{n^k}{k!}p^k(1-p)^n\to\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}.$$

- $C_n^k p^k (1-p)^{n-k}$ 与 $\frac{\lambda^k}{k!}e^{-\lambda}$.
  - 当$n \to \infty$, $np \to \lambda$ 时, $\sum_{k=0}^{\infty} |\star - \star| \to 0$. (命题A.0.1)
  - 当$np = \lambda$ 时, $\sum_{k=0}^{\infty} |\star - \star| \leqslant 2np^2 = 2\lambda^2/n$. (例2.5.16)
- 例2.1.6. 每人独立地以10^{-4} 的概率患某种重大疾病, 6 万人买保险. 求保险公司恰好赔偿7 人的概率.
  - 约为$\frac{\lambda^7}{7!}e^{-6} = 0.1377$, 误差 $\leqslant 0.0012$.
- §2.1 习题8. $B(n,p)$, $P(\lambda)$ 都是单峰的, 最大值点$\approx np, \lambda$.
- §2.1 习题10. $k \cdot p_k = \lambda \cdot p_{k-1}$, $k = 1, 2, \cdots$.
- $P(0)$ 指$X \equiv 0$ 的分布列.

总结:

- 离散型随机变量X:

$$\mathrm{P}(X = x_i) = p_i, \quad i \in I.$$

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

或

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

- 其中, $x_1, x_2, \cdots$ 互不相等; $p_i \geqslant 0, \forall i;\ \sum\limits_{i \in I} p_i = 1.$
- 分布列, $\{(x_i, p_i) : i \in I\}$,

$$\mu(\{x_i\}) = p_i, \quad i \in I$$

二、离散型随机变量的函数及其分布列

例2.1.7 (两点分布).

- 随机变量$X \sim B(1,p)$, 可能值: $0,1$.
- 函数$f:\{0,1\}\to\mathbb{R}$, $f(0)=a$, $f(1)=b$.
- $Y=f(X)$, 称为随机变量X 的函数.
- Y 的分布列:

$$\mathrm{P}(Y=a)=\mathrm{P}(X=0)=1-p,\quad \mathrm{P}(Y=b)=\mathrm{P}(X=1)=p.$$

- Y 服从两点分布.

例2.1.8 (帕斯卡分布).

- $X \sim NB(r,p)$, $f(x) = x + r$.
- $Y = X + r$.
- $Y \sim P(r,p)$, 分布列:

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

例2.1.9 (几何分布的尾分布与无记忆性). $X \sim G(p)$,

- $\mathrm{P}(X > n) = \sum\limits_{i=n+1}^{\infty} (1-p)^{i-1} p = (1-p)^n$.
- $f(k) = k - n$, $Y = X - n$.
- Y 的分布列:

$$\mathrm{P}(Y = k) = (1-p)^{n+k-1} p, \quad k = -(n-1), \cdots, 1, 0, 1, 2, \cdots.$$

- 若事件$\{X > n\}$ 发生, 则$Y \geqslant 1$.
- Y 的条件分布列: $k \geqslant 1$,

$$\mathrm{P}(Y = k | X > n) = \frac{\mathrm{P}(X = n + k)}{\mathrm{P}(X > n)} = \frac{(1-p)^{n+k-1} p}{(1-p)^n} = (1-p)^{k-1} p.$$

- 无记忆性:

$$\mathrm{P}(Y = k | X > n) = \mathrm{P}(X = k), \quad k = 1, 2, \cdots$$

§2.2 连续型随机变量

- 密度函数:

$$p(x) \geqslant 0, \quad \int_{-\infty}^{\infty} p(x)\mathrm{d}x = 1.$$

- 若

$$\mathrm{P}(X \leqslant x) = \int_{-\infty}^{x} p(y)\mathrm{d}y, \quad \forall x \in \mathbb{R}.$$

则称$p(\cdot)$ 为X 的密度函数, 记为$p_X(\cdot)$.

- 一般地,

$$\mathrm{P}(X > x) = \int_{x}^{\infty} p(y)\mathrm{d}y, \quad \mathrm{P}(a < X \leqslant b) = \int_{a}^{b} p(x)\mathrm{d}x;$$

$$p(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{P}(X \leqslant x) = -\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{P}(X > x).$$

一、几个常见的密度函数

- 均匀(Uniform)分布, $X \sim U(a,b)$:

$$\mathrm{P}(c < X \leqslant d) = (d-c)/(b-a), \quad a \leqslant c < d \leqslant b;$$

$$p(x) = \begin{cases} \dfrac{1}{b-a}, & a \leqslant x \leqslant b, \\ 0, & \text{其他}. \end{cases}$$

- 指数(Exponential)分布, $X \sim \mathrm{Exp}(\lambda)$:

$$p(x)=\begin{cases}\lambda \mathrm{e}^{-\lambda x}, & x>0;\\ 0, & x\leqslant 0.\end{cases}$$

- 例2.2.1. $X=$ 出现第一个放射粒子的时刻.
    - 以$1/n$ 为微观单位时间.

(b) 蓝×表示放射粒子的时刻

    - $X_n=$ 第一个H 出现的时刻$\sim G(\lambda/n)$. $X\approx X_n/n$.

$$\mathrm{P}(X>x)=\lim_{n\to\infty}\mathrm{P}(X_n>nx)=\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{[nx]}=\mathrm{e}^{-\lambda x}.$$

    - 密度函数: $p(x)=(\mathrm{e}^{-\lambda x})'=\lambda\mathrm{e}^{-\lambda x}$, $x>0$.

- 无记忆性: 对任意$T,t\geqslant 0$,

$$\mathrm{P}(X>T+t|X>T)=\mathrm{P}(X>t).$$

- 标准正态(Normal)分布, $X \sim N(0,1)$:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}.$$

- $\phi(\cdot)$ 确实是密度函数: 往证 $\int_{-\infty}^{\infty} \phi(x)\mathrm{d}x = 1$.

$$\star^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}} \mathrm{d}x\mathrm{d}y = \int_0^{2\pi}\int_0^{\infty} \frac{1}{2\pi}\mathrm{e}^{-\frac{r^2}{2}} r\mathrm{d}r\mathrm{d}\theta$$
$$= \left(\int_0^{2\pi} \frac{1}{2\pi}\mathrm{d}\theta\right) \cdot \left(\int_0^{\infty} \mathrm{e}^{-\frac{r^2}{2}} \mathrm{d}\frac{r^2}{2}\right) = 1.$$

- 引理2.2.3(局部极限定理). 设$S_n \sim B(n, 1/2)$, 则

$$\lim_{n\to\infty} \max_{k:a\leqslant x_k\leqslant b} \left| \frac{\mathrm{P}(S_n = k)}{\phi(x_k)\Delta x_k} - 1 \right| = 0, \quad 其中\ x_k = \frac{k - n/2}{\sqrt{n}}.$$

- 例2.2.4(SRW), 高尔顿板, $\phi(\cdot)$ 的性质, 图2.2.

二、连续型随机变量的函数及其密度函数

例2.2.5 (线性变换). X 是连续型, $f(x) = a + bx$, $(b \neq 0)$, $Y = f(X)$, 则

$$p_Y(y) = \frac{1}{|b|} \cdot p_X\left(\frac{y-a}{b}\right).$$

- $\mathrm{P}(Y \leqslant y) = \mathrm{P}(a + bX \leqslant y) \overset{b>0}{=} \mathrm{P}\left(X \leqslant \frac{y-a}{b}\right)$
- $\int_{-\infty}^{\frac{y-a}{b}} p_X(x)\mathrm{d}x = \int_{-\infty}^{y} p_Y(z)\mathrm{d}z.$
- $=: z = a + bx, \quad p_Y(z) := \frac{1}{b} \cdot p_X\left(\frac{z-a}{b}\right).$
- $b < 0$ 的情形类似.

例2.2.6 & 2.2.7 (正态的线性变换).

- $Z \sim N(0,1)$, $\mu \in \mathbb{R}$, $\sigma > 0$, 则$X = \mu + \sigma Z$ 的密度函数为:

$$p_X(x) = \frac{1}{\sigma} \cdot p_Z\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

- $X \sim N(\mu, \sigma^2)$, 参数为μ, σ^2 的正态分布.
- $X \sim N(\mu, \sigma^2)$, $a \in \mathbb{R}$, $b \neq 0$, $Y = a + bX$, 则

$$p_Y(y) = \frac{1}{\sqrt{2\pi(b\sigma)^2}} \exp\left\{-\frac{(y-(a+b\mu))^2}{2(b\sigma)^2}\right\}.$$

- 取$a = -\mu/\sigma$, $b = 1/\sigma$, 得$X^* = (X-\mu)/\sigma \sim N(0,1)$.
- $X = \mu + \sigma X^*$,

  $Y = a + bX = (a + b\mu) + (b\sigma)X^* \sim N(a + b\mu, (b\sigma)^2)$.

例2.2.8 & 2.2.9. $X \sim \text{Exp}(\lambda)$.

- $aX \sim \text{Exp}(\lambda/a)$, $(a > 0)$. 特别地, $\lambda X \sim \text{Exp}(1)$:

$$\text{P}(aX > x) = \text{P}(X > x/a) = \text{e}^{-(\lambda/a)x}.$$

- $Y \sim \text{Exp}(1) \Rightarrow Y/\lambda \sim \text{Exp}(\lambda)$.
- $\lambda = 1/2$ 时, $\sqrt{X}$ 服从瑞利分布:

$$\text{P}(\sqrt{X} > y) = \text{P}(X > y^2) = \text{e}^{-y^2/2} \Rightarrow p(y) = y\text{e}^{-y^2/2},\ y > 0.$$

- $\lambda = 1$ 时, $\alpha X^{1/m}$ 服从威布尔分布$W(m, \alpha)$:

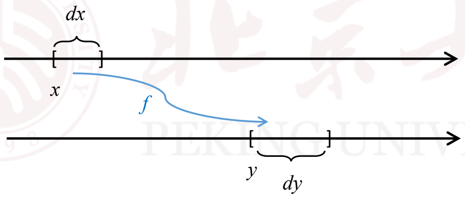
$$p(z) = \frac{m}{\alpha^m} z^{m-1} \text{e}^{-(z/\alpha)^m},\ \ z > 0.$$

- 验证: $m > 1$ 时, $\text{P}(X > T + t | X > T) < \text{P}(X > t)$.

定理2.2.10. 设X 是连续型, $\mathrm{P}(a < X < b) = 1$, $f : (a, b) \to (c, d)$ 一对一, $f'(x) \neq 0$, 则$Y = f(X)$ 的密度函数为

$$p_Y(y) = p_X\big(g(y)\big) \cdot |g'(y)|,\ y \in (c, d). \quad (g = f^{-1})$$

- 确定x, y 的取值范围. $p_X(x)|\mathrm{d}x| = p_Y(y)|\mathrm{d}y|$,



- 多对一:

$$p_Y(y) = \sum_{i \in I} p_X\big(g_i(y)\big) \cdot |g_i'(y)|,\ y \in (c, d). \quad (g_i = f_i^{-1})$$

例2.2.11. $X \sim U(-\pi/2, \pi/2)$, $Y = \tan X$, $W = \mu + \alpha Y$, $(\alpha \neq 0)$. 求Y 与W 的密度函数.

- $p_Y(y) = p_X(x) \cdot \dfrac{1}{1+y^2} = \dfrac{1}{\pi} \cdot \dfrac{1}{1+y^2}$.
- 设$\alpha > 0$ ($\alpha < 0$ 类似):

$$p_W(w) = \frac{1}{\alpha} p_Y\left(\frac{w-\mu}{\alpha}\right) = \frac{1}{\pi} \cdot \frac{\alpha}{\alpha^2 + (w-\mu)^2}.$$

- 注: $C(\mu, \alpha) = C(\mu, -\alpha)$.
- $W \sim C(\mu, \alpha)$, 参数为μ, α 的柯西分布. $Y \sim C(0, 1)$.
- §2.2 习题4(1).
  - $W \sim C(\mu, \alpha) \Rightarrow Y = (W - \mu)/\alpha \sim C(0, 1)$.
  - $a + bW = a + b\mu + b\alpha Y \sim C(a + b\mu, b\alpha)$.
- $X \sim U(0, \pi)$ 或$U(0, 2\pi) \Rightarrow \tan X \sim C(0, 1)$.

例2.2.12. 设$X \sim N(0,1)$, 求$Y = X^2$ 的密度函数.

- $f(x) = x^2$, 二对一. 范围: $y > 0$, 密度函数:

$$\begin{aligned} p_Y(y) &= \phi(\sqrt{y}) \cdot (\sqrt{y})' - \phi(-\sqrt{y}) \cdot (-\sqrt{y})' \\ &= \frac{1}{\sqrt{2\pi}} y^{-1/2} \mathrm{e}^{-y/2}. \end{aligned}$$

伽马分布$\Gamma(\alpha,\lambda)$, $\alpha>0$, $\lambda>0$.

