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：$P(A\,|\,B) = \dfrac{P(AB)}{P(B)}.$

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

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

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

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

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

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

  - 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 (r.v.) 与分布函数 (c.d.f.)：$F(x) = P\left\{\xi(\omega) < x\right\}, \quad -\infty < x < \infty$

    - 分布函数是单调不降、取值 $[0, 1]$ 的左连续函数. 它完整描述了随机变量，是研究的主要对象.
    - 随机变量依取值
      - 离散型：分布律 (mass function)、分布列
      - 连续型：概率密度 (p.d.f.)
    - 概率计算：
      - 离散型：$P\{x \in D\} = \sum\limits_{x_i \in D} p_i, \quad P\left\{(x, y) \in D\right\} = \sum\limits_{(x_i, y_j) \in D} p_{ij}$
      - 连续型：$P\{x \in D\} = \int\limits_D p(x)\mathrm{d}x, \quad P\left\{(x, y) \in D\right\} = \iint\limits_D p(x, y)\mathrm{d}x\mathrm{d}y$
    - 主要分布：
      - 离散型：Bernoulli, binomial, Poisson, hyper-geometric, geometric
      - 连续型：uniform, exponential, normal, Γ

已学知识点

- 第三章 随机变量与分布函数

  - 随机向量，随机变量的独立性：

    - 随机向量即多元随机变量
      - 联合分布： 联合分布函数、联合分布律、联合密度
      - 边际分布： 边际分布函数、边际分布律、边际密度
      - 条件分布： 条件分布函数、条件分布律、条件密度
      - 独立性： 与事件独立性几乎完全相同
    - 主要分布：
      - 离散型： 多项分布、多元超几何分布
      - 连续型： 二元均匀分布、二元正态分布

$$(\xi, \eta) \sim N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right) \sim N\left(\begin{pmatrix}\mu_1 \\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}\right) \Longrightarrow \begin{cases} \xi \sim N\left(\mu_1, \sigma_1^2\right), \quad \eta \sim N\left(\mu_2, \sigma_2^2\right) \\ \left(\eta \mid \xi = x\right) \sim N\left(\mu_2 + \rho\dfrac{\sigma_2}{\sigma_1}\left(x-\mu_1\right), \sigma_2^2\left(1-\rho^2\right)\right) \\ \left(\xi \mid \eta = y\right) \sim N\left(\mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}\left(y-\mu_2\right), \sigma_1^2\left(1-\rho^2\right)\right) \\ \xi \text{ 与 } \eta \text{ 相互独立} \iff \rho = 0 \end{cases}$$

已学知识点

- 第三章　随机变量与分布函数
  - 随机变量的函数及其分布：
    - 随机变量的函数什么情况下还是随机变量
    - 离散型易：$\begin{cases} \eta = g(\xi): & \text{对应法} \\ \zeta = g(\xi, \eta): & \text{表上作业法，独立情形和的卷积公式} \end{cases}$
    - 连续型难：$\eta = g(\xi)$，直接法 $F_\eta(y) = P\{\eta < y\} = \int_{g(x)<y} p_\xi(x)\mathrm{d}x$

3.3 随机变量的函数及其分布 (续)

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

三、随机向量的函数的分布律

- 设 $\xi=(\xi_1,\xi_2,\ \dots,\xi_n)$ 的密度函数为 $p(x_1,x_2,\ \dots,x_n)$，若随机变量 $\eta=g(\xi_1,\xi_2,\ \dots,\xi_n)$，则 η 的分布函数为

$$G(y)=P\{\eta<y\}=P\left\{g(\xi_1,\xi_2,\ \dots,\xi_n)<y\right\}$$

$$=\underset{g(\xi_1,\xi_2,\ \dots,\xi_n)<y}{\int\cdots\int} p(x_1,x_2,\ \dots,x_n)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n$$

$$\xrightarrow{\eta=g(\xi_1,\xi_2,\ \dots,\xi_n)\ \text{的密度函数为}} g(y)=\frac{\mathrm{d}}{\mathrm{d}y}G(y)=G'(y)\,,\quad \mathrm{a.e.}\ y\in\mathbb{R}$$

三、随机向量的函数的分布律

- 和的分布

定理：设 (X, Y) 的密度函数为 $p(x, y)$，则 $Z = X + Y$ 的密度函数为

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

特别，当 X, Y 相互独立时则有

$$p_Z(z) = \int_{-\infty}^{+\infty} p_X(x)\, p_Y(z - x)\mathrm{d}x = \int_{-\infty}^{+\infty} p_X(z - y)\, p_Y(y)\mathrm{d}y$$

三、随机向量的函数的分布律

● 和的分布

定理：设 (X, Y) 的密度函数为 $p(x, y)$，则 $Z = X + Y$ 的密度函数为

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

特别，当 X, Y 相互独立时则有

$$p_Z(z) = \int_{-\infty}^{+\infty} p_X(x)\, p_Y(z-x)\mathrm{d}x = \int_{-\infty}^{+\infty} p_X(z -$$

$Z = X + Y$ 的分布函数

$$G(z) = P\{Z < z\} = P\{X + Y < z\} = \iint_{x+y<z} p(x, y)\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z-x} p(x, y)\,\mathrm{d}y\right)\mathrm{d}x$$

$Z = X + Y$ 的密度函数

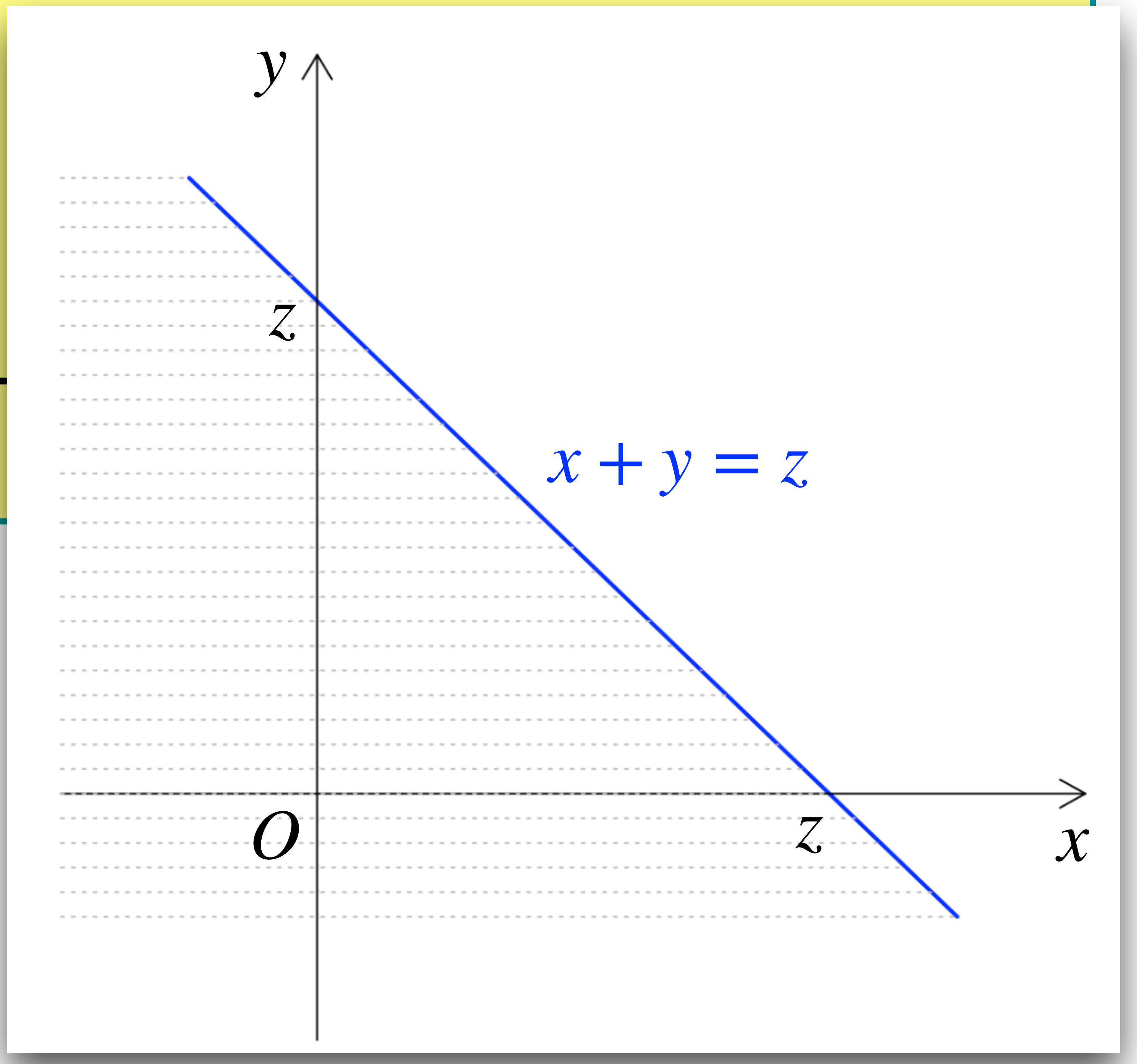
$$p_Z(z) = \frac{\mathrm{d}}{\mathrm{d}z}G(z) = \int_{-\infty}^{+\infty} p(x, z-x)\,\mathrm{d}x$$

