

国际金融试验班 2022 年秋 · 时间序列

## 第 2 讲：概率基础

授课人：刘 岩

武汉大学经管学院金融系

2022 年 9 月 5 日

# 本讲内容

- ① 概率空间
- ② 1-元随机变量
- ③ 多元随机变量
- ④ 独立性

## 符号体系

- 数字:  $a, b, c$  或  $\alpha, \beta, \gamma$
- 数系: 实数  $\mathbb{R}$ , 有理数  $\mathbb{Q}$ , 整数  $\mathbb{Z}$ , 自然数  $\mathbb{N}$
- 向量-矩阵: 列向量  $x$ , 矩阵  $X$ , 转置  $x^T, X^T$ , 行列式  $\det X$
- 集合: 简单情形时, 如  $A, B$ ; 多类符号混用时, 如  $\mathcal{A}, \mathcal{B}$
- 集合运算: 交  $A \cap B$ , 并  $A \cup B$ , 补  $A^c$ , 空集  $\emptyset$
- 函数:  $f(\cdot), F(\cdot)$  或  $\Phi(\cdot)$
- 概率算符: 概率  $\mathbb{P}$ , 期望  $\mathbb{E}$ , 方差  $\text{var}$ , 协方差  $\text{cov}$
- 一般算符: 如滞后算符  $\mathcal{L}$
- 数学表达: 任意  $\forall$ , 存在  $\exists$ , 属于  $\in$ , 包含于  $\subset$

## 本节内容

- 1 概率空间
- 2 1-元随机变量
- 3 多元随机变量
- 4 独立性

## 概率空间：样本空间

概率空间 (probability space), 即三元组  $(\Omega, \mathcal{F}, \mathbb{P})$

- ①  $\Omega$ : 样本空间 (sample space), 一个集合, 其元素为各种可能发生的随机状况
  - 如抛硬币,  $\Omega = \{H, T\}$ , 正面、反面; 又如 GDP 增速 (百分比),  $\Omega = [-100, \infty)$
  - 样本空间中的点一般记做  $\omega \in \Omega$ , 称为样本点 (sample point) 或随机元 (random point)
  - 从经济、金融角度看,  $\omega$  称为状态 (state),  $\Omega$  称为状态空间 (state space) 更合适

## 概率空间：事件域

- ②  $\mathcal{F}$ : 事件域 (event field), 是  $\Omega$  的部分或全部子集的集合; 每个这样的子集表示一个随机事件 (event)
  - 抛硬币,  $\mathcal{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$
  - GDP 增速,  $\mathcal{F} = \{\Omega \text{ 中的开、闭区间及其交、并、补集}\}$ ; 具体例子: 高增长事件  $= \{\omega > 8\}$ , 低增长事件  $= \{0 \leq \omega < 6\}$ , 适度增长  $= \{6 \leq \omega \leq 8\}$

注 数学上, 事件域  $\mathcal{F}$  称为  $\sigma$ -域, 满足下列要求

- $\emptyset, \Omega \in \mathcal{F}$
- $\forall A \in \mathcal{F}, A^c \in \mathcal{F}$
- $\forall \{A_i\}_{i=1}^{\infty} \subset \mathcal{F}, \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

## 事件域：集合运算的涵义

事件域  $\mathcal{F}$  是集合——随机事件——的集合

- 并集：给定  $A, B \in \mathcal{F}$ ,  $A \cup B$  表示事件  $A$  或  $B$  至少发生其一的事件
  - 高增长事件  $\cup$  适度增长事件  $= \{\omega \geq 6\}$ , 表示 GDP 增速处于适度或高增速状态
- 交集： $A \cap B$  表示事件  $A$  和  $B$  同时发生的事件
- 补集： $A^c$  表示事件  $A$  没有发生

**练习** 解释  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$  和  $A \cup (B \cap C) = (A \cap B) \cup (A \cap C)$  的涵义

## 概率空间：概率测度

③  $\mathbb{P}$ : 概率测度 (probability measure), 简称概率; 给出事件域  $\mathcal{F}$  中每个事件的可能性大小; 可看做函数  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ , 满足两条性质

- $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$
- 给定  $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ , 且  $A_i \cap A_j = \emptyset \ \forall i, j = 1, \dots, \infty$ , 则

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

这一性质又称概率测度的可数可加性 (countably additive)

**练习** 考虑  $\Omega = [0, 1]$ ,  $\mathcal{F}$  为其所有区间及其可数并、补集的集合, 令  $\mu : \mathcal{F} \rightarrow [0, 1]$  为  $\mathcal{F}$  上的函数, 取值为  $\mu([a, b]) = (b - a)^2, \forall 0 \leq a < b \leq 1$ 。请问  $\mu$  是概率测度吗?

