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：$\left(\Omega, \mathscr{F}, P\right)$
    - 难点和要点：事件域 $\mathscr{F}$ 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：$P(A\mid B)=\dfrac{P(AB)}{P(B)}.$
    - 乘法公式：$P\left(A_1A_2\cdots A_n\right)=P\left(A_1\right)\cdot P\left(A_2\middle|A_1\right)\cdot P\left(A_3\middle|A_1A_2\right)\cdot\cdots\cdot P\left(A_n\middle|A_1A_2\cdots A_{n-1}\right)$
    - 全概率公式：$P(B)=\sum_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\middle|A_i\right)$
    - Bayes 公式：$P\left(A_k\middle|B\right)=\dfrac{P\left(A_k\right)\cdot P\left(B\middle|A_k\right)}{\sum_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\middle|A_i\right)}$

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

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 $(\Omega, \mathscr{F}, P)$，其中

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

    - Bernoulli 分布
    - Pascal 分布
    - 二项分布
    - 多项分布
    - 几何 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：设 $\xi(\omega)$ 是定义于概率空间 $(\Omega, \mathscr{F}, P)$ 上的单值实函数，如果对于直线上任一 Borel 点集 B，有 $\{\omega: \xi(\omega) \in B\} \in \mathscr{F}$，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.

  - 分布函数：称 $F(x) = P\{\xi(\omega) < x\}, -\infty < x < \infty$ 为随机变量 $\xi(\omega)$ 的 (累积) 分布函数. 记作 $\xi(\omega) \sim F(x)$.

  - 分布函数的性质与概率计算：

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

    ② 有界性：$0 \leq F(x) \leq 1$，且 $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$, $F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$.

    ③ 右连续性：$F(x+0) = \lim\limits_{t \to x+0} F(t) = F(x)$.

    ④ 分布函数可以唯一确定概率分布.

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 的全部可能取值是有限个或可列无限多个
    - 概率分布 (分布律)：ξ 的所有可能值 x_k，$k=1,2,\ \ldots$，则其概率分布 $P\{\xi=x_k\}\triangleq p_k$，$k=1,2,\ \ldots$
    - 分布列：

ξ	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

    - 概率分布的性质：

① $p_k\geq 0$，$k=1,2,\ \ldots$　　② $\sum_{k=1}^{\infty}p_k=1$

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

已学知识点

- 第三章 随机变量与分布函数
  - 常见的离散型随机变量及其分布

① 退化 (degenerate) 分布：$P\left\{\xi = c\right\} = 1$

② Bernoulli 分布：$P\left\{\xi = k\right\} = b(k;\ 1, p) = p^k(1-p)^{1-k}\ ,\quad k = 0,\ 1$

③ 二项 (binomial) 分布：$P\left\{\xi = k\right\} = b(k;\ n, p) = \binom{n}{k} p^k(1-p)^{n-k}\ ,\quad k = 0,\ 1,\ 2,\ \dots,\ n$

- 二项分布可以分解为 n 个独立同分布的 Bernoulli 分布的随机变量之和.

④ 超几何 (hyper-geometric) 分布：$h_x = P\left\{\xi = x\right\} = \dfrac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}\ ,\quad x = 0,\ 1,\ 2,\ \dots,\ k$

⑤ Poisson 分布：$P\left\{\xi = x\right\} = \dfrac{\lambda^x}{x!}\mathrm{e}^{-\lambda}\ ,\quad x = 0,\ 1,\ 2,\ \dots$

已学知识点

- 第三章 随机变量与分布函数
  - 常见的离散型随机变量及其分布

⑥ 几何 (geometric) 分布：$g(x,\ p) = P\left\{\eta = x\right\} = (1-p)^{x-1}p\ ,\quad x = 1,\ 2,\ \ldots$

- 几何分布的无记忆性：$P\left\{\eta > m+n \,\middle|\, \eta > m\right\} = P\left\{\eta > n\right\}\ ,\quad \forall\ m,\ n \geq 1$

⑦ Pascal 分布：$P\left\{\zeta = x\right\} = \binom{x-1}{r-1} p^r q^{x-r}\ ,\quad x = r,\ r+1,\ \ldots$

- Pascal 分布可以分解为 r 个独立同分布的几何分布的随机变量之和.

⑧ 负二项分布：$Nb(x;\ r,\ p) = \binom{-r}{x} p^r\,(-q)^x\ ,\quad x = 0,\ 1,\ 2,\ \ldots$

已学知识点

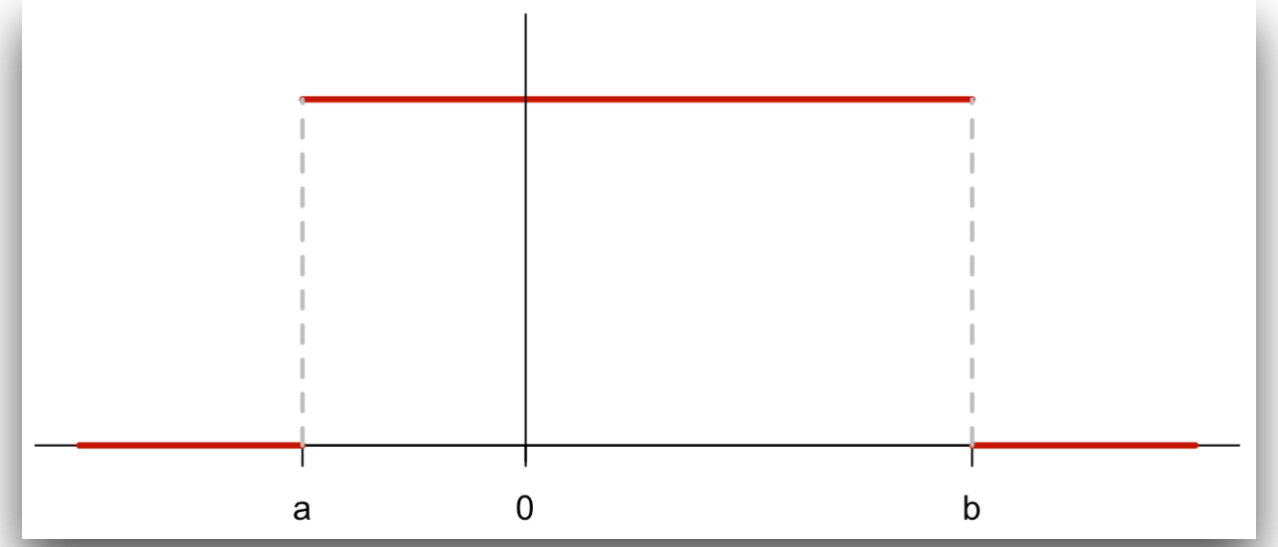
- 第三章 随机变量与分布函数
  - 连续型随机变量：存在非负可积函数 $p(x)$，使得 $F(x) = \int_{-\infty}^{x} p(t)\, \mathrm{d}t$
    - 其分布函数 $F(x)$ 绝对连续、几乎处处可导：$p(x) = F'(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x)$, for a.e. $x \in \mathbb{R}$
    - 概率密度的性质：

      ① $p(x) \geq 0$　　② $\int_{-\infty}^{+\infty} p(x)\, \mathrm{d}x = 1$
    - 概率计算：对 $\mathbb{R}$ 中任一 Borel 点集 B 有 $P\left\{\xi \in B\right\} = \int_B p(x)\, \mathrm{d}x$.

已学知识点

- 第三章 随机变量与分布函数
  - 常见的连续型随机变量及其分布

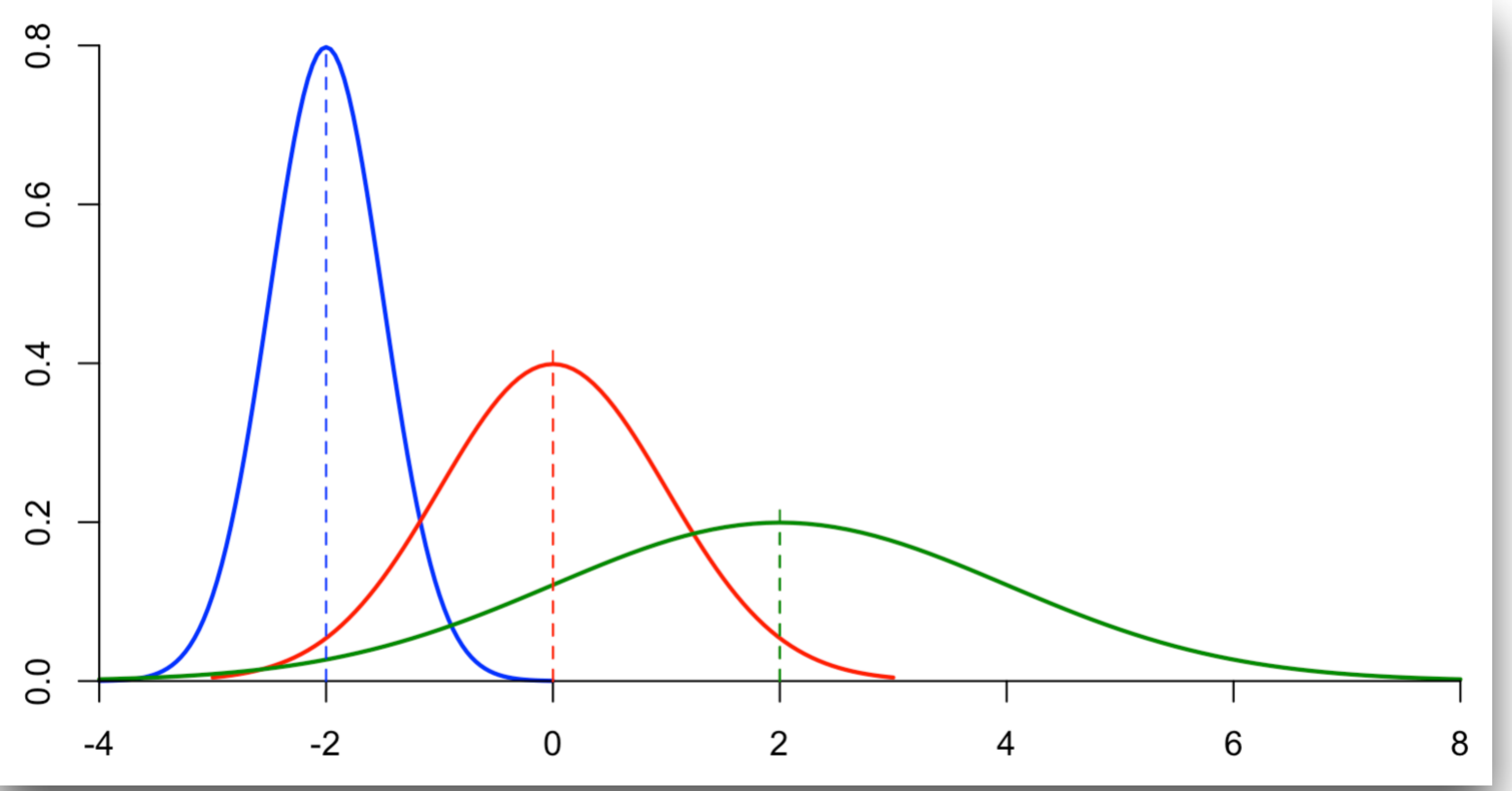
① 均匀 (uniform) 分布：$\xi \sim U\left[a,\ b\right]$. p.d.f. $p(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{其它} \end{cases}$



