

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

- 退化的: $P(X = c) = 1$.

- 伯努利分布:

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

- 二项(Binomial)分布, $X \sim B(n, p)$:

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

- 几何分布, $X \sim G(p)$:

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

- 泊松分布, $X \sim P(\lambda)$:

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

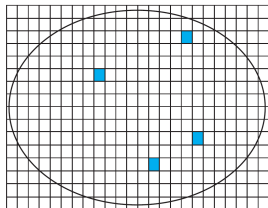
- §2.1 习题10. $k \cdot p_k = \lambda \cdot p_{k-1}$, $k = 1, 2, \dots$.

- 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 $p_n = \frac{\lambda}{n}$;
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.



(a) 蓝色小块为放射粒子小块



(b) 蓝×表示放射粒子的时刻

图 2.1

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

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

$$C_n^k p^k (1-p)^{n-k} \text{ 与 } \frac{\lambda^k}{k!} e^{-\lambda}.$$

- 当 $n \rightarrow \infty$, $np \rightarrow \lambda$ 时, $\sum_{k=0}^{\infty} |\star - \star| \rightarrow 0$. (命题A.0.1)
- 当 $np = \lambda$ 时, $\sum_{k=0}^{\infty} |\star - \star| \leq 2np^2 = 2\lambda^2/n$. (例2.5.16)
- 应用(例2.1.6)
- §2.1 习题8. 都是单峰的, 最大值点 $\approx np, \lambda$.

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

- $P(X > n) = \sum_{i=n+1}^{\infty} (1-p)^{i-1}p = (1-p)^n.$

- $f(k) = k - n, Y = X - n.$

- Y 的分布列:

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

- 若事件 $\{X > n\}$ 发生, 则 $Y \geq 1$.

- Y 的条件分布列: $k \geq 1$,

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

- 无记忆性:

$$P(Y = k | X > n) = P(X = k), \quad k = 1, 2, \dots$$

§2.2 连续型随机变量

- 密度函数:

$$p(x) \geq 0, \quad \int_{-\infty}^{\infty} p(x) dx = 1.$$

- 若

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

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

- 一般地,

$$P(X > x) = \int_x^{\infty} p(y) dy, \quad P(a < X \leq b) = \int_a^b p(x) dx;$$

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

一、几个常见的密度函数

- 均匀(Uniform)分布, $X \sim U(a, b)$:

$$p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{其他.} \end{cases} \quad \tilde{p}(x) = \frac{1}{b-a} \mathbf{I}_{\{a < x < b\}}.$$

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

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

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

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

- 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 $\sim G(\lambda/n)$. $X \approx X_n/n$.

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

- 密度函数: $p(x) = -(e^{-\lambda x})' = \lambda e^{-\lambda x}$, $x > 0$.
- 无记忆性: 对任意 $T, t \geq 0$,

$$P(X > T + t | X > T) = P(X > t).$$

- 标准正态(Normal)分布, $X \sim N(0, 1)$:

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

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

$$\begin{aligned} \star^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} \underbrace{dx dy}_{\text{极坐标}} = \int_0^{2\pi} \int_0^{\infty} \frac{1}{2\pi} e^{-\frac{r^2}{2}} \underbrace{r dr d\theta}_{\text{极坐标}} \\ &= \left(\int_0^{2\pi} \frac{1}{2\pi} d\theta \right) \cdot \left(\int_0^{\infty} e^{-\frac{r^2}{2}} d\frac{r^2}{2} \right) = 1. \end{aligned}$$

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

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

- 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

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

$$\bullet \quad P(Y \leq y) = P(a + bX \leq y) \stackrel{b \geq 0}{=} P\left(X \leq \frac{y-a}{b}\right)$$

$$\bullet \quad \int_{-\infty}^{\frac{y-a}{b}} p_X(x) dx = \int_{-\infty}^y p_Y(z) dz.$$

$$\bullet \quad = : z = a + bx, \quad p_Y(z) := \frac{1}{b} \cdot p_X\left(\frac{z-a}{b}\right).$$

• $b < 0$ 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:

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

- $X \sim N(\mu, \sigma^2)$, 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, $\lambda X \sim \text{Exp}(1)$:

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

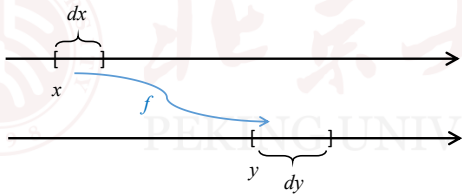
- $Y \sim \text{Exp}(1) \Rightarrow Y/\lambda \sim \text{Exp}(\lambda)$.
- $\alpha X^{1/m} \sim W(m, \alpha)$:

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

定理2.2.10. 设 X 是连续型, $P(a < X < b) = 1$, $f : (a, b) \rightarrow (c, d)$ 一对一, $f'(x) \neq 0$, $Y = f(X)$, 则

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

- 确定 x, y 的取值范围. $p_X(x)|dx| = p_Y(y)|dy|$,



- 多对一:

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

例2.2.11. $X \sim U(-\pi/2, \pi/2)$, $Y = \tan X$, 求 p_Y .

