

概 率 论

Probability

已学知识点

● 第一章 事件与概率

▶ 随机现象与统计规律性

- ① 概率的频率解释依然是当今最通行的解释.
- ② 描述频率趋近于概率的大数定律总是概率论的第一大数定律.
- ③ 实际当中用频率作为概率的估计是十分自然的.

▶ 样本空间与事件

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω
$A \subset B$	A 是 B 的子集	A 发生则 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A - B$	差集	A 发生而 B 不发生
\bar{A}	余集	A 不发生

已学知识点

● 第一章 事件与概率

- ▶ 古典概型 (等可能概率模型): (1) 样本空间样本点有限; (2) 每个样本点等可能出现.
 - 计数方法: 排列组合.
 - 三个基本性质: 非负性、规范性、有限可加性.
- ▶ 几何概率: 以等可能性定义概率, 处理无限场合, 概率是几何体的测度之比.
 - 三个基本性质: 非负性、规范性、可列可加性.
- ▶ 概率空间: (Ω, \mathcal{F}, P)
 - 难点和要点: 事件域 \mathcal{F} 的选择, 太小不能满足需要, 太大难以定义概率.
 - 选择包含我们关注的所有事件的 σ 域, 保证事件对交、并、逆、差作可列次运算的封闭性.
 - 在这种 σ 域上, 能定义满足非负、规范和可列可加性的概率测度.

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 条件概率: $P(A|B) = \frac{P(AB)}{P(B)}$.

- 乘法公式: $P(A_1A_2\cdots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1A_2) \cdot \cdots \cdot P(A_n|A_1A_2\cdots A_{n-1})$

- 全概率公式: $P(B) = \sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i)$

- Bayes 公式: $P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i)}$

$$\left. \begin{array}{l} A_i \cap A_j = \Phi \quad (i \neq j) \\ \bigcup_{i=1}^{\infty} A_i = \Omega \end{array} \right\}$$

已学知识点

● 第二章 条件概率与统计独立性

- ▶ 事件独立性：两个事件 $P(AB) = P(A) \cdot P(B)$. 三个事件
- $$\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$$

- ▶ 试验独立性：一个试验的结果对其它各试验的可能结果的概率都无影响.

- ▶ Bernoulli 试验 E : 概率空间 (Ω, \mathcal{F}, P) , 其中

$$A \subset \Omega, \quad \mathcal{F} = \{\Phi, A, \bar{A}, \Omega\}, \quad P(A) = p, \quad P(\bar{A}) = q, \quad (p > 0, q > 0, p + q = 1)$$

○ Bernoulli 分布

○ 二项分布

○ 几何分布

○ Pascal 分布

○ 多项分布

○ Poisson 分布

已学知识点

- 第三章 随机变量与分布函数

- ▶ **随机变量与分布函数**: 定义于概率空间 (Ω, \mathcal{F}, P) 上的单值实函数 $\xi(\omega)$, 若对直线上任一 Borel 点集 B , 有 $\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$, 则称 $\xi(\omega)$ 为**随机变量**, 称 $F(x) = P\{\xi(\omega) \leq x\}$, $-\infty < x < \infty$ 为 $\xi(\omega)$ 的**(累积) 分布函数**. 记作 $\xi(\omega) \sim F(x)$.

- 分布函数的性质:

- ① **单调性**: 若 $a < b$, 则 $F(a) \leq F(b)$.

- ② **有界性**: $0 \leq F(x) \leq 1$, 且 $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$, $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$.

- ③ **右连续性**: $F(x+0) = \lim_{t \rightarrow x+0} F(t) = F(x)$.

- ④ **分布函数可以唯一确定概率分布**.

已学知识点

- 第三章 随机变量与分布函数

- ▶ **随机变量与分布函数**: 定义于概率空间 (Ω, \mathcal{F}, P) 上的单值实函数 $\xi(\omega)$, 若对直线上任一 Borel 点集 B , 有 $\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$, 则称 $\xi(\omega)$ 为**随机变量**, 称 $F(x) = P\{\xi(\omega) \leq x\}$, $-\infty < x < \infty$ 为 $\xi(\omega)$ 的**(累积) 分布函数**. 记作 $\xi(\omega) \sim F(x)$.

- 利用分布函数计算概率:

- ① $P\{\xi(\omega) \leq a\} = F(a)$

- ② $P\{\xi(\omega) = a\} = F(a) - F(a - 0)$

- ③ $P\{\xi(\omega) \geq a\} = 1 - F(a - 0)$

- ④ $P\{\xi(\omega) > a\} = 1 - F(a)$

- ⑤ $P\{a < \xi(\omega) \leq b\} = F(b) - F(a)$

已学知识点

● 第三章 随机变量与分布函数

▶ **离散型随机变量**: 随机变量 ξ 的全部可能取值为有限个或可列无限多个 $x_1, x_2, \dots, x_k, \dots$

○ 分布律 (列): $P\{\xi = x_k\} \triangleq p_k, k = 1, 2, \dots$

ξ	x_1	x_2	\dots	x_k	\dots
p_k	p_1	p_2	\dots	p_k	\dots

○ 已知分布列求分布函数: $F(x) = \sum_{x_k \leq x} p_k, -\infty < x < +\infty$

$p_k \geq 0, \sum_{k=1}^{\infty} p_k = 1$

○ 已知分布函数求分布列: $P\{\xi = x\} = F(x) - F(x-0), \forall x \in \mathbb{R}$

○ 常见的离散型分布:

① 退化分布

② Bernoulli 分布

③ 二项分布

④ 超几何分布

⑤ Poisson 分布

⑥ 几何分布

⑦ Pascal 分布

⑧ 负二项分布

已学知识点

- 第三章 随机变量与分布函数

- ▶ **连续型随机变量**: 存在非负可积函数 $p(x)$ (概率密度函数), 使得 $F(x) = \int_{-\infty}^x p(t) dt$ $p(x) \geq 0, \int_{-\infty}^{+\infty} p(x)dx = 1$

- 分布函数 $F(x)$ 绝对连续、几乎处处可导: $p(x) = F'(x) = \frac{d}{dx}F(x)$, for a.e. $x \in \mathbb{R}$

- 概率计算: 对 \mathbb{R} 中任一 Borel 点集 B 有 $P\{\xi \in B\} = \int_B p(x) dx$.