X, Y 相互独立

$$p(x, y) = p_X(x) \cdot p_Y(y)$$

$$\Longrightarrow \quad p_Z(z) = \int_{-\infty}^{+\infty} p_X(x) \cdot p_Y(z-x)\,\mathrm{d}x$$

卷积公式



三、随机向量的函数的分布律

- 商、差、积的分布

定理：设 (X, Y) 的密度函数为 $p(x, y)$，则 $Z_1 = \dfrac{X}{Y}, Z_2 = X - Y, Z_3 = XY$ 的密度函数分别为

$$p_{Z_1}(z) = \int_{-\infty}^{+\infty} |y| \cdot p(zy, y)\,\mathrm{d}y$$

$$p_{Z_2}(z) = \int_{-\infty}^{+\infty} p(x, x-z)\,\mathrm{d}x = \int_{-\infty}^{+\infty} p(y+z, y)\,\mathrm{d}y$$

$$p_{Z_3}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} \cdot p\left(x, \frac{z}{x}\right)\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{|y|} \cdot p\left(\frac{z}{y}, y\right)\mathrm{d}y$$

三、随机向量的函数的分布律

- 商、差、积的分布

定理：设 (X, Y) 的密度函数为 $p(x, y)$，则 $Z_1 = \frac{X}{Y}, Z_2 = X - Y, Z_3 = XY$ 的密度函数分别为

$$p_{Z_1}(z) = \int_{-\infty}^{+\infty} |y| \cdot p(zy, y)\,\mathrm{d}y$$

$$p_{Z_2}(z) = \int_{-\infty}^{+\infty} p(x, x-z)\,\mathrm{d}x = \int_{-\infty}^{+\infty} p(y+z, y)$$

$$p_{Z_3}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} \cdot p\left(x, \frac{z}{x}\right)\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{|y|} \cdot$$

$Z_1 = \frac{X}{Y}$ 的分布函数

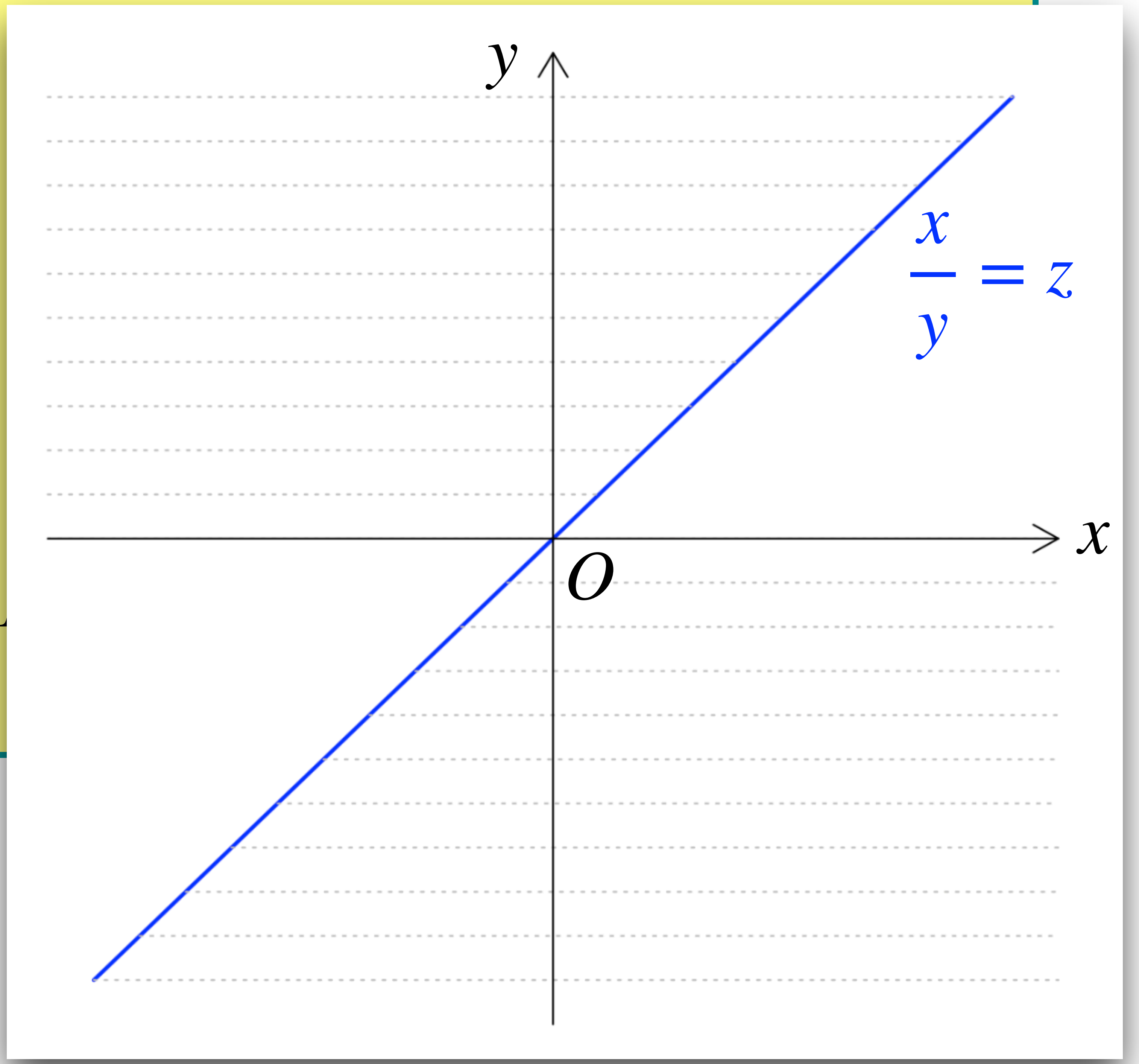
$$G(z) = P\{Z_1 < z\} = P\left\{\frac{X}{Y} < z\right\} = \iint_{\frac{x}{y}<z} p(x, y)\,\mathrm{d}x\mathrm{d}y$$

$$= \int_0^{+\infty}\left(\int_{-\infty}^{zy} p(x, y)\,\mathrm{d}x\right)\mathrm{d}y + \int_{-\infty}^{0}\left(\int_{zy}^{+\infty} p(x, y)\,\mathrm{d}x\right)\mathrm{d}y$$

$Z_1 = \frac{X}{Y}$ 的密度函数

$$p_{Z_1}(z) = \frac{\mathrm{d}}{\mathrm{d}z}G(z) = \int_0^{+\infty} y \cdot p(zy, y)\,\mathrm{d}y - \int_{-\infty}^{0} y \cdot p(zy, y)\,\mathrm{d}y$$

$$= \int_{-\infty}^{+\infty} |y| \cdot p(zy, y)\,\mathrm{d}y$$



三、随机向量的函数的分布律

- 顺序统计量的分布

设 $X_1, X_2, \ldots, X_n$ 是 n 个相互独立的随机变量，具有相同的分布函数 $F(x)$ 和密度函数 $p(x)$，

对 $X_1, X_2, \ldots, X_n$ 的一组样本 $x_1, x_2, \ldots, x_n$，将其由小到大排序为

$$x_{\min} = x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} = x_{\max}$$

定义随机变量 $X_{(k)}$ 取值 $x_{(k)}$，$k = 1, 2, \ldots, n$，称 $X_{(k)}$ 为第 k 个顺序统计量.