- $p_Y(y) = p_X(x) \cdot \frac{1}{1+y^2} = \frac{1}{\pi} \cdot \frac{1}{1+y^2}.$

- $X \sim U(0, \pi)$ 或 $U(0, 2\pi) \Rightarrow \tan X \sim C(0, 1).$

- $W = \mu + \alpha Y \sim C(\mu, \alpha).$

设 $\alpha > 0$ ($\alpha < 0$ 类似),

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

§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.

$$X : \Omega \rightarrow \mathbb{R},$$

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

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

$$\{X \leq x\} \in \mathcal{F}, \quad \forall x \in \mathbb{R},$$

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

二、随机变量的分布函数

- 定义2.3.6. X 的分布函数:

$$F_X(x) = P(X \leq x),$$

X 的尾分布函数:

$$G_X(x) = P(X > x).$$

- 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2.3.8. $F = F_X$ 满足三条性质:

- (1) 单调上升性: 若 $x \leq y$, 则 $F(x) \leq F(y)$;
- (2) 规范性: $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$;
- (3) 右连续性: 对任意 $x \in \mathbb{R}$, $\lim_{y \searrow x} F(y) = F(x)$.

- (1) & (2): 概率的非负性& 规范性.
- 单调函数总可以定义右极限与左极限:

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

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

三、分布

- (概率)分布: 称 $(\mathbb{R}, \mathcal{B})$ 上的概率. 记为 μ .
- 若 X 是随机变量, 则 $\{X \in B\} \in \mathcal{F}, \forall B \in \mathcal{B}$. (引理2.3.9)
- X 的分布:

$$B \mapsto P(X \in B),$$

记为 μ_X 或 $\mathcal{L}(X)$.

- 若 $F_X \equiv F_Y$, 则 $\mu_X \equiv \mu_Y$, 此时称 X 与 Y 同分布, 记为 $X \stackrel{d}{=} Y$. (推论1.5.16, 定义2.3.10)
- $X \stackrel{d}{=} Y$ 当且仅当 $\mathcal{L}(X) = \mathcal{L}(Y)$.
- 若 $P(X = Y) = 1$, 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).

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

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

四、广义随机变量、可测映射与随机元*

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

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

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

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

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

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

$$\int f(x)dx := \lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} \frac{k}{n} \cdot \lambda \left(\left\{ x : \frac{k}{n} \leq f(x) < \frac{k+1}{n} \right\} \right).$$

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

- 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:

$$P(X \in B) = \int_B p(x)dx, \quad \forall B \in \mathcal{B}.$$

- 一般情形, $a \in \mathbb{R}$,

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

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

$$\int f(x)dx := \int f^+(x)dx - \int f^-(x)dx,$$

- 积分有线性, 高维类似.
- 例2.3.4. 对于连续型随机变量, 密度函数 $p(\cdot)$ 是非负Borel 函数, 积分是Lebesgue 积分.

§2.4 随机向量

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

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

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

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

- $\sigma(\vec{X}) = \{\{\vec{X} \in B\} : B \in \mathcal{B}^n\} \subseteq \mathcal{F}$.
- 研究 $\vec{X}$ 的联合分布

$$\mu_{\vec{X}}, \mathcal{L}(\vec{X}) : \mathcal{B}^n \rightarrow \mathbb{R}, \quad B \mapsto P(\vec{X} \in B).$$

一、离散型随机向量

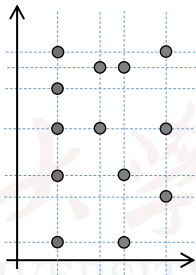
- X 与 Y 都是离散型, 联合分布(列):

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

- $X_1, \dots, X_n$ 都是离散型. 联合分布列:

$$P(X_1 = x_{i_1}^{(1)}, \dots, X_n = x_{i_n}^{(n)}),$$

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



- 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. $(X_1, \dots, X_{k-1})$ 的联合分布列.