- 密度函数:
$$p(x)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x},\ x>0.$$
- $\Gamma(\alpha)=\displaystyle\int_0^{\infty} y^{\alpha-1}\mathrm{e}^{-y}\mathrm{d}y,\ \ (y=\lambda x).$
- $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$, $\Gamma(1)=1$, $\Gamma(1/2)=\sqrt{\pi}$.
- 例2.2.12. 若$X\sim N(0,1)$, 则$X^2=\Gamma(1/2,1/2)$.
- 自由度为n 的卡方分布: $\chi^2(n)=\Gamma(n/2,1/2)$.
- 设$X\sim\mathrm{Exp}(\lambda)=\Gamma(1,\lambda)$, $S_n\sim\Gamma(n,\lambda)$,
  则$2\lambda X\sim\mathrm{Exp}(1/2)=\chi^2(2)$, $2\lambda S_n\sim\chi^2(2n)$.

§2.3 随机变量的定义与分布函数

一、随机变量的定义

- 函数、完全反像.

$$X: \quad \Omega \to \mathbb{R},$$
$$X^{-1}(D) = \{\omega : X(\omega) \in D\} = \{X \in D\}, \quad \forall D \subseteq \mathbb{R}.$$

- 定义2.3.1. 设$(\Omega, \mathscr{F})$ 是可测空间. 若$X : \Omega \to \mathbb{R}$ 满足

$$\{X \leqslant x\} \in \mathscr{F}, \quad \forall x \in \mathbb{R},$$

则称X 为$(\Omega, \mathscr{F})$ 上的随机变量(random variable).

- 例2.3.2. 设$\Omega = \{\mathrm{R}, \mathrm{O}, \mathrm{Y}, \mathrm{G}, \mathrm{B}, \mathrm{P}\}$, X 将样本分别映为$1 \sim 6$.

- 定义2.3.3. f 是Borel 函数指$\{x : f(x) \leqslant a\} \in \mathscr{B}, \forall a \in \mathbb{R}$.
- 例2.3.4 (连续型). 密度函数$p(\cdot)$ 是非负的Borel 函数, 以下积分是Lebesgue 积分,

$$\mathrm{P}(X \leqslant x) = \int_{-\infty}^{x} p(y)\mathrm{d}y, \quad \mathrm{P}(X \in B) = \int_{B} p(y)\mathrm{d}y.$$

  - 密度函数不唯一. 单独谈论一个点x 的$p(x)$ 是没有意义的.
  - 密度: 假设$p(\cdot)$ 在x_0 点连续, 则$p(x_0)$ 可确定:

$$p(x_0) = \lim_{\delta \to 0+} \frac{\mathrm{P}(|X - x_0| \leqslant \delta)}{2\delta}.$$

  - 密度不是概率: $\mathrm{P}(X = x) \neq p(x)$.
  - 若X 是连续型随机变量, 则对任意$x \in \mathbb{R}$, $\mathrm{P}(X = x) = 0$.

例2.3.5. 设$X \sim U(0,b)$, 其中$b>0$ 为待定参数. 重复独立试验得到观测值$x_1,\cdots,x_n$. 试估值b.

- $p(b;x)=\frac{1}{b}\cdot \mathbf{I}_{\{0\leqslant x\leqslant b\}}$ 和$\frac{1}{b}\cdot \mathbf{I}_{\{0<x<b\}}$ 都是密度函数.
- 第i 次小试验得到观察值X_i, 则$A_i=\{|X_i-x_i|<\delta\}$, $i=1,\cdots,n$ 相互独立.
- $\mathrm{P}(A_1\cdots A_n)\approx p(b;x_1)\cdots p(b;x_n)\delta^n$.
- 求$L(b)=p(b;x_1)\cdots p(b;x_n)$ 的最大值点.
- $\hat{b}=\max\{x_1,\cdots,x_n\}$.

二、随机变量的分布函数

- 定义2.3.6. X 的分布函数: $F_X(x) := \mathrm{P}(X \leqslant x)$,
  尾分布函数: $G_X(x) = \mathrm{P}(X > x)$.
- 连续型: $F'(x) = p(x)$, $G'(x) = -p(x)$.
- 离散型: $F(x_i) - F(x_i-) = p_i$, 跳点: x_i, 跳跃幅度: p_i.
- 奇异型: F 连续但没有“导数”.
- 命题2.3.7. $N(0,1)$, $\phi(x) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$,

$$\frac{1}{1+1/x^2} \leqslant \frac{G(x)}{x^{-1}\phi(x)} \leqslant 2,\ \forall x > 0; \qquad \lim_{x\to\infty}\frac{G(x)}{x^{-1}\phi(x)} = 1.$$

命题2.3.8. $F = F_X$ 满足三条性质:

(1) 单调上升性: 若$x \leqslant y$, 则$F(x) \leqslant F(y)$;

(2) 规范性: $\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to\infty} F(x) = 1$;

(3) 右连续性: 对任意$x \in \mathbb{R}$, $\lim\limits_{y\searrow x} F(y) = F(x)$.

- 概率的(1) 非负性; (2) 规范性.
- 单调函数总可以定义左极限与右极限:

$$g(x+) \lim_{y\searrow x} g(y), \quad g(x-) := \lim_{y\nearrow x} g(y).$$

- 概率的(3) 连续性: $\{X \leqslant x + 1/n\} \searrow \{X \leqslant x\}$.
  - $F(x-) = \mathrm{P}(X < x)$: $\{X \leqslant x - 1/n\} \nearrow \{X < x\}$.
  - 跳点: $\mathrm{P}(X = x) = F_X(x) - F_X(x-)$.

三、分布

- 引理2.3.9. (1) X 为随机变量$\Leftrightarrow \{X \in B\} \in \mathscr{F}, \forall B \in \mathscr{B}$.
  (2) f 为Borel 函数$\Leftrightarrow \{x : f(x) \in B\} \in \mathscr{B}, \forall B \in \mathscr{B}$.
- X 生成的σ 代数: $\sigma(X) = \{\{X \in B\} : B \in \mathscr{B}\}$,
  f 生成的σ 代数: $\sigma(f) = \{\{x : f(x) \in B\} : B \in \mathscr{B}\}$.
- (概率)分布: 称$(\mathbb{R}, \mathscr{B})$ 上的概率. 记为μ.
- X 的分布: $B \mapsto \mathrm{P}(X \in B)$. 记为$\mu_X$ 或$\mathscr{L}(X)$.
- 定义2.3.10. 若$F_X(x) = F_Y(x), \forall x \in \mathbb{R}$, 则称$X$ 与Y 同分布, 记为$X \overset{\mathrm{d}}{=} Y$.
- $X \overset{\mathrm{d}}{=} Y$ 当且仅当$\mathscr{L}(X) = \mathscr{L}(Y)$.
- 若$\mathrm{P}(X = Y) = 1$, 则称X 与Y 几乎必然相等,
  记为$X = Y$ a.s..

例2.3.11 (示性函数).

$$\mathbf{I}_A(\omega) = \mathbf{I}_{\{\omega\in A\}} := \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A. \end{cases}$$

- $\mathbf{I}_A \sim B(1,p)$, $p = \mathrm{P}(A)$.
- 设$X \sim B(1,p)$.
    - 令$A = \{X = 1\}$, $B = \{X = 0\}$.
    - 令$C = (A \cup B)^c = \{\omega : X(\omega) \neq 0$ 或 $1\}$.
    - $\mathrm{P}(C) = 0$, $\{X \neq \mathbf{I}_A\} = C$. 因此$X = \mathbf{I}_A$ a.s..

- $X = Y$ a.s. $\Rightarrow X \overset{\mathrm{d}}{=} Y$:
    - $$\begin{aligned} \mathrm{P}(X \leqslant x) &= \mathrm{P}(X \leqslant x, X = Y) \\ &= \mathrm{P}(Y \leqslant x, X = Y) = \mathrm{P}(Y \leqslant x). \end{aligned}$$
- 反之不然.
    - 例, $X \sim N(0,1)$. $X \overset{\mathrm{d}}{=} -X$, 但$\mathrm{P}(X = -X) = 0$.
    - 例2.3.12. 抛硬币与掷色子.
- 若$X = Y$ a.s., 则可视为同一个随机变量.
- X 可只几乎必然有定义. (例, 几何分布.)
- $X \overset{\mathrm{d}}{=} Y$, 研究分布时可采用任意一个, 可称$Y$ 为X 的复制.

总结:

- 顺序:

$$\Omega \to \mathscr{F} \to \begin{array}{ccc} \nearrow & \mathrm{P} & \searrow \\ \searrow & X & \nearrow \end{array} \to \mu_X .$$

- 谈及随机变量X 时, 只需要$(\Omega, \mathscr{F})$, 不需要P.
  - 一般先默认P, 研究X 的分布.
  - 同一个X, 在不同的概率(如, P_A) 下, 服从不同的分布.
- 概率P 与随机变量X.

	含义	定义域	自变量	要求
P	权重	$\mathscr{F}$	事件A	三条
X	观测值	Ω	样本ω	$\{X \leqslant x\} \in \mathscr{F}$

抽象	Ω: 集合	$\mathscr{F}$: σ 代数	ω: 符号	P: 概率
具体	$\mathbb{R}$: 实轴	$\mathscr{B}$: 区间生成	x: 实数	μ_X: 分布

四、广义随机变量、可测映射与随机元*

- 可测映射$\xi : (\mathbb{X}, \mathscr{F}) \to (\mathbb{Y}, \mathscr{G})$.
  - $(\mathbb{X}, \mathscr{F})$ 和$(\mathbb{Y}, \mathscr{G})$ 是可测空间.
  - $\xi : \mathbb{X} \to \mathbb{Y}$ 满足可测性:

$$\{x \in \mathbb{X} : \xi(x) \in B\} \in \mathscr{F}, \quad \forall B \in \mathscr{G}.$$

- 随机变量$X : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$,
  Borel函数/可测函数$(\mathbb{Y}, \mathscr{G}) = (\mathbb{R}, \mathscr{B})$.
- 广义实数集与广义随机变量:
  - $\hat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, $\hat{\mathscr{B}} = \{B \cup S : B \in \mathscr{B}, S \subseteq \{-\infty, \infty\}\}$.
  - $X : (\Omega, \mathscr{F}) \to (\hat{\mathbb{R}}, \hat{\mathscr{B}})$.
- 随机元: $X : (\Omega, \mathscr{F}) \to (S, \mathscr{S})$.

补充知识：Lebesgue 积分

- 考虑Borel 函数/可测函数.
- f, g 几乎处处(almost everywhere)相等指

$$\lambda(\{x : f(x) \neq g(x)\}) = 0, \quad \text{记为 } f = g \text{ a.e..}$$

- $f \geqslant 0$ a.e., $f \geqslant g$ a.e., 同理.
- 设$f \geqslant 0$ a.e., 令

$$\int f(x)\mathrm{d}x := \lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{k}{n} \cdot \lambda\left(\left\{x : \frac{k}{n} \leqslant f(x) < \frac{k+1}{n}\right\}\right).$$

- $\int_B f(x)\mathrm{d}x := \int f(x)\mathbf{I}_{\{x\in B\}}\mathrm{d}x$, $B \in \mathscr{B}$.
- $f = g$ a.e. $\Leftrightarrow \int_B f(x)\mathrm{d}x = \int_B g(x)\mathrm{d}x$, $\forall B \in \mathscr{B}$.

- 例. X 的密度函数指一族几乎处处相等的非负Borel 函数, 它们都满足:

$$\mathrm{P}(X \in B) = \int_B p(x)\mathrm{d}x, \quad \forall B \in \mathscr{B}.$$

- 一般情形, $a \in \mathbb{R}$,

$$a^+ := a \vee 0,\ a^- := -(a \wedge 0),\ a = a^+ - a^-,\ |a| = a^+ + a^-.$$

- $f = f^+ - f^-$, (设下式右边不是$\infty - \infty$)

$$\int f(x)\mathrm{d}x := \int f^+(x)\mathrm{d}x - \int f^-(x)\mathrm{d}x,$$

- 积分有线性, 高维类似.