三、随机向量的函数的分布律

- 顺序统计量的分布

定理：设 X, Y 相互独立，分布函数分别为 $F_X(x), F_Y(y)$，则 $\xi = \max\{X, Y\}$ 与 $\eta = \min\{X, Y\}$ 的分布函数分别为

$$F_\xi(z) = F_X(z) \cdot F_Y(z)$$

$$F_\eta(z) = 1 - \left[1 - F_X(z)\right] \cdot \left[1 - F_Y(z)\right]$$

$\xi = \max\{X, Y\}$ 的分布函数

$$F_\xi(z) = P\left\{\xi < z\right\} = P\{\max\{X, Y\} < z\} = P\{X < z, Y < z\}$$

X, Y 相互独立

$$= P\{X < z\} \cdot P\{Y < z\} = F_X(z) \cdot F_Y(z)$$

$\eta = \min\{X, Y\}$ 的分布函数

$$F_\eta(z) = P\left\{\eta < z\right\} = P\{\min\{X, Y\} < z\} = 1 - P\{\min\{X, Y\} \geq z\}$$

X, Y 相互独立

$$= 1 - P\{X \geq z, Y \geq z\} = 1 - P\{X \geq z\} \cdot P\{Y \geq z\}$$

$$= 1 - \left[1 - P\{X < z\}\right] \cdot \left[1 - P\{Y < z\}\right] = 1 - \left[1 - F_X(z)\right] \cdot \left[1 - F_Y(z)\right]$$

三、随机向量的函数的分布律

- 顺序统计量的分布

▶ 推广到 n 个相互独立的随机变量的情形

设 $X_1, X_2, \ldots, X_n$ 是 n 个相互独立的随机变量，它们的分布函数分别为 $F_{X_i}(x_i), i=1, \ldots, n$.

定义 $\xi = \max\limits_{1\le i\le n}\{X_i\}, \eta = \min\limits_{1\le i\le n}\{X_i\}$，则 ξ, η 的分布函数分别为

$$F_\xi(z) = F_{X_1}(z)\cdot F_{X_2}(z)\cdot \cdots \cdot F_{X_n}(z)$$

$$F_\eta(z) = 1-\left[1-F_{X_1}(z)\right]\cdot\left[1-F_{X_2}(z)\right]\cdot \cdots \cdot\left[1-F_{X_n}(z)\right]$$

特别，当 $X_1, X_2, \ldots, X_n$ 相互独立且具有相同的分布函数 $F_X(x)$ 时，则有

$$F_\xi(z) = \left[F_X(z)\right]^n$$

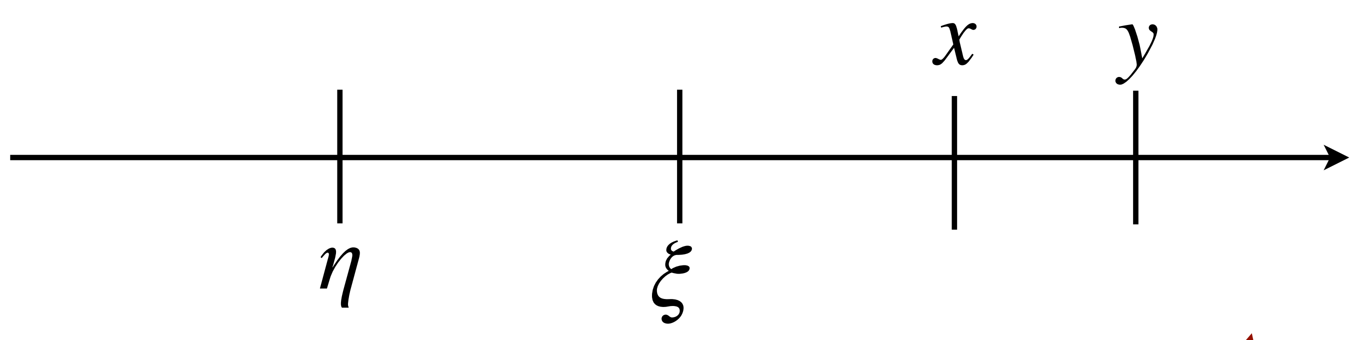
$$F_\eta(z) = 1-\left[1-F_X(z)\right]^n$$

三、随机向量的函数的分布律

- 顺序统计量的分布

$$\xi=\max_{1\le i\le n}\{X_i\},\ \eta=\min_{1\le i\le n}\{X_i\}\ \Longrightarrow\ \eta\le\xi$$

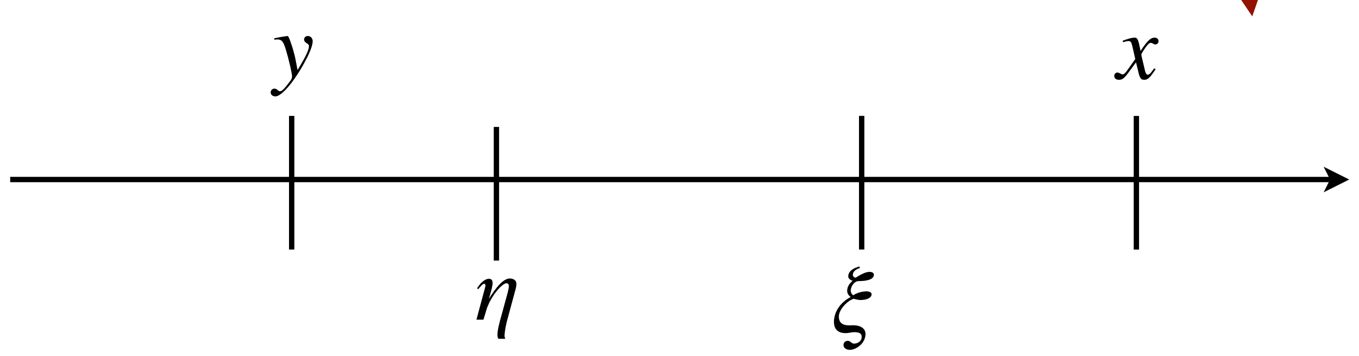
  - (ξ,η) 的联合分布



(ξ,η) 的联合分布函数

$$G(x,y)=P\{\xi<x,\eta<y\}=\begin{cases}P\{\xi<x\}\ , & x\le y\\ P\{\xi<x\}-P\{\xi<x,\eta\ge y\}\ , & x>y\end{cases}$$

$$=\begin{cases}[F(x)]^n\ , & x\le y\\ [F(x)]^n-[F(x)-F(y)]^n\ , & x>y\end{cases}$$



(ξ,η) 的联合概率密度函数

$$q(x,y)=\frac{\partial^2}{\partial x\partial y}G(x,y)=\begin{cases}0\ , & x\le y\\ n(n-1)[F(x)-F(y)]^{n-2}p(x)\,p(y)\ , & x>y\end{cases}$$