## 本节内容

- 1 概率空间
- 2 1-元随机变量
- 3 多元随机变量
- 4 独立性

## 随机变量：定义

给定概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$

- 随机变量 (random variable, r.v.)  $X$  是从  $\Omega$  到实数的一个函数,  
 $X : \Omega \rightarrow \mathbb{R} \cup \{\pm\infty\}$ 
  - $\mathbb{R} \cup \{\pm\infty\}$  表示随机变量可以取值正负无穷；但后续不考虑
  - 抛硬币示例： $\omega = H, T$ ，定义  $X(H) = 0, X(T) = 1$ ，则  $X$  的 0-1 取值反映正负面
  - GDP 增速示例： $\omega \in [-100, \infty)$ ，定义  $X(\omega) = \omega$ ，则  $\omega$  即表示状态又表示 GDP 增速随机变量的取值
- 这个函数需要满足可测 (measurable) 的要求：

$$\underbrace{\{\omega : X(\omega) \leq x\}}_{\text{事件}} \in \mathcal{F}, \quad \forall x \in \mathbb{R}$$

练习 上面两个例子满足可测性吗？

## 随机变量：分布函数

给定概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  及其上的一个随机变量  $X$

- 根据定义,  $\forall x \in \mathbb{R}$ , 事件  $\{\omega : X(\omega) \leq x\} \in \mathcal{F}$ , 故  $\mathbb{P}$  可给出该事件的概率
- $\forall x \in \mathbb{R}$ , 称

$$F(x) = \mathbb{P}(\{\omega : X(\omega) \leq x\})$$

为  $X$  的累积分布函数 (cumulative distribution function, CDF), 简称为分布函数

- $\{\omega : X(\omega) \leq x\}$  一般简记为  $\{X \leq x\}$
- 分布函数单调递增:  $\forall x, y \in \mathbb{R}$ , 若  $x < y$ , 则  $F(x) \leq F(y)$ 
  - $F$  可能出现跳跃; 但利用概率  $\mathbb{P}$  的可数可加性, 可说明  $F$  右连续、左极限存在 (稍微复杂些的[练习](#))
- 分布函数左 0 右 1:  $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$

## 随机变量：离散型

给定概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  及其上的一个随机变量  $X$

- 如果  $X$  取值是离散的 (discrete), 如抛硬币中  $X = 0, 1$ , 则称  $X$  为 离散型 r.v.;  
 $X$  的取值集合可以有限, 也可以是可数集 (无穷)
  - 若  $X$  取值集合为  $\{x_1, x_2, \dots\} \subset \mathbb{R}$ , 则其分布函数  $F$  是一个阶梯函数, 在  $x_i$  处往上跳跃
  - 此时, 通常用  $p_i = \mathbb{P}(X = x_i) = F(x_i) - F(x_i-)$  表示  $X$  的分布;  
 $F(x_i-) = \lim_{x \uparrow x_i} F(x)$  表示左极限
- 除此之外,  $X$  取值可以是连续的 (continuous), 如 GDP 增速; 或更一般的离散、连续混合

## 随机变量：连续型

给定概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  及其上的一个随机变量  $X$

- 一般来说，当  $X$  取值连续时，其分布函数也可以有跳跃
  - 高速公路上汽车行使速度的分布在限速位置就有跳跃——很多汽车保持在限速，但也有部分超过或者低于限速
- 一类特别的连续取值 r.v.：  $X$  称为连续型 r.v.，如果其分布函数  $F$  可写为另一个非负函数  $f$  的积分

$$F(x) = \int_{-\infty}^x f(y) dy$$

此时称  $f$  为  $X$  的密度函数 (density function, DF)

## 随机变量：密度函数与分布函数

给定概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上的连续型 r.v.  $X$ , CDF  $F$ , DF  $f$

- 此时  $F$  一定是连续函数
- 在非精确意义下,  $f(x) = F'(x)$ 
  - $F$  不一定是处处可微, 但一定是“几乎”处处可微
- $\forall a < b$ ,

$$\mathbb{P}(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(y)dy$$

特别的,  $1 = \mathbb{P}(\Omega) = \mathbb{P}(-\infty \leq X \leq \infty) = \int_{-\infty}^{\infty} f(y)dy$

**练习** 继续第 8 页练习中的  $\Omega, \mathcal{F}$ , 令  $\mathbb{P}([a, b]) = b^2 - a^2$ , 请问  $\mathbb{P}$  是概率测度? 并定义  $X(\omega) = \omega \in \Omega$ , 求  $X$  的 CDF 及 DF