§2.4 随机向量

- 随机向量: 同一个$(\Omega,\mathscr{F})$ 中的多个随机变量(一起考虑).
- n 维随机向量: $\vec{X}=(X_1,\cdots,X_n)$

$$\vec{X}:\Omega\to\mathbb{R}^n,\ \ \omega\mapsto(X_1(\omega),\cdots,X_n(\omega)).$$

- $\{\vec{X}\leqslant\vec{x}\}$:

$$\begin{aligned}&\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\}\\ =&\{\vec{X}\in D\},\quad D=(-\infty,x_1]\times\cdots\times(-\infty,x_n].\end{aligned}$$

- $\sigma(\vec{X})=\big\{\{\vec{X}\in B\},\ \forall B\in\mathscr{B}^n\big\}\subseteq\mathscr{F}.$
- 研究$\vec{X}$ 的联合分布

$$\mu_{\vec{X}},\ \mathscr{L}(\vec{X}):\mathscr{B}^n\to\mathbb{R},\ B\mapsto \mathrm{P}(\vec{X}\in B).$$

一、离散型随机向量

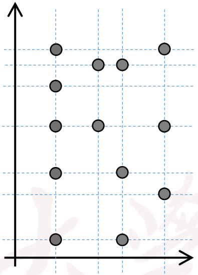
- 二维. X 与Y 都是离散型. 联合分布(列):

$$\mathrm{P}(X=x_i, Y=y_j)=p_{ij}, \quad i\in I,\ j\in J.$$

- $X_1,\cdots,X_n$ 都是离散型. 联合分布列:

$$\mathrm{P}\left(X_1=x_{i_1}^{(1)},\cdots,X_n=x_{i_n}^{(n)}\right),$$

$$i_1\in I_1,\cdots,i_n\in I_n.$$



- 例2.4.2 (多项分布). 某骰子有k 个面, 第i 个面的概率为p_i. 掷n 次. 第i 个面X_i 次. $(X_1,\cdots,X_{k-1})$ 的联合分布列.
- 例2.4.3. (多项超几何分布). k 种等级的产品共N 个, 第i 种N_i 个. 任取n 个, 第i 种取到X_i 个. $(X_1,\cdots,X_{k-1})$ 的联合分布列.

例2.4.1. $X=$ 首次成功的时刻, $Y=$ 前n 次中的成功次数. 求(X,Y) 的联合分布列.

- $X\sim G(p)$, $Y\sim B(n,p)$. 下面求p_{ij}.
- $i\leqslant n$:

$$p_{ij}=C_{n-i}^{j-1}p^j(1-p)^{n-j},\quad i=1,\cdots,n;\ j=1,\cdots,n-i+1.$$

- $i\geqslant n+1$:

$$p_{i0}=(1-p)^{i-1}p,\quad i=n+1,n+2,\cdots$$

二、连续型随机向量

- 二维. 联合(概率)密度函数:
  $p(\cdot,\cdot):\mathbb{R}^2\to\mathbb{R}$, 非负、可测, 且 $\iint_{\mathbb{R}^2} p(x,y)\mathrm{d}x\mathrm{d}y=1.$
- 联合分布函数: $F_{X,Y}(x,y)=\mathrm{P}(X\leqslant x, Y\leqslant y)$.
- 记$D_{x,y}=(-\infty,x]\times(-\infty,y]$. 若

$$F_{X,Y}(x,y)=\iint_{D_{x,y}} p(u,v)\mathrm{d}u\mathrm{d}v,\quad \forall x,y\in\mathbb{R},$$

  则称(X,Y) 为连续型, 称$p(\cdot,\cdot)$ 为(X,Y) 的联合密度函数, 记为$p_{X,Y}(\cdot,\cdot)$.
- 若为$p(\cdot,\cdot)$ 的连续点, 则

$$p(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}=\lim_{\mathrm{Vol}_2(D)\to 0}\frac{\mathrm{P}\big((X,Y)\in D\big)}{\mathrm{Vol}_2(D)}.$$

- $\forall D \in \mathscr{B}^2$,

$$P\big((X,Y) \in D\big) = \iint\limits_D p_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

- 若(X,Y) 为连续型, 则$P(X = Y) = 0$.
- n 维类似.
- 均匀分布$(X,Y) \sim U(S)$: $P\big((X,Y) \in D\big)$ 正比于D 的体积.
  - 例2.4.7. $S = \{(x,y) \in \mathbb{R}^2 : x,y \geqslant 0 \text{ 且 } x+y \leqslant 1\}$.

$$P(X \leqslant Y) = \frac{1/4}{1/2} = \frac{1}{2}.$$

  - 例. $U \sim U(0,2\pi)$, $(\cos\Theta, \sin\Theta) \sim U(S^1)$.

二维正态分布, $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

- 参数范围: $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$; $\rho \in (-1, 1)$.
- 联合密度函数形如

$$p(x, y) = C \exp\left\{-\frac{1}{2(1-\rho^2)} \cdot I\right\},$$

- 其中 $I = u^2 - 2\rho uv + v^2$, $\quad u = \dfrac{x - \mu_1}{\sigma_1}$, $v = \dfrac{y - \mu_2}{\sigma_2}$.
- 例2.4.6. $C = \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$:

$$\int_{-\infty}^{\infty} p(x, y)\mathrm{d}y = C \int_{-\infty}^{\infty} \exp\left\{-\frac{(v-\rho u)^2 + (1-\rho^2)u^2}{2(1-\rho^2)}\right\} \sigma_2 \mathrm{d}v$$

$$= C\sigma_2 \mathrm{e}^{-\frac{u^2}{2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(v-\rho u)^2}{2(1-\rho^2)}\right\} \mathrm{d}v$$

$$= C\sigma_2 \mathrm{e}^{-\frac{u^2}{2}} \cdot \sqrt{2\pi(1-\rho^2)} = \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathrm{e}^{-\frac{u^2}{2}}.$$

- 二元标准正态分布: $\mu_1=\mu_2=0,\ \sigma_1=\sigma_2=1;\ \rho=0.$

$$q(x,y)=\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}.$$

- 一般情形, $u=\frac{x-\mu_1}{\sigma_1}$, $v=\frac{y-\mu_2}{\sigma_2}$,

$$\begin{aligned} I =& u^2-2\rho uv+v^2 \\ =&(v-\rho u)^2+(\sqrt{1-\rho^2}u)^2. \end{aligned}$$

- 于是

$$\begin{aligned} p(x,y) =& C\exp\left\{-\frac{1}{2(1-\rho^2)}\cdot I\right\} \\ =&\tilde{C}\cdot q\left(\frac{v-\rho u}{\sqrt{1-\rho^2}},\ u\right) \end{aligned}$$

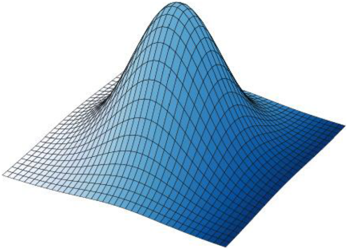


图 **2.3** 二维正态分布的联合密度函数的图像

三、一般形式的随机向量

- $\vec{X}=(X_1,\cdots,X_n)$. $\vec{X}$ 的联合分布:

$$\mu_{\vec{X}}(B):=\mathrm{P}(\vec{X}\in B),\quad \forall B\in\mathscr{B}^n.$$

- 联合分布函数$F_{\vec{X}}(\cdot)=\mathrm{P}(\vec{X}\leqslant\vec{x})$.
- $\vec{X}$ 与$\vec{Y}$ 同分布, $\vec{X}\overset{\mathrm{d}}{=}\vec{Y}$ 指: $\mu_{\vec{X}}=\mu_{\vec{Y}}$, $F_{\vec{X}}=F_{\vec{Y}}$.
- 离散型: 联合分布列相同.
  连续型: 联合密度函数几乎处处相等.
- 命题2.8.1. 设$\vec{X}$, $\vec{Y}$ 是n 维随机向量, $\vec{X}\overset{\mathrm{d}}{=}\vec{Y}$,
  $f:\mathbb{R}^n\to\mathbb{R}^m$ 是Borel 函数, 则$f(\vec{X})\overset{\mathrm{d}}{=}f(\vec{Y})$.
- $\vec{X}$ 与$\vec{Y}$ 几乎必然相等, $\vec{X}=\vec{Y}$ a.s. 指: $\mathrm{P}(\vec{X}=\vec{Y})=1$.
- $\vec{X}=\vec{Y}$ a.s. $\Rightarrow\vec{X}\overset{\mathrm{d}}{=}\vec{Y}$, 反之不然.

均匀分布$U(S)$.

- 例2.4.8. $U \sim U(0, 2\pi)$, $X = \cos U$, $Y = \sin U$.
    - $(X, Y) \sim U(S^1)$, 不是离散型也不是连续型.
- 例2.4.9. $S = \{(x, y, z) : x, y, z \geqslant 0 \text{ 且 } x + y + z = 1\}$.
    - $(X, Y, Z) \sim U(S) \Rightarrow (X, Y, Z)$ 等等都服从$U(S)$.
    - $(X, Y, Z) \sim U(S) \Rightarrow (X, Y), (X, Z), (Y, Z) \sim U(\hat{S})$, 其中$\hat{S} = \{(x, y) : x, y \geqslant 0 \text{ 且 } x + y \leqslant 1\}$. $U(\hat{S})$ 为连续型.
    - $(X, Y) \sim U(\hat{S}) \Rightarrow (X, Y, Z) \sim U(S) \Rightarrow (X, Z) \sim U(S)$. (例2.8.2 & 例2.8.3(高维))
- 例2.4.10. $S = \{\vec{x} : x_1^2 + \cdots + x_n^2 = 1\} \subseteq \mathbb{R}^n$. 列向量.
    - $\mathbf{A}$ 为n 阶正交矩阵, $\vec{Y} = \mathbf{A}\vec{X}$.
    - $\vec{X} \sim U(S) \Rightarrow \vec{Y} \sim U(S)$.

例2.4.12 (随机图). 完全图K_n 中的每条边, 独立地以概率p 保留, 以概率$1-p$ 删除. 得到随机子图$G(n,p)$.

- 建模一、$\Omega =$ 完全图的所有子图. $|\Omega| = 2^{C_N^2}$.
    - 若样本ω 中有k 条边, 则令$p_\omega = p^k(1-p)^{C_n^2-k}$, 得到P.
    - $G(n,p)$ 指随机样本.
- 建模二、$(\Omega, \mathscr{F}, \mathrm{P})$ 是成功率为p 的伯努利试验.
    - 将顶点编号$1,2,\cdots,n$. $\forall i>j$, (i,j) 表示连接顶点i,j 的边.
    - 所有边排序: $(2,1),(3,1),(3,2),(4,1),(4,2),(4,3),\cdots$.
    - H_n 发生, 则保留第k 条边, 否则删除第k 条边. $X_k = \mathbf{I}_{H_n}$.
    - 样本ω 对应着K_n 的一个子图$G(n,p;\omega)$, 简记为$G(n,p)$.
    - $K_n \subseteq K_{n+1}$, $G(n,p) \subseteq G(n+1,p)$.

一、边缘分布

- $\vec{X} = (X_1, \cdots, X_n)$.
  称$\vec{Y} := (X_{i_1}, \cdots, X_{i_m})$ 为$\vec{X}$ 的m $(< n)$ 维边缘,
  称$\mu_{\vec{Y}}$, $F_{\vec{Y}}$ 为$\vec{X}$ 的m 维边缘(联合)分布/分布函数.
- 离散型, 边缘分布(列). 例, $\mathrm{P}(X = x_i) = \sum_{j \in J} p_{ij}$, $i \in I$.
- 例2.5.1. 多项分布与多项超几何分布.
- 连续型, 边缘密度函数. 例, $p_X(x) = \int_{-\infty}^{\infty} p(x, y) \mathrm{d}y$.
- 例2.5.2. 二维正态分布的边缘为正态分布.
- 例2.5.3. $\vec{X} \stackrel{\mathrm{d}}{=} \vec{Y} \Rightarrow (X_{i_1}, \cdots, X_{i_m}) \stackrel{\mathrm{d}}{=} (Y_{i_1}, \cdots, Y_{i_m})$.
  反之不然: $X = \mathbf{I}_{\{H\}}$, $Y = \mathbf{I}_{\{T\}}$, (X, X) 与(X, Y) 不同分布.

二、随机变量的独立性

- 若$\forall x_1, \cdots, x_n$,

$$\mathrm{P}(X_1 \leqslant x_1, \cdots, X_n \leqslant x_n) = \mathrm{P}(X_1 \leqslant x_1) \cdots \mathrm{P}(X_n \leqslant x_n),$$

则称$X_1, \cdots, X_n$ (相互)独立.

- ⋆⋆ 当且仅当$\forall B_1, \cdots, B_n \in \mathscr{B}$,

$$\mathrm{P}(X_1 \in B_1, \cdots, X_n \in B_n) = \mathrm{P}(X_1 \in B_1) \cdots \mathrm{P}(X_n \in B_n).$$

- ⋆⋆ 当且仅当⋆⋆, $\{X_1 \in B_1\}, \cdots, \{X_n \in B_n\}$ 相互独立.
- 若$\forall i \neq j$, X_i 与X_j 独立, 则称$X_1, \cdots, X_n$ 两两独立.
- 若$X_1, \cdots, X_n$ 独立且$X_i \overset{\mathrm{d}}{=} X_1$, $\forall i$, 则称它们独立同分布。

independent and identically distributed (i.i.d.).

离散型.