二、连续型随机向量

- 联合分布函数: $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$.
- (X,Y) 为连续型指:

$$F_{X,Y}(x,y) = \int_{-\infty}^x \int_{-\infty}^y p(u,v) dv du, \quad \forall x,y \in \mathbb{R}.$$

称 $p(\cdot, \cdot)$: (X,Y) 的联合密度函数, 记为 $p_{X,Y}(\cdot, \cdot)$.

- ****: 非负、可测, 且 $\iint_{\mathbb{R}^2} p(x,y) dx dy = 1$.
- 若 (x,y) 为 $p(\cdot, \cdot)$ 的连续点, 则

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

- $\forall D \in \mathcal{B}^2, P((X,Y) \in D) = \iint_D p_{X,Y}(x,y) dx dy$.
- 若 (X,Y) 为连续型, 则 $P(X=Y) = 0$.

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

- 例2.4.6. $C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$:

$$\begin{aligned} \int_{-\infty}^{\infty} p(x, y) dy &= C \int_{-\infty}^{\infty} \exp \left\{ -\frac{\boxed{(v - \rho u)^2} + \boxed{(1 - \rho^2)u^2}}{2(1 - \rho^2)} \right\} \sigma_2 dv \\ &= C \sigma_2 e^{-\frac{u^2}{2}} \cdot \sqrt{2\pi(1 - \rho^2)} = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{u^2}{2}}. \end{aligned}$$

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

$$I = \boxed{(v - \rho u)^2} + \boxed{(\sqrt{1 - \rho^2} u)^2}.$$

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

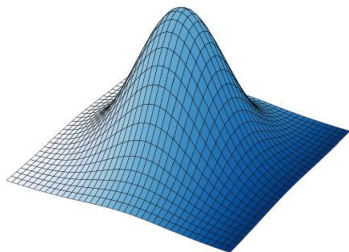


图 2.3 二维正态分布的联合密度函数的图像

三、一般形式的随机向量

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

$$F_{\vec{X}}(\cdot) = P(\vec{X} \leq \vec{x}), \quad \mu_{\vec{X}}(B) := P(\vec{X} \in B), \quad \forall B \in \mathcal{B}^n.$$

- $\vec{X}$ 与 $\vec{Y}$ 同分布, $\vec{X} \stackrel{d}{=} \vec{Y}$ 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, $\vec{X} \stackrel{d}{=} \vec{Y}$,
 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是 Borel 函数, 则 $f(\vec{X}) \stackrel{d}{=} f(\vec{Y})$.
- $\vec{X}$ 与 $\vec{Y}$ 几乎必然相等, $\vec{X} = \vec{Y}$ a.s. 指: $P(\vec{X} = \vec{Y}) = 1$.
- $\vec{X} = \vec{Y}$ a.s. $\Rightarrow \vec{X} \stackrel{d}{=} \vec{Y}$, 反之不然.

均匀分布: $\vec{X} \sim U(S)$.

- 例2.4.8. 设 $U \sim U(0, 2\pi)$, 则 $(\cos U, \sin U) \sim U(S^1)$.
- 例2.4.7 & 2.4.9 & 2.8.2 & 2.8.3.

$$S = \{\vec{x} \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0 \text{ 且 } x_1 + x_2 + x_3 = 1\}.$$

- 可交换性: 重排后, $(X_i, X_j, X_k) \sim U(S)$.
- $(X_i, X_j) \sim U(\hat{S})$, 是连续型, 其中

$$\hat{S} = \{(x, y) \in \mathbb{R}^2 : x, y \geq 0 \text{ 且 } x + y \leq 1\}.$$

- $(X, Y) \sim U(\hat{S}) \Rightarrow (X, 1 - X - Y) \sim U(S)$.
- 例2.4.10. $S = \{\vec{x} : x_1^2 + \cdots + x_n^2 = 1\} \subseteq \mathbb{R}^n$. 列向量.
则 $\vec{Y} = \mathbf{A}\vec{X} \sim U(S)$, 其中 $\mathbf{A}$ 为 n 阶正交矩阵.

§2.5 边缘分布与独立性

一、边缘分布

- $\vec{X} = (X_1, \dots, X_n)$.

称 $\vec{Y} := (X_{i_1}, \dots, X_{i_m})$ 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, $P(X = x_i) = \sum_{j \in J} p_{ij}, i \in I$.
- 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, $p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy$.
- 例2.5.2. 二维正态分布的边缘为正态分布.
- 例2.5.3. $\vec{X} \stackrel{d}{=} \vec{Y} \Rightarrow (X_{i_1}, \dots, X_{i_m}) \stackrel{d}{=} (Y_{i_1}, \dots, Y_{i_m})$.