- 常见的连续型分布:

- ① 均匀分布

- ② 正态分布

- ③ 指数分布

- ④ Erlang 分布

- ⑤ Γ 分布

已学知识点

● 第三章 随机变量与分布函数

▶ **随机向量**: $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$ 是定义在同一概率空间 (Ω, \mathcal{F}, P) 上的随机变量, 称

$\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega))$ 为 **n 维随机向量**, 称

$F(x_1, x_2, \dots, x_n) = P\{\xi_1(\omega) \leq x_1, \xi_2(\omega) \leq x_2, \dots, \xi_n(\omega) \leq x_n\}$ 为 **(联合) 分布函数**.

○ (联合) 分布函数的性质:

① **单调性**: $F(x_1, x_2, \dots, x_n)$ 关于每个一元变量 x_i 都是单调不降函数.

② **规范性**: $F(+\infty, +\infty, \dots, +\infty) = 1$, $F(-\infty, x_2, x_3, \dots, x_n) = 0$,

$$F(x_1, -\infty, x_3, \dots, x_n) = 0, \dots, F(x_1, x_2, \dots, x_{n-1}, -\infty) = 0$$

③ **右连续性**: $F(x_1, x_2, \dots, x_n)$ 关于每个一元变量 x_i 都是右连续的函数.

④ **$n = 2$ 时**: $\forall a_1 \leq b_1, a_2 \leq b_2$ 有 $F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \geq 0$.

已学知识点

● 第三章 随机变量与分布函数

▶ **随机向量**: $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$ 是定义在同一概率空间 (Ω, \mathcal{F}, P) 上的随机变量, 称

$\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega))$ 为 **n 维随机向量**, 称

$F(x_1, x_2, \dots, x_n) = P\{\xi_1(\omega) \leq x_1, \xi_2(\omega) \leq x_2, \dots, \xi_n(\omega) \leq x_n\}$ 为 **(联合) 分布函数**.

○ 常见的多元离散型分布:

① 联合分布律: $p\{x_1, x_2, \dots, x_n\} = P\{\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n\}$

② 概率计算: $P\{\xi \in D\} = \sum_{x \in D} p(x_1, x_2, \dots, x_n)$

③ 多项分布: $P(\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_r = k_r) = \frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$

④ 多元超几何分布: $P(\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_r = n_r) = \frac{\binom{N_1}{n_1} \binom{N_2}{n_2} \dots \binom{N_r}{n_r}}{\binom{N}{n}}$

$\xi \backslash \eta$	y_1	y_2	\dots	y_j	\dots
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	\dots	$p(x_1, y_j)$	\dots
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	\dots	$p(x_2, y_j)$	\dots
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$	\dots	$p(x_i, y_j)$	\dots
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots

已学知识点

● 第三章 随机变量与分布函数

▶ **随机向量**: $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$ 是定义在同一概率空间 (Ω, \mathcal{F}, P) 上的随机变量, 称

$\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega))$ 为 **n 维随机向量**, 称

$F(x_1, x_2, \dots, x_n) = P\{\xi_1(\omega) \leq x_1, \xi_2(\omega) \leq x_2, \dots, \xi_n(\omega) \leq x_n\}$ 为 **(联合) 分布函数**.

○ 常见的多元连续型分布:

① 联合 p.d.f.: 非负可积函数 $p(x_1, \dots, x_n)$ 使 $F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(y_1, y_2, \dots, y_n) dy_1 \cdots dy_n$

② 概率计算: $P\{\xi \in D\} = \int_D \cdots \int p(x_1, \dots, x_n) dx_1 \cdots dx_n$

③ 多元均匀分布: $p(x_1, x_2, \dots, x_n) = \begin{cases} \frac{1}{S}, & (x_1, x_2, \dots, x_n) \in G \\ 0, & (x_1, x_2, \dots, x_n) \notin G \end{cases}$

④ 多元正态分布: $p(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}) \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})^T\right\}$, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\boldsymbol{\mu} \in \mathbb{R}^n$, $\Sigma > 0$

已学知识点

● 第三章 随机变量与分布函数

▶ 边际分布：仅讨论二维情形

$$p(x_i, y_j) \geq 0, \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p(x_i, y_j) = 1$$

○ **离散型**：联合分布律 $p(x_i, y_j) \triangleq P\{\xi = x_i, \eta = y_j\}$, $i, j = 1, 2, \dots$

○ ξ, η 的**边际 (边缘) 分布**：

○ 若 (ξ, η) 的联合分布函数为 $F(x, y)$, 则

$$F_{\xi}(x) = F(x, +\infty), \quad F_{\eta}(y) = F(+\infty, y)$$

$\xi \backslash \eta$	y_1	y_2	\dots	y_j	\dots	$p_1(\cdot)$
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	\dots	$p(x_1, y_j)$	\dots	$p_1(x_1)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	\dots	$p(x_2, y_j)$	\dots	$p_1(x_2)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$	\dots	$p(x_i, y_j)$	\dots	$p_1(x_i)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
$p_2(\cdot)$	$p_2(y_1)$	$p_2(y_2)$	\dots	$p_2(y_j)$	\dots	1

已学知识点

- 第三章 随机变量与分布函数

- ▶ 边际分布：仅讨论二维情形

- 连续型： $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y p(u, v) du dv$, $p(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$, a.e.

- 概率计算： $P \left\{ (\xi, \eta) \in D \right\} = \iint_D p(x, y) dx dy$

- 边际概率密度函数： $p_{\xi}(x) = \int_{-\infty}^{+\infty} p(x, y) dy$, $p_{\eta}(y) = \int_{-\infty}^{+\infty} p(x, y) dx$

已学知识点

- 第三章 随机变量与分布函数

- ▶ 条件分布：仅讨论二维情形

- 离散型： η 关于 ξ 的条件分布 $P\left\{\eta = y_j \mid \xi = x_i\right\} = \frac{p(x_i, y_j)}{p_1(x_i)}, j = 1, 2, 3, \dots$

ξ 关于 η 的条件分布 $P\left\{\xi = x_i \mid \eta = y_j\right\} = \frac{p(x_i, y_j)}{p_2(y_j)}, i = 1, 2, 3, \dots$

$\xi \backslash \eta$	y_1	y_2	\dots	y_j	\dots	$p_1(\cdot)$
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	\dots	$p(x_1, y_j)$	\dots	$p_1(x_1)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	\dots	$p(x_2, y_j)$	\dots	$p_1(x_2)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$	\dots	$p(x_i, y_j)$	\dots	$p_1(x_i)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
$p_2(\cdot)$	$p_2(y_1)$	$p_2(y_2)$	\dots	$p_2(y_j)$	\dots	1