- $X_1, \cdots, X_n$ 独立, 则$\forall x_i \in D_i$ (X_i 的可能值),

$$\mathrm{P}(X_1 = x_1, \cdots, X_n = x_n) \stackrel{(*)}{=} \mathrm{P}(X_1 = x_1) \cdots \mathrm{P}(X_n = x_n).$$

- 反过来, 若$\star\star$ 形如$f_1(x_1) \cdots f_n(x_n)$, ($f_i \geqslant 0$), 则
  - $\mathrm{P}(X_i = x_i) = \frac{1}{C_i} f_i(x_i)$, $\forall x_i \in D_i$. $C_i = \sum_{x_i \in D_i} f_i(x_i)$.
  - $C_1 \cdots C_n = 1$, 故 $\stackrel{(*)}{=}$ 成立.
  - $\mathrm{P}(X_1 \in B_1, \cdots, X_n = B_n)$

$$\begin{aligned} &= \sum_{x_1 \in B_1, \cdots x_n \in B_n} \mathrm{P}(X_1 = x_1) \cdots \mathrm{P}(X_n = x_n) = \cdots \\ &= \mathrm{P}(X_1 \in B_1) \cdots \mathrm{P}(X_n = B_n). \end{aligned}$$

故$X_1, \cdots, X_n$ 独立.

例2.5.7. 考虑成功率为p 的Bernoulli 试验.

- 令$X_i = \mathbf{I}_{H_i}$. $X_1, \cdots, X_n$ 独立同分布, $\forall n$.
- 此时称$X_1, X_2, \cdots$ 独立同分布.
- 令$Y_i =$ 第$i-1$ 次成功与第i 次成功之间的失败次数.
- $\{Y_1 = k_1, Y_2 = k_2, \cdots, Y_n = k_n\}$:
  k_1 个T 接一个H; k_2 个T 接一个H; $\cdots$, k_n 个T 接一个H.
- 概率为$(1-p)^{k_1}p \cdot (1-p)^{k_2}p \cdots (1-p)^{k_n}p$.
- $Y_1, \cdots, Y_n$ 独立同分布, $Y_1 + 1 \sim G(p)$.

连续型.

- 设$X_1, \cdots, X_n$ 都是连续型随机变量. 若它们独立, 则

$$\mathrm{P}(X_1 \leqslant x_1, \cdots, X_n \leqslant x_n) = \mathrm{P}(X_1 \leqslant x_1) \cdots \mathrm{P}(X_n \leqslant x_n)$$

$$= \int_{-\infty}^{x_1} p_1(y_1) \mathrm{d}y_1 \cdots \int_{-\infty}^{x_n} p_n(y_n) \mathrm{d}y_n = \int_{\vec{y} \leqslant \vec{x}} p_1(y_1) \cdots p_n(y_n) \mathrm{d}y_1 \cdots \mathrm{d}y_n.$$

故$(X_1, \cdots, X_n)$ 是连续型, $\star\star$ 是联合密度函数.

- 反过来, 若$\star\star$ 形如$f_1(x_1) \cdots f_n(x_n)$, $(f_i \geqslant 0)$, 则
  - $p_i(x_i) = \dfrac{1}{C_i} f_i(x_i)$, $C_i = \displaystyle\int_{-\infty}^{\infty} f_i(x_i) \mathrm{d}x_i$.
  - $C_1 \cdots C_n = 1$, 故$\star\star = p_1(x_1) \cdots p_n(x_n)$.
  - $\star\star = \star\star_{\mathrm{RHS}} = \star\star_{\mathrm{LHS}} = \star\star$. 故$X_1, \cdots, X_n$ 独立.

例2.5.8 (二维正态分布).$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.
X 与Y 相互独立当且仅当$\rho = 0$.

- $u = \dfrac{x-\mu_1}{\sigma_1}$, $v = \dfrac{y-\mu_2}{\sigma_2}$, 联合密度函数:

$$p(x,y) := \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(u^2 - 2\rho uv + v^2\right)\right\}.$$

- $\Leftarrow$: 若$\rho = 0$, 联合密度如下, 故独立.

$$\frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \cdot e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}.$$

- $\Rightarrow$: $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 独立. 故$\star\star$ 为联合密度.
- $\star\star$ 与$\star\star$ 均为(X, Y) 的联合密度, 均连续, 故$\star\star \equiv \star\star$.
- 取$x = \mu_1$, $y = \mu_2$, 得$\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} = \dfrac{1}{2\pi\sigma_1\sigma_2}$. 从而$\rho = 0$.

例2.5.9. 设$X_1,\cdots,X_n$ 独立; $Y_1,\cdots,Y_n$ 独立; $X_i \stackrel{d}{=} Y_i$, $\forall i$.

- $\vec{X}=(X_1,\cdots,X_n)$, $\vec{Y}=(Y_1,\cdots,Y_n)$.
  $\vec{X}$ 与$\vec{Y}$ 的联合分布函数相等, 均为$F_1(z_1)\cdots F_n(z_n)$.
- 若$X_1,\cdots,X_n$ 独立同分布, 则对于$1,\cdots,n$ 的任意全排列$i_1,\cdots,i_n$,

$$(X_{i_1},\cdots,X_{i_n}) \stackrel{d}{=} (X_1,\cdots,X_n).$$

三、随机向量的独立性

- $\vec{X}$ 与$\vec{Y}$ 相互独立指

$$\mathrm{P}\left(\vec{X} \preceq \vec{x},\ \vec{Y} \preceq \vec{y}\right) = \mathrm{P}\left(\vec{X} \preceq \vec{x}\right) \mathrm{P}\left(\vec{Y} \preceq \vec{y}\right),\ \forall \vec{x}, \vec{y}.$$

- $\vec{X}^{(1)}, \cdots, \vec{X}^{(n)}$ 相互独立指

$$\mathrm{P}\left(\vec{X}^{(i)} \preceq \vec{x}^{(i)}, i = 1, \cdots, n\right) = \prod_{i=1}^{n} \mathrm{P}\left(\vec{X}^{(i)} \preceq \vec{x}^{(i)}\right), \quad \forall \vec{x}^{(1)}, \cdots, \vec{x}^{(n)}.$$

- ⋆⋆ 当且仅当对任意Borel 集B_i's, $\{\vec{X}^{(i)} \in B_i\}$'s 相互独立.
- 命题2.5.11. 若⋆⋆, 则∀ Borel 函数/可测映射, $f_i(\vec{X}^{(i)})$'s 独立.
- ⋆⋆ 不能推出X_1, X_2 相互独立.
- 独立同分布, 两两独立, 类似.

一、离散型随机向量的条件分布

- 在$\{X=x_i\}$ 发生的条件下, Y 的条件分布(列):

$$p_{Y|X}(y_j|x_i)=\mathrm{P}(Y=y_j|X=x_i)=\frac{\mathrm{P}(Y=y_j,X=x_i)}{\mathrm{P}(X=x_i)},\ j\in J.$$

- 计算: $p_{Y|X}(\cdot|x_i)\propto \mathrm{P}(Y=\cdot, X=x_i)$, (固定$x_i$).
- $\mathrm{P}(X=x_i,Y=y_j)=\mathrm{P}(X=x_i)\mathrm{P}(Y=y_j|X=x_i)$.
- $\mathrm{P}(Y=y_j)=\sum\limits_{i\in I}\mathrm{P}(X=x_i)p_{Y|X}(y_j|x_i)$.
- X 与Y 相互独立的
  - 必要条件: $\mathscr{L}(Y|X=x_i)=\mathscr{L}(Y), \forall i$.
  - 充分条件: $\mathscr{L}(Y|X=x_i)$ 不依赖于i.

例2.6.1. 某昆虫所产的卵的数目$X \sim P(\lambda)$. 每个卵独立地以概率p 孵化为幼虫. Y: 该昆虫的幼虫数. $Z = X - Y$, 则$Y \sim P(\lambda p)$, $Z \sim P(\lambda(1-p))$, 且Y 与Z 相互独立.

- 边缘分布: $X \sim P(\lambda)$.
- 条件分布: $\mathscr{L}(Y|X = n) = B(n, p)$.
- 联合分布: $\forall k, \ell = 0, 1, \cdots,$

$$
\begin{aligned}
& \mathrm{P}(Y = k, Z = \ell) = \mathrm{P}(X = k + \ell, Y = k) \\
= \quad & \frac{\lambda^{k+\ell}}{(k+\ell)!} \mathrm{e}^{-\lambda} \cdot \frac{(k+\ell)!}{k!\ell!} p^k q^\ell = \frac{(\lambda p)^k}{k!} \cdot \frac{(\lambda q)^\ell}{\ell!} \mathrm{e}^{-\lambda}.
\end{aligned}
$$

故结论成立. ($\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda p} \cdot \mathrm{e}^{-\lambda q}$.)

二、连续型随机向量的条件分布

- 条件分布函数:

$$\mathrm{P}(Y \leqslant y|X = x) := \lim_{\delta\to 0+} \mathrm{P}(Y \leqslant y|x-\delta < X \leqslant x+\delta)$$

$$= \lim_{\delta\to 0+} \frac{\int_{x-\delta}^{x+\delta}\int_{-\infty}^{y} p(u,v)dvdu}{\int_{x-\delta}^{x+\delta} p_X(u)du} = \int_{-\infty}^{y} \frac{p(x,v)}{p_X(x)}dv.$$

- 在$\{X = x\}$ 的条件下, Y 的条件密度函数:

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}.$$

- 计算: 固定x, $p_{Y|X}(\cdot|x) \propto p(x,\cdot)$.
- $p(x,y)$, $p_Y(y)$, 独立性的充分条件、必要条件与离散型类似.

例2.6.2 (二维正态分布) $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

$$p(x,y) = C\exp\left\{-\frac{1}{2(1-\rho^2)}\cdot I\right\},$$

其中, $C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$,

$$I = u^2 - 2\rho uv + v^2, \quad u = \frac{x-\mu_1}{\sigma_1}, \ v = \frac{y-\mu_2}{\sigma_2}.$$

- 结论: $X \sim N(\mu_1,\sigma_1^2)$, $\mathscr{L}(Y|X) = N(\mu,\sigma^2)$, 其中

$$\mu = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \quad \sigma^2 = (1-\rho^2)\sigma_2^2.$$

- 关键点: $I = (v-\rho u)^2 + (1-\rho^2)u^2$. $\mathrm{d}x = \sigma_1\mathrm{d}u$, $\mathrm{d}y = \sigma_2\mathrm{d}v$.

方法一、先求边缘密度, 再求条件密度.

- 例2.4.6 & 例2.5.3.

$$C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}, \quad X \sim N(\mu_1, \sigma_1^2).$$

- $p_{Y|X}(y|x) = \dfrac{p(x,y)}{p_X(x)} = \dfrac{C \exp\left\{-\frac{1}{2(1-\rho^2)} \cdot I\right\}}{\frac{1}{\sqrt{2\pi}\sigma_1}\exp\{-\frac{1}{2}u^2\}} = \cdots$
- $-\frac{1}{2(1-\rho^2)} \cdot I + \frac{1}{2}u^2$:

$$-\frac{(v-\rho u)^2 + (1-\rho^2)u^2}{2(1-\rho^2)} + \frac{u^2}{2} = -\frac{(v-\rho u)^2}{2(1-\rho^2)} = -\frac{(y-\mu)^2}{2\sigma^2}.$$

- 故$\mathscr{L}(Y|X) = N(\star, \star)$

方法二、先求条件密度, 再求边缘密度.

- $f_x(y) = \exp\left\{-\frac{(v-\rho u)^2}{2(1-\rho^2)}\right\}$, $C_2(x) = C\mathrm{e}^{-\frac{u^2}{2}}$.

$$\begin{aligned} p(x,y) &= C\exp\left\{-\frac{(v-\rho u)^2+(1-\rho^2)u^2}{2(1-\rho^2)}\right\} \\ &= C_2(x)f_x(y). \end{aligned}$$

- $\mu = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$, $\sigma^2 = (1-\rho^2)\sigma_2^2$, $C_3(x) = \frac{1}{\sqrt{2\pi\sigma^2}}$.

$$p_{Y|X}(y|x) = C_3(x)f_x(y) = C_3(x)\exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\},$$

- $p_X(x) = \frac{p(x,y)}{p_{Y|X}(y|x)} = \frac{C_2(x)}{C_3(x)} = C\sqrt{2\pi\sigma^2}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$.

  故$X \sim N(\mu_1, \sigma_2^2)$, $C = \frac{1}{\sqrt{2\pi\sigma_1^2}\sqrt{2\pi\sigma^2}} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$.

三、一般形式的条件分布

- 固定x. 令$\tilde{F}_{Y|X}(y|x) = \lim\limits_{\varepsilon\to 0+} \mathrm{P}(Y \leqslant y|x-\varepsilon < X \leqslant x+\varepsilon)$.
- $\tilde{F}_{Y|X}(y|x)$ 关于y 单调上升, $F_{Y|X}(y|x) := \tilde{F}_{Y|X}(y+|x)$.
- 例. $U \sim U(0,1)$, $X = Y = U$.
  - 直观: $\mathscr{L}(Y|X=x)$ 为x 对应的单点分布.
  - $\tilde{F}_{Y|X}(y|x) = \begin{cases} 0, & y < x; \\ 1/2, & y = x; \\ 1, & y > x. \end{cases}$ $F_{Y|X}(y|x) = \begin{cases} 0, & y < x; \\ 1, & y \geqslant x. \end{cases}$
- 例2.6.3. $(X,Y) \sim U(S^1)$.

$$\mathscr{L}(Y|X=x) \text{ 为 } \pm\sqrt{1-x^2} \text{ 上的等概率分布.}$$

例2.6.4. 在$\mathbb{R}^3$ 的单位球面S 上取四点, 求这四点共半球的概率.