② 正态 (normal) 分布：$\xi \sim N\left(\mu,\ \sigma^2\right)$. p.d.f. $p(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\ \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < +\infty$

标准正态分布：$N(0,\ 1)$

p.d.f. $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\ \mathrm{e}^{-\frac{x^2}{2}}$, $-\infty < x < +\infty$

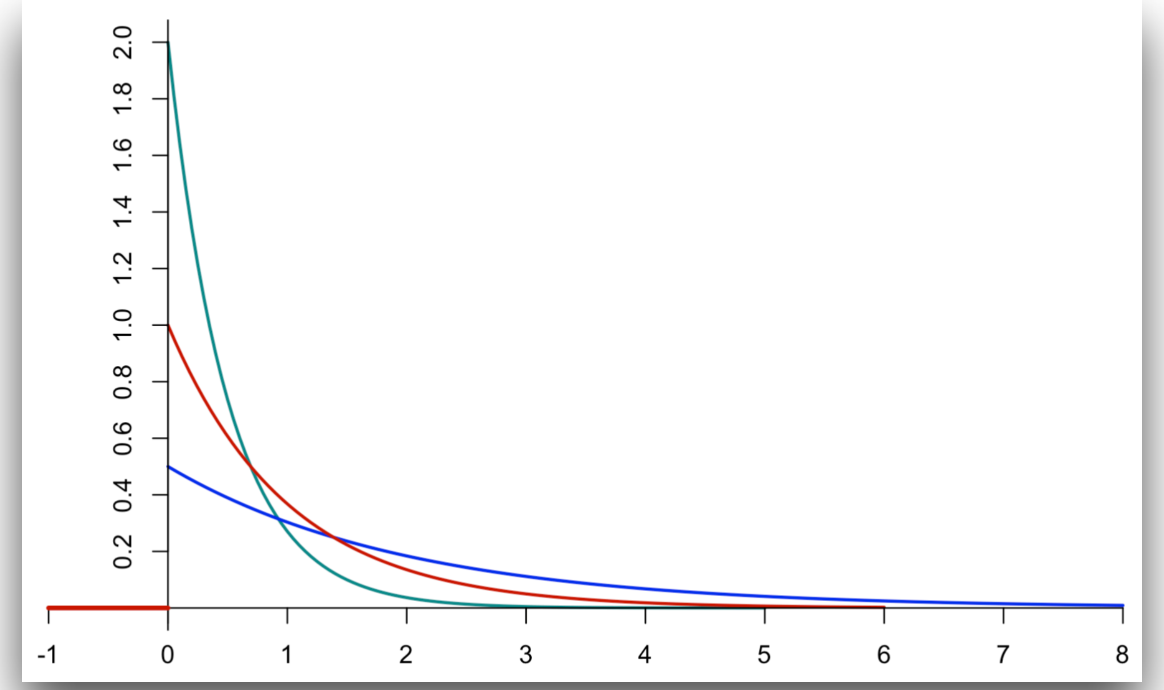


▸ 性质：$\xi \sim N\left(\mu,\ \sigma^2\right) \implies \dfrac{\xi - \mu}{\sigma} \sim N(0,\ 1)$.

已学知识点

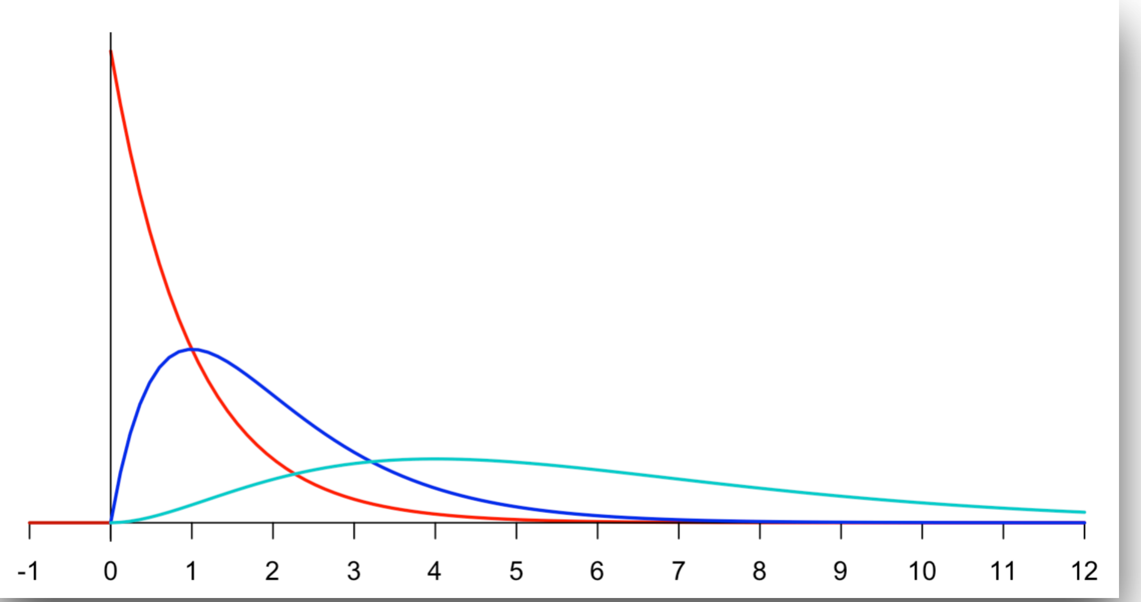
- 第三章 随机变量与分布函数
  - 常见的连续型随机变量及其分布

③ 指数 (exponential) 分布：$\xi \sim \text{Exp}(\lambda)$. p.d.f. $p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

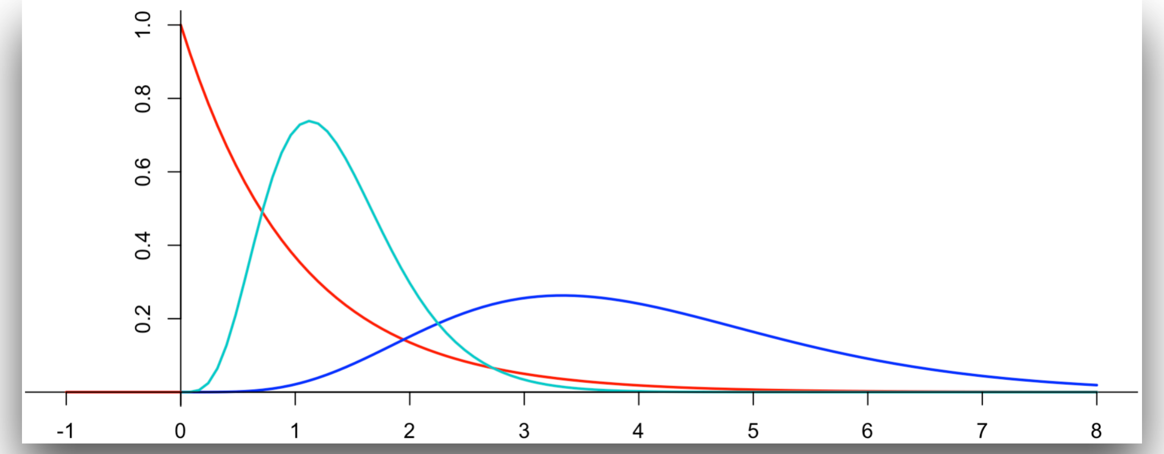


▸ 定理：取非负实值的随机变量 ξ 服从指数分布，当且仅当它具有无记忆性.

④ Erlang (埃尔朗) 分布： p.d.f. $p(x) = \begin{cases} \dfrac{\lambda^r}{(r-1)!} x^{r-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$



⑤ Γ (伽马) 分布：$\xi \sim \Gamma(r, \lambda)$. p.d.f. $p(x) = \begin{cases} \dfrac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$



▸ 特别：$\Gamma(1, \lambda) = \text{Exp}(\lambda)$.

当 $r \in \mathbb{N}^+$ 时， $\Gamma(r, \lambda)$ 就是 Erlang 分布.

当 $n \in \mathbb{N}^+$ 时， $\Gamma\left(\dfrac{n}{2}, \dfrac{1}{2}\right)$ 就是 $\chi^2(n)$ 分布.

已学知识点

- 第三章 随机变量与分布函数

  - 随机向量：若 $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$ 是定义在同一概率空间 $(\Omega, \mathscr{F}, P)$ 上的随机变量，则称

    $\boldsymbol{\xi}(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$ 构成一个 n 维随机向量，亦称 n 维随机变量.