三、随机向量的函数的分布律

- 顺序统计量的分布

$$\xi = \max_{1\le i\le n}\{X_i\},\ \eta = \min_{1\le i\le n}\{X_i\}$$

- (ξ,η) 的联合分布

$$q(x,y) = \begin{cases} 0\ , & x \le y \\ n(n-1)\left[F(x)-F(y)\right]^{n-2} p(x)\,p(y)\ , & x > y \end{cases}$$

- 极差 $R = \xi - \eta$ 的分布

定理：设 (X,Y) 的密度函数为 $p(x,y)$，则 $Z_1 = \dfrac{X}{Y}$, $Z_2 = X - Y$, $Z_3 = XY$ 的密度函数分别为

$$p_{Z_1}(z) = \int_{-\infty}^{+\infty} |y| \cdot p(zy, y)\,\mathrm{d}y$$

$$p_{Z_2}(z) = \int_{-\infty}^{+\infty} p(x, x-z)\,\mathrm{d}x = \int_{-\infty}^{+\infty} p(y+z, y)\,\mathrm{d}y$$

$$p_{Z_3}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|}\cdot p\left(x, \frac{z}{x}\right)\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{|y|}\cdot p\left(\frac{z}{y}, y\right)\mathrm{d}y$$

$$p_R(r) = \begin{cases} 0\ , & r \le 0 \\ \int_{-\infty}^{+\infty} p(x, x-r)\,\mathrm{d}x\ , & r > 0 \end{cases}$$

$$= \begin{cases} 0\ , & r \le 0 \\ \int_{-\infty}^{+\infty} n(n-1)\left[F(x)-F(x-r)\right]^{n-2} p(x)\cdot p(x-r)\,\mathrm{d}x\ , & r > 0 \end{cases}$$

定理：设 X, Y 相互独立，分布函数分别为 $F_X(x), F_Y(y)$，则 $\xi = \max\{X, Y\}$ 与 $\eta = \min\{X, Y\}$ 的分布函数分别为

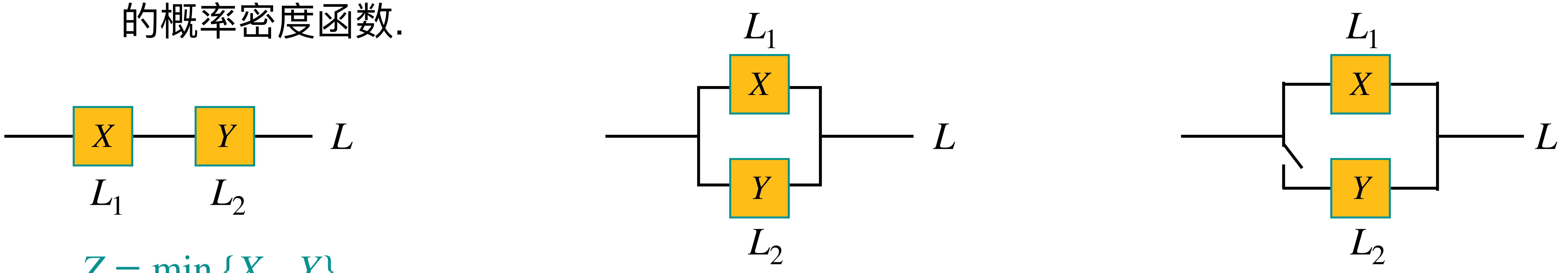
$$F_\xi(z) = F_X(z) \cdot F_Y(z)$$

$$F_\eta(z) = 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)]$$

三、随机向量的函数的分布律

- 顺序统计量的分布

▶ 例：设系统 L 由两个独立的子系统 L_1, L_2 连接而成，共有三种连接方式：(1) 串联，(2) 并联，(3) 备用 (当系统 L_1 损坏时 L_2 开始工作)，如下图所示. 设 L_1, L_2 的寿命分别为 X, Y，已知它们的概率密度分别为 $p_X(x)$, $p_Y(y)$，分别就以上三种连接方式给出系统 L 的寿命 Z 的概率密度函数.



$Z = \min\{X , Y\}$

$$\Longrightarrow \quad p_Z(z) = p_X(z) \cdot [1 - F_Y(z)] + p_Y(z) \cdot [1 - F_X(z)]$$

三、随机向量的函数的分布律

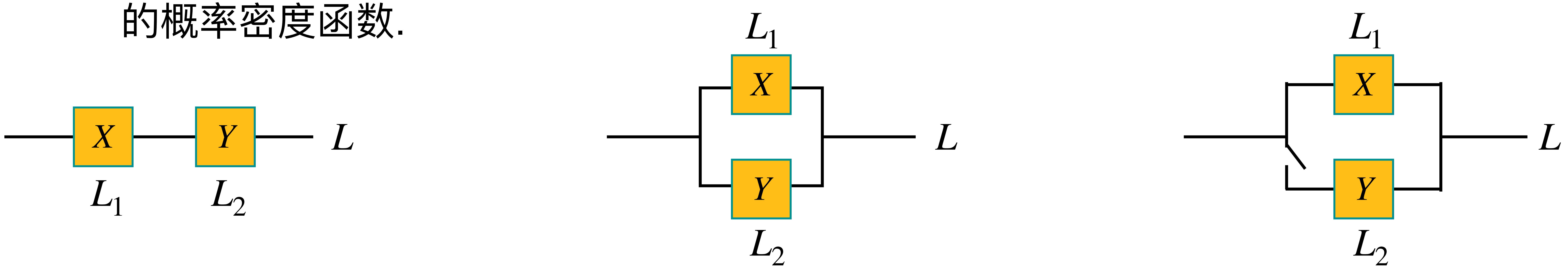
定理：设 X, Y 相互独立，分布函数分别为 $F_X(x), F_Y(y)$，则 $\xi = \max\{X, Y\}$ 与 $\eta = \min\{X, Y\}$ 的分布函数分别为

$$F_\xi(z) = F_X(z) \cdot F_Y(z)$$

$$F_\eta(z) = 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)]$$

- 顺序统计量的分布

例：设系统 L 由两个独立的子系统 L_1, L_2 连接而成，共有三种连接方式：(1) 串联，(2) 并联，(3) 备用 (当系统 L_1 损坏时 L_2 开始工作)，如下图所示. 设 L_1, L_2 的寿命分别为 X, Y，已知它们的概率密度分别为 $p_X(x)$, $p_Y(y)$，分别就以上三种连接方式给出系统 L 的寿命 Z 的概率密度函数.



$$Z = \max\{X , Y\}$$

$$\Longrightarrow \quad p_Z(z) = p_X(z) \cdot F_Y(z) + p_Y(z) \cdot F_X(z)$$

定理：设 (X, Y) 的密度函数为 $p(x, y)$，则 $Z = X + Y$ 的密度函数为

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

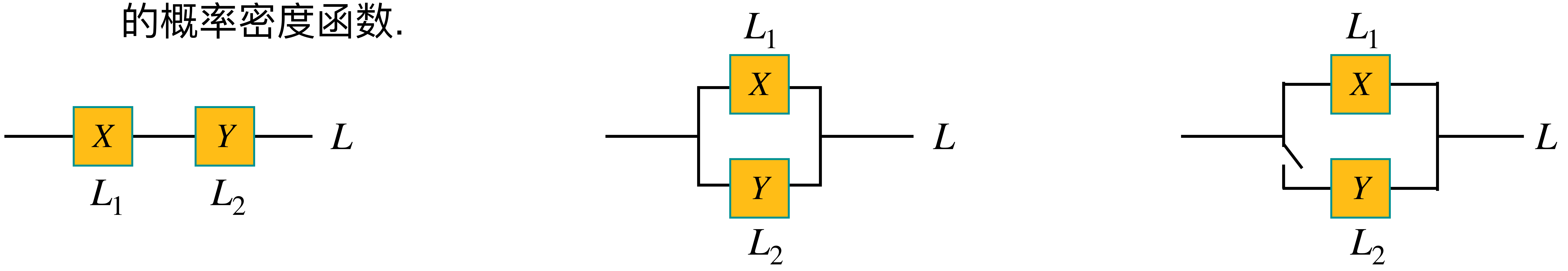
特别，当 X, Y 相互独立时则有

$$p_Z(z) = \int_{-\infty}^{+\infty} p_X(x)\, p_Y(z-x)\mathrm{d}x = \int_{-\infty}^{+\infty} p_X(z-y)\, p_Y(y)\mathrm{d}y$$