3.2 随机向量， 随机变量的独立性 (续)

一、随机向量及其分布

二、边际 (边缘) 分布

三、条件分布

四、随机变量的独立性

三、条件 (conditional) 分布

- **连续情形**：对连续型随机向量 (ξ, η) ，我们来定义给定 $\xi = x$ 时 η 的条件分布函数

$$F(y|x) \triangleq P\{\eta \leq y | \xi = x\} = \frac{P\{\xi = x, \eta \leq y\}}{P\{\xi = x\}}$$

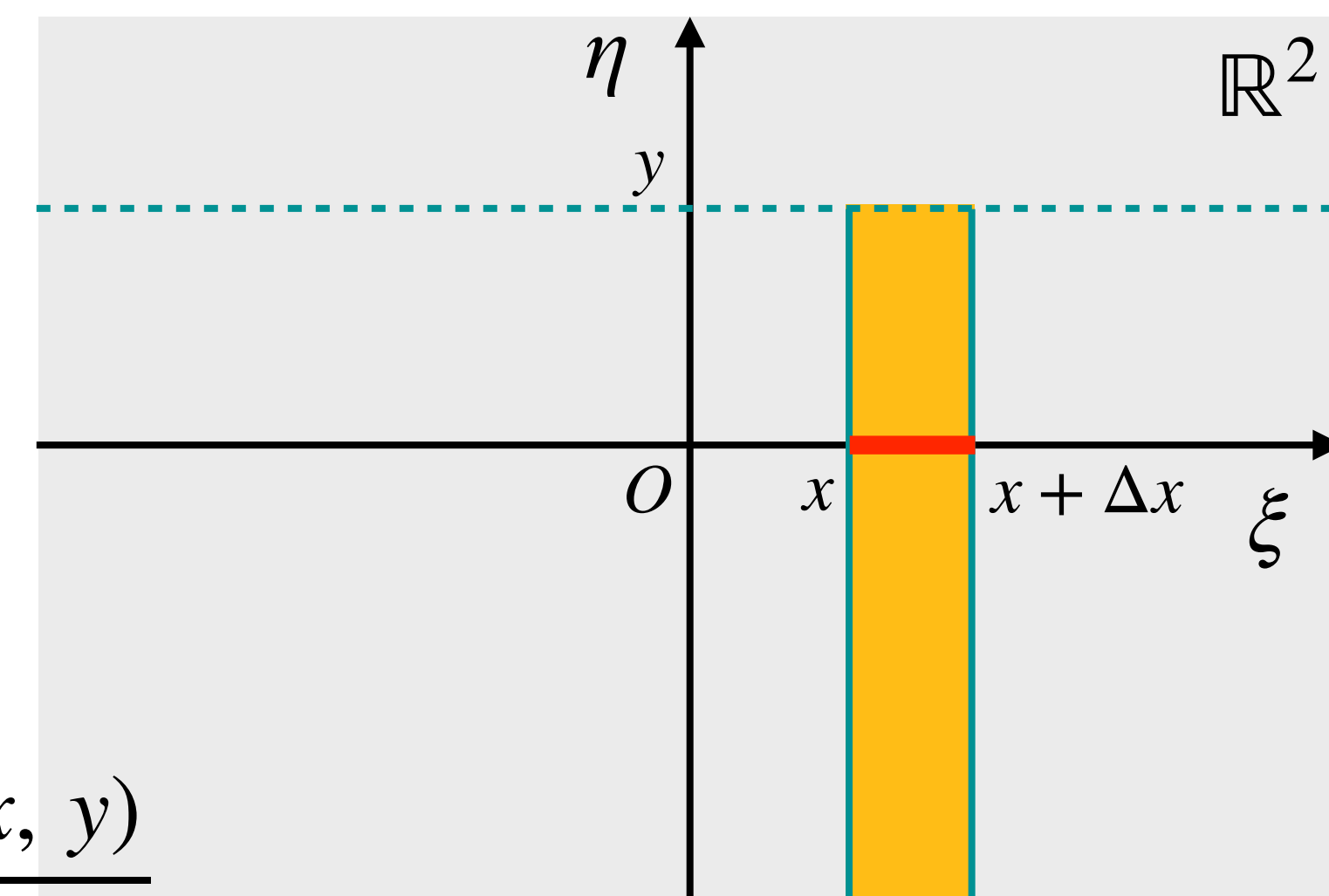
→ $P\{\xi = x\} = 0$, 无法定义!

$$= \lim_{\Delta x \rightarrow 0} P\{\eta \leq y | x \leq \xi < x + \Delta x\} = \lim_{\Delta x \rightarrow 0} \frac{P\{x \leq \xi < x + \Delta x, \eta \leq y\}}{P\{x \leq \xi < x + \Delta x\}}$$

$$= \lim_{\Delta x \rightarrow 0} \frac{\int_x^{x+\Delta x} \left(\int_{-\infty}^y p(u, v) dv \right) du}{\int_x^{x+\Delta x} p_\xi(u) du}$$

洛必达法则 →

$$= \lim_{\Delta x \rightarrow 0} \frac{\int_{-\infty}^y p(x + \Delta x, v) dv}{p_\xi(x + \Delta x)}$$



若 $p(x, y)$ 为连续函数且 $p_\xi(x) \neq 0$ →

$$= \int_{-\infty}^y \frac{p(x, v)}{p_\xi(x)} dv \implies p(y|x) = \frac{d}{dy} F(y|x) = \frac{p(x, y)}{p_\xi(x)}$$

→ $p_\xi(x) \neq 0$

三、条件 (conditional) 分布

- **连续情形**：对连续型随机向量 (ξ, η) ，给定 $\xi = x$ 的条件下， η 的条件密度函数为

$$p(y|x) = \frac{p(x, y)}{p_\xi(x)}, \quad p_\xi(x) \neq 0$$

同理，给定 $\eta = y$ 的条件下， ξ 的条件密度函数为

$$p(x|y) = \frac{p(x, y)}{p_\eta(y)}, \quad p_\eta(y) \neq 0$$

三、条件 (conditional) 分布

- 连续情形 $\implies (\eta | \xi = x) \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2)\right)$
- 例: 二元正态分布 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的条件分布仍然是正态分布