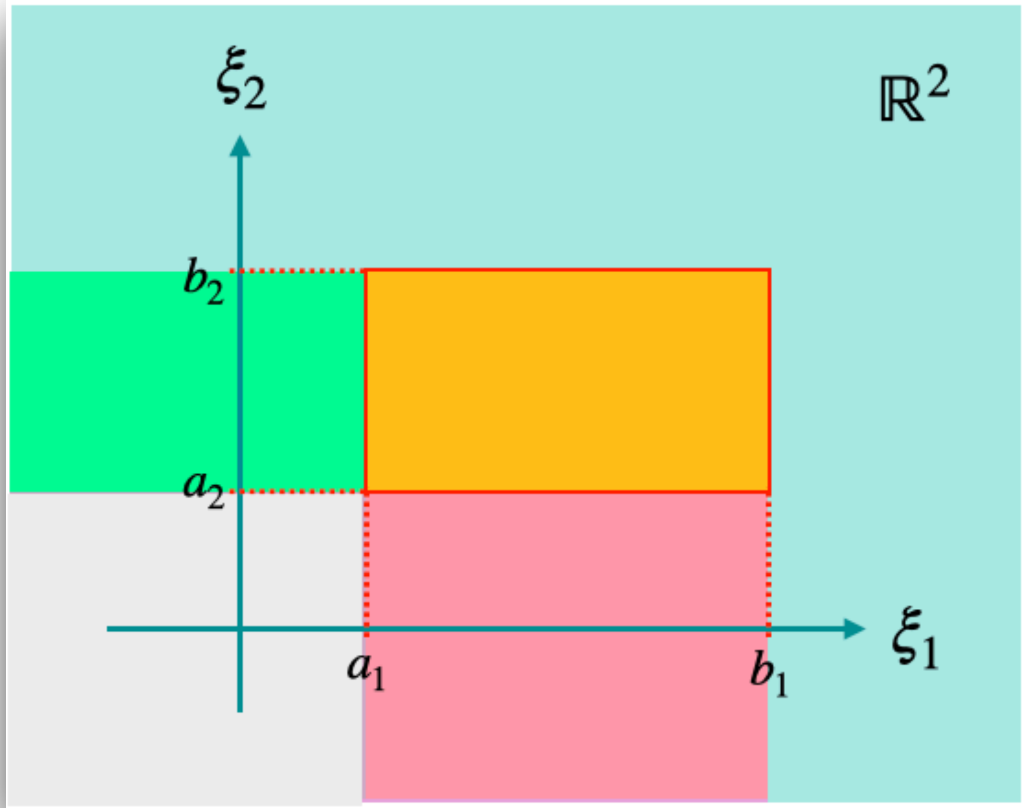
  - (联合)分布函数 (joint cumulative distribution function)

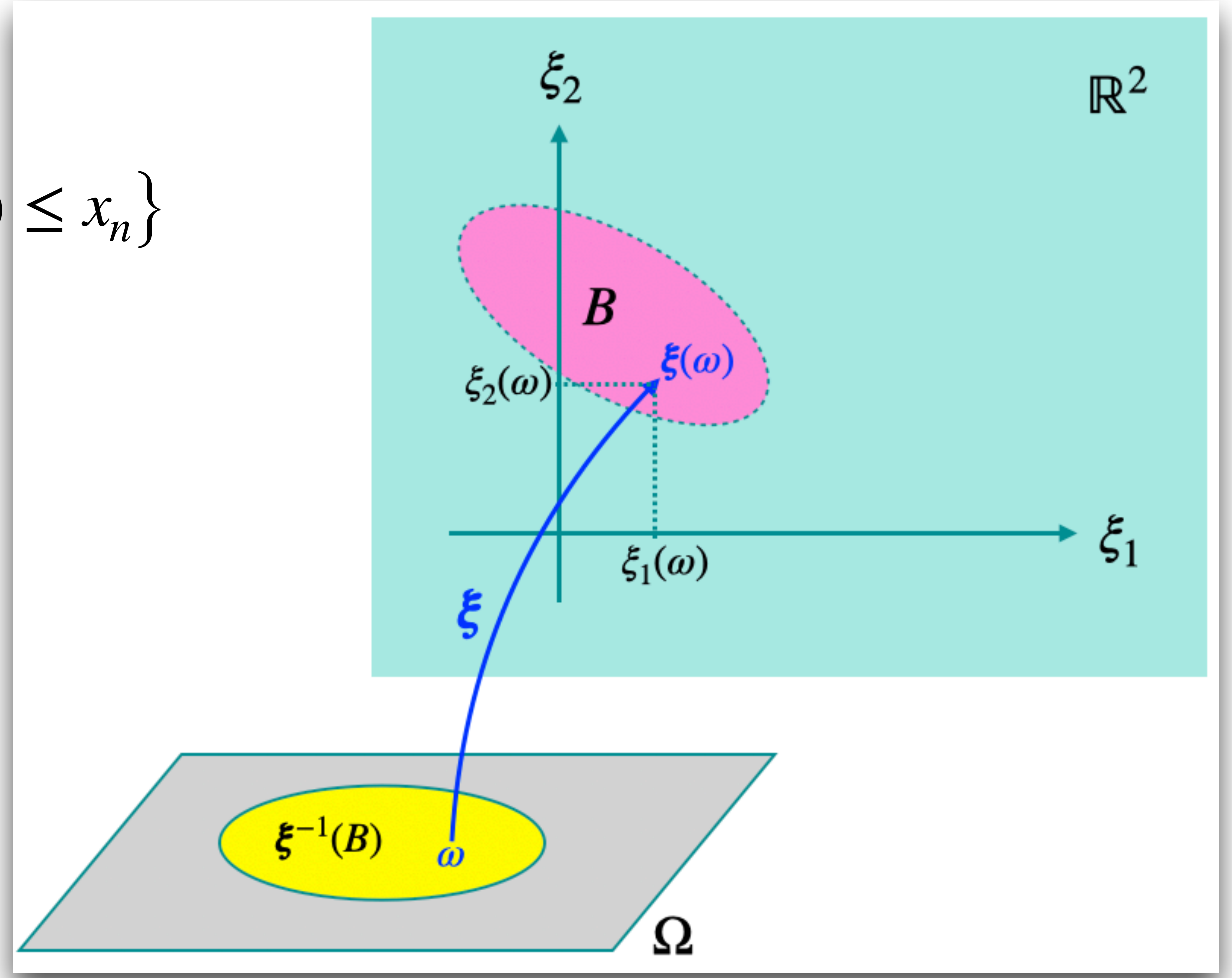
    $$F(x_1, x_2, \ldots, x_n) = P\{\xi_1(\omega) \le x_1, \xi_2(\omega) \le x_2, \ldots, \xi_n(\omega) \le x_n\}$$

    - 事件概率的计算：

$$\begin{aligned} P\{a_1 < \xi_1 \le b_1, a_2 < \xi_2 \le b_2\} = & P\{\xi_1 \le b_1, \xi_2 \le b_2\} \\ & - P\{\xi_1 \le a_1, \xi_2 \le b_2\} \\ & - P\{\xi_1 \le b_1, \xi_2 \le a_2\} \\ & + P\{\xi_1 \le a_1, \xi_2 \le a_2\} \end{aligned}$$

$$= F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2)$$





已学知识点

- 第三章 随机变量与分布函数
  - (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.

④ 成立则 ① 成立，但 ① 成立不能推出 ④ 成立

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

④ $n=2$ 时：$\forall\ a_1 \le b_1\ , a_2 \le b_2$ 都有 $F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \ge 0.$

后三个性质是二元函数为某二元随机向量的分布函数应满足的条件

已学知识点

- 常见的多元离散型分布

① 多项 (multinomial) 分布：试验的可能结果为 $A_1, A_2, \dots, A_r$，记 $P(A_i) = p_i$, $i = 1, 2, \dots, r$，且

$$p_1 + p_2 + \cdots + p_r = 1$$

独立重复该试验 n 次，定义 $\xi_1, \xi_2, \dots, \xi_r$ 分别为 $A_1, A_2, \dots, A_r$ 出现的次数，则

$$P(\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_r = k_r) = \frac{n!}{k_1!\, k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

$k_i \geq 0$ 取整数，且 $k_1 + k_2 + \cdots + k_r = n$

多项分布用于有放回抽样的场合.

已学知识点

- 常见的多元离散型分布

袋中装有编号为 i 的球 N_i 只，$i=1, 2, \ldots, r$，$N_1+N_2+\cdots+N_r=N$

从中随机取出 n 只球，定义 $\xi_1, \xi_2, \ldots, \xi_r$ 分别表示取出的球中编号为 $1, 2, \ldots, r$ 号球的个数，则

$$P(\xi_1=n_1, \xi_2=n_2, \ldots, \xi_r=n_r)=\frac{\binom{N_1}{n_1}\binom{N_2}{n_2}\cdots\binom{N_r}{n_r}}{\binom{N}{n}}$$

$n_i \geq 0$ 取整数，且 $n_1+n_2+\cdots+n_r=n$

多元超几何分布用于不放回抽样的场合.

已学知识点

- 常见的多元连续型分布
  - 联合概率密度函数 (joint probability density function)： 存在非负函数 $p(x_1, x_2, \ldots, x_n)$ 使得

$$F(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} p(y_1, y_2, \ldots, y_n)\, \mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n$$

  - 联合概率密度函数满足：

(1) 非负性： $p(x_1, x_2, \ldots, x_n) \geq 0$

(2) 规范性： $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} p(x_1, x_2, \ldots, x_n)\, \mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n = 1$