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 $\forall x_1, \dots, x_n$,

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = P(X_1 \leq x_1) \cdots P(X_n \leq x_n),$$

则称 $X_1, \dots, X_n$ (相互)独立.

- $\star\star$ 当且仅当 $\forall B_1, \dots, B_n \in \mathcal{B}$,

$$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n).$$

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

$X_1, \dots, X_n$ 独立当且仅当 “ $=$ ”, 只需验证 “ $=$ ” .

- 离散型:

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n).$$

$$P(X_1 = x_1, \dots, X_n = x_n) = f_1(x_1) \cdots f_n(x_n).$$

- 连续型:

$$p_{\vec{X}}(\vec{x}) = p_{X_1}(x_1) \cdots p_{X_n}(x_n).$$

$$p_{\vec{X}}(\vec{x}) = f_1(x_1) \cdots f_n(x_n).$$

例2.5.7. 考虑成功率为 p 的Bernoulli 试验.

- 令 $X_i = \mathbf{I}_{H_i}$. $X_1, \dots, X_n$ 独立同分布, $\forall n$.
- 注: 此时称 $X_1, X_2, \dots$ 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.
- $\{Y_1 = k_1, Y_2 = k_2, \dots, Y_n = k_n\}$:

k_1 个T 接一个H; k_2 个T 接一个H; $\dots$, k_n 个T 接一个H.

概率为

$$(1-p)^{k_1} p \cdot (1-p)^{k_2} p \cdots (1-p)^{k_n} p.$$

- $Y_1, \dots, Y_n$ 独立同分布, $Y_1 + 1 \sim G(p)$.

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 $\rho = 0$.

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

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

- $\Leftarrow$: 若 $\rho = 0$, 则独立:

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

- $\Rightarrow$: 若独立, 则 $\star\star$ 为联合密度.
- $p(x, y)$ 与 $\tilde{p}(x, y)$ 均为 (X, Y) 的联合密度, 均连续, 故相等.
- 取 $x = \mu_1, y = \mu_2$, 得 $\rho = 0$.

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

- 记 $\vec{X} = (X_1, \dots, X_n), \vec{Y} = (Y_1, \dots, Y_n)$.

- $\vec{X} \stackrel{d}{=} \vec{Y}$:

$$F_{\vec{X}}(\vec{z}) = F_1(z_1) \cdots F_n(z_n) = F_{\vec{Y}}(\vec{z}).$$

- 设 $X_1, \dots, X_n$ 独立同分布. 可交换性:

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

三、随机向量的独立性

- $\vec{X}$ 与 $\vec{Y}$ 相互独立指:

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

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

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

- ★★ 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 $\forall$ Borel 函数/可测映射, $f_i(\vec{X}^{(i)})$'s 独立.
- 独立同分布, 两两独立, 类似.

§2.6. 条件分布

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):

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

- 计算: $p_{Y|X}(\cdot|x_i) \propto P(Y = \cdot, X = x_i)$, (固定 x_i).
- $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j|X = x_i)$.
- $P(Y = y_j) = \sum_{i \in I} P(X = x_i)p_{Y|X}(y_j|x_i)$.
- X 与 Y 相互独立的
 - 必要条件: $\mathcal{L}(Y|X = x_i) = \mathcal{L}(Y), \forall i$.
 - 充分条件: $\mathcal{L}(Y|X = x_i)$ 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: $\mathcal{L}(Y|X = n) = B(n, p)$.
- 联合分布: $\forall k, \ell = 0, 1, \dots$,

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

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

二、连续型随机向量的条件分布

- 条件分布函数:

$$\begin{aligned} P(Y \leq y | X = x) &:= \lim_{\varepsilon \rightarrow 0+} P(Y \leq y | x - \varepsilon < X \leq x + \varepsilon) \\ &= \lim_{\varepsilon \rightarrow 0+} \frac{\int_{x-\varepsilon}^{x+\varepsilon} \int_{-\infty}^y p(u, v) dv du}{\int_{x-\varepsilon}^{x+\varepsilon} p_X(u) du} = \int_{-\infty}^y \frac{p(x, v)}{p_X(x)} dv. \end{aligned}$$

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

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

- 结论: $X \sim N(\mu_1, \sigma_1^2)$, $\mathcal{L}(Y|X = x) = N(\mu, \sigma^2)$, 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: $I = (v - \rho u)^2 + (1 - \rho^2) u^2$.

$$dx = \sigma_1 du, \quad dy = \sigma_2 dv.$$

方法一、先求边缘密度, 再求条件密度.