- 设ξ_i, $i = 1, 2, 3, 4$ 独立同分布, $\sim U(S)$.
  $A = \{\xi_4 \in \triangle_{-\xi_1, -\xi_2, -\xi_3}\}$.
- 再独立地取W_1, W_2, W_3 独立同分布, 等可能取± 1.
- $\eta_i = W_i \xi_i$, 则$\eta_1, \eta_2, \eta_3, \xi_4$ 独立同分布, $\sim U(S)$.
- $B =$ "这四点共半球" $= \{\xi_4 \in \triangle_{-\eta_1, -\eta_2, -\eta_3}\}$.
- $\forall \alpha_1, \alpha_2, \alpha_3$, $\mathrm{P}(B | \xi_1 = \alpha_1, \xi_2 = \alpha_2, \xi_3 = \alpha_3) \equiv 1/8$.
- 故$\mathrm{P}(B) = 1/8$.   (B 与ξ_1, ξ_2, ξ_3 独立.)

§2.7 随机变量的函数

一、随机变量的函数

- X 为离散型, $f: x_i \to f(x_i)$. $Y = f(X)$.
  找出所有可能值y_j,
  $$\mathrm{P}(Y = y_j) = \sum_{i \in I: f(x_i) = y_j} p_i.$$
- 设X 是随机变量.
  一般地, 要求f 为Borel 函数, 以保证$f(X)$ 是随机变量.
- 命题2.7.1. 若$X \stackrel{\mathrm{d}}{=} Y$, 则$f(X) \stackrel{\mathrm{d}}{=} f(Y)$.

二、分布函数及其广义逆函数

- 若以下三条成立, 则称$F(\cdot)$ 为分布函数.

  (1) 单调上升性: 若$x \leqslant y$, 则$F(x) \leqslant F(y)$;

  (2) 规范性: $\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to\infty} F(x) = 1$;

  (3) 右连续性: 对任意$x \in \mathbb{R}$, $\lim\limits_{y\searrow x} F(y) = F(x)$.

- 例. $F_X(\cdot)$ 是分布函数.
- F 的广义反函数:

$$F^{-1}(u) := \inf\{x \in \mathbb{R} : F(x) \geqslant u\}, \quad 0 < u < 1.$$

- 分位数: $F^{-1}(u)$. 例, $X \sim N(0,1)$, $x_u = \Phi^{-1}(u)$, 则$\Phi(x_u) = u$.

引理2.7.2. F 与F^{-1} 都是Borel 函数, 并且

$$F^{-1}(u) \leqslant x \Leftrightarrow u \leqslant F(x), \quad x < F^{-1}(u) \Leftrightarrow F(x) < u.$$

- $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geqslant u\} \triangleq x_0.$
- 若$x > x_0$, 则$F(x) \geqslant u$;
  若$x < x_0$, 则$F(x) < u$.
- 若$x = x_0$, 则$F(x) \geqslant u$.
- F 是Borel 函数:

$$\{x : F(x) < u\} = (-\infty, F^{-1}(u)) \in \mathscr{B}.$$

- F^{-1} 是Borel 函数:

$$\{u : F^{-1}(u) \leqslant x\} = \big(0, F(x)\big].$$

例2.7.3. 令$U = F_X(X)$. F_X 为连续函数当且仅当$U \sim U(0,1)$.

- $\Rightarrow$: $F = F_X$.

$$F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geqslant u\} \triangleq x_0.$$

$$F^{-1}(u) \leqslant x \Leftrightarrow u \leqslant F(x), \quad x < F^{-1}(u) \Leftrightarrow F(x) < u.$$

- $\mathrm{P}(F(X) < u) = \mathrm{P}(X < x_0) = F(x_0)$.
- $F(x_0) \geqslant u$; $F(x_0-) \leqslant u$. 因此$F(x_0) = u$.
- $\mathrm{P}(U < u) = u \Rightarrow U \sim U(0,1)$:

$$\mathrm{P}(U = u) \leqslant \mathrm{P}(u \leqslant U < u + \frac{1}{n}) = (u + \frac{1}{n}) - u \to 0.$$

- $\Leftarrow$: 反证法. 取x_0 使得$a = \mathrm{P}(X < x_0) < b = F_X(x_0)$.
- $\mathrm{P}(U \in (a,b)) = 0$. 因此$U$ 不服从$U(0,1)$.

例2.7.4. 设F 是分布函数, $U \sim U(0,1)$.
令$X = F^{-1}(U)$, 则$F_X = F$.

- $$F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geqslant u\} \triangleq x_0.$$
  $$F^{-1}(u) \leqslant x \Leftrightarrow u \leqslant F(x), \quad x < F^{-1}(u) \Leftrightarrow F(x) < u.$$
- $\forall x$,
  $$F_X(x) = \mathrm{P}(X \leqslant x) = \mathrm{P}(U \leqslant F(x)) = F(x).$$
- 注: 任意分布函数都是某随机变量的分布函数.

- 若$1-G$ 是分布函数, 则称G 为尾分布函数.
- 广义反函数:

$$G^{-1}(u)=\inf\{x:G(x)\leqslant u\}=F^{-1}(1-u).$$

- 推论2.7.5. 若G 为尾分布函数, $U\sim U(0,1)$, $X=G^{-1}(U)$, 则$G_X=G$.
- 例2.7.3 的推论: F_X 为连续函数当且仅当$G_X(X)\sim U(0,1)$.
- 例. $X\sim \mathrm{Exp}(\lambda)$, $G(x)=\mathrm{e}^{-\lambda x}$, $x>0$.

$$\mathrm{e}^{-\lambda X}\sim U(0,1),\quad -\frac{1}{\lambda}\ln U\sim \mathrm{Exp}(\lambda).$$

- F_1, F_2 是分布函数.
- 例2.7.7. $F(x) = \min\{F_1(x), F_2(x)\}$ 是分布函数.
  - $X = \max\{F_1^{-1}(U), F_2^{-1}(U)\}$. $\forall x \in \mathbb{R}$,

$$F_X(x) = \mathrm{P}(U \leqslant F_1(x), U \leqslant F_2(x)) = F(x).$$

- 例2.7.8. $F(x) = pF_1(x) + (1-p)F_2(x)$ 是分布函数.
  - 取U, W 独立同分布, 令

$$Z = 1_{\{W \leqslant p\}} F_1^{-1}(U) + 1_{\{W > p\}} F_2^{-1}(U).$$

  - 则

$$\begin{aligned} F_Z(z) &= \mathrm{P}\left(W \leqslant p, F_1^{-1}(U) \leqslant z\right) + \mathrm{P}\left(W > p, F_2^{-1}(U) \leqslant z\right) \\ &= pF_1(z) + qF_2(z) = F(z). \end{aligned}$$

三、独立和

例2.7.10. 设X, Y 相互独立, 都取整数值, 分布列为$\mathrm{P}(X=n)=p_n$, $\mathrm{P}(Y=n)=q_n$, $n\in\mathbb{Z}$. 求$X+Y$ 的分布列.

- $\forall n$,

$$\mathrm{P}(X+Y=n)=\sum_{k\in\mathbb{Z}}\mathrm{P}(X=k,Y=n-k)=\sum_{k\in\mathbb{Z}}p_k q_{n-k}=\sum_{\ell\in\mathbb{Z}}p_{n-\ell}q_\ell.$$

- 序列$\{c_n\}$ 为$\{a_n\}$ 与$\{b_n\}$ 的卷积:

$$c_n=\sum_k a_k b_{n-k}=\sum_\ell a_{n-\ell}b_\ell.$$

例2.7.11. 设X, Y 均为连续型, 相互独立, 求$X+Y$ 的密度.

- 联合密度: $p_{X,Y}(x,y) = p_X(x)p_Y(y)$.
- 记$D_z = \{(x,y) \in \mathbb{R}^2 : x+y \leqslant z\}$. 则

$$\mathrm{P}(X+Y \leqslant z) = \iint_{D_z} p_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y$$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z-x} p_Y(y)\mathrm{d}y\right)p_X(x)\mathrm{d}x = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z} p_Y(y-x)\mathrm{d}y\right)p_X(x)\mathrm{d}x.$$

$$= \int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} p_X(x)p_Y(y-x)\mathrm{d}x\right)\mathrm{d}y.$$

- $\forall z$,

$$p_{X+Y}(z) = \int_{-\infty}^{\infty} p_X(x)p_Y(z-x)\mathrm{d}x = \int_{-\infty}^{\infty} p_X(z-y)p_Y(y)\mathrm{d}y.$$

- 函数h 为f 与g 的卷积, 记作$h = f * g$:

$$h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)\mathrm{d}x = \int_{-\infty}^{\infty} f(z-y)g(y)\mathrm{d}y.$$

- 设μ, ν 是分布. μ 与ν 的卷积指$\mathscr{L}(X+Y)$, 记为$\mu * \nu$, 其中

$$X \sim \mu, \quad Y \sim \nu, \quad X\text{与}Y\text{相互独立}.$$

- 设$\{\mu_\alpha : \alpha \in I\}$ 是一族分布. 若$\forall \alpha, \beta \in I$, 存在$\gamma \in I$ 使得

$$\mu_\alpha * \mu_\beta = \mu_\gamma,$$

则称该分布族具有可加性或再生性.

成功率为p 的Bernoulli 试验.

- 例2.7.12. $B(n,p) * B(m,p) = B(n+m,p)$.
  - $X_i = \mathbf{I}_{H_i}$, $X_1, \cdots, X_{n+m}$ 独立同分布.
  - $S_n := X_1 + \cdots + X_n \sim B(n,p)$,
    $\hat{S}_m := X_{n+1} + \cdots + X_{n+m} \sim B(m,p)$, 且它们相互独立.
  - $S_n + \hat{S}_m = X_1 + \cdots + X_{n+m} \sim B(n+m,p)$.
- 例2.7.13. $NB(r,p) * NB(s,p) = NB(r+s,p)$,
  $$P(r,p) * P(s,p) = P(r+s,p).$$
  - Y_i: 第$i-1$ 个H与第i 个H 之间T 的数目.
  - $Y_1, \cdots, Y_{r+s}$ 独立同分布.

例2.7.14. $P(\lambda_1) * P(\lambda_2) = P(\lambda_1 + \lambda_2)$.

- 设$X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$, 且X 与Y 相互独立, 则

$$\begin{aligned}\mathrm{P}(X+Y=n) &= \sum_{k=0}^{n} \mathrm{P}(X=k, Y=n-k) \\ &= \sum_{k=0}^{n} \frac{\lambda_1^k}{k!}\mathrm{e}^{-\lambda_1} \frac{\lambda_2^{n-k}}{(n-k)!}\mathrm{e}^{-\lambda_2} = \mathrm{e}^{-(\lambda_1+\lambda_2)} \frac{1}{n!} \sum_{k=0}^{n} C_n^k \lambda_1^k \lambda_2^{n-k} \\ &= \mathrm{e}^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!}, \quad n = 0, 1, 2, \cdots\end{aligned}$$

- 令$\lambda = \lambda_1 + \lambda_2$, $p = \lambda_1/\lambda$.
  - 例2.6.1. $W \sim P(\lambda p)$ (幼虫数), $Z \sim P(\lambda(1-p))$, 相互独立.
  - $(X, Y) \overset{\mathrm{d}}{=} (W, Z) \Rightarrow X + Y \overset{\mathrm{d}}{=} W + Z \sim P(\lambda)$.

例2.7.15. $\Gamma(\alpha_1,\lambda)*\Gamma(\alpha_2,\lambda)=\Gamma(\alpha_1+\alpha_2,\lambda)$.