(1) 与 (2) 构成联合概率密度函数的充分必要条件

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

一、随机向量及其分布

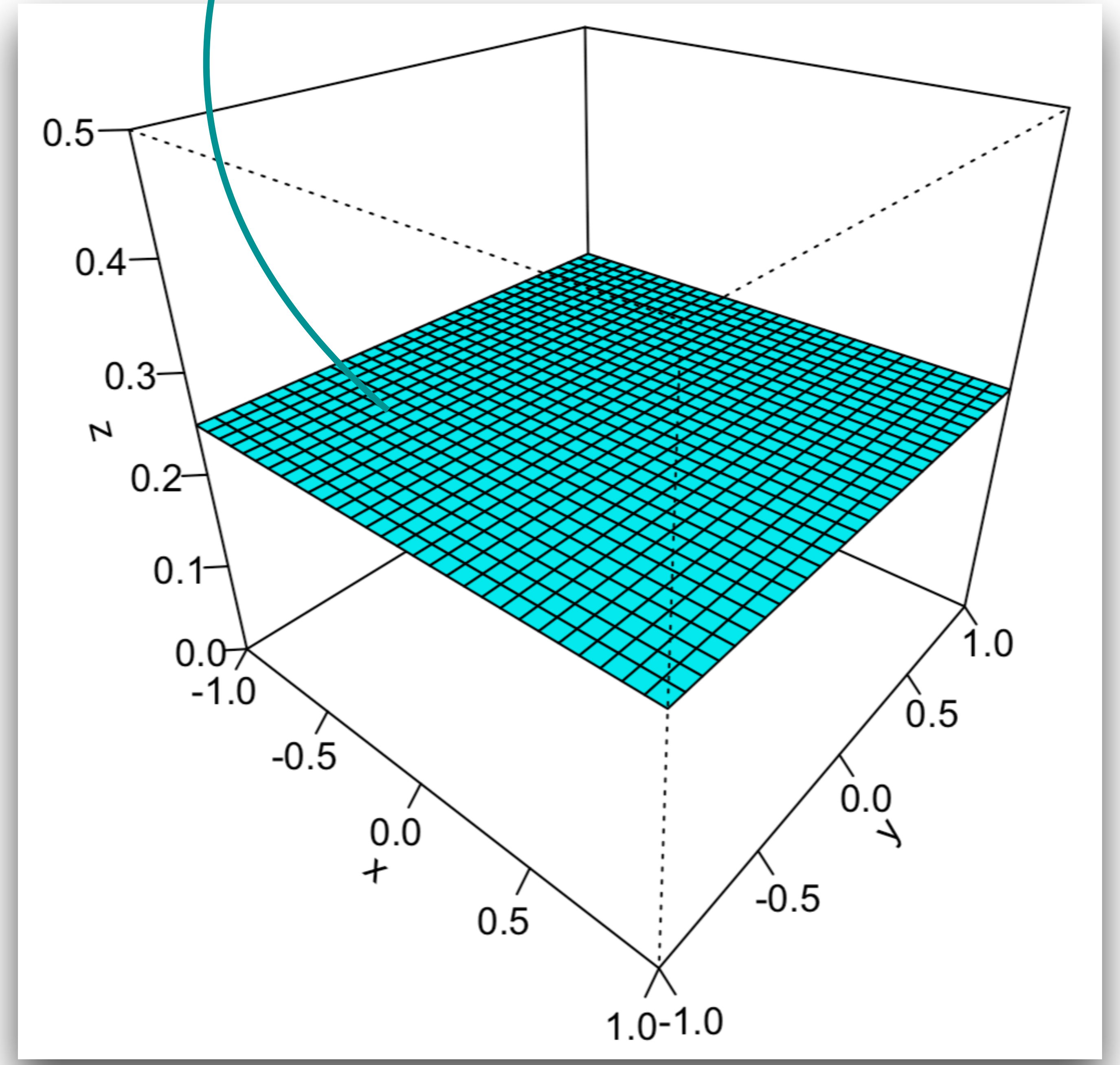
- 常见的多元连续型分布

① 均匀 (uniform) 分布：设 G 是 $\mathbb{R}^n$ 中的有限区域，其测度 $S>0$，则由密度函数

$$p(x_1,x_2,\ldots,x_n)=\begin{cases}\dfrac{1}{S}, & (x_1,x_2,\ldots,x_n)\in G\\ 0, & (x_1,x_2,\ldots,x_n)\notin G\end{cases}$$

给出的分布称为区域 G 上的均匀分布.

$$p(x,y)=\begin{cases}\dfrac{1}{4}, & -1\le x,y\le 1\\ 0, & 其它\end{cases}$$



▸ 二维均匀分布的背景：

向平面上有界区域 G 上任投一质点，若质点落在 G 内任一小区域 B 的概率与小区域的面积成正比，而与 B 的位置无关，则质点的坐标 (X,Y) 在 G 上服从均匀分布.

一、随机向量及其分布

- 常见的多元连续型分布

② 多元正态 (multinormal) 分布：设 $\Sigma = (\sigma_{ij})$ 是 n 阶正定对称矩阵，$\Sigma^{-1} = (\gamma_{ij})$ 是 Σ 的逆矩阵，

$\det\Sigma$ 是 Σ 的行列式值，$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)$ 是任意实值行向量，

则由密度函数

$$p(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}(\det\Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\sum_{j,k=1}^{n}\gamma_{jk}(x_j-\mu_j)(x_k-\mu_k)\right\}, \quad -\infty < x_1, x_2, \ldots, x_n < \infty$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}(\det\Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\right\}, \quad \boldsymbol{x} = (x_1, x_2, \ldots, x_n)$$

定义的分布称为 n 元正态分布，记为 $N(\boldsymbol{\mu}, \Sigma)$.

一、随机向量及其分布

- 常见的多元连续型分布

② 多元正态 (multinormal) 分布：

$$p(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}(\det \Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\sum_{j,\,k=1}^{n} \gamma_{jk}\left(x_j - \mu_j\right)\left(x_k - \mu_k\right)\right\}, \quad -\infty < x_1, x_2, \dots, x_n < \infty$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}(\det \Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\,\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\right\}, \quad \boldsymbol{x} = (x_1, x_2, \dots, x_n)$$

▸ 二元正态分布：设二维随机向量 (X, Y) 具有概率密度函数

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

则称 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 的二元正态分布. 记为

$$(X, Y) \sim N\left(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho\right)$$

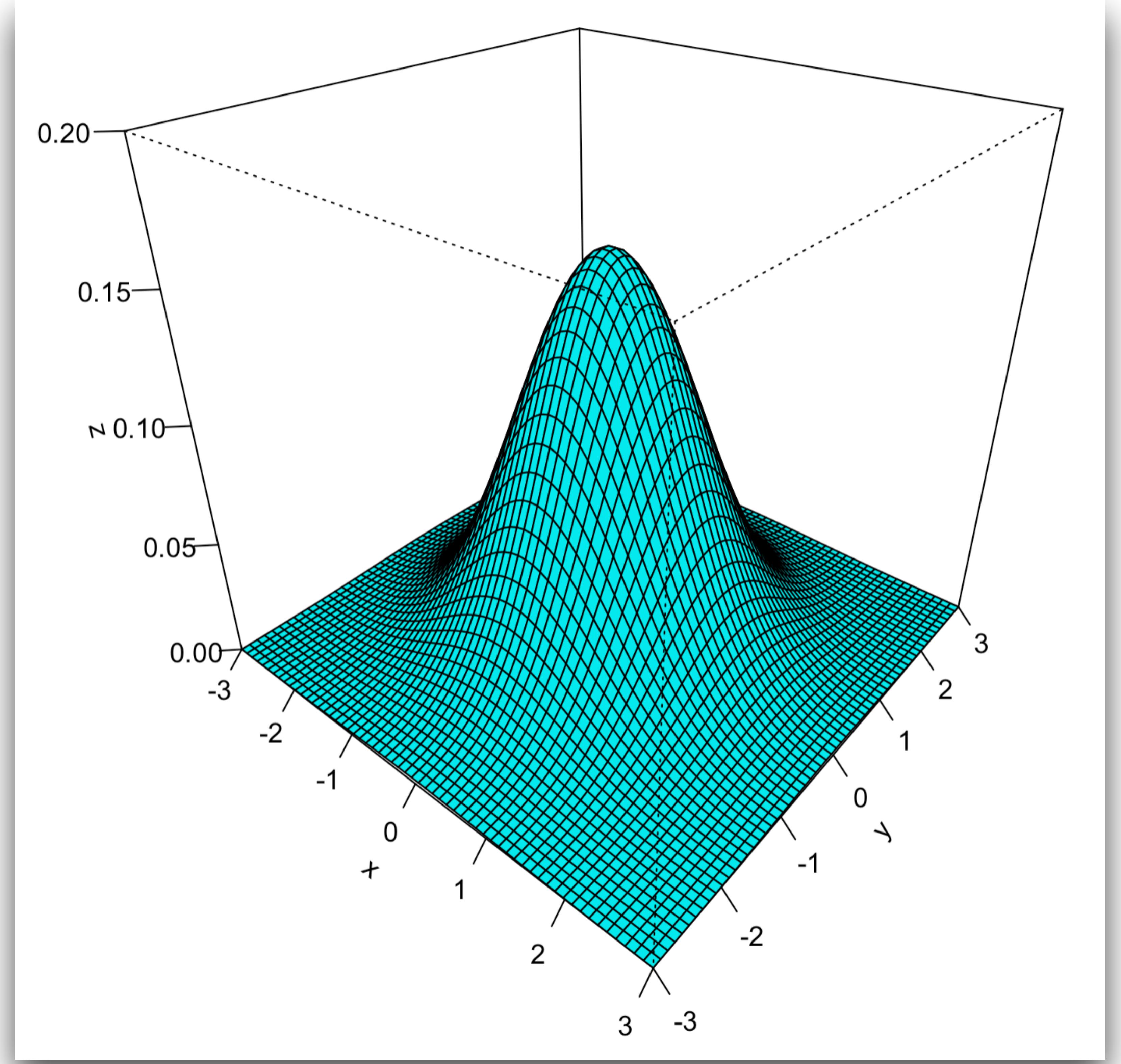
一、随机向量及其分布

● 常见的多元连续型分布

② 多元正态 (multinormal) 分布：

▸ 二元正态分布： $p(x, y) = \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\dfrac{1}{2\left(1-\rho^2\right)}\left[\left(\dfrac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\dfrac{x-\mu_1}{\sigma_1}\right)\left(\dfrac{y-\mu_2}{\sigma_2}\right) + \left(\dfrac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}$

二维正态分布的联合密度函数 $p(x, y)$ 的图形是一个钟形曲面，它与平行于坐标平面 OXY 的水平平面相交的截面曲线为椭圆，而与平行于另外两个坐标平面的竖直平面相交的截面曲线为一元正态密度函数曲线.



二、边际 (marginal) 分布： 为方便起见，讨论将就二维情形进行.

- 离散情形

设离散型随机变量 ξ 取值 $x_1, x_2, \ldots, x_i, \ldots$，离散型随机变量 η 取值 $y_1, y_2, \ldots, y_j, \ldots$

$$p\left(x_i, y_j\right) \triangleq P\left\{\xi = x_i, \eta = y_j\right\}, \qquad i, j = 1, 2, \ldots$$

称 $\left\{p\left(x_i, y_j\right)\right\}$ 为随机向量 $\left(\xi, \eta\right)$ 的联合分布律.

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

- 联合分布律的性质：

(1) $p\left(x_i, y_j\right) \geq 0$

(2) $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p\left(x_i, y_j\right) = 1$

二、边际 (marginal) 分布：为方便起见，讨论将就二维情形进行.