- 例2.4.6 & 例2.5.3.

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

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

- $-\frac{1}{2(1-\rho^2)} \cdot I + \frac{1}{2}u^2:$

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

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

方法二、先求条件密度, 再求边缘密度.

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

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

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

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

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

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

三、一般形式的条件分布

- 固定 x . 令 $\tilde{F}_{Y|X}(y|x) = \lim_{\varepsilon \rightarrow 0+} P(Y \leq y | x - \varepsilon < X \leq x + \varepsilon)$.
- $\tilde{F}_{Y|X}(y|x)$ 关于 y 单调上升, $F_{Y|X}(y|x) := \tilde{F}_{Y|X}(y + |x)$.
- 例2.6.3. $(X, Y) \sim U(S^1)$.

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

§2.7 随机变量的函数

一、随机变量的函数

- X 为离散型, $f: x_i \rightarrow f(x_i)$. $Y = f(X)$.

找出所有可能值 y_j ,

$$P(Y = y_j) = \sum_{i \in I: f(x_i) = y_j} p_i.$$

- 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 $X \stackrel{d}{=} Y$, 则 $f(X) \stackrel{d}{=} f(Y)$.

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 $x \leq y$, 则 $F(x) \leq F(y)$;

(2) 规范性: $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$;

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

- F 的广义反函数:

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

- 尾分布函数类似.

引理2.7.2. F 与 F^{-1} 都是Borel 函数, 并且

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

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

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

- F^{-1} 是Borel 函数:

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

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

• $\Rightarrow: F = F_X$.

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

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

- $P(F(X) < u) = P(X < x_0) = F(x_0)$.
- $F(x_0) \geq u; F(x_0-) \leq u$. 因此 $F(x_0) = u$.
- $P(U < u) = u \Rightarrow U \sim U(0, 1)$:

$$P(U = u) \leq P(u \leq U < u + \frac{1}{n}) = (u + \frac{1}{n}) - u \rightarrow 0.$$

- $\Leftarrow$: 反证法. 取 x_0 使得 $a = P(X < x_0) < b = F_X(x_0)$.
- $P(U \in (a, b)) = 0$. 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 $F_X = F$.

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

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

- $\forall x,$

$$F_X(x) = P(X \leq x) = P(U \leq F(x)) = F(x).$$

- 注: 任意分布函数都是某随机变量的分布函数.

- 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.
 - $X = \max\{F_1^{-1}(U), F_2^{-1}(U)\}. \forall x \in \mathbb{R},$

$$F_X(x) = P(U \leq F_1(x), U \leq F_2(x)) = F(x).$$

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

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

- $F_Z = F$:

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

三、独立和

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

- 联合密度: $p_{X,Y}(x, y) = p_X(x)p_Y(y)$.

- 记 $D_z = \{(x, y) \in \mathbb{R}^2 : x + y \leq z\}$, 则

$$\begin{aligned} P(X + Y \leq z) &= \iint_{D_z} p_X(x)p_Y(y)dx dy \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-x} p_Y(y)dy \right) p_X(x)dx = \int_{-\infty}^{\infty} \left(\int_{-\infty}^z p_Y(y-x)dy \right) p_X(x)dx. \\ &= \int_{-\infty}^z \left(\int_{-\infty}^{\infty} p_X(x)p_Y(y-x)dx \right) dy. \end{aligned}$$

- $\forall z,$

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

- 卷积 $h = f * g$:

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

- 设 μ, ν 是分布. μ 与 ν 的卷积指 $\mathcal{L}(X+Y)$, 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.

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

$$\mathcal{L}(S_n) * \mathcal{L}(S_m) = \mathcal{L}(S_{n+m}), \quad (X_1 + \dots + X_n) + (X_{n+1} + \dots + X_{n+m}).$$

例2.7.15 ~ 17. 伽马分布 $\Gamma(\alpha, \lambda)$, $\alpha > 0$, $\lambda > 0$. 密度函数:

$$p(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0, \quad \text{其中 } \Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy.$$

• $\Gamma(\alpha_1, \lambda) * \Gamma(\alpha_2, \lambda) = \Gamma(\alpha_1 + \alpha_2, \lambda): \forall z > 0,$

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

例2.7.15 ~ 17(续). 伽马分布 $\Gamma(\alpha, \lambda)$, $\alpha > 0$, $\lambda > 0$. 密度函数:

$$p(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0, \quad \text{其中 } \Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy.$$