## 1-元随机变量：常见分布

很多情况下，有了 r.v. 的 CDF 或 DF 后，我们不必再关心  $(\Omega, \mathcal{F}, \mathbb{P})$  本身；r.v. 的概率/统计信息都在分布中

- 典型离散分布，r.v.  $X$ ，取值集合记为  $\mathcal{V}$ 
  - 两点分布，又称 Bernoulli 分布： $\mathcal{V} = \{0, 1\}$ ， $\mathbb{P}(X = 0) = p$
  - 二项式分布： $\mathcal{V} = \{0, 1, \dots, n\}$ ， $\mathbb{P}(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ ， $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ ， $p \in (0, 1)$
  - Poisson 分布： $\mathcal{V} = \mathbb{N} \equiv \{0, 1, \dots\}$ ， $\mathbb{P}(X = i) = \frac{\lambda^i e^{-\lambda}}{i!}$ ， $\lambda > 0$
- 典型连续分布
  - 均匀分布  $U([a, b])$ ： $\mathcal{V} = [a, b]$ ， $f(x) = \frac{1}{b-a}$
  - 正态分布  $N(\mu, \sigma^2)$ ： $\mathcal{V} = \mathbb{R}$ ， $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$
  - 指数分布  $\exp(\lambda)$ ： $\mathcal{V} = \mathbb{R}_+ \equiv [0, \infty)$ ， $f(x) = \lambda e^{-\lambda x}$ ， $\lambda > 0$

## 1-元随机变量：矩

给定 r.v.  $X$  及其 CDF  $F$

- $\forall \alpha \in \mathbb{R}$ , 定义  $X$  的  $\alpha$ -阶矩 (moment):

$$\mathbb{E}[X^\alpha] = \int x^\alpha dF(x)$$

积分限由  $X$  的取值范围  $\mathcal{V}$  决定

注 上面的积分可理解为 Riemann-Stieltjes 积分

- 与  $g(x)$  为被积函数的 Riemann 积分定义中对  $x$  进行分划, 令  $\Delta x_k = x_k - x_{k-1} \rightarrow 0$ , 取离散和  $\sum_k g(x_k) \Delta x_k$  极限类似, 此处对  $\sum_k g(x_k) \Delta F(x_k) = \sum_k g(x_k) [F(x_k) - F(x_{k-1})]$  取极限
- 对离散型 r.v., 上述积分自然转化为离散和  $\sum_i x_i^\alpha p_i$
- 对连续型 r.v., 形式上  $dF(x) = f(x)dx$ , 上述积分化为常规 Riemann 积分形式  $\int x^\alpha f(x)dx$

## 1-元随机变量：期望与方差

- $X$  的 1-阶矩称为  $X$  的期望 (expectation), 记做  $\mathbb{E}X$  或  $\mu_X$
- $X$  的方差 (variance) 定义为

$$\text{var}(X) = \mathbb{E}[(X - \mathbb{E}X)^2]$$

又称为 2-阶中心矩

练习 证明  $\text{var}(X) = \mathbb{E}[X^2] - (\mathbb{E}X)^2$

- $X$  的标准差 (standard deviation) 定义为  $\sigma_X = \sqrt{\text{var}(X)}$

练习 计算上述 6 个典型分布的期望、方差和标准差

## 本节内容

- 1 概率空间
- 2 1-元随机变量
- 3 多元随机变量**
- 4 独立性

## 多元随机变量：联合分布

给定  $(\Omega, \mathcal{F}, \mathbb{P})$  及其上定义的  $n$  个 r.v.  $X_1, \dots, X_n$

- 多元随机事件： $\{X_1 \leq x_1, \dots, X_n \leq x_n\}$  表示  $\{X_1 \leq x_1\}, \dots, \{X_n \leq x_n\}$   $n$  个事件同时发生，即

$$\{X_1 \leq x_1, \dots, X_n \leq x_n\} = \bigcap_{1 \leq i \leq n} \{X_i \leq x_i\} \in \mathcal{F}$$

- $\forall x_1, \dots, x_n \in \mathbb{R}$ ,  $X_1, \dots, X_n$  的联合累积分布函数 (joint CDF) 定义为

$$F(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$$

简称为联合分布 (joint distribution)

## 多元随机变量：离散型

给定  $n$  个 r.v.  $X_1, \dots, X_n$

- 离散型： $\forall i, X_i$  为离散取值；此时  $X_1, \dots, X_n$  的联合分布就是一个  $n$ -维数组  