三、随机向量的函数的分布律

- 顺序统计量的分布

  - 例：设系统 L 由两个独立的子系统 L_1, L_2 连接而成，共有三种连接方式：(1) 串联，(2) 并联，(3) 备用 (当系统 L_1 损坏时 L_2 开始工作)，如下图所示. 设 L_1, L_2 的寿命分别为 X, Y，已知它们的概率密度分别为 $p_X(x)$，$p_Y(y)$，分别就以上三种连接方式给出系统 L 的寿命 Z 的概率密度函数.



$$Z = X + Y$$

$$\Longrightarrow \quad p_Z(z) = \int_{-\infty}^{\infty} p_X(x) \cdot p_Y(z-x)\ \mathrm{d}x = \int_{-\infty}^{\infty} p_X(z-y) \cdot p_Y(y)\ \mathrm{d}y$$

四、连续型随机向量的变换

- 问题：设 $(\xi_1, \xi_2, \ldots, \xi_n)$ 的联合概率密度函数为 $p(x_1, x_2, \ldots, x_n)$，确定 $(\eta_1, \eta_2, \ldots, \eta_m)$ 的联合分布，其中

$$\begin{cases} \eta_1 = g_1(\xi_1, \xi_2, \ldots, \xi_n) \\ \eta_2 = g_2(\xi_1, \xi_2, \ldots, \xi_n) \\ \vdots \quad \vdots \\ \eta_m = g_m(\xi_1, \xi_2, \ldots, \xi_n) \end{cases}$$

$(\eta_1, \eta_2, \ldots, \eta_m)$ 的联合分布函数

$$G(y_1, y_2, \ldots, y_m) = P\{\eta_1 < y_1, \eta_2 < y_2, \ldots, \eta_m < y_m\}$$

$$= \underset{\substack{g_1(x_1, x_2, \ldots, x_n) < y_1 \\ g_2(x_1, x_2, \ldots, x_n) < y_2 \\ \vdots \quad \vdots \\ g_m(x_1, x_2, \ldots, x_n) < y_m}}{\int \cdots \int} p(x_1, x_2, \ldots, x_n)\, \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n$$

$m = 1$： 随机向量的函数的情形.

$m = n = 1$： 单个随机变量的函数的情形.

四、连续型随机向量的变换

$$\begin{cases} \eta_1 = g_1(\xi_1, \xi_2, \ldots, \xi_n) \\ \eta_2 = g_2(\xi_1, \xi_2, \ldots, \xi_n) \\ \vdots \quad \vdots \\ \eta_m = g_m(\xi_1, \xi_2, \ldots, \xi_n) \end{cases}$$

- 考虑一个重要的特殊情形

p.d.f. $q(y_1, y_2, \ldots, y_n)$

设 $(\xi_1, \xi_2, \ldots, \xi_n)$ 的联合概率密度函数为 $p(x_1, x_2, \ldots, x_n)$，确定 $(\eta_1, \eta_2, \ldots, \eta_m)$ 的联合分布.

考虑 $m = n$ 且 $(\xi_1, \xi_2, \ldots, \xi_n)$ 与 $(\eta_1, \eta_2, \ldots, \eta_n)$ 有一一对应变换关系的情形.

$$\text{记} \begin{cases} y_1 = g_1(x_1, x_2, \ldots, x_n) \\ y_2 = g_2(x_1, x_2, \ldots, x_n) \\ \vdots \qquad \vdots \\ y_n = g_n(x_1, x_2, \ldots, x_n) \end{cases} \xrightarrow{\text{一一对应关系，存在唯一反函数}} \begin{cases} x_1 = h_1(y_1, y_2, \ldots, y_n) \\ x_2 = h_2(y_1, y_2, \ldots, y_n) \\ \vdots \qquad \vdots \\ x_n = h_n(y_1, y_2, \ldots, y_n) \end{cases}$$

$$\xrightarrow{(\eta_1, \eta_2, \ldots, \eta_n)\text{ 的分布函数}} G(y_1, y_2, \ldots, y_n) = P\{\eta_1 < y_1, \eta_2 < y_2, \ldots, \eta_n < y_n\} = \underset{\substack{u_1 < y_1 \\ u_2 < y_2 \\ \vdots \quad \vdots \\ u_n < y_n}}{\int \cdots \int} q(u_1, u_2, \ldots, u_n)\, \mathrm{d}u_1 \mathrm{d}u_2 \cdots \mathrm{d}u_n$$

四、连续型随机向量的变换

- 考虑一个重要的特殊情形

$$\xrightarrow{\text{变量变换}} \begin{cases} x_1 = h_1(y_1, y_2, \dots, y_n) \\ x_2 = h_2(y_1, y_2, \dots, y_n) \\ \vdots \qquad \vdots \\ x_n = h_n(y_1, y_2, \dots, y_n) \end{cases}$$

$$G(y_1, y_2, \dots, y_n) = \underset{\substack{g_1(x_1, x_2, \dots, x_n) < y_1 \\ g_2(x_1, x_2, \dots, x_n) < y_2 \\ \vdots \qquad \vdots \\ g_n(x_1, x_2, \dots, x_n) < y_n}}{\int \cdots \int} p(x_1, x_2, \dots, x_n)\, dx_1 dx_2 \cdots dx_n = \underset{\substack{u_1 < y_1 \\ u_2 < y_2 \\ \vdots \quad \vdots \\ u_n < y_n}}{\int \cdots \int} q(u_1, u_2, \dots, u_n)\, du_1 du_2 \cdots du_n$$

$$\Longrightarrow \quad q(y_1, y_2, \dots, y_n) = \begin{cases} p(x_1, x_2, \dots, x_n) \cdot |J|, & (y_1, y_2, \dots, y_n) \text{ 属于 } (g_1, g_2, \dots, g_n) \text{ 的值域} \\ 0, & \text{其它} \end{cases}$$

坐标变换的 Jacobian 行列式： $J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial y_n} & \frac{\partial x_2}{\partial y_n} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$

假定偏导数存在且连续

四、连续型随机向量的变换

- 例：设 (ξ_1, ξ_2) 的密度函数为 $p(x_1, x_2)$，而

$$\begin{cases} \eta_1 = a\xi_1 + b\xi_2 \\ \eta_2 = c\xi_1 + d\xi_2 \end{cases} \quad 且 \quad \Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$$

求 (η_1, η_2) 的密度函数.