- 若 $X \sim \text{Exp}(\lambda) = \Gamma(1, \lambda)$, 则 $S_n \sim \Gamma(n, \lambda)$.
- 若 $Z \sim N(0, 1)$, 则 $Z^2 \sim \Gamma(1/2, 1/2)$ (例2.2.12). 于是

$$Z_1^2 + \cdots + Z_n^2 \sim \Gamma(n/2, 1/2) \triangleq \chi^2(n).$$

- $Z_1^2 + Z_2^2 \sim \Gamma(1, 1/2) = \text{Exp}(1/2)$.
- 若 $X \sim \text{Exp}(\lambda)$, 则 $2\lambda X \sim \text{Exp}(1/2)$, 故 $2\lambda S_n \sim \chi^2(2n)$.

§2.8 随机向量的变换

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 .
- $f: D_1 \rightarrow D_2$ 足够光滑, 一对一.
- $I(\vec{x}) = \det(\partial y_i / \partial x_j)_{i,j \leq n}$, $J(\vec{y}) = \det(\partial x_i / \partial y_j)_{i,j \leq n}$.
- $\forall \vec{y} \in D_2$,

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

- 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 ~ 2.8.7. 设 X, Y 相互独立, 都服从 $N(0, 1)$.

$$X = R \cos \Theta, \quad Y = R \sin \Theta.$$

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

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

- Θ, R 相互独立, $\Theta \sim U(0, 2\pi), R^2 = X^2 + Y^2 \sim \text{Exp}(1/2):$

$$p_W(w) = p_R(r) \frac{dr}{dw} = r e^{-r^2/2} \cdot \frac{1}{2r} = \frac{1}{2} e^{-\frac{w}{2}}, \quad \forall w > 0.$$

- $X/Y = \tan \Theta \sim C(0, 1).$
- 正交变换: $\mathbf{A}(X, Y)^T \stackrel{d}{=} (X, Y)^T. R = R, \Theta = \Theta + \alpha.$
- $W \stackrel{d}{=} -2 \ln U_1, \Theta \stackrel{d}{=} 2\pi U_2,$ 故

$$(X, Y) \stackrel{d}{=} (\sqrt{-2 \ln U_1} \cos(2\pi U_2), \sqrt{-2 \ln U_1} \sin(2\pi U_2)).$$

四、顺序统计量

- 设 $X_1, \dots, X_n$ 独立同分布.
- $\forall \omega$, 从小到大排序: $X_{(1)}(\omega) \leq \dots \leq X_{(n)}(\omega)$.
- 称 $X_{(1)}, \dots, X_{(n)}$ 为 $X_1, \dots, X_n$ 的顺序统计量.
- 设 X_1 为连续型, 此时

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

- 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 & 2.8.11. $(X_{(1)}, \dots, X_{(n)})$ 的联合密度函数为

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

- 设 $X_{(k)} = X_{I_k}$, $k = 1, \dots, n$.
- $\vec{I} = (I_1, \dots, I_n)$ 是随机全排列, 且与 $\star\star$ 相互独立:

$$\forall 0 < x_1 < y_1 < \cdots < x_n < y_n < t,$$

$$\begin{aligned} & \mathbb{P}(x_r < X_{(r)} \leq y_r, \forall r; \vec{I} = (i_1, \dots, i_n)) \\ &= \mathbb{P}(x_r < X_{i_r} \leq y_r, \forall r) = \frac{1}{n!} \int_D \star\star d\vec{x}. \end{aligned}$$

- $X_{(k)}$ 的密度函数为 $n C_{n-1}^{k-1} F(x)^{k-1} G(x)^{n-k} p(x)$.
 - $\mathbb{P}(X_{(k)} \leq x) = \sum_{i=k}^n C_n^i F(x)^i (1 - F(x))^{n-i}$.
 - 对联合密度积分.

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

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

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

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

- 可交换性: $(X_{i_1}, \dots, X_{i_{n+1}}) \stackrel{d}{=} \vec{X}$.