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

$$\begin{cases} \xi \sim N(\mu_1, \sigma_1^2) \\ \eta \sim N(\mu_2, \sigma_2^2) \end{cases}, \quad p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}, \quad p_\eta(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}$$

$$p(y|x) = \frac{p(x, y)}{p_\xi(x)} = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{\left[y - \left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1)\right)\right]^2}{2\sigma_2^2 (1 - \rho^2)}\right\}$$

三、条件 (conditional) 分布

- 连续情形

$$\Rightarrow \left(\begin{array}{l} (\eta | \xi = x) \sim N \left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2) \right) \\ (\xi | \eta = y) \sim N \left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2), \sigma_1^2 (1 - \rho^2) \right) \end{array} \right)$$

- 例: 二元正态分布 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的条件分布仍然是正态分布

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right] \right\}$$

$$\left\{ \begin{array}{l} \xi \sim N(\mu_1, \sigma_1^2) \\ \eta \sim N(\mu_2, \sigma_2^2) \end{array} \right., \quad p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left\{ -\frac{(x-\mu_1)^2}{2\sigma_1^2} \right\}, \quad p_\eta(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left\{ -\frac{(y-\mu_2)^2}{2\sigma_2^2} \right\}$$

$$p(x|y) = \frac{p(x, y)}{p_\eta(y)} = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp \left\{ -\frac{\left[x - \left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2) \right) \right]^2}{2\sigma_1^2 (1 - \rho^2)} \right\}$$

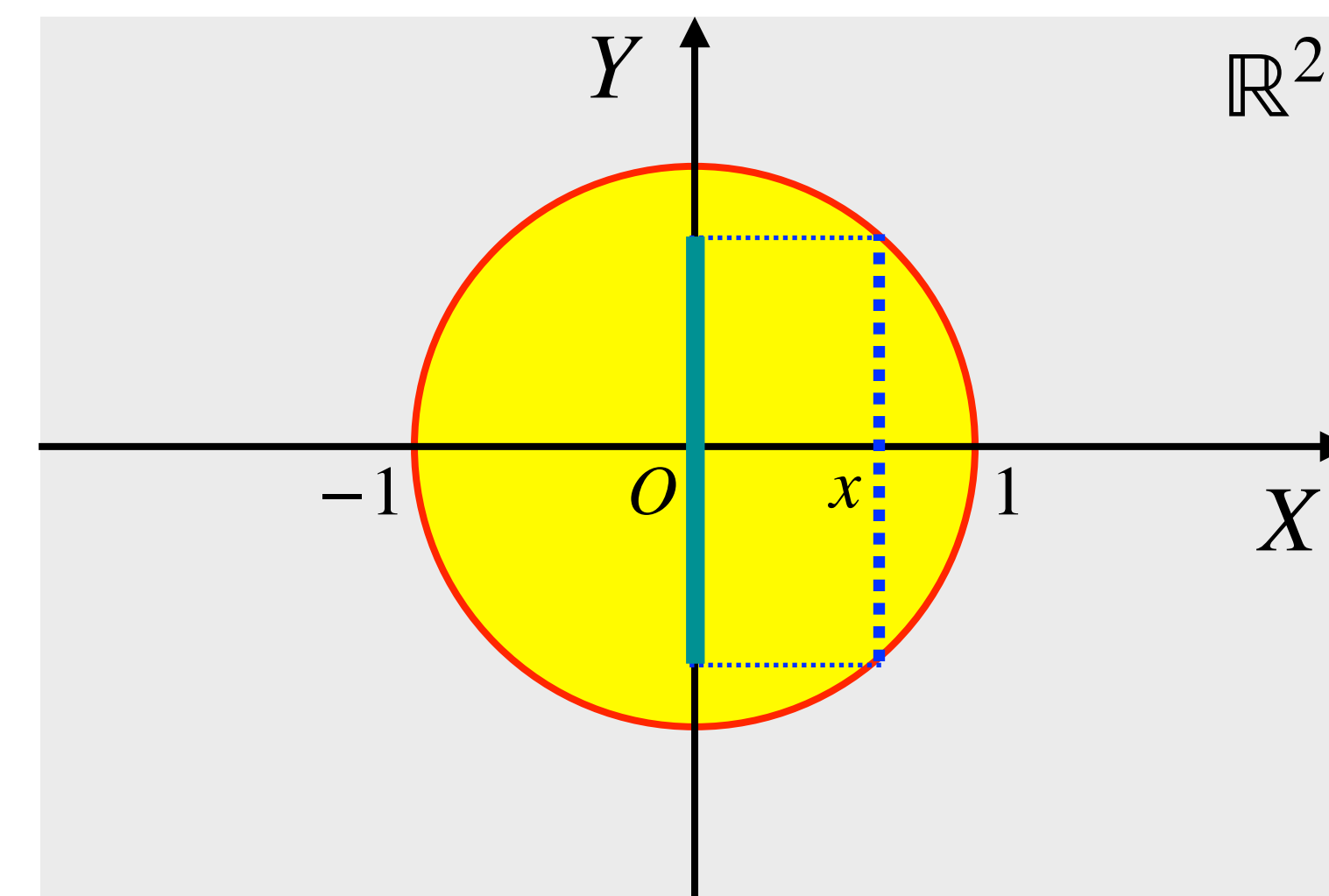
三、条件 (conditional) 分布

- 连续情形 $\implies (\eta | \xi = x) \sim U(-\sqrt{1-x^2}, \sqrt{1-x^2})$, $(\xi | \eta = y) \sim U(-\sqrt{1-y^2}, \sqrt{1-y^2})$
- 例: 二维随机向量 (X, Y) 服从区域 $D = \{(x, y) | x^2 + y^2 \leq 1\}$ 上的均匀分布.

$$p(x, y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \leq 1 \\ 0, & \text{其它} \end{cases}$$

$$p_X(x) = \begin{cases} \frac{2}{\pi} \sqrt{1-x^2}, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases}, \quad p_Y(y) = \begin{cases} \frac{2}{\pi} \sqrt{1-y^2}, & |y| \leq 1 \\ 0, & |y| > 1 \end{cases}$$

$$p(y | x) = \frac{p(x, y)}{p_X(x)} = \begin{cases} \frac{1}{2\sqrt{1-x^2}}, & -\sqrt{1-x^2} < y < \sqrt{1-x^2} \\ 0, & \text{其它} \end{cases}$$