设 $\begin{cases} y_1 = ax_1 + bx_2 \\ y_2 = cx_1 + dx_2 \end{cases} \Longrightarrow \begin{cases} x_1 = \dfrac{d}{\Delta}y_1 - \dfrac{b}{\Delta}y_2 \\ x_2 = -\dfrac{c}{\Delta}y_1 + \dfrac{a}{\Delta}y_2 \end{cases} \xrightarrow{\text{变换的 Jacobian}} J = \begin{vmatrix} \dfrac{d}{\Delta} & -\dfrac{b}{\Delta} \\ -\dfrac{c}{\Delta} & \dfrac{a}{\Delta} \end{vmatrix} = \dfrac{ad-bc}{\Delta^2} = \dfrac{1}{ad-bc}$

$$\Longrightarrow \quad q(y_1, y_2) = p(x_1, x_2) \cdot |J| = \frac{1}{|ad-bc|} \cdot p\left(\frac{d}{\Delta}y_1 - \frac{b}{\Delta}y_2,\ -\frac{c}{\Delta}y_1 + \frac{a}{\Delta}y_2\right)$$

四、连续型随机向量的变换

- 例：设 ξ, η 相互独立，且 $\xi \sim \chi_m^2, \eta \sim \chi_n^2$，求 $\alpha = \xi + \eta$ ，$\beta = \dfrac{\xi}{\eta} \cdot \dfrac{n}{m}$ 的联合和密度函数 $q(u, v)$.

$$\begin{cases} \xi \sim \chi_m^2 \Longrightarrow p_\xi(x) = \dfrac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} x^{\frac{m}{2}-1} \mathrm{e}^{-\frac{x}{2}} \\ \eta \sim \chi_n^2 \Longrightarrow p_\eta(y) = \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} \mathrm{e}^{-\frac{y}{2}} \end{cases} \xrightarrow{\xi,\ \eta \text{ 相互独立}} p(x, y) = p_\xi(x) \cdot p_\eta(y)$$

$$= \frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} x^{\frac{m}{2}-1} y^{\frac{n}{2}-1} \mathrm{e}^{-\frac{x+y}{2}}, \quad x > 0 , y > 0$$

$$\xrightarrow{\text{作变换}} \begin{cases} u = x + y \\ v = \dfrac{x}{y} \cdot \dfrac{n}{m} \end{cases} \Longrightarrow \begin{cases} x = \dfrac{muv}{n + mv} \\ y = \dfrac{nu}{n + mv} \end{cases} \xrightarrow{\text{变换的 Jacobian}} J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{mv}{n + mv} & \dfrac{n}{n + mv} \\ \dfrac{mu(n + mv) - m^2uv}{(n + mv)^2} & -\dfrac{mnu}{(n + mv)^2} \end{vmatrix}$$

$$= -\frac{m}{n} \cdot \frac{u}{\left(1 + \frac{m}{n}v\right)^2}$$

四、连续型随机向量的变换

- 例：设 ξ, η 相互独立，且 $\xi \sim \chi^2_m, \eta \sim \chi^2_n$，求 $\alpha = \xi + \eta$, $\beta = \dfrac{\xi}{\eta} \cdot \dfrac{n}{m}$ 的联合和密度函数 $q(u, v)$.

由相同的随机向量构成的不同函数也可能是独立的

(α , β) 的联合密度为 $\longrightarrow$ $q(u, v) = p(x, y) \cdot |J|$

$$= \frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{\frac{m}{n}uv}{1+\frac{m}{n}v}\right)^{\frac{m}{2}-1} \left(\frac{u}{1+\frac{m}{n}v}\right)^{\frac{n}{2}-1} e^{-\frac{u}{2}} \cdot \frac{m}{n} \cdot \frac{u}{\left(1+\frac{m}{n}v\right)^2}, \quad u>0 , v>0$$

$$= \left[\underbrace{\frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)} u^{\frac{m+n}{2}-1} e^{-\frac{u}{2}}}_{\alpha \sim \chi^2_{m+n}}\right] \cdot \left[\overbrace{\frac{\Gamma\left(\frac{m+n}{2}\right)\cdot\left(\frac{m}{n}\right)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{v^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}v\right)^{\frac{m+n}{2}}}}^{\beta \sim F_{m,n}}\right], \quad u>0 , v>0$$

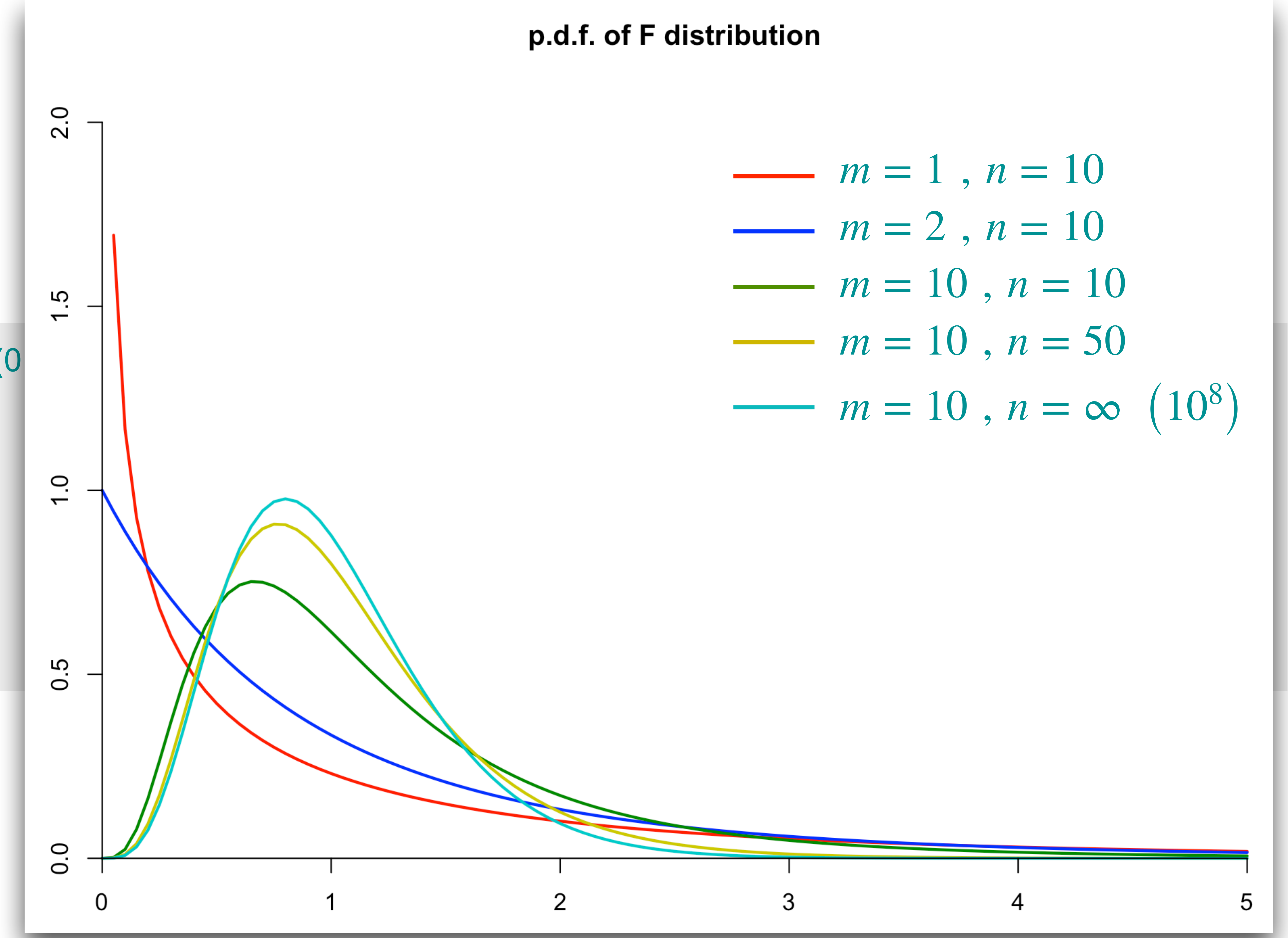
$\Longrightarrow$ α, β 相互独立

四、连续型随机向量的变换

- **例**：设 ξ, η 相互独立，且 $\xi \sim \chi_m^2, \eta \sim \chi_n^2$，求 $\alpha = \xi + \eta$, $\beta = \dfrac{\xi}{\eta} \cdot \dfrac{n}{m}$ 的联合和密度函数 $q(u, v)$.