- 密度函数:

$$p_\alpha(x)=C_\alpha x^{\alpha-1}\mathrm{e}^{-\lambda x},\ x>0.\quad \left(C_\alpha=\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)$$

- ⋆⋆ 对应的密度: $\forall z>0$,

$$\begin{aligned}
p(z) &= \int_{-\infty}^{\infty} p_{\alpha_1}(x)p_{\alpha_2}(z-x)\mathrm{d}x \\
&= C_{\alpha_1}C_{\alpha_2}\int_0^z x^{\alpha_1-1}\mathrm{e}^{-\lambda x}\cdot(z-x)^{\alpha_2-1}\mathrm{e}^{-\lambda(z-x)}\mathrm{d}x \\
&= C\mathrm{e}^{-\lambda z}\int_0^z x^{\alpha_1-1}(z-x)^{\alpha_2-1}\mathrm{d}x \\
&= C\mathrm{e}^{-\lambda z}\cdot\int_0^1 (tz)^{\alpha_1-1}\big((1-t)z\big)^{\alpha_2-1}z\mathrm{d}t\quad (x=zt) \\
&= \underline{C\mathrm{e}^{-\lambda z}\cdot z^{\alpha_1+\alpha_2-1}}\hat{C}.
\end{aligned}$$

设$X_1, X_2, \cdots$ 独立同分布, $S_n = X_1 + \cdots + X_n$.

$$\mathscr{L}(S_n) * \mathscr{L}(S_m) = \mathscr{L}(S_{n+m}).$$

- 例2.7.16. $X_1 \sim \text{Exp}(\lambda) \sim \Gamma(1, \lambda)$. $S_n \sim \Gamma(n, \lambda)$.
- 例2.7.17. $Z \sim N(0,1)$, $X_1 = Z^2 \sim \Gamma(1/2, 1/2)$,

$$S_n \sim \Gamma(n/2, 1/2) = \chi^2(n), \quad Z_1^2 + Z_2^2 \sim \text{Exp}(1/2).$$

- 例2.7.12. $X_1 \sim B(1,p)$, $S_n \sim B(n,p)$.
- 例2.7.13. $X_1 \sim G(p)$, $S_n \sim P(n,p)$, $S_n - n \sim NB(n,p)$.
- 例2.7.14. $X_1 \sim P(1)$, $S_n \sim P(n)$.
- 例. $X_1 \sim N(0,1)$, $S_n \sim N(0,n)$. 一般地, §2.7 习题2.

$$N(\mu_1, \sigma_1^2) * N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

一、随机向量的变换

- $\vec{X}$ 是n 维连续型随机向量，取值于D_1.
- $f: D_1 \to D_2$ 足够光滑，一对一.
- $I(\vec{x}) = \det(\partial y_i/\partial x_j)_{i,j\leqslant n}$, $J(\vec{y}) = \det(\partial x_i/\partial y_j)_{i,j\leqslant n}$.
- $\forall \vec{y} \in D_2$,

$$p_{\vec{Y}}(\vec{y}) = p_{\vec{X}}(\vec{x}) \cdot \frac{1}{|I(\vec{x})|} = p_{\vec{X}}(f^{-1}(\vec{y})) \cdot |J(\vec{y})|,$$

- f 多对一，$g_i(y) = x_i$, $i \in I$,

$$\sum_{i\in I} p_X(\vec{x}^{(i)}) \cdot \frac{1}{|I(\vec{x}^{(i)})|} = \sum_{i\in I} p_X(g_i(\vec{y})) \cdot |J_i(\vec{y})|.$$

二、正态情形

例2.8.4. 设X, Y 相互独立, 都服从$N(0,1)$. 令

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix},$$

则Z_1, Z_2 相互独立, 都服从$N(0,1)$.

- $f:(x,y)\to(z_1,z_2)$, 一对一, Jacobi 行列式为1.
- $p_{Z_1,Z_2}(z_1,z_2) = p_{X,Y}(x,y) = \dfrac{1}{2\pi}\mathrm{e}^{-(x^2+y^2)/2} = \dfrac{1}{2\pi}\mathrm{e}^{-(z_1^2+z_2^2)/2}$.
- $\mathbf{A}$ 为正交矩阵即可.
- 一般地, 若(X,Y) 服从二维正态分布, $\mathbf{A}$ 非退化, 则(Z_1,Z_2) 服从二维正态分布.

例. $(X,Y)\sim N(0,0,\sigma_1^2,\sigma_2^2,\rho)$, 研究$(Z_1,Z_2)^T=\mathbf{A}(X,Y)^T$.

- $p_{Z_1,Z_2}(z_1,z_2)=p_{X,Y}(x,y)$.
- $p_{X,Y}(x,y)=C\exp\{-\frac{1}{2}\cdot I\}$,

$$I=(x,y)\mathbf{\Sigma}^{-1}(x,y)^T=\frac{1}{(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-2\rho\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right).$$

- I 为z_1,z_2 的二次型, 且
  $\hat\sigma_{12}=\rho\sigma_1\sigma_2(\cos^2\alpha-\sin^2\alpha)-(\sigma_1^2-\sigma_2^2)\cos\alpha\sin\alpha.$
- 取α 使得$\hat\sigma_{12}=0$:

$$\begin{cases}\alpha=\pi/4, & 若\sigma_1^2=\sigma_2^2;\\ \tan(2\alpha)=2\rho\sigma_1\sigma_2/(\sigma_1^2-\sigma_2^2), & 若\sigma_1^2\neq\sigma_2^2.\end{cases}$$

- 此时 Z_1,Z_2 相互独立, $Z_1\sim N(0,\hat\sigma_1^2)$, $Z_2\sim N(0,\hat\sigma_2^2)$.

例2.8.5. 极坐标: $X = R\cos\Theta$, $Y = R\sin\Theta$, 则$\Theta \sim U(0, 2\pi)$.

- $Z_1 = R\cos\hat{\Theta}_\alpha$, $Z_2 = R\sin\hat{\Theta}_\alpha$,
  其中$\hat{\Theta}_a = \Theta + \alpha \pmod{2\pi}$.
- $(Z_1, Z_2) \stackrel{d}{=} (X, Y) \Rightarrow \hat{\Theta}_a \stackrel{d}{=} \Theta$, $\forall\alpha$.
- $\Theta \in U(0, 2\pi)$, 即$(X, Y)/R \sim U(S^1)$.
- 推论: $X/Y = \tan\Theta \sim C(0, 1)$.

例2.8.5. Θ, R 相互独立.

- $r > 0$, $\theta \in (0, 2\pi)$

$$p_{R,\Theta}(r,\theta) = \frac{1}{2\pi}\mathrm{e}^{-(x^2+y^2)/2} \cdot r = \frac{1}{2\pi} r\mathrm{e}^{-r^2/2}.$$

- 结论1: Θ, R 相互独立. (本例)
- 结论2: $\Theta \sim U(0, 2\pi)$. (例2.8.5)
- 结论3: $W = R^2 = X^2 + Y^2 \sim \mathrm{Exp}(1/2)$ (例2.7.17): $\forall w > 0$,

$$p_W(w) = p_R(r)\frac{dr}{dw} = r\mathrm{e}^{-r^2/2} \cdot \frac{1}{2r} = \frac{1}{2}\mathrm{e}^{-\frac{w}{2}}.$$

高维情形:

- 设$X_1, \cdots, X_n$ 相互独立, 都服从$N(0,1)$. 联合密度:

$$p_{\vec{X}}(\vec{x}) = \frac{1}{\sqrt{2\pi}^n} \exp\left\{-\frac{r^2}{2}\right\}, \quad r^2 = x_1^2 + \cdots + x_n^2.$$

- $R = \|\vec{X}\|$, $R^2 \sim \chi^2(n)$.
- $\Theta = \vec{X}/R \sim U(S^{n-1})$:
  - 做正交变换$\vec{Z} = \mathbf{A}\vec{X}$, 则$\|\vec{Z}\| = \|\vec{X}\| = R$.
  - $p_{\vec{Z}}(\vec{z}) = p_{\vec{X}}(\vec{x})$, 故$\vec{Z} \stackrel{d}{=} \vec{X}$.
  - $\hat{\Theta} = \vec{Z}/R$, 则$(R, \Theta) \stackrel{d}{=} (R, \hat{\Theta}) \Rightarrow \Theta \stackrel{d}{=} \hat{\Theta}$, $\forall \mathbf{A}$.
- Θ 与R 相互独立.
  - $|d\vec{z}| = r^{n-1} dr |d\theta|$,

$$p_{R,\Theta}(r, \theta) = C r^{n-1} \exp\left\{-\frac{r^2}{2}\right\}, \quad r > 0, \ \theta \in S^{n-1}.$$

例2.8.7. 设U_1, U_2 相互独立, 都服从$U(0,1)$. 令

$$Z_1 = \sqrt{-2\ln U_1}\cos(2\pi U_2), \quad Z_1 = \sqrt{-2\ln U_1}\sin(2\pi U_2),$$

则Z_1, Z_2 相互独立, 都服从$N(0,1)$.

- $-2\ln U_1 \sim \mathrm{Exp}(1/2)$, $\hat{R} = \sqrt{-2\ln U_1} \stackrel{\mathrm{d}}{=} R$.
- $\hat{\Theta} := 2\pi U_2 \stackrel{\mathrm{d}}{=} \Theta$.
- 由独立性知$(\hat{R}, \hat{\Theta}) \stackrel{\mathrm{d}}{=} (R, \Theta)$.
- $Z_1 = \hat{R}\cos\hat{\Theta}$, $Z_2 = \hat{R}\sin\hat{\Theta}$.

三、最大值、最小值

例2.8.8. 设$X_1, \cdots, X_n$ 相互独立. 令

$$W := \max\{X_1, \cdots, X_n\}, \quad V := \min\{X_1, \cdots, X_n\}.$$

- 记$F_i = F_{X_i}$, $G_i = 1 - F_i$.
- $F_W(w) = \mathrm{P}(W \leqslant w)$:

$$\mathrm{P}(X_1 \leqslant w, \cdots, X_n \leqslant w) = \prod_{i=1}^{n} \mathrm{P}(X_i \leqslant w) = \prod_{i=1}^{n} F_i(w).$$

- $G_V(v) = \mathrm{P}(V > v)$:

$$\mathrm{P}(X_1 > v, \cdots, X_n > v) = \prod_{i=1}^{n} \mathrm{P}(X_i > v) = \prod_{i=1}^{n} G_i(v).$$

- 若都有密度, 则

$$p_W(w) = \sum_{i=1}^{n} p_i(x) \prod_{1 \leqslant j \leqslant n, j \neq i} F_j(w), \quad p_V(v) = \sum_{i=1}^{n} p_i(x) \prod_{1 \leqslant j \leqslant n, j \neq i} G_j(v).$$

例2.8.9. 设$X_1, \cdots, X_n$ 相互独立, $X_i \sim \mathrm{Exp}(\lambda_i)$,

$$Y = \min\{X_1, \cdots, X_n\} = X_W.$$

记$\lambda = \lambda_1 + \cdots + \lambda_n$, 则

$$Y \sim \mathrm{Exp}(\lambda), \quad \mathrm{P}(W = i) = \lambda_i/\lambda, \quad Y \text{ 与} W \text{ 相互独立.}$$

- $\forall y > 0, \forall i, \{Y > y, W = i\} = \{X_i > y; X_j > X_i, \forall j \neq i\}$
- 概率为:

$$\int_y^{\infty} \left(\int_{x_j' s > x_i} \lambda_1 \cdots \lambda_n \mathrm{e}^{-(\lambda_1 x_1 + \cdots + \lambda_n x_n)} \mathrm{d}x_j' s \right) \mathrm{d}x_i$$

$$= \int_y^{\infty} \lambda_i \mathrm{e}^{-\lambda x_i} \mathrm{e}^{-(\lambda - \lambda_i) x_i} \mathrm{d}x_i = \frac{\lambda_i}{\lambda} \mathrm{e}^{-\lambda y}.$$

四、顺序统计量

- 设$X_1, \cdots, X_n$ 独立同分布.
- $\forall\omega$, 从小到大排序: $X_{(1)}(\omega) \leqslant \cdots \leqslant X_{(n)}(\omega)$.
- 称$X_{(1)}, \cdots, X_{(n)}$ 为$X_1, \cdots, X_n$ 的顺序统计量.
- 例. X_1 为离散型, $\mathrm{P}(X_{(k)} = x, \forall k) = \mathrm{P}(X_1 = x)^n$.
- 下设X_1 为连续型, 此时

$$X_{(1)}(\omega) < \cdots < X_{(n)}(\omega) \quad \text{a.s.}.$$

- X_1 的密度/分布/尾分布函数分别记作$p(\cdot)$, $F(\cdot)$, $G(\cdot)$.

例2.8.10. $X_{(k)}$ 的密度函数为$nC_{n-1}^{k-1}F(x)^{k-1}G(x)^{n-k}p(x)$.