§2.9 随机变量序列

一、随机变量序列

- 设 $(\Omega, \mathcal{F}, P)$ 是概率空间, $X_1, X_2, \dots$ 为其上一列随机变量.
- 定义2.9.1. 两两独立、相互独立、独立同分布.
- 例2.9.2. 以下都是广义随机变量:

$$\sup_{n \geq 1} X_n, \inf_{n \geq 1} X_n, \limsup_{n \rightarrow \infty} X_n, \liminf_{n \rightarrow \infty} X_n.$$

二、无穷维随机向量& 四、随机过程

- $\mathbf{X} = (X_1, X_2, \dots)$: 无穷维随机向量, 取值于 $\mathbb{R}^\infty$ 的随机元.
- 只需交代有限维边缘分布 $\mathcal{L}(X_{i_1}, \dots, X_{i_n})$.
- $\mathbf{X}$ 与 $\mathbf{Y}$ 同分布, $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$:

$$(X_1, \dots, X_n) \stackrel{d}{=} (Y_1, \dots, Y_n), \quad \forall n.$$

- $\mathbf{X}$ 与 $\mathbf{Y}$ 相互独立:

$$(X_1, \dots, X_n) \text{ 与 } (Y_1, \dots, Y_m) \text{ 相互独立, } \forall n, m.$$

- 若 $\mathbf{X}$ 与 $\mathbf{Y}$ 相互独立, 则

$$\{\mathbf{X} \in D_1\} \text{ 与 } \{\mathbf{Y} \in D_2\} **, \quad f(\mathbf{X}) \text{ 与 } g(\mathbf{Y}) **.$$

- 独立同分布类似.

泊松过程

例2.9.5. 设 $\xi_1, \xi_2, \dots$ 独立同分布, 都服从 $\text{Exp}(\lambda)$.

令 $S_n = \xi_1 + \dots + \xi_n$, 则

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

称 $\{X_t : t \geq 0\}$ 为泊松过程.

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

$$\begin{aligned} P(X_t = k) &= \int_0^t \int_{t-x}^{\infty} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x} \cdot \lambda e^{-\lambda y} dy dx \\ &= \int_0^t \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x} \cdot e^{-\lambda(t-x)} dx \\ &= \frac{\lambda^k}{(k-1)!} e^{-\lambda t} \cdot \frac{t^k}{k} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad \forall k \geq 1. \end{aligned}$$

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

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

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

$$\begin{aligned} & P(x_i < S_i \leq y_i, i = 1, \dots, n \text{ 且 } \xi_{n+1} > t - S_n) \\ &= \int_{x_1}^{y_1} \dots \int_{x_n}^{y_n} \int_{t-u_n}^{\infty} \lambda^n e^{-\lambda u_n} \cdot \lambda e^{-\lambda z} dz du_n \dots du_1 \\ &= \int_{x_1}^{y_1} \dots \int_{x_n}^{y_n} \lambda^n e^{-\lambda u_n} \cdot e^{-\lambda(t-u_n)} du_n \dots du_1 \\ &= \lambda^n e^{-\lambda t} (y_1 - x_1) \dots (y_n - x_n). \end{aligned}$$

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

例2.8.9. 设 η, ζ 相互独立, $\eta \sim \text{Exp}(\lambda_1)$, $\zeta \sim \text{Exp}(\lambda_2)$.

- 记 $\xi = \min\{\eta, \zeta\}$, $\lambda = \lambda_1 + \lambda_2$, $p = \lambda_1/\lambda$, 则

$\xi \sim \text{Exp}(\lambda)$, $P(\xi = \eta) = p$, A 与 ξ 相互独立.

- $\forall y > 0$, $\{\xi > x, A\} = \{\zeta > \eta > x\}$.
- $\forall x > 0$,

$$\begin{aligned} P(\xi > x, A) &= \int_x^\infty \int_u^\infty \lambda_1 \lambda_2 e^{-(\lambda_1 u + \lambda_2 v)} dv du \\ &= \int_x^\infty \lambda_1 e^{-\lambda_1 u} e^{-\lambda_2 u} du = \frac{\lambda_1}{\lambda} e^{-\lambda x}. \end{aligned}$$

例2.7.14. 设 $Y \sim P(\lambda_1)$, $Z \sim P(\lambda_2)$, 且 Y 与 Z 相互独立.

- 记 $\lambda = \lambda_1 + \lambda_2$, $p = \lambda_1/\lambda$.

$$X = Y + Z \sim P(\lambda), \quad \mathcal{L}(Y|X = n) = B(n, p).$$

- $\hat{Y} \sim P(\lambda p)$ (幼虫数), $\hat{Z} \sim P(\lambda(1 - p))$, 相互独立. (例2.6.1)
- $(Y, Z) \stackrel{d}{=} (\hat{Y}, \hat{Z})$. 故

$$X \stackrel{d}{=} \hat{X}, \quad \mathcal{L}(Y|X = n) = \mathcal{L}(\hat{Y}|\hat{X} = n).$$

例2.9.7 & §2.6 习题3. V 与 $\{\xi_1, \xi_2, \dots\}$ 相互独立, $V \sim G(p)$, 则

$$W := S_V \sim \text{Exp}(\lambda p), \quad \mathcal{L}(V-1|W=x) = P(\lambda q x).$$