四、随机变量的独立性

- **定义**: 设 $\xi_1, \xi_2, \dots, \xi_n$ 是定义在概率空间 (Ω, \mathcal{F}, P) 上的 n 个随机变量, 若对于任意的

x_1, x_2, \dots, x_n 都有

$$P \{ \xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n \} = P \{ \xi_1 \leq x_1 \} \cdot P \{ \xi_2 \leq x_2 \} \cdot \dots \cdot P \{ \xi_n \leq x_n \}$$

成立, 则称随机变量 $\xi_1, \xi_2, \dots, \xi_n$ **相互独立** (independent).

$$F(x_1, x_2, \dots, x_n) = F(x_1) \cdot F(x_2) \cdot \dots \cdot F(x_n)$$

边缘分布唯一确定联合分布

- ▶ 对于离散型随机变量 $\xi_1, \xi_2, \dots, \xi_n$, 它们**相互独立**的定义等价于对于任意一组可能取值

x_1, x_2, \dots, x_n 都有

$$P \{ \xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n \} = P \{ \xi_1 = x_1 \} \cdot P \{ \xi_2 = x_2 \} \cdot \dots \cdot P \{ \xi_n = x_n \}$$

四、随机变量的独立性

- ▶ 对于离散型随机变量 $\xi_1, \xi_2, \dots, \xi_n$, 当 $n = 2$ 时, 则要求对任意的可能取值 x, y 都有

$$P\{\xi_1 = x, \xi_2 = y\} = P\{\xi_1 = x\} \cdot P\{\xi_2 = y\}$$

$$\begin{aligned} \forall (x = x_i, y = y_j) &\rightarrow P\{\xi_1 = x_i, \xi_2 = y_j\} \\ &= p(x_i, y_j) \\ &= p_1(x_i) \cdot p_2(y_j) \end{aligned}$$

$\xi_1 \backslash \xi_2$	y_1	y_2	\dots	y_j	\dots	$p_1(\bullet)$
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	\dots	$p(x_1, y_j)$	\dots	$p_1(x_1)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	\dots	$p(x_2, y_j)$	\dots	$p_1(x_2)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$	\dots	$p(x_i, y_j)$	\dots	$p_1(x_i)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
$p_2(\bullet)$	$p_2(y_1)$	$p_2(y_2)$	\dots	$p_2(y_j)$	\dots	1

四、随机变量的独立性

- 定义:** 设 $\xi_1, \xi_2, \dots, \xi_n$ 是定义在概率空间 (Ω, \mathcal{F}, P) 上的 n 个随机变量, 若对于任意的 x_1, x_2, \dots, x_n 都有

$$P\{\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n\} = P\{\xi_1 \leq x_1\} \cdot P\{\xi_2 \leq x_2\} \cdot \dots \cdot P\{\xi_n \leq x_n\}$$

成立, 则称随机变量 $\xi_1, \xi_2, \dots, \xi_n$ **相互独立** (independent).

$$F(x_1, x_2, \dots, x_n) = F(x_1) \cdot F(x_2) \cdot \dots \cdot F(x_n)$$

边缘分布唯一确定联合分布

- ▶ 此时条件分布化为无条件分布**

$$P(\eta \leq y \mid \xi = x) = \frac{P(\xi = x, \eta \leq y)}{P(\xi = x)} = \frac{P(\xi = x) \cdot P(\eta \leq y)}{P(\xi = x)} = P(\eta \leq y)$$

四、随机变量的独立性

▶ 例：袋中装有 2 只白球及 3 只黑球，定义

$$\xi = \begin{cases} 1, & \text{第 1 次摸出白球} \\ 0, & \text{第 1 次摸出黑球} \end{cases}, \quad \eta = \begin{cases} 1, & \text{第 2 次摸出白球} \\ 0, & \text{第 2 次摸出黑球} \end{cases}$$

则 (ξ, η) 的联合概率分布与边际分布为

有放回摸球

$\eta \backslash \xi$	0	1	$p_2(y_j)$
0	$\frac{3}{5} \cdot \frac{3}{5}$	$\frac{2}{5} \cdot \frac{3}{5}$	$\frac{3}{5}$
1	$\frac{3}{5} \cdot \frac{2}{5}$	$\frac{2}{5} \cdot \frac{2}{5}$	$\frac{2}{5}$
$p_1(x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	1

$\implies \xi, \eta$ 相互独立

不放回摸球

$\eta \backslash \xi$	0	1	$p_2(y_j)$
0	$\frac{3}{5} \cdot \frac{2}{4}$	$\frac{2}{5} \cdot \frac{3}{4}$	$\frac{3}{5}$
1	$\frac{3}{5} \cdot \frac{1}{4}$	$\frac{2}{5} \cdot \frac{2}{4}$	$\frac{2}{5}$
$p_1(x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	1

$\implies \xi, \eta$ 不相互独立

四、随机变量的独立性

- **定义**: 设 $\xi_1, \xi_2, \dots, \xi_n$ 是定义在概率空间 (Ω, \mathcal{F}, P) 上的 n 个随机变量, 若对于任意的 x_1, x_2, \dots, x_n 都有

$$P\{\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n\} = P\{\xi_1 \leq x_1\} \cdot P\{\xi_2 \leq x_2\} \cdot \dots \cdot P\{\xi_n \leq x_n\}$$

成立, 则称随机变量 $\xi_1, \xi_2, \dots, \xi_n$ **相互独立** (independent).

$$F(x_1, x_2, \dots, x_n) = F(x_1) \cdot F(x_2) \cdot \dots \cdot F(x_n)$$

边缘分布唯一确定联合分布

- ▶ 对于连续型随机变量 $\xi_1, \xi_2, \dots, \xi_n$, 它们**相互独立**的定义等价于对于任意的 x_1, x_2, \dots, x_n 几乎处处成立

$$p(x_1, x_2, \dots, x_n) = p(x_1) \cdot p(x_2) \cdot \dots \cdot p(x_n)$$

$\xi_1, \xi_2, \dots, \xi_n$ 的联合概率密度

$\xi_1, \xi_2, \dots, \xi_n$ 各自的边缘概率密度

四、随机变量的独立性

▶ 例: 对于 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, 其联合概率密度函数为

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right] \right\}$$

边缘分布 \rightarrow
$$\begin{cases} \xi \sim N(\mu_1, \sigma_1^2) \\ \eta \sim N(\mu_2, \sigma_2^2) \end{cases}, \quad p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right), \quad p_\eta(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right)$$