$$\beta \sim F_{m,\,n} \quad : \quad p_\beta(v) = \frac{\Gamma\left(\frac{m+n}{2}\right) \cdot \left(\frac{m}{n}\right)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{v^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}v\right)^{\frac{m+n}{2}}} , \qquad v > 0$$

```
curve(df(x, 1, 10), 0, 5, axes = FALSE, xlab = '', ylab = '', lwd = 2, col = 'red', ylim = c(0
curve(df(x, 2, 10), 0, 5, add = TRUE, lwd = 2, col = 'blue')
curve(df(x, 10, 10), 0, 5, add = TRUE, lwd = 2, col = 'green4')
curve(df(x, 10, 50), 0, 5, add = TRUE, lwd = 2, col = 'yellow3')
curve(df(x, 10, 10^8), 0, 5, add = TRUE, lwd = 2, col = 'cyan3')
axis(1, pos = 0, at = 0:5)
axis(2, pos = 0, at = c(0, 0.5, 1.0, 1.5, 2.0))
```



四、连续型随机向量的变换

- 增补变量法

▶ 例：设 ξ, η 相互独立，$\xi \sim N(0,\ 1)$，$\eta \sim \chi_n^2$，求 $T = \dfrac{\xi}{\sqrt{\dfrac{\eta}{n}}}$ 的密度函数.

$$\begin{cases} \xi \sim N(0,\ 1) \Longrightarrow p_\xi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ \eta \sim \chi_n^2 \Longrightarrow p_\eta(y) = \dfrac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} \end{cases} \xrightarrow{\text{引入增补变量 } S=\eta} \begin{cases} s = y \\ t = \dfrac{x}{\sqrt{\dfrac{y}{n}}} \end{cases} \Longrightarrow \begin{cases} x = t \cdot \sqrt{\dfrac{s}{n}} \\ y = s \end{cases}$$

$$\xrightarrow{\text{变换的 Jacobian}} J = \begin{vmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial y}{\partial s} \\ \dfrac{\partial x}{\partial t} & \dfrac{\partial y}{\partial t} \end{vmatrix} = \begin{vmatrix} \dfrac{t}{2n}\left(\dfrac{s}{n}\right)^{-\frac{1}{2}} & 1 \\ \left(\dfrac{s}{n}\right)^{\frac{1}{2}} & 0 \end{vmatrix} = -\left(\dfrac{s}{n}\right)^{\frac{1}{2}}$$

四、连续型随机向量的变换

$$p_T(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}},\quad -\infty<t<+\infty$$

- 增补变量法

  - 例：设 ξ,η 相互独立，$\xi\sim N(0,1)$，$\eta\sim\chi_n^2$，求 $T=\dfrac{\xi}{\sqrt{\frac{\eta}{n}}}$ 的密度函数.

(S,T) 的联合密度

$$q(s,t)=p(x,y)\cdot|J|=\frac{1}{\sqrt{2\pi}}e^{-\frac{st^2}{2n}}\cdot\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}s^{\frac{n}{2}-1}e^{-\frac{s}{2}}\cdot\left(\frac{s}{n}\right)^{\frac{1}{2}},\quad s>0,\ -\infty<t<+\infty$$

T 的边际密度

$$p_T(t)=\int_0^{\infty}q(s,t)\mathrm{d}s=\frac{1}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^{\infty}e^{-\left(1+\frac{t^2}{n}\right)\frac{s}{2}}\cdot s^{\frac{n+1}{2}-1}\mathrm{d}s$$

变量变换 $u=\left(1+\dfrac{t^2}{n}\right)\dfrac{s}{2}$

$$=\frac{1}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^{\infty}e^{-u}\cdot\left(\frac{2u}{1+\frac{t^2}{n}}\right)^{\frac{n+1}{2}-1}\cdot\frac{2}{1+\frac{t^2}{n}}\mathrm{d}u=\frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\boxed{\int_0^{\infty}e^{-u}\cdot u^{\frac{n+1}{2}-1}\mathrm{d}u}$$

$$\Gamma\left(\frac{n+1}{2}\right)$$

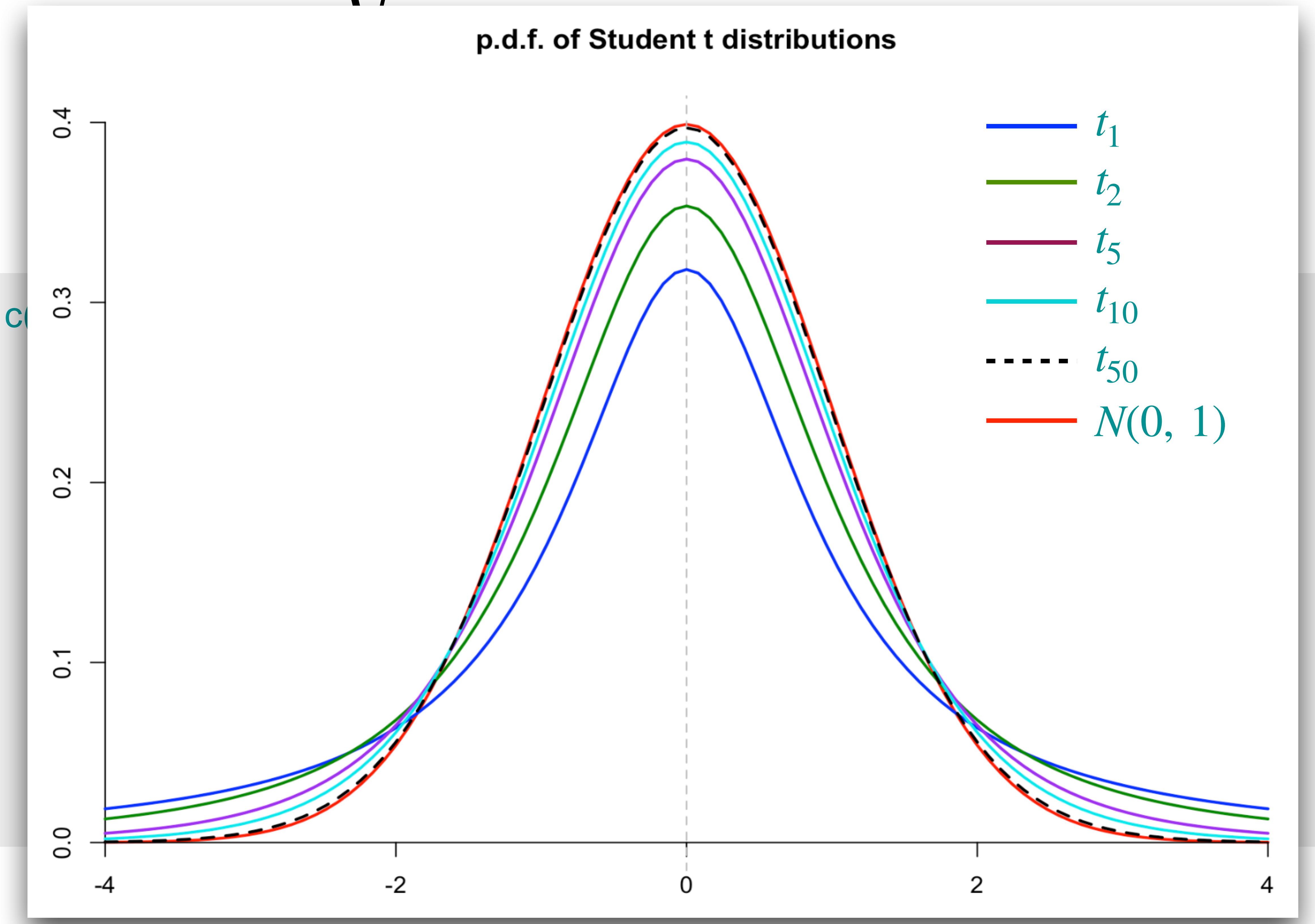
四、连续型随机向量的变换

- 增补变量法

  - 例：设 ξ, η 相互独立，$\xi \sim N(0,\ 1)$，$\eta \sim \chi_n^2$，求 $T = \dfrac{\xi}{\sqrt{\frac{\eta}{n}}}$ 的密度函数.