- $P(W > x) = \sum_{n=1}^{\infty} P(V = n, S_n > x) = \dots$
- $\forall 0 < \delta < x,$

$$\begin{aligned} P(V = n, x - \delta \leq W \leq x + \delta) &= P(V = n, x - \delta \leq S_n \leq x + \delta) \\ &= \int_{x-\delta}^{x+\delta} q^{n-1} p \frac{\lambda^n}{(n-1)!} u^{n-1} e^{-\lambda u} du, \quad n = 1, 2, \dots \end{aligned}$$

- $P(V-1 = m|W=x): m = 0, 1, 2, \dots,$

$$\frac{\int_{x-\delta}^{x+\delta} \lambda p \frac{(\lambda q u)^m}{m!} e^{-\lambda u} du}{\int_{x-\delta}^{x+\delta} \lambda p e^{-\lambda p u} du} \xrightarrow{\delta \rightarrow 0} \frac{(\lambda q x)^m}{m!} e^{-\lambda q x}.$$

§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: $X \sim \text{Exp}(1), X/\lambda \sim \text{Exp}(\lambda)$.
 - 无效耦合: $X \sim \text{Exp}(\lambda_1), Y \sim \text{Exp}(\lambda_2), X, Y$ 相互独立.

例2.7.6 (耦合). 设 $F_1(x) \leq F_2(x)$, $\forall x$. 构造 (X_1, X_2) 使得

$$F_i = F_{X_i}, \quad i = 1, 2 \text{ 且 } X_1 \geq X_2.$$

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

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

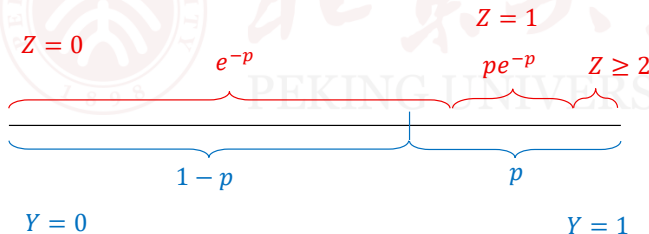
- 取 $U \sim U(0, 1)$, $F_1^{-1}(U) \geq F_2^{-1}(U).$

例2.5.16 (小数定律). 设 $X_1, \dots, X_n$ 相互独立, $X_i \sim B(1, p_i)$.

$\lambda = \sum_{i=1}^n p_i$, $\delta = \max_{1 \leq i \leq n} p_i$, $X = \sum_{i=1}^n X_i$, 则

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

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



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

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

$$U_i \rightarrow (Y_i, Z_i), \quad \mathbf{P}(Y_i \neq Z_i) \leq p_i^2.$$

- $Y = \sum_{i=1}^n W_i \stackrel{d}{=} X, Z = \sum_{i=1}^n Z_i \sim P(\lambda).$

- $d_{\text{TV}}(\mathcal{L}(X), P(\lambda))$

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

补充

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

找 g , 使得 $(x, y) \mapsto (f(x, y), g(x, y))$ 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:

$$F_W(w) = P(W \leq w) = P((X, Y) \in D_w).$$

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 $\mathcal{L}(\sigma_1 X + \sigma_2 Y)$.
- 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 $X_1, X_2, \dots$ 独立, 都 $\sim \text{Exp}(\lambda)$.

求 $S_n = X_1 + \dots + X_n$ 的分布.

- 记 $\vec{X} = (X_1, \dots, X_n)$, $\vec{S} = (S_1, \dots, S_n)$.

- $0 < s_1 < \dots < s_n$,

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

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

$$\int_{\star} \lambda^n e^{-\lambda s_n} ds_1 \cdots ds_{n-1} = \lambda^n e^{-\lambda s_n} \frac{s_n^{n-1}}{(n-1)!} = \frac{\lambda^n}{(n-1)!} s_n^{n-1} e^{-\lambda s_n}.$$

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

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

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 :

$$F_W(w) = \underbrace{\mathbb{P}(\underbrace{X + Y}_{\text{red}} \leq \underbrace{w}_{\text{red}})} = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{w-x} p(x,y) dy \right) dx.$$

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

$$\underline{\star\star} = \int_{-\infty}^{\infty} p_X(x) \mathbb{P}(\underbrace{x + Y}_{\text{red}} \leq \underbrace{w}_{\text{red}} | X = x) dx.$$

- 做变量替换 $z = x + y$.

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

- 求导:

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