- **离散情形**

设离散型随机变量 ξ 取值 $x_1, x_2, \ldots, x_i, \ldots$，离散型随机变量 η 取值 $y_1, y_2, \ldots, y_j, \ldots$

$$p\left(x_i, y_j\right) \triangleq P\left\{\xi = x_i, \eta = y_j\right\}, \qquad i, j = 1, 2, \ldots$$

称 $\left\{p\left(x_i, y_j\right)\right\}$ 为随机向量 (ξ, η) 的联合分布律.

▸ ξ, η 的边际 (边缘) 分布：

$$p_1\left(x_i\right) \triangleq P\left\{\xi = x_i\right\} = \sum_{j=1}^{\infty} p\left(x_i, y_j\right), \quad i = 1, 2, \ldots$$

$$p_2\left(y_j\right) \triangleq P\left\{\eta = y_j\right\} = \sum_{i=1}^{\infty} p\left(x_i, y_j\right), \quad j = 1, 2, \ldots$$

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

二、边际 (marginal) 分布：为方便起见，讨论将就二维情形进行.

- 离散情形

▸ 例：袋中装有 2 只白球及 3 只黑球，进行不放回摸球，定义

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

联合分布不能由边际分布唯一确定，还需考虑它们之间的联系。研究多维随机变量的原因！

则 (ξ,η) 的联合概率分布与边际分布为

$\eta \backslash \xi$	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

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

二、边际 (marginal) 分布：为方便起见，讨论将就二维情形进行.

- 离散情形

  ▸ 例：从 1, 2, 3, 4 中随机取一个数 ξ，再从 1，…, ξ 中随机地取一个数 η，确定 (ξ, η) 的联合分布律.

ξ 的取值范围：1, 2, 3, 4, $\quad P\{\xi = i\} = \dfrac{1}{4}, \quad i = 1, 2, 3, 4$

η 的取值范围：1, 2, 3, 4 且 $\eta \le \xi$, $\quad P\left\{\eta = j \middle| \xi = i\right\} = \dfrac{1}{i}, \quad j = 1, \ldots, i$

$$P(\xi = i, \eta = j) = P(\xi = i) \cdot P\left\{\eta = j \middle| \xi = i\right\} = \frac{1}{4i}, \quad 1 \le j \le i \le 4$$

乘法公式

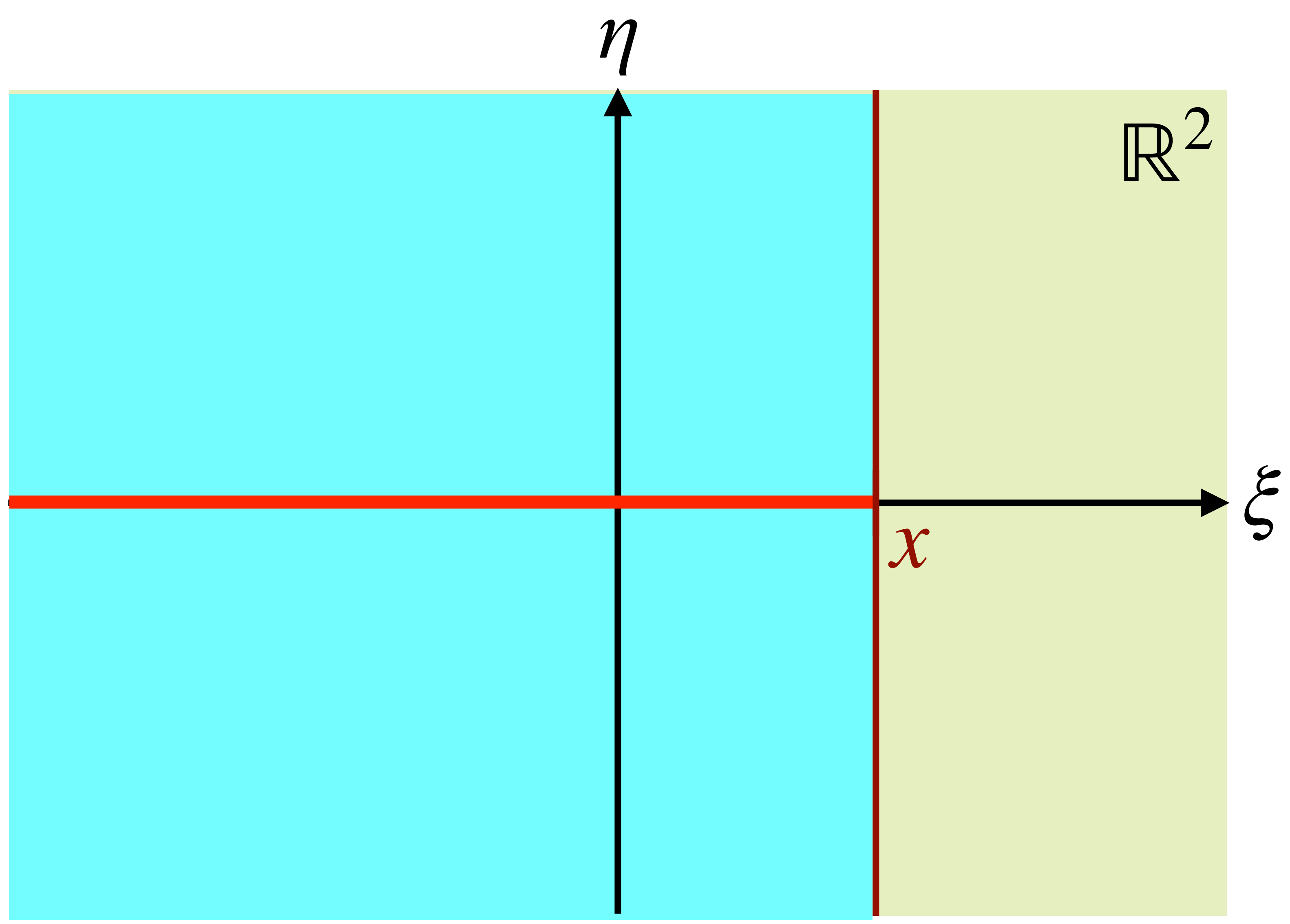
$\xi \backslash \eta$	1	2	3	4
1	$\frac{1}{4}$	0	0	0
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

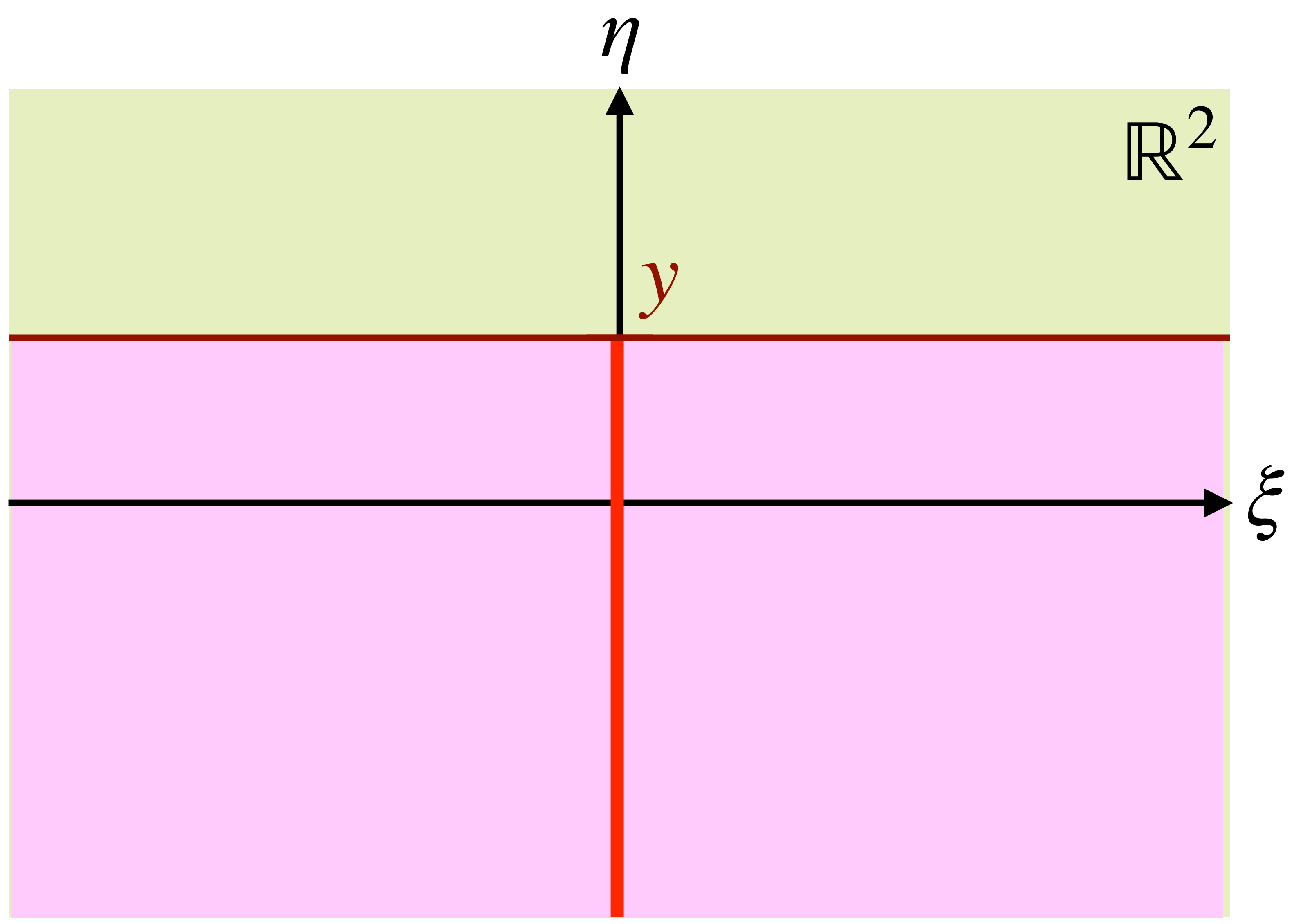
二、边际 (marginal) 分布： 为方便起见，讨论将就二维情形进行.