$$p_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < +\infty$$

```
curve(dnorm(x, 0, 1), -4, 4, axes = FALSE, xlab = '', ylab = '', lwd = 2, col = 'red', xlim = c
axis(1, pos = 0)
axis(2, pos = -4)
abline(v = 0, lty = 2, col = 'grey')
curve(dt(x, 1), -4, 4, add = TRUE, lwd = 2, col = 'blue')
curve(dt(x, 2), -4, 4, add = TRUE, lwd = 2, col = 'green4')
curve(dt(x, 5), -4, 4, add = TRUE, lwd = 2, col = 'purple')
curve(dt(x, 10), -4, 4, add = TRUE, lwd = 2, col = 'cyan2')
curve(dt(x, 50), -4, 4, add = TRUE, lwd = 2, col = 'black', lty = 2)
title(main = 'p.d.f. of Student t distributions')
```



五、随机变量的函数的独立性

定理：若 $\xi_1, \xi_2, \ldots, \xi_n$ 是相互独立的随机变量，则它们各自的函数构成的随机变量

$f_1(\xi_1), f_2(\xi_2), \ldots, f_n(\xi_n)$ 也是相互独立的随机变量，这里的 $f_i\ (i=1, 2, \ldots, n)$ 是任意的

一元 Borel 函数.

该定理可以推广到随机向量的场合.

对任意的一元 Borel 点集 $A_1, A_2, \ldots, A_n$

$$P\left\{f_1(\xi_1)\in A_1,\ f_2(\xi_2)\in A_2,\ \ldots,\ f_n(\xi_n)\in A_n\right\}$$

$$=P\left\{\xi_1\in f_1^{-1}(A_1),\ \xi_2\in f_2^{-1}(A_2),\ \ldots,\ \xi_n\in f_n^{-1}(A_n)\right\}$$

$$=P\left\{\xi_1\in f_1^{-1}(A_1)\right\}\cdot P\left\{\xi_2\in f_2^{-1}(A_2)\right\}\cdot\cdots\cdot P\left\{\xi_n\in f_n^{-1}(A_n)\right\}$$

$$=P\left\{f_1(\xi_1)\in A_1\right\}\cdot P\left\{f_2(\xi_2)\in A_2\right\}\cdot\cdots\cdot P\left\{f_n(\xi_n)\in A_n\right\}$$

$\Longrightarrow$ $f_1(\xi_1), f_2(\xi_2), \ldots, f_n(\xi_n)$ 是相互独立的随机变量

五、随机变量的函数的独立性

- 例：设 ξ, η 相互独立，均服从标准正态分布 $N(0,\ 1)$，试证化为极坐标后，$\rho = \sqrt{\xi^2 + \eta^2}$ 及 $\varphi = \arctan\left(\dfrac{\eta}{\xi}\right)$ （取值于 $[0,\ 2\pi]$） 是相互独立的随机变量.

设 $\begin{cases} r = \sqrt{x^2 + y^2} \\ \tan\theta = \dfrac{y}{x} \end{cases} \Longrightarrow \begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \quad (0 \le \theta \le 2\pi)$

$$\xrightarrow{\text{变换的 Jacobian}} J = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial y}{\partial r} \\ \dfrac{\partial x}{\partial \theta} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{vmatrix} = r$$

五、随机变量的函数的独立性

- 例：设 ξ, η 相互独立，均服从标准正态分布 $N(0,1)$，试证化为极坐标后，$\rho=\sqrt{\xi^2+\eta^2}$ 及 $\varphi=\arctan\left(\dfrac{\eta}{\xi}\right)$（取值于 $[0,2\pi]$）是相互独立的随机变量.

$$\begin{cases} \xi \sim N(0,1) \Longrightarrow p_\xi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \\ \eta \sim N(0,1) \Longrightarrow p_\eta(y)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \end{cases} \xrightarrow{\xi,\,\eta\text{ 相互独立}} p(x,y)=p_\xi(x)\cdot p_\eta(y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}},\quad -\infty<x,y<+\infty$$

$$\xrightarrow{(\rho,\,\varphi)\text{ 的联合密度}} q(r,\theta)=p(x,y)\cdot|J| =\frac{1}{2\pi}e^{-\frac{r^2}{2}}\cdot r=\frac{1}{2\pi}\cdot re^{-\frac{r^2}{2}},\quad 0\le r<+\infty,\ 0\le\theta\le 2\pi$$

$$\xrightarrow{\rho,\,\varphi\text{ 的边际密度}} \begin{cases} q_\rho(r)=\displaystyle\int_0^{2\pi} q(r,\theta)d\theta=re^{-\frac{r^2}{2}},\quad r\ge 0 \\ q_\varphi(\theta)=\displaystyle\int_0^{+\infty} q(r,\theta)dr=\dfrac{1}{2\pi},\quad 0\le\theta\le 2\pi \end{cases} \Longrightarrow q(r,\theta)=q_\rho(r)\cdot q_\varphi(\theta),\quad \text{相互独立}$$

五、随机变量的函数的独立性

- 例：设 U_1, U_2 相互独立，均服从 $[0, 1]$ 上的均匀分布，试证 $V_1 = \sqrt{-2\ln U_1}\cdot\cos(2\pi U_2)$ 及 $V_2 = \sqrt{-2\ln U_1}\cdot\sin(2\pi U_2)$ 相互独立且均服从标准正态分布.

$$\begin{cases} U_1 \sim U[0, 1] \implies p_{U_1}(x_1) = 1, \quad x_1 \in [0, 1] \\ U_2 \sim U[0, 1] \implies p_{U_2}(x_2) = 1, \quad x_2 \in [0, 1] \end{cases} \xrightarrow{U_1,\, U_2\text{ 相互独立}} p(x_1, x_2) = p_{U_1}(x_1)\cdot p_{U_2}(x_2) = 1, \quad 0 \le x_1, x_2 \le 1$$

设 $$\begin{cases} y_1 = \sqrt{-2\ln x_1}\cdot\cos(2\pi x_2) \\ y_2 = \sqrt{-2\ln x_1}\cdot\sin(2\pi x_2) \end{cases} \implies \begin{cases} y_1^2 + y_2^2 = -2\ln x_1 \\ y_2 = y_1\tan(2\pi x_2) \end{cases} \implies \begin{cases} x_1 = e^{-\frac{y_1^2+y_2^2}{2}} \\ x_2 = \dfrac{1}{2\pi}\arctan\dfrac{y_2}{y_1} \end{cases}$$

$$\xrightarrow{\text{变换的 Jacobian}} J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_2}{\partial y_1} \\ \dfrac{\partial x_1}{\partial y_2} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}$$

五、随机变量的函数的独立性

- 例：设 U_1, U_2 相互独立，均服从 $[0, 1]$ 上的均匀分布，试证 $V_1 = \sqrt{-2\ln U_1}\cdot\cos(2\pi U_2)$ 及 $V_2 = \sqrt{-2\ln U_1}\cdot\sin(2\pi U_2)$ 相互独立且均服从标准正态分布.

(V_1, V_2) 的联合密度

$$q(y_1, y_2) = p(x_1, x_2)\cdot|J| = 1\cdot\frac{1}{2\pi}\mathrm{e}^{-\frac{y_1^2+y_2^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_1^2}{2}}\right)\cdot\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_2^2}{2}}\right)$$

该结果常用来产生相互独立的 $N(0, 1)$ 随机数

V_1, V_2 的边际密度

$$\begin{cases} q_{V_1}(y_1) = \int_{-\infty}^{+\infty} q(y_1, y_2)\,\mathrm{d}y_2 = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_1^2}{2}} \\ q_{V_2}(y_2) = \int_{-\infty}^{+\infty} q(y_1, y_2)\,\mathrm{d}y_1 = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_2^2}{2}} \end{cases}$$

$\Longrightarrow$ $q(y_1, y_2) = q_{V_1}(y_1)\cdot q_{V_2}(y_2)$， 相互独立