\Rightarrow X, Y 相互独立的充分必要条件是 $\rho = 0$

条件分布 \rightarrow
$$\begin{cases} (\eta | \xi = x) \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2)\right) \\ (\xi | \eta = y) \sim N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2), \sigma_1^2 (1 - \rho^2)\right) \end{cases} \xrightarrow{\rho=0} \begin{cases} (\eta | \xi = x) \sim N(\mu_2, \sigma_2^2) \\ (\xi | \eta = y) \sim N(\mu_1, \sigma_1^2) \end{cases}$$

四、随机变量的独立性

▶ 例: 设 (ξ, η) 服从矩形 $G = \{(x, y) ; a \leq x \leq b, c \leq y \leq d\}$ 上的均匀分布,

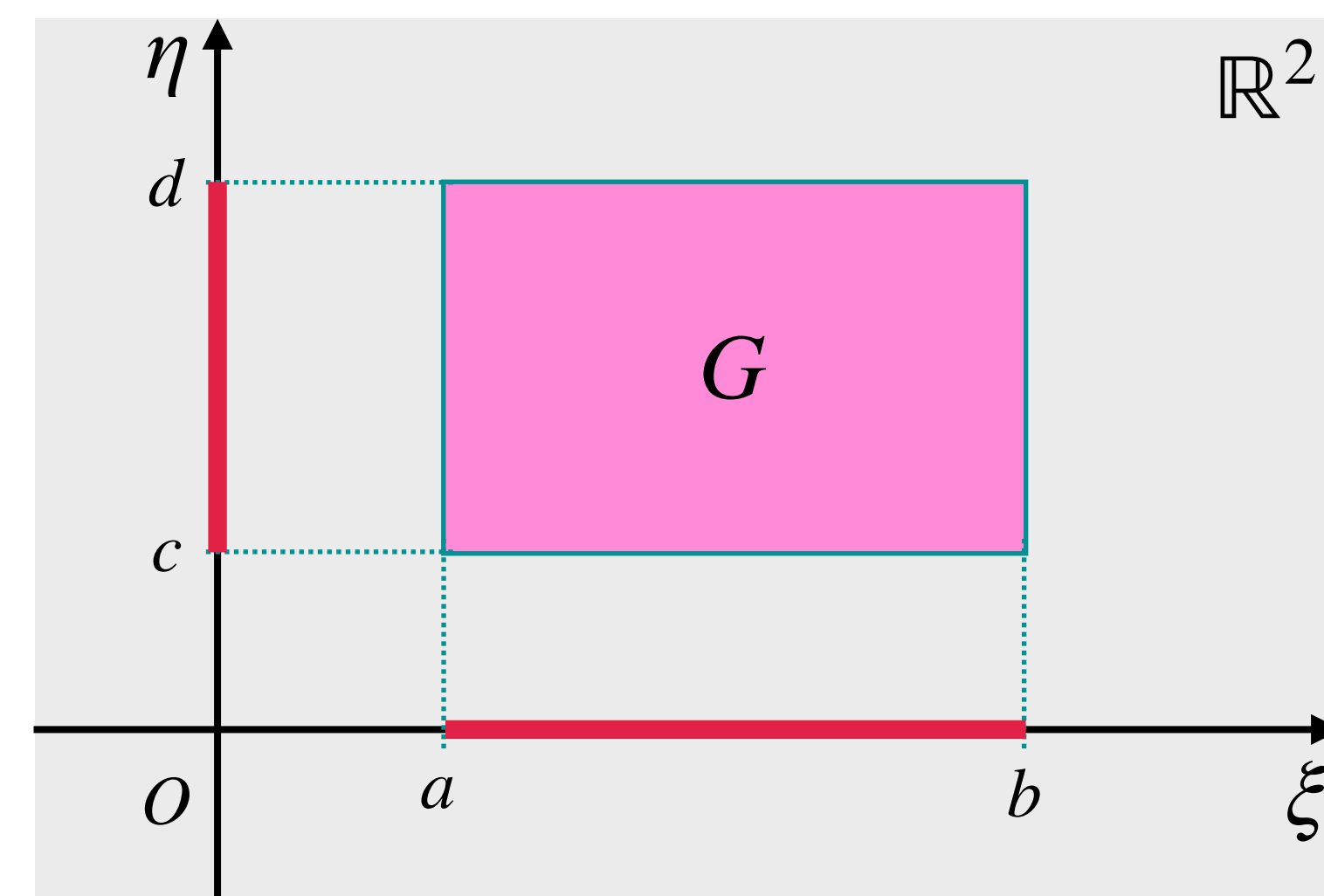
则其联合概率密度函数为
$$p(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)}, & a \leq x \leq b, c \leq y \leq d \\ 0, & \text{其它} \end{cases}$$

$\Rightarrow \xi \sim U[a, b], \eta \sim U[c, d]$

且 ξ, η 相互独立

反之, 若 $\xi \sim U[a, b], \eta \sim U[c, d]$, 且它们相互独立,

则 (ξ, η) 服从矩形 G 上的均匀分布.



四、随机变量的独立性

定理: 如果随机变量 $\xi_1, \xi_2, \dots, \xi_n$ 相互独立, 则其中任何一部分随机变量仍独立.

- 于是, 整体独立的多个随机变量是两两独立的, 但其逆命题不成立.
- 拓展到无穷多个随机变量的情形: 一族无穷多个随机变量称作相互独立, 是指其中任意有限个随机变量均相互独立.

四、随机变量的独立性

- 我们给出二维情形 ($n = 2$) 中独立性定义的等价形式

定理: 随机变量 ξ, η 相互独立当且仅当 $\forall B_1, B_2 \in \mathcal{B}_1$ 都有

$$P \{ \xi \in B_1, \eta \in B_2 \} = P \{ \xi \in B_1 \} \cdot P \{ \eta \in B_2 \}.$$

定理: 若 (ξ, η) 为离散型随机向量, 则 ξ, η 相互独立当且仅当 $\forall i, j = 1, 2, 3, \dots$ 都有

$$p(x_i, y_j) = p_\xi(x_i) \cdot p_\eta(y_j).$$

定理: 若 (ξ, η) 为连续型随机向量, 则 ξ, η 相互独立当且仅当 $\forall x, y \in \mathbb{R}$ 都有

$$p(x, y) = p_\xi(x) \cdot p_\eta(y), \text{ a.e.}$$

定理: 若随机变量 ξ, η 相互独立, 则条件分布化为无条件分布 $P \{ \eta \leq y \mid \xi = x \} = P \{ \eta \leq y \}.$

四、随机变量的独立性

- 随机向量之间的独立性:

- 定义: 对于 n 维随机向量 ξ 和 m 维随机向量 η , 如果 $\forall A \in \mathcal{B}_n, B \in \mathcal{B}_m$ 都有

$$P\{\xi \in A, \eta \in B\} = P\{\xi \in A\} \cdot P\{\eta \in B\}$$

则称 ξ 与 η 相互独立.