- $\forall x$,

$$\mathrm{P}(X_{(k)} \leqslant x) = \sum_{i=k}^{n} C_n^i F(x)^i \left(1 - F(x)\right)^{n-i}.$$

- 记$f(y) = \sum_{i=k}^{n} C_n^i y^i (1-y)^{n-i}$, 则$F_{(k)}(x) = f(F(x))$.
- $p_{(i)}(x) = f'(F(x))p(x)$,

$$f'(y) = \cdots = nC_{n-1}^{k-1} y^{k-1} (1-y)^{n-k}.$$

例2.8.11. $(X_{(1)},\cdots,X_{(n)})$ 的联合密度函数为

$$n!p(x_1)\cdots p(x_n),\quad x_1<\cdots<x_n.$$

- $D=\{\vec{x}\in\mathbb{R}^n: x_i\neq x_j,\forall i\neq j\}$,
  $\hat{D}=\{\vec{x}\in\mathbb{R}^n: x_1<\cdots<x_n\}$.
- $f: D\to\hat{D}$, $n!$ 到1 的映射, Jacobi行列式为± 1.
- $\forall\vec{x}\in\hat{D}$, $n!$ 个原像的联合密度均为$p(x_1)\cdots p(x_n)$.
- 进一步, 设$X_{(k)}=X_{I_k}$, $k=1,\cdots,n$, 则
  $\vec{I}=(I_1,\cdots,I_n)\sim$ 全排中的均匀分布, 且与★★ 相互独立.

$$\begin{aligned}
& \mathrm{P}\Big(x_r<X_{(r)}\leqslant x_r+\Delta x_r,\forall r;\ \vec{I}=(i_1,\cdots,i_n)\Big)\\
= & \mathrm{P}(x_r<X_{i_r}\leqslant x_r+\Delta x_r,\forall r)\\
= & \mathrm{P}\Big(x_r<X_{(r)}\leqslant x_r+\Delta x_r,\forall r\Big)\cdot\frac{1}{n!}.
\end{aligned}$$

例2.8.12. 设$U_1, \cdots, U_n$ 独立同分布, $\sim U(0,1)$. 记

$$X_i = U_{(i)} - U_{(i-1)}, \quad (U_{(0)} = 0,\ U_{(n+1)} = 1).$$

- $(U_{(1)}, \cdots, U_{(n)})$ 的联合密度: $n! \cdot \mathbf{I}_{\{0<u_1<u_2<\cdots<u_n<1\}}$.
- $(X_1, \cdots, X_n)$ 的联合密度: $n! \cdot \mathbf{I}_{\{x_1\cdots,x_n>0;\ x_1+\cdots+x_n<1\}}$.
- $(X_1, \cdots, X_n) \sim U(S)$, 其中

$$S = \left\{(x_1, \cdots, x_n) \in \mathbb{R}^n : x_i > 0, \forall i \text{ 且 } \sum_{i=1}^n x_i < 1\right\}.$$

- $(X_1, \cdots, X_{n+1}) \sim U(D)$, 其中

$$D := \left\{\vec{x} \in \mathbb{R}^{n+1} : x_i > 0, \forall i \text{ 且 } \sum_{i=1}^{n+1} x_i = 1\right\}.$$

- $\forall$ 全排列, $(X_{i_1}, \cdots, X_{i_{n+1}}) \sim U(D)$. 可交换.

补充

- $f:\mathbb{R}^n\to\mathbb{R}^m$: 以$\mathbb{R}^2\to\mathbb{R}^1$ 为例, $W=f(X,Y)$.
- 方法一、补变量法:

  找g, 使得$(x,y)\mapsto\big(f(x,y),g(x,y)\big)$ 是一对一的.

  (1) 令$V=g(X,Y)$, 求联合密度:

$$p_{W,V}(w,v)=p_{X,Y}(x,y)\cdot\left|\frac{\partial(x,y)}{\partial(w,v)}\right|.$$

  (2) 求边缘密度:

$$p_W(w)=\int p_{W,V}(w,v)dv.$$

- 方法二、分布函数法:

$$F_W(w)=P(W\leqslant w)=\mathrm{P}\big((X,Y)\in D_w\big).$$

例(补变量法). $N(0,\sigma_1^2) * N(0,\sigma_2^2) = N(0,\sigma_1^2+\sigma_2^2)$.

- 取X, Y 相互独立, 都服从$N(0,1)$. 求$\mathscr{L}(\sigma_1 X+\sigma_2 Y)$.
- 记$\sigma^2=\sigma_1^2+\sigma_2^2$. $\mathbf{A}=\frac{1}{\sigma}\begin{pmatrix}\sigma_1 & \sigma_2\\ -\sigma_2 & \sigma_1\end{pmatrix}$ 是正交矩阵.
- $(Z_1,Z_2)^T=\mathbf{A}(X,Y)^T$, 则$Z_1,Z_2$ 相互独立, 都服从$N(0,1)$.
- $\sigma_1 X+\sigma_2 Y=\sigma Z_1 \sim N(0,\sigma^2)$.

例(补变量法). 设$X_1, X_2, \cdots$ 独立, 都$\sim \mathrm{Exp}(\lambda)$.

求$S_n = X_1 + \cdots + X_n$ 的分布.

- 记$\vec{X} = (X_1, \cdots, X_n)$, $\vec{S} = (S_1, \cdots, S_n)$.
- $0 < s_1 < \cdots < s_n$,

$$p_{\vec{S}}(\vec{s}) = p_{\vec{X}}(\vec{x}) = \lambda^n \mathrm{e}^{-\lambda(x_1+\cdots+x_n)} = \lambda^n \mathrm{e}^{-\lambda s_n}.$$

- $S_n \sim \Gamma(n, \lambda)$: $\forall s_n > 0$, $p_{S_n}(s_n) =$

$$\int_{\star} \lambda^n \mathrm{e}^{-\lambda y} \mathrm{d}s_1 \cdots \mathrm{d}s_{n-1} = \lambda^n \mathrm{e}^{-\lambda s_n} \frac{s_n^{n-1}}{(n-1)!} = \frac{\lambda^n}{(n-1)!} s_n^{n-1} \mathrm{e}^{-\lambda s_n}.$$

- $\mathscr{L}(\vec{S}|S_{n+1} = t) = \mathscr{L}(U_{(1)}, \cdots, U_{(n)})$, 其中$U_1 \sim U(0, t)$:

$$p_{\vec{S}|S_{n+1}}(\vec{s}|t) = \frac{\lambda^{n+1}\mathrm{e}^{-\lambda t}}{\frac{\lambda^{n+1}}{n!} t^n \mathrm{e}^{-\lambda t}} = n! \cdot \frac{1}{t^n}, \quad 0 < s_1 < \cdots < s_n < t.$$

例(分布函数法). 设$p_{X,Y}(x,y)=p(x,y)$, 令$W=X+Y$, 求p_W.

- 求F_W:

$$F_W(w)=\underset{\sim\sim\sim\sim\sim\sim}{\mathrm{P}(X+Y\leqslant w)}=\int_{-\infty}^{\infty}\underline{\left(\int_{-\infty}^{w-x}p(x,y)\mathrm{d}y\right)}\mathrm{d}x.$$

- 全概率公式: $\underline{(\star\star)}=p_X(x)\int_{-\infty}^{w-x}p_{Y|X}(y|x)\mathrm{d}y$, 故

$$\underset{\sim\sim}{\star\star}=\int_{-\infty}^{\infty}p_X(x)\mathrm{P}(x+Y\leqslant w|X=x)dx.$$

- 做变量替换$z=x+y$.

$$=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{w}p(x,z-x)\mathrm{d}z\right)\mathrm{d}x=\int_{-\infty}^{w}\left(\int_{-\infty}^{\infty}p(x,z-x)\mathrm{d}x\right)\mathrm{d}z.$$

- 求导:

$$p_W(w)=\int_{-\infty}^{\infty}p(x,w-x)\mathrm{d}x=\int_{-\infty}^{\infty}p_X(x)p_{Y|X}(w-x|x)dx.$$

§2.9 随机变量序列

一、随机变量序列

- 设$(\Omega, \mathscr{F}, \mathrm{P})$ 是概率空间, $X_1, X_2, \cdots$ 为其上一列随机变量.
- 定义2.9.1. 相互独立: $X_1, \cdots, X_n$ 相互独立, $\forall n \geqslant 2$.
  独立同分布: 又$X_i \stackrel{\mathrm{d}}{=} X_1$, $\forall i$.
  两两独立: X_i 与X_j 相互独立, $\forall i \neq j$.
- 例2.9.2. 以下都是广义随机变量:

$$\sup_{n\geqslant 1} X_n, \quad \inf_{n\geqslant 1} X_n, \quad \limsup_{n\to\infty} X_n, \quad \liminf_{n\to\infty} X_n.$$

二、无穷维随机向量

- $\mathbb{R}^\infty := \{\mathbf{x} = (x_1, x_2, \cdots) : x_i \in \mathbb{R}, i = 1, 2, \cdots\}$,

$$\mathbf{X} : \Omega \to \mathbb{R}^\infty,\ \omega \mapsto (X_1(\omega), X_2(\omega), \cdots).$$

- 记$C_a^{(i)} := \{\mathbf{x} : x_i \leqslant a\}$, $\forall i \geqslant 1,\ a \in \mathbb{R}$.
- X_i 是随机变量指$\{X_i \leqslant a\} = \{\mathbf{X} \in C_a^{(i)}\} \in \mathscr{F}$, $\forall a$.
- $C_{a_1\cdots a_n}^{(i_1\cdots i_n)} := C_{a_1}^{(i_1)} \cdots C_{a_n}^{(i_n)} = \{\mathbf{x} : x_{i_1} \leqslant a_1, \cdots, x_{i_n} \leqslant a_n\}$.
- $X_1, X_2, \cdots$ 是随机变量序列指

$$\{\mathbf{X} \in C_{a_1\cdots a_n}^{(i_1\cdots i_n)}\} \in \mathscr{F}, \quad \forall n \geqslant 1; 1 \leqslant i_1 < \cdots < i_n; a_r's \in \mathbb{R}.$$

- $\mathscr{B}^\infty := \sigma(\{\star\star : \star\star\}) = \sigma(\mathscr{P})$, 其中$\mathscr{P} := \{\star\star : \star\star\}$.
- $\star\star$ 当且仅当

$\mathbf{X} : (\Omega, \mathscr{F}) \to (\mathbb{R}^\infty, \mathscr{B}^\infty)$ 为可测映射/无穷维随机向量.

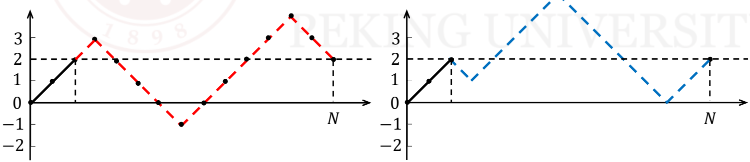
- 称$\mu_{\mathbf{X}}(D) = \mathrm{P}(\mathbf{X} \in D)$, $D \in \mathscr{B}^{\infty}$ 为$\mathbf{X}$ 的分布.
- $\mathscr{P}$ 为π 系, 故只需交代有限维边缘分布$\mathscr{L}(X_{i_1}, \cdots, X_{i_n})$.
- $\mathbf{X}$ 与$\mathbf{Y}$ 同分布, $\mathbf{X} \overset{\mathrm{d}}{=} \mathbf{Y}$: $(X_1, \cdots, X_n) \overset{\mathrm{d}}{=} (Y_1, \cdots, Y_n)$, $\forall n$
- $\mathbf{X}$ 与$\mathbf{Y}$ 相互独立: $\forall n, m$, $(X_1, \cdots, X_n)$ 与$(Y_1, \cdots, Y_m)$ 相互独立.
- 若$\mathbf{X}$ 与$\mathbf{Y}$ 相互独立, 则
  $\{\mathbf{X} \in D_1\}$ 与$\{\mathbf{Y} \in D_2\}$ 相互独立, $f(\mathbf{X})$ 与$g(\mathbf{Y})$ 相互独立.
- 独立同分布类似.

三、随机变量族 & 四、随机过程

- $\mathbf{X} = \{X_\alpha : \alpha \in I\}$ 类似, 有限维边缘.
- 将I 视为时间.
  离散型: $I = \mathbb{Z}_+ = \{0, 1, 2, \cdots\}, \{0, 1, \cdots, N\}$, 等等.
  连续型: $\mathbb{R}_+ = [0, \infty), [0, T]$, 等等.
  高维: $\mathbb{Z}^2, \mathbb{R}^2$ 等等.
- 称$\mathbf{X} = \{X_\alpha : \alpha \in I\}$ 为随机过程.
- 一族随机元$X_\alpha : (\Omega, \mathscr{F}) \to (S, \mathscr{S})$, 类似.

例2.9.4 (随机游动). 设$\xi_1, \xi_2, \cdots$ 独立同分布, $S_n = \xi_1 + \cdots + \xi_n$. 称$\{S_n : n = 0, 1, 2, \cdots\}$ 为随机游动.

- 简单随机游动: $\mathrm{P}(\xi_1 = 1) = \mathrm{P}(\xi_1 = -1) = 1/2$.
- 投票定理. 甲、乙两人竞选学生会主席, N 位同学投票, 甲比乙多i 票. 求P(甲保持领先). (例1.7.6)
- 反射原理, 答案为i/N.



例2.9.5 (泊松过程). 设$\xi_1, \xi_2, \cdots$ 独立同分布, 都服从Exp(λ). 令$S_n = \xi_1 + \cdots + \xi_n$, 则

$$X_t := |\{n \geqslant 1 : S_n \leqslant t\}| \sim P(\lambda t).$$

称$\{X_t : t \geqslant 0\}$ 为泊松过程.

- $\{X_t = k\} = \{S_k \leqslant t < S_k + \xi_{k+1}\}$.
- $k = 0$: $\mathrm{P}(X_t = 0) = \mathrm{P}(\xi_1 > t) = \mathrm{e}^{-\lambda t}$.
- $k \geqslant 1$:

$$\begin{aligned}
\mathrm{P}(X_t = k) &= \int_0^t \int_{t-x}^{\infty} \frac{\lambda^k}{(k-1)!} x^{k-1} \mathrm{e}^{-\lambda x} \cdot \lambda \mathrm{e}^{-\lambda y} \mathrm{d}y \mathrm{d}x \\
&= \int_0^t \frac{\lambda^k}{(k-1)!} x^{k-1} \mathrm{e}^{-\lambda x} \cdot \mathrm{e}^{-\lambda(t-x)} \mathrm{d}x \\
&= \frac{\lambda^k}{(k-1)!} \mathrm{e}^{-\lambda t} \cdot \frac{t^k}{k} = \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t}.
\end{aligned}$$

例2.9.6. $\mathscr{L}((S_1,\cdots,S_n)|X_t=n)=\mathscr{L}(U_{(1)},\cdots,U_{(n)})$，其中$U_1,\cdots,U_n$ 独立同分布，$\sim U(0,t)$.