- 一般地，若 (ξ, η) 是二维 (离散或连续) 随机向量，其分布函数为 $F(x, y)$，则 ξ, η 的分布函数为

$$F_\xi(x) = P\{\xi \le x\} = P\{\xi \le x, \eta < +\infty\} = F(x, +\infty)$$

ξ 的边际分布函数





$$F_\eta(y) = P\{\eta \le y\} = P\{\xi < +\infty, \eta \le y\} = F(+\infty, y)$$

η 的边际分布函数

二、边际 (marginal) 分布： 为方便起见，讨论将就二维情形进行.

- 连续情形

设连续型随机向量 (ξ, η) 的联合分布函数为 $F(x, y)$，联合概率密度函数为 $p(x, y)$，则

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u, v)\, \mathrm{d}u\, \mathrm{d}v \,, \qquad p(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \,, \quad \text{a.e.}$$

▸ 在连续点 (x, y) 处，(ξ, η) 落在小的矩形区域 $(x\,,\, x+\Delta x) \times (y\,,\, y+\Delta y)$ 上的概率近似等于

$$p(x, y) \cdot \Delta x \cdot \Delta y$$

▸ 于是，(ξ, η) 落在平面上任意区域 D 上的概率为：

$$P\left\{ (\xi, \eta) \in D \right\} = \iint\limits_{D} p(x, y)\, \mathrm{d}x\, \mathrm{d}y$$

二、边际 (marginal) 分布：为方便起见，讨论将就二维情形进行.

- 连续情形

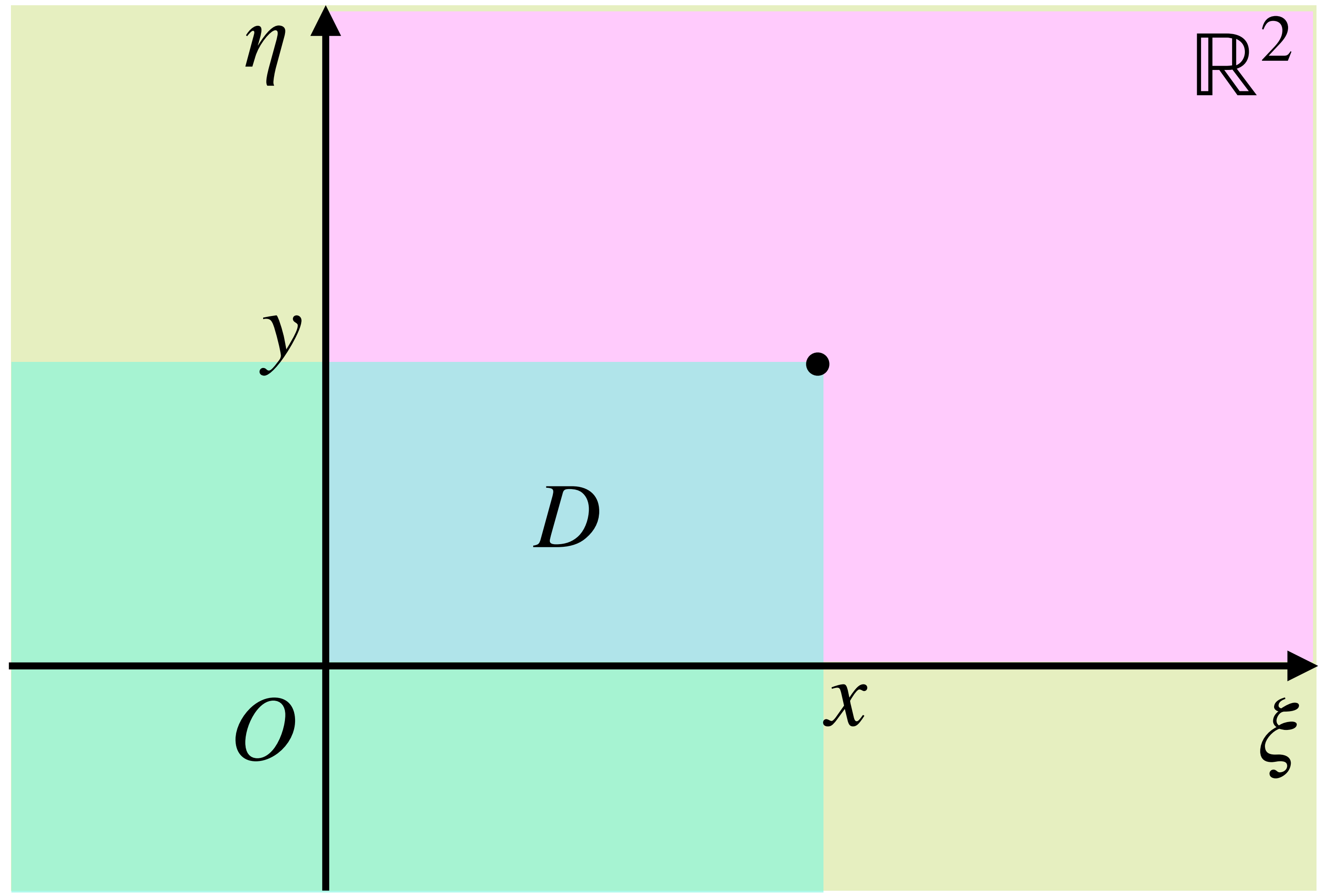
▸ 例：已知 (ξ, η) 的联合概率密度函数为 $p(x, y) = \begin{cases} 2e^{-(2x+y)}, & x, y \geq 0 \\ 0, & \text{其它} \end{cases}$

① 求 (ξ, η) 的联合分布函数 $F(x, y)$.

$\forall\, x > 0,\ y > 0$

$$F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u, v)\,\mathrm{d}u\mathrm{d}v = \iint\limits_{D} 2e^{-(2u+v)}\mathrm{d}u\mathrm{d}v$$

$$= \int_0^y e^{-v}\left[\int_0^x 2e^{-u}\mathrm{d}u\right]\mathrm{d}v = \left(1 - e^{-2x}\right)\left(1 - e^{-y}\right)$$



$$\Longrightarrow \quad F(x, y) = \begin{cases} \left(1 - e^{-2x}\right)\left(1 - e^{-y}\right), & x > 0,\ y > 0 \\ 0, & \text{其它} \end{cases}$$

二、边际 (marginal) 分布：为方便起见，讨论将就二维情形进行.

- 连续情形

▸ 例：已知 (ξ, η) 的联合概率密度函数为 $p(x, y) = \begin{cases} 2e^{-(2x+y)}, & x, y \ge 0 \\ 0, & \text{其它} \end{cases}$

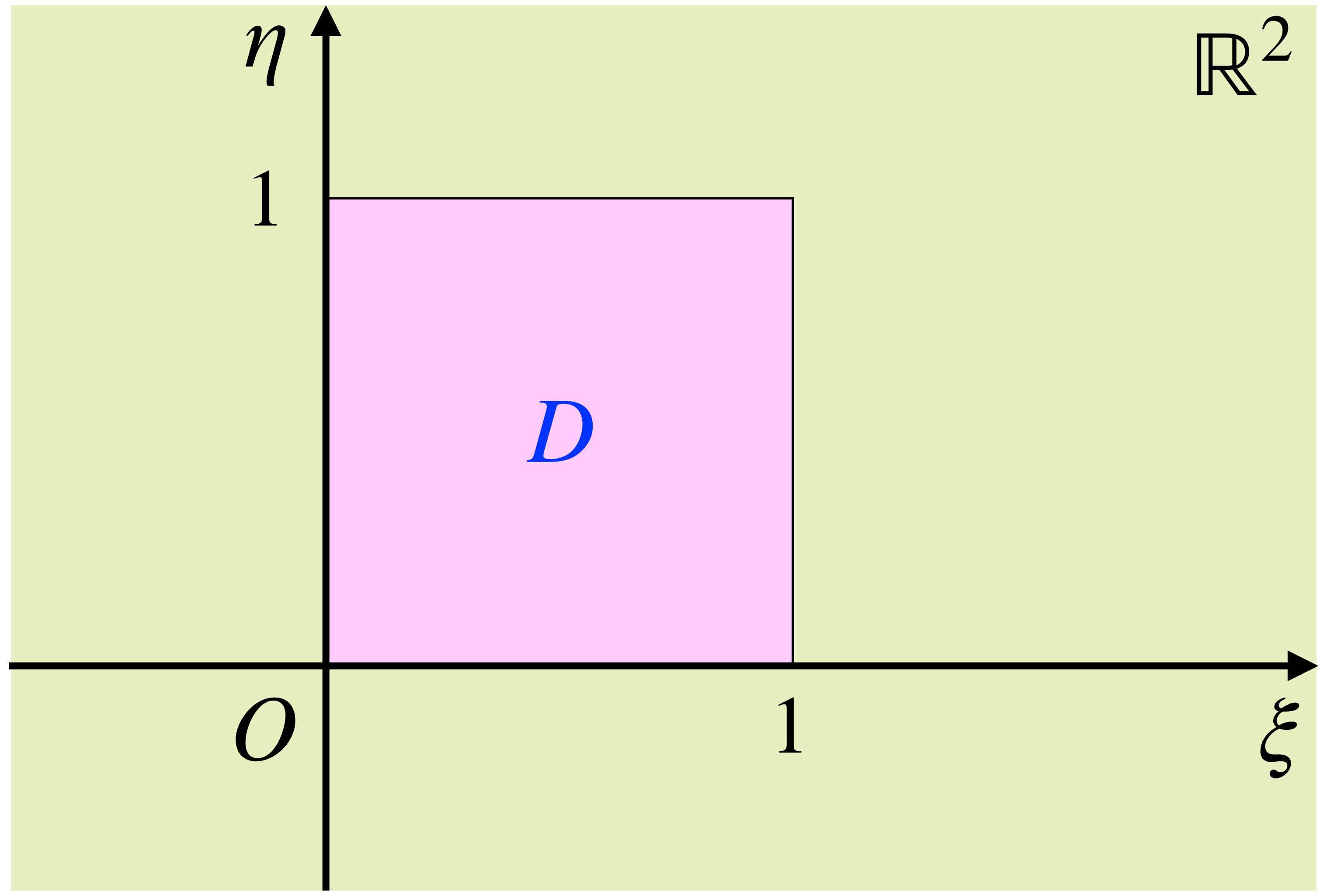
① 求 (ξ, η) 的联合分布函数 $F(x, y)$.

$$F(x, y) = \begin{cases} \left(1 - e^{-2x}\right)\left(1 - e^{-y}\right), & x > 0, y > 0 \\ 0, & \text{其它} \end{cases}$$

② 求 $P\left\{0 < \xi \le 1, 0 < \eta \le 1\right\}$.

$$P\left\{0 < \xi \le 1, 0 < \eta \le 1\right\} = F(1, 1) - F(1, 0) - F(0, 1) + F(0, 0)$$

$$= \left(1 - e^{-2}\right)\left(1 - e^{-1}\right)$$



$$P\left\{0 < \xi \le 1, 0 < \eta \le 1\right\} = \iint\limits_{D} p(x, y)\mathrm{d}x\mathrm{d}y = \int_0^1 \int_0^1 2e^{-(2x+y)}\mathrm{d}x\mathrm{d}y$$

$$= \left(1 - e^{-2}\right)\left(1 - e^{-1}\right)$$

二、边际 (marginal) 分布：为方便起见，讨论将就二维情形进行.