- ▶ 显然, 若 ξ 与 η 相互独立, 则 ξ 的子向量与 η 的子向量亦相互独立.

n 维 Borel 点集

m 维 Borel 点集

3.3 随机变量的函数及其分布

- 一、Borel 函数与随机变量的函数
- 二、单个随机变量的函数的分布
- 三、随机向量的函数的分布
- 四、随机向量的变换
- 五、随机变量的函数的独立性

一、Borel 函数与随机变量的函数

定义： 设 $y = g(x)$ 是 \mathbb{R}^1 到 \mathbb{R}^1 上的一个映射，若对于一切 \mathbb{R}^1 中的 Borel 点集 B_1 ，都有

$$\{x : g(x) \in B_1\} \in \mathcal{B}_1$$

其中 \mathcal{B}_1 为 \mathbb{R}^1 上的 Borel σ 域，则称 $g(x)$ 是一元 Borel (可测) 函数.

- ▶ Borel 函数是很广泛的一类函数，我们所碰到的大部分函数都是 Borel 函数.
 - 所有的分段连续函数都是 Borel 函数.
 - 所有的分段单调函数都是 Borel 函数.

一、Borel 函数与随机变量的函数

命题：若 ξ 是概率空间 (Ω, \mathcal{F}, P) 上的随机变量，而 $g(x)$ 是一元 Borel 函数，则 $g(\xi)$ 是 (Ω, \mathcal{F}, P) 上的随机变量.

$$\forall B_1 \in \mathcal{B}_1, \quad g(x) \text{ 是一元 Borel 函数} \implies A_1 \triangleq \{x : g(x) \in B_1\} \in \mathcal{B}_1$$

• **定义：**设 $\xi(\omega)$ 是定义于概率空间 (Ω, \mathcal{F}, P) 上的单值实函数，如果对于直线上任一 Borel 点集 B ，有

$$\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$$

则称 $\xi(\omega)$ 为**随机变量**，而 $P(\xi(\omega) \in B)$ 称为随机变量 $\xi(\omega)$ 的**概率分布**.

$$\implies \{\omega : \xi(\omega) \in A_1\} \in \mathcal{F}$$

$$\implies \{\omega : g(\xi(\omega)) \in B_1\} = \{\omega : \xi(\omega) \in A_1\} \in \mathcal{F}$$

$$\implies g(\xi) \text{ 是 } (\Omega, \mathcal{F}, P) \text{ 上的随机变量}$$

一、Borel 函数与随机变量的函数

- 随机向量的函数

定义： 设 $y = g(x_1, x_2, \dots, x_n)$ 是 \mathbb{R}^n 到 \mathbb{R}^1 上的一个映射，若对于一切 \mathbb{R}^1 中的 Borel 点集 B_1 都有

$$\left\{ (x_1, x_2, \dots, x_n) : g(x_1, x_2, \dots, x_n) \in B_1 \right\} \in \mathcal{B}_n$$

则称 $g(x_1, x_2, \dots, x_n)$ 是 **n 元 Borel (可测) 函数**.

命题： 若 $(\xi_1, \xi_2, \dots, \xi_n)$ 是概率空间 (Ω, \mathcal{F}, P) 上的随机向量，而 $g(x_1, x_2, \dots, x_n)$ 是 n 元 Borel 函数，则 $g(\xi_1, \xi_2, \dots, \xi_n)$ 是 (Ω, \mathcal{F}, P) 上的随机变量.

► 更一般地， n 维随机向量 $(\xi_1, \xi_2, \dots, \xi_n)$ 的 m 个函数

$$g_1(\xi_1, \xi_2, \dots, \xi_n), g_2(\xi_1, \xi_2, \dots, \xi_n), \dots, g_m(\xi_1, \xi_2, \dots, \xi_n)$$

g_1, g_2, \dots, g_m 都是 n 元 Borel 函数

则可构成一个 m 维的随机向量 (g_1, g_2, \dots, g_m) .

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

▶ 设 ξ 的分布列为

ξ	x_1	x_2	\dots	x_n	\dots
$p_i = P\{\xi = x_i\}$	p_1	p_2	\dots	p_n	\dots

则 $g(\xi)$ 的分布列为

$g(\xi)$	$g(x_1)$	$g(x_2)$	\dots	$g(x_n)$	\dots
$P\{g(\xi) = g(x_i)\}$	p_1	p_2	\dots	p_n	\dots

如果某些 $g(x_i)$ 相等, 将它们作适当并项且将对应的概率 p_i 相加即可.

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

▶ 例：随机变量 ξ 的分布列为

ξ	0	1	2	3	4	5
p_i	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{9}$	$\frac{1}{9}$

则 $\eta = 2\xi + 1$ 的分布列为

$\eta = 2\xi + 1$	1	3	5	7	9	11
$p_i = P\{\eta = y_i\}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{9}$	$\frac{1}{9}$

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

▶ 例：随机变量 ξ 的分布列为

ξ	0	1	2	3	4	5
p_i	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{9}$	$\frac{1}{9}$

则 $\zeta = (\xi - 2)^2$ 的分布列为

$\zeta = (\xi - 2)^2$	0	1	4	9
$p_i = P\{\zeta = z_i\}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{11}{36}$	$\frac{1}{9}$

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

▶ 二元情形的表上作业法：一般分三步进行，举例如下

▶ 例：设 (ξ, η) 的联合分布列为

$\xi \backslash \eta$	-1	0	2
0	0.1	0.2	0
1	0.3	0.05	0.1
2	0.15	0	0.1

确定 $\zeta = \xi + \eta$ 的分布.

① 将 (ξ, η) 的联合分布列表达如右上表.

② 求每一对 (x_i, y_j) 对应的函数值 $g(x_i, y_j)$ 并标在对应小方框的右下角.