- $\mathrm{P}(X_t=n)=\mathrm{P}(0<S_1<\cdots<S_n<t<S_n+\xi_{n+1})$
$$=\mathrm{e}^{-\lambda t}\lambda^n/n!.$$

- $0<x_1<y_1<\cdots<x_n<y_n<t.$

$$\begin{aligned}
&\mathrm{P}(x_i<S_i\leqslant y_i, i=1,\cdots,n \text{ 且 } \xi_{n+1}>t-S_n)\\
=&\int_{x_1}^{y_1}\cdots\int_{x_n}^{y_n}\int_{t-u_n}^{\infty}\lambda^n\mathrm{e}^{-\lambda u_n}\cdot\lambda\mathrm{e}^{-\lambda z}\mathrm{d}z\mathrm{d}u_n\cdots\mathrm{d}u_1\\
=&\int_{x_1}^{y_1}\cdots\int_{x_n}^{y_n}\lambda^n\mathrm{e}^{-\lambda u_n}\cdot\mathrm{e}^{-\lambda(t-u_n)}\mathrm{d}u_n\cdots\mathrm{d}u_1\\
=&\lambda^n\mathrm{e}^{-\lambda t}(y_1-x_1)\cdots(y_n-x_n).
\end{aligned}$$

- $\mathscr{L}((S_1,\cdots,S_n)|X_t=n)$ 的联合密度: $n!\dfrac{1}{t^n}\cdot\mathbf{I}_{\{0<u_1<\cdots<u_n<t\}}$.

例2.9.7. Y 与$\{\xi_1, \xi_2, \cdots\}$ 相互独立, $Y \sim G(p)$.
令$W = S_Y$. 求$\mathscr{L}(W)$.

- 范围: $x > 0$. 记$q = 1 - p$. $W \sim \text{Exp}(\lambda p)$:

$$\begin{aligned}
\mathrm{P}(W > x) &= \sum_{n=1}^{\infty} \mathrm{P}(Y = n, S_n > x) \\
&= \sum_{n=1}^{\infty} q^{n-1} p \int_x^{\infty} \frac{\lambda^n}{(n-1)!} u^{n-1} \mathrm{e}^{-\lambda u} \mathrm{d}u \\
&= \lambda p \int_x^{\infty} \sum_{m=0}^{\infty} \frac{(\lambda q u)^m}{m!} \mathrm{e}^{-\lambda u} \mathrm{d}u = \lambda p \int_x^{\infty} \mathrm{e}^{\lambda q u} \mathrm{e}^{-\lambda u} \mathrm{d}u \\
&= \int_x^{\infty} \lambda p \mathrm{e}^{-\lambda p u} \mathrm{d}u = \mathrm{e}^{-\lambda p x}.
\end{aligned}$$

§2.6 习题3. 求$\mathscr{L}(Y|W=x)$.

- $W=S_Y$. $\forall 0<\delta<x$,

$$\mathrm{P}(Y=n,x-\delta\leqslant W\leqslant x+\delta)=P(Y=n,x-\delta\leqslant S_n\leqslant x+\delta)$$
$$=q^{n-1}p\int_{x-\delta}^{x+\delta}\frac{\lambda^n}{(n-1)!}u^{n-1}\mathrm{e}^{-\lambda u}du$$
$$=\int_{x-\delta}^{x+\delta}\lambda p\frac{(\lambda qu)^m}{m!}\mathrm{e}^{-\lambda u}du.\quad (m=n-1)$$

- $\mathscr{L}(Y|W=x)=P(\lambda qx)$:

$$\mathrm{P}(Y=n|x-\delta\leqslant W\leqslant x+\delta)$$
$$=\frac{\displaystyle\int_{x-\delta}^{x+\delta}\lambda p\frac{(\lambda qu)^m}{m!}\mathrm{e}^{-\lambda u}du}{\displaystyle\int_{x-\delta}^{x+\delta}\lambda p\mathrm{e}^{-\lambda pu}\mathrm{d}u}\xrightarrow{\delta\to 0}\frac{(\lambda qx)^m}{m!}\mathrm{e}^{-\lambda qx}.$$

例. 设$\mathscr{L}(N|S_n=x)=P(\mu x)$. 求$\mathscr{L}(N)$ 和$\mathscr{L}(S_n|N=k)$.

- $N\sim NB(n,\dfrac{\mu}{\lambda+\mu})$: $k=0,1,2\cdots$,

$$\mathrm{P}(N=k)=\int_0^\infty \frac{\lambda^n}{(n-1)!}x^{n-1}\mathrm{e}^{-\lambda x}\frac{(\mu x)^k}{k!}\mathrm{e}^{-\mu x}\mathrm{d}x$$

$$=\frac{\lambda^n\mu^k}{(n-1)!k!}\int_0^\infty x^{n+k-1}\mathrm{e}^{-(\lambda+\mu)x}\mathrm{d}x=\frac{\lambda^n\mu^k}{(n-1)!k!}\cdot\frac{(n+k-1)!}{(\lambda+\mu)^{n+k}}$$

$$=C_{n+k-1}^k q^k p^n,\quad p=\frac{\lambda}{\lambda+\mu},\ q=1-p.$$

- $\mathscr{L}(S_n|N=k)=\Gamma(n+k,\lambda+\mu)$:

$$\begin{aligned}\mathrm{P}(S_n\leqslant s,N=k) &= \int_0^s \star\star\mathrm{d}x\\ &= \frac{(\lambda+\mu)^{n+k}}{(n+k-1)!}\int_0^s x^{n+k-1}\mathrm{e}^{-(\lambda+\mu)x}\mathrm{d}x.\end{aligned}$$

补§2.5 四、耦合

- 定义2.5.12. 设μ_1, μ_2 是分布. 若μ 是联合分布, 以μ_1 和μ_2 为边缘分布, 则称μ 为μ_1 与μ_2 的耦合.
- 构造二维随机元$(X, Y) \sim \mu$, 则$X \sim \mu_1$, $Y \sim \mu_2$.
  用X, Y 的函数值的关系揭示μ_1 与μ_2 的关系.
- 例. 指数分布. $\lambda_1 > \lambda_2 > 0$.
  - 直观: λ 越大, 对应的随机变量越小.
  - 印证: $X \sim \mathrm{Exp}(1)$, $X/\lambda \sim \mathrm{Exp}(\lambda)$.
  - 无效耦合: $X \sim \mathrm{Exp}(\lambda_1)$, $Y \sim \mathrm{Exp}(\lambda_2)$, X, Y 相互独立.

例2.7.6 (耦合). 设$F_1(x) \leqslant F_2(x)$, $\forall x$. 构造(X_1, X_2) 使得F_i 是X_i的分布函数且$X_1 \geqslant X_2$.

- $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geqslant u\}$.
- $F_2^{-1}(u) \leqslant F_1^{-1}(u)$:

  $\forall x \geqslant F_1^{-1}(u)$, $F_1(x) \geqslant u$. 于是 $F_2(x) \geqslant u$, 从而 $x \geqslant F_2^{-1}(u)$.

- 取$U \sim U(0,1)$, $F_1^{-1}(U) \geqslant F_2^{-1}(U)$.

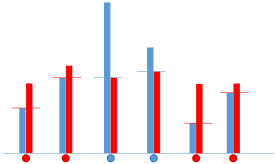
例2.5.13. $B(n,p)$.

- 取$U_1,\cdots,U_n$ 独立同分布, $\sim U(0,1)$. $X_{i,p}=\mathbf{I}_{\{U_i\leqslant p\}}$.
- $X_{1,p},\cdots,X_{n,p}$ 独立同分布, $\sim B(1,p)$.
- $Y_p=X_{1,p}+\cdots+X_{n,p}\sim B(n,p)$.
- 若$p<q$, 则$X_{i,p}\leqslant X_{i,q}$, $Y_p\leqslant Y_q$.
- 直观: p 越大, 正面次数越多.
- $Z_i=X_{i,q}-X_{i,p}=\mathbf{I}_{\{p<U\leqslant q\}}$,
  $Z=Y_q-Y_p=Z_1+\cdots+Z_n\sim B(n,q-p)$,
  $Y_q=Y_p+Z$, Z 与Y_p 不是相互独立的.

例2.5.14 (随机图). $G(n,p)$.

- $\Omega = K_n$ 的所有子图, P_p.
- 增事件A 指: 若$G \in A$ 且$G \subseteq \hat{G}$, 则$\hat{G} \in A$.
- 例. $A =$ “顶点i 与顶点j 连通”, 或 “图中存在三角形”.
- 随机变量: $U_1, U_2, \cdots$ 独立同分布, $\sim U(0,1)$.
  $G(n,p)$: 第k 条保留当且仅当$\{U_k \leqslant p\}$.
- 若$p < q$, 则$G(n,p) \subseteq G(n,q)$.
- 若A 是增事件, 则$G(n,p) \in A \Rightarrow G(n,q) \in A$.
  故$\mathrm{P}_p(A) \leqslant \mathrm{P}_q(A)$.

例2.5.15 (输运问题/最优耦合). n 个城市, 产量比列p_i, 销量比例:q_i. 设运输单价相同. 该如何运输?

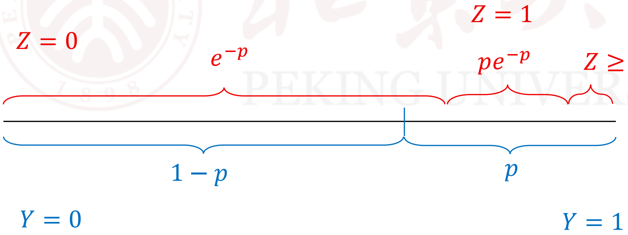


- $X \sim \mu$, 产地, $Y \sim \nu$: 销售地.
- 往证: 存在耦合使得$\mathrm{P}(X \neq Y)$ 最小,
  最小值为全变差距离
  $$d = d_{\mathrm{TV}}(\mu, \nu) := \frac{1}{2}\sum_i |p_i - q_i|.$$
- $c_i = p_i \wedge q_i$. $D_1 = \{k : p_i > q_k\}$, $D_2 = \{\ell : q_\ell > p_\ell\}$.
- $a_k = p_k - c_k$, $b_\ell = q_\ell - c_\ell$. $\sum\limits_{D_1} a_k = \sum\limits_{D_2} b_\ell = 1 - \sum\limits_i c_i = d$.
- $\mathrm{P}(X = Y) = \sum\limits_i \mathrm{P}(X = Y = i) \leqslant \sum\limits_i c_i$. $\quad \mathrm{P}(X \neq Y) \geqslant d$.
- 存在耦合使得等号成立:
  - 切割$[p, 1]$, 第i 个区间长c_i, 其上令$X = Y = i$.
    切割$[0, p]$, 第$k \in D_1$ 个区间长a_k, 其上令$X = k$.
    切割$[0, p]$, 第$\ell \in D_2$ 个区间长b_ℓ, 其上令$Y = \ell$.

例2.5.16 (小数定律). 设$X_1, \cdots, X_n$ 相互独立, $X_i \sim B(1, p_i)$. $\lambda = \sum_{i=1}^{n} p_i$, $\delta = \max_{1 \leqslant i \leqslant n} p_i$, $X = \sum_{i=1}^{n} X_i$, 则

$$d_{\mathrm{TV}}(\mathscr{L}(X), P(\lambda)) = \frac{1}{2} \sum_{k=0}^{\infty} \left| \mathrm{P}(X = k) - \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \right| \leqslant \lambda\delta.$$

- $U \sim U(0,1)$, $Y := f_p(U) \sim B(1,p)$, $Z := g_p(U) \sim P(p)$.



- $\mathrm{P}(Y \neq Z) = p - pe^{-p} \leqslant p^2$, $(1 - p \leqslant \mathrm{e}^{-p})$.

- $U_1, \cdots, U_n$ i.i.d. $\sim U(0,1)$.

$$U_i \to (Y_i, Z_i), \quad \mathrm{P}(Y_i \neq Z_i) \leqslant p_i^2.$$

- $Y = \sum_{i=1}^n W_i \overset{d}{=} X$, $Z = \sum_{i=1}^n Z_i \sim P(\lambda)$.
- $d_{\mathrm{TV}}(\mathscr{L}(X), P(\lambda))$

$$\leqslant \mathrm{P}(Y \neq Z) \leqslant \sum_{i=1}^n \mathrm{P}(Y_i \neq Z_i) \leqslant \sum_{i=1}^n p_i^2 \leqslant \lambda\delta.$$