- 连续情形

▶ 例：已知 (ξ, η) 的联合概率密度函数为 $p(x, y) = \begin{cases} 2\mathrm{e}^{-(2x+y)}, & x, y \geq 0 \\ 0, & \text{其它} \end{cases}$

① 求 (ξ, η) 的联合分布函数 $F(x, y)$.

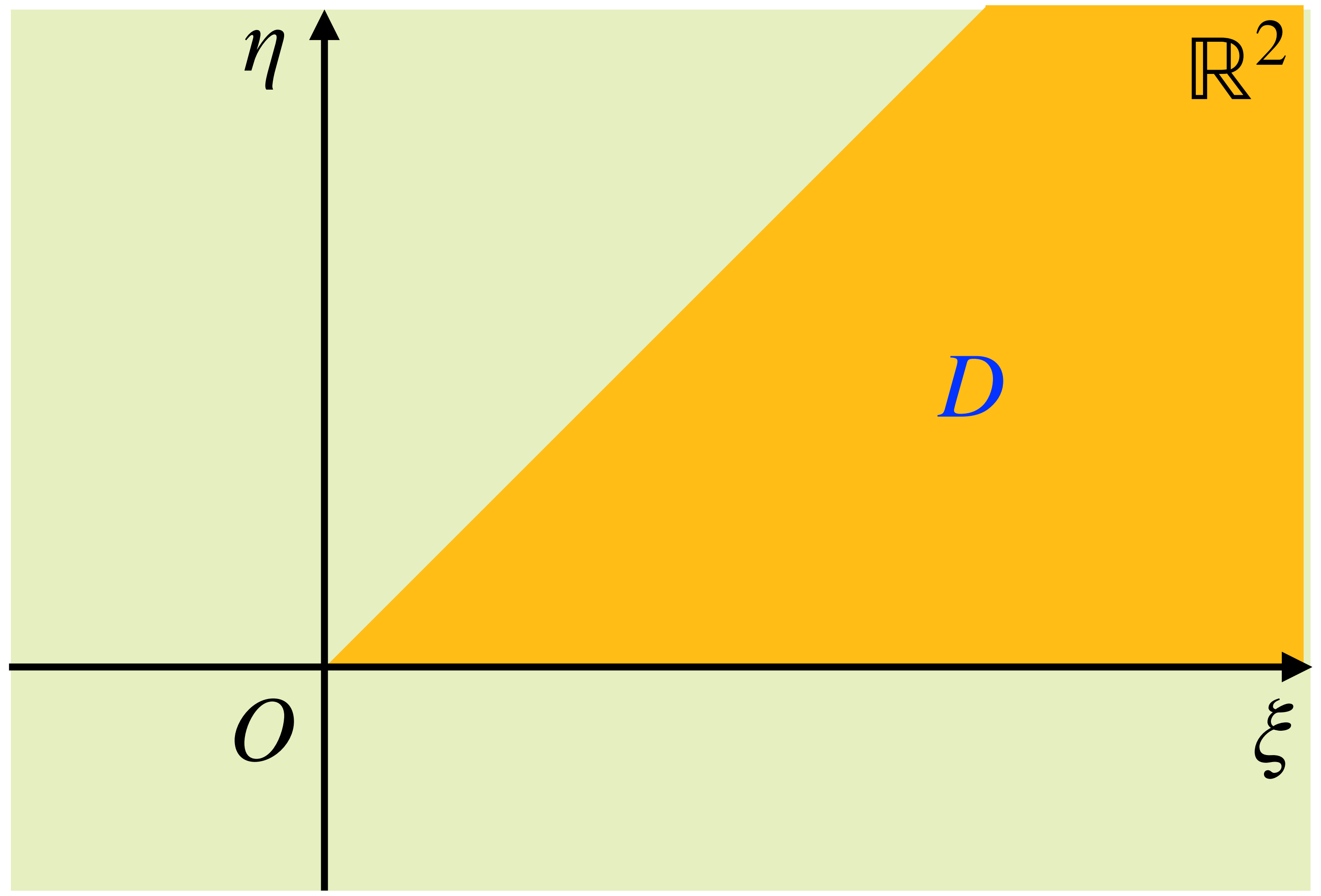
$F(x, y) = \begin{cases} \left(1-\mathrm{e}^{-2x}\right)\left(1-\mathrm{e}^{-y}\right), & x>0, y>0 \\ 0, & \text{其它} \end{cases}$

② 求 $P\left\{0 \leq \xi < 1, 0 \leq \eta < 1\right\}$.

③ 求 $P\left\{\xi \geq \eta\right\}$.

$$P\left\{\xi \geq \eta\right\} = \iint_D p(x, y)\mathrm{d}x\mathrm{d}y = \iint_D 2\mathrm{e}^{-(2x+y)}\mathrm{d}x\mathrm{d}y$$

$$= \int_0^{+\infty} \mathrm{e}^{-y}\left(\int_y^{+\infty} 2\mathrm{e}^{-2x}\mathrm{d}x\right)\mathrm{d}y = \int_0^{+\infty} \mathrm{e}^{-3y}\mathrm{d}y = \frac{1}{3}$$



二、边际 (marginal) 分布

- 连续情形

若连续型随机向量 (ξ, η) 的联合分布函数为 $F(x, y)$，联合概率密度函数为 $p(x, y)$，则

$$F_\xi(x) = P\{\xi \le x\} = P\{\xi \le x\,, \eta < +\infty\} = F(x\,,\ +\infty)$$

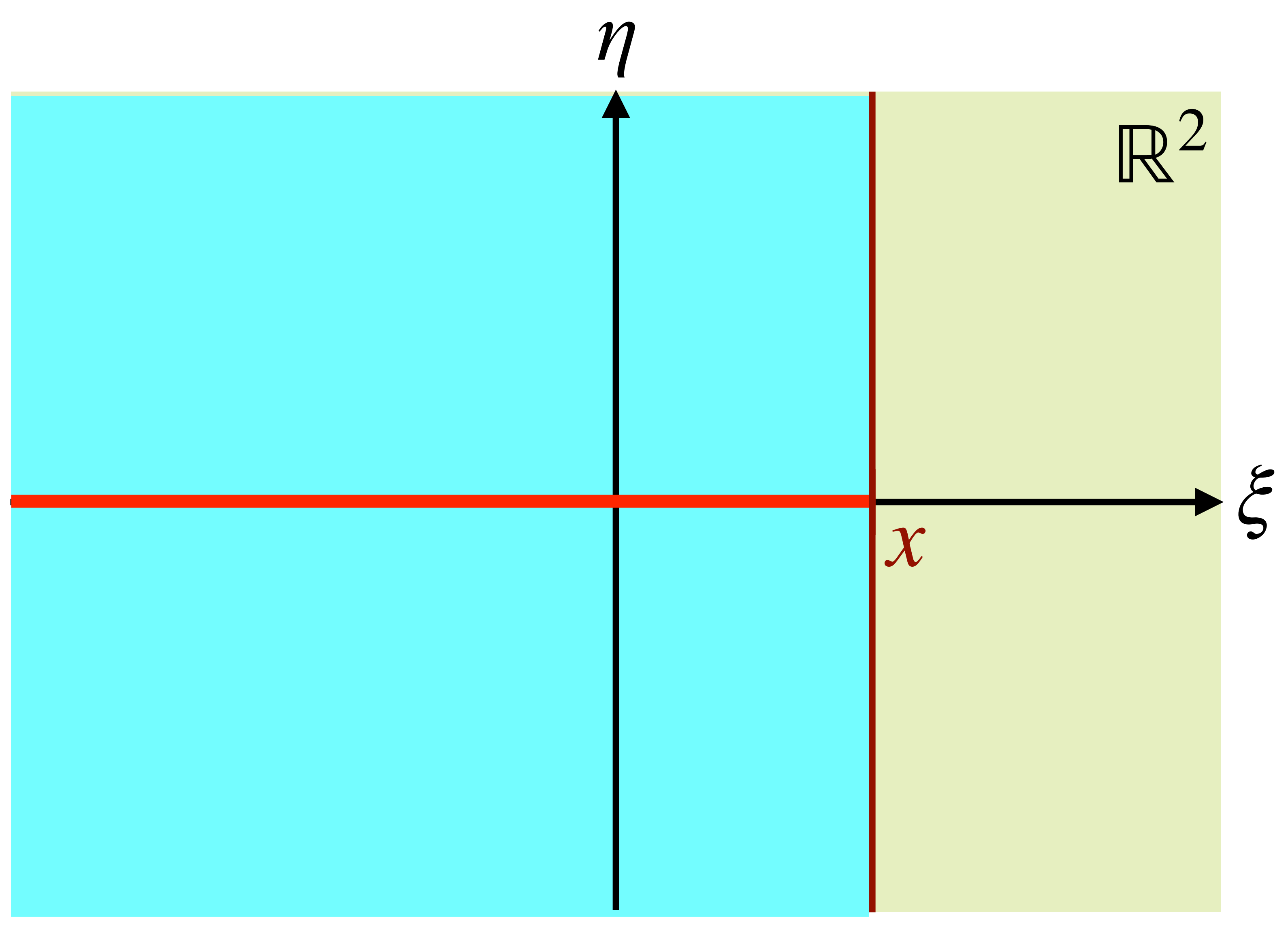
$$= \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} p(u, v)\mathrm{d}v \right) \mathrm{d}u$$

ξ 的边际分布函数

ξ 也是连续型随机变量

$$\Longrightarrow\quad p_\xi(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_\xi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} p(u, v)\mathrm{d}v \right) \mathrm{d}u = \int_{-\infty}^{+\infty} p(x, y)\mathrm{d}y$$

ξ 的边际概率密度函数



- 同理， $F_\eta(y) = P\{\eta \le y\} = F(+\infty, y)$ 也是连续型分布函数，η 的边际概率密度函数为

$$p_\eta(y) = \int_{-\infty}^{+\infty} p(x, y)\mathrm{d}x$$

二、边际 (marginal) 分布

- 连续情形

  ▸ 二元正态分布的边际分布仍为正态分布. $(\xi,\eta)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right) \Longrightarrow \begin{cases}\xi\sim N\left(\mu_1,\sigma_1^2\right) ✓\\ \eta\sim N\left(\mu_2,\sigma_2^2\right)\end{cases}$

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

$$p_\xi(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\exp\left[-\frac{u^2-2\rho uv+v^2}{2\left(1-\rho^2\right)}\right]\mathrm{d}v$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\exp\left[-\frac{(v-\rho u)^2+u^2-\rho^2u^2}{2(1-\rho^2)}\right]\mathrm{d}v$$

$N\left(\rho u,\sqrt{1-\rho^2}\right)$ 的密度函数

$$=\frac{1}{\sqrt{2\pi}\,\sigma_1}\exp\left(-\frac{u^2}{2}\right)\cdot\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\exp\left[-\frac{(v-\rho u)^2}{2(1-\rho^2)}\right]\mathrm{d}v=\frac{1}{\sqrt{2\pi}\,\sigma_1}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)$$

二、边际 (marginal) 分布

- 连续情形

▶ 注意：$\begin{cases} \xi \sim N\left(\mu_1, \sigma_1^2\right) \\ \eta \sim N\left(\mu_2, \sigma_2^2\right) \end{cases} \not\Rightarrow \quad (\xi, \eta) \sim N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$

▶ 例：(ξ, η) 的联合概率密度函数 $p(x, y) = \dfrac{1}{2\pi}(1 + xy) \cdot \mathrm{e}^{-\frac{x^2+y^2}{2}}, \quad -\infty < x, y < +\infty$

$$p_\xi(x) = \int_{-\infty}^{+\infty} p(x, y)\mathrm{d}y = \int_{-\infty}^{+\infty} \frac{1}{2\pi}(1 + xy) \cdot \mathrm{e}^{-\frac{x^2+y^2}{2}} \mathrm{d}y = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}} \Longrightarrow \xi \sim N(0, 1)$$

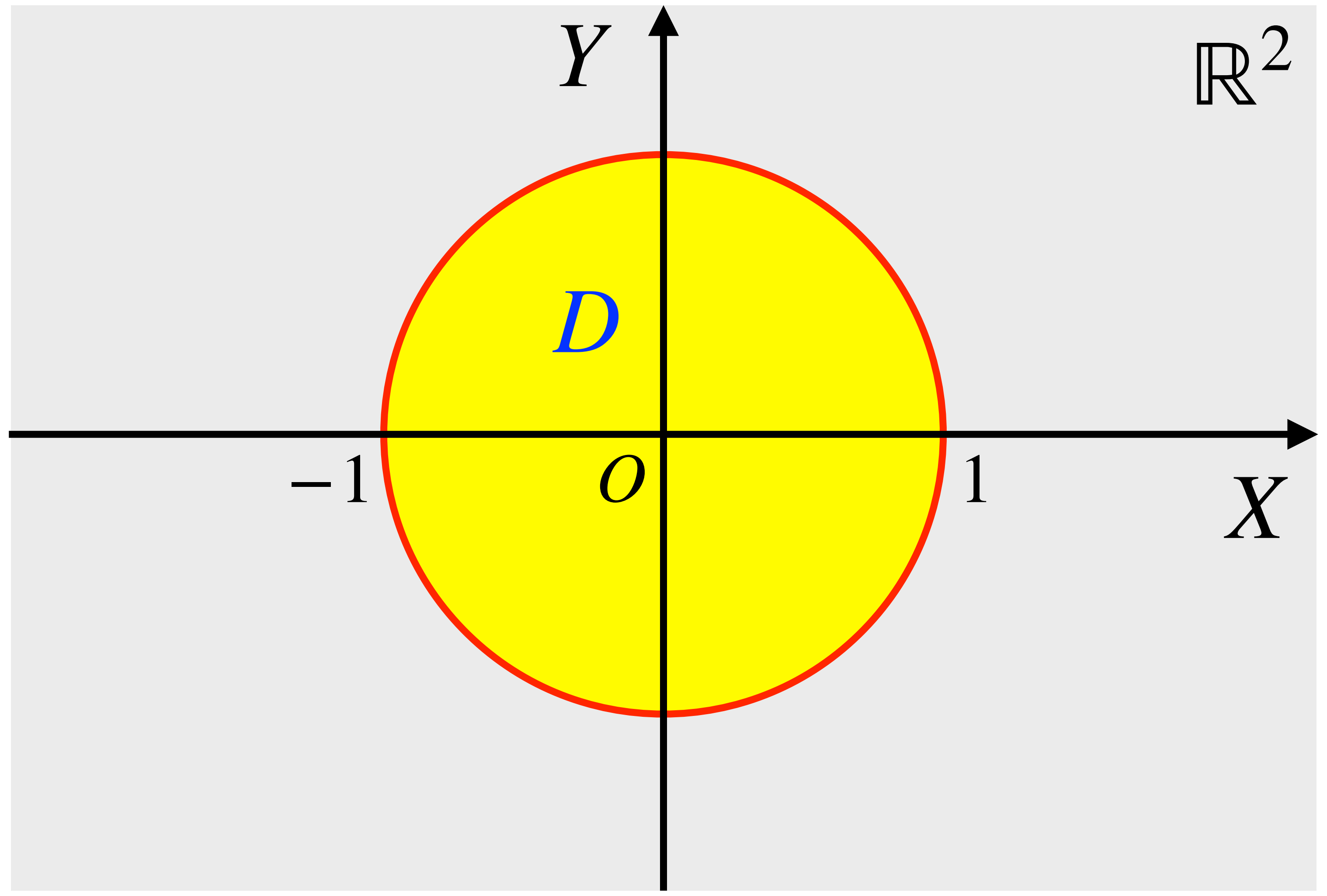
$$p_\eta(y) = \int_{-\infty}^{+\infty} p(x, y)\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{2\pi}(1 + xy) \cdot \mathrm{e}^{-\frac{x^2+y^2}{2}} \mathrm{d}x = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}} \Longrightarrow \eta \sim N(0, 1)$$

但显然 (ξ, η) 不服从二元正态分布

二、边际 (marginal) 分布

- 连续情形　　问题：均匀分布的边际分布是否还是均匀分布?

▸ 例：二维随机向量 (X, Y) 服从区域 $D = \left\{(x, y) \,\middle|\, x^2 + y^2 \le 1\right\}$ 上的均匀分布，求随机变量 X, Y 的边际概率密度.



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

圆面上均匀分布的边际分布不是一维均匀分布

$$p_X(x) = \int_{-\infty}^{+\infty} p(x, y)\mathrm{d}y = \begin{cases} \displaystyle\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} \frac{1}{\pi}\mathrm{d}y, & -1 \le x \le 1 \\ 0, & |x| > 1 \end{cases}$$

$$= \begin{cases} \dfrac{2}{\pi}\sqrt{1 - x^2}, & |x| \le 1 \\ 0, & |x| > 1 \end{cases}$$

$$p_Y(y) = \int_{-\infty}^{+\infty} p(x, y)\mathrm{d}x = \begin{cases} \dfrac{2}{\pi}\sqrt{1 - y^2}, & |y| \le 1 \\ 0, & |y| > 1 \end{cases}$$

三、条件 (conditional) 分布

- 对于多个随机事件可以讨论它们的条件概率，同样地，对于多个随机变量也可以讨论它们的条件分布. 仍就二维的场合进行讨论.

- 离散情形：给定 (ξ,η) 的联合分布律，可求得 ξ,η 的边际分布律，

已知 $\xi=x_i$ 时，事件 $\{\eta=y_j\}$ 的条件概率为

$$P\{\eta=y_j \mid \xi=x_i\}=\frac{P\{\xi=x_i,\eta=y_j\}}{P\{\xi=x_i\}}=\frac{p(x_i,y_j)}{p_1(x_i)}$$

随机变量 η 关于 ξ 的条件分布

随机变量 $(\eta \mid \xi=x_i)$ 的分布

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

三、条件 (conditional) 分布

- 对于多个随机事件可以讨论它们的条件概率，同样地，对于多个随机变量也可以讨论它们的条件分布. 仍就二维的场合进行讨论.

- 离散情形：给定 (ξ, η) 的联合分布律，可求得 ξ, η 的边际分布律，

  已知 $\eta = y_j$ 时，事件 $\{\xi = x_i\}$ 的条件概率为

$$P\left\{\xi = x_i \middle| \eta = y_j\right\} = \frac{P\left\{\xi = x_i, \eta = y_j\right\}}{P\left\{\eta = y_j\right\}} = \frac{p\left(x_i, y_j\right)}{p_2\left(y_j\right)}$$

随机变量 ξ 关于 η 的条件分布

随机变量 $\left(\xi \middle| \eta = y_j\right)$ 的分布

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

三、条件 (conditional) 分布

- 例：从 1, 2, 3, 4 中随机取一个数 ξ，再从 1，…, ξ 中随机地取一个数 η，(ξ, η) 的联合分布为

ξ, η 的边际分布律为

$\xi \backslash \eta$	1	2	3	4	$p_{i\cdot}$
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$	1

① 随机变量 $\left(\xi \,\middle|\, \eta = 1\right)$ 的条件分布律为

$\xi \mid \eta = 1$	1	2	3	4
$p_{i\cdot}$	$\frac{12}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{3}{25}$

② 随机变量 $\left(\xi \,\middle|\, \eta = 2\right)$ 的条件分布律为

$\xi \mid \eta = 2$	2	3	4
$p_{i\cdot}$	$\frac{6}{13}$	$\frac{4}{13}$	$\frac{3}{13}$

③ 随机变量 $\left(\eta \,\middle|\, \xi = 3\right)$ 的条件分布律为

$\eta \mid \xi = 3$	1	2	3
$p_{i\cdot}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$