③ 将第二个表格中 $g(x_i, y_j)$ 值相同的项合并
即得 $\zeta = g(\xi, \eta)$ 的分布.

$\xi \backslash \eta$	-1	0	2
0	0.1 -1	0.2 0	0 2
1	0.3 0	0.05 1	0.1 3
2	0.15 1	0 2	0.1 4

$\zeta = \xi + \eta$	-1	0	1	2	3	4
P_i	0.1	0.5	0.2	0	0.1	0.1

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

$$P\{\xi + \eta = r\} = \sum_{k=0}^r P\{\xi = k\} \cdot P\{\eta = r - k\}$$

▶ 离散卷积 (convolution) 公式: 相互独立的随机变量之和的分布.

设 ξ, η 相互独立, 取值均为非负整数, 且概率分布分别为

ξ	x_1	x_2	...	x_k	...	η	y_1	y_2	...	y_k	...
$p_i = P\{\xi = x_i\}$	a_1	a_2	...	a_k	...	$p_i = P\{\eta = y_i\}$	b_1	b_2	...	b_k	...

令 $\zeta = \xi + \eta$, 则

$$\begin{aligned} P\{\zeta = r\} &= P\{\xi + \eta = r\} = P\left\{\{\xi = 0, \eta = r\} \cup \{\xi = 1, \eta = r - 1\} \cup \dots \cup \{\xi = r, \eta = 0\}\right\} \\ &= P\{\xi = 0, \eta = r\} + P\{\xi = 1, \eta = r - 1\} + \dots + P\{\xi = r, \eta = 0\} \\ &= P\{\xi = 0\} \cdot P\{\eta = r\} + P\{\xi = 1\} \cdot P\{\eta = r - 1\} + \dots + P\{\xi = r\} \cdot P\{\eta = 0\} \\ &= a_0 b_r + a_1 b_{r-1} + \dots + a_r b_0 = \sum_{k=0}^r a_k b_{r-k}, \quad r = 0, 1, 2, \dots \end{aligned}$$

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

定理： 若 $\xi \sim B(m, p)$, $\eta \sim B(n, p)$ 相互独立, 则 $\xi + \eta \sim B(m + n, p)$.

$$\xi \sim B(m, p) \implies P\{\xi = k\} = C_m^k p^k (1-p)^{m-k}, \quad k = 0, 1, 2, \dots, m$$

$$\eta \sim B(n, p) \implies P\{\eta = k\} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n$$

卷积公式 \rightarrow

$$P\{\xi + \eta = r\} = \sum_{k=0}^r P\{\xi = k\} \cdot P\{\eta = r - k\} = \sum_{k=0}^r C_m^k p^k (1-p)^{m-k} \cdot C_n^k p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^r C_m^k p^k (1-p)^{m-k} \cdot C_n^{r-k} p^{r-k} (1-p)^{n-(r-k)} = \sum_{k=0}^r C_m^k C_n^{r-k} p^r (1-p)^{m+n-r}$$

$$= C_{m+n}^r p^r (1-p)^{m+n-r}, \quad r = 0, 1, 2, \dots, m+n$$

可以证明: $\sum_{k=0}^r C_m^k C_n^{r-k} = C_{m+n}^r$

$$\implies \xi + \eta \sim B(m + n, p)$$

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

定理：若 $\xi \sim B(m, p)$, $\eta \sim B(n, p)$ 相互独立，则 $\xi + \eta \sim B(m + n, p)$.

▶ 从二项分布的背景很容易理解这一结论:

$\xi \sim B(m, p) \implies$ 独立重复掷一枚硬币 m 次, ξ 为正面向上的次数

$\eta \sim B(n, p) \implies$ 独立重复掷一枚硬币 n 次, η 为正面向上的次数

参数 p 为硬币出现正面向上的概率

ξ, η 相互独立 $\implies \xi + \eta$ 则为独立重复掷一枚硬币 $m + n$ 次当中正面向上的次数

$\implies \xi + \eta \sim B(m + n, p)$

一、Borel 函数与随机变量的函数

- 离散型随机变量的函数的分布

定理： 若 $\xi \sim P(\lambda_1)$, $\eta \sim P(\lambda_2)$ 相互独立, 则 $\xi + \eta \sim P(\lambda_1 + \lambda_2)$.

$$\xi \sim P(\lambda_1) \implies P\{\xi = k\} = \frac{\lambda_1^k}{k!} e^{-\lambda_1}, \quad k = 0, 1, 2, \dots$$

$$\eta \sim P(\lambda_2) \implies P\{\eta = k\} = \frac{\lambda_2^k}{k!} e^{-\lambda_2}, \quad k = 0, 1, 2, \dots$$

卷积公式 \rightarrow

$$\begin{aligned} P\{\xi + \eta = r\} &= \sum_{k=0}^r P\{\xi = k\} \cdot P\{\eta = r - k\} = \sum_{k=0}^r \frac{\lambda_1^k}{k!} e^{-\lambda_1} \cdot \frac{\lambda_2^{r-k}}{(r-k)!} e^{-\lambda_2} \\ &= \sum_{k=0}^r \frac{1}{k! (r-k)!} \lambda_1^k \lambda_2^{r-k} e^{-(\lambda_1 + \lambda_2)} = \frac{1}{r!} \sum_{k=0}^r \frac{r!}{k! (r-k)!} \lambda_1^k \lambda_2^{r-k} e^{-(\lambda_1 + \lambda_2)} \\ &= \left(\sum_{k=0}^r C_r^k \lambda_1^k \lambda_2^{r-k} \right) \frac{1}{r!} e^{-(\lambda_1 + \lambda_2)} = \frac{(\lambda_1 + \lambda_2)^r}{r!} e^{-(\lambda_1 + \lambda_2)}, \quad r = 0, 1, 2, \dots \end{aligned}$$

二、单个随机变量的函数的分布

- **问题**: 已知 ξ 的分布函数 $F(x)$ 或密度函数 $p(x)$, 求 $\eta = g(\xi)$ 的分布函数 $G(y)$ 或密度函数 $q(y)$.

分布函数的定义 $\rightarrow G(y) = P\{\eta \leq y\} = P\{g(\xi) \leq y\} = \int_{g(\xi) \leq y} p(x) dx$

- ▶ 积分计算的难易程度既与被积函数 $p(x)$ 有关, 更与积分区域 $\{x : g(\xi) \leq y\}$ 的形状相关.
- ▶ 具体问题具体分析, 大致可分为两类: 直接法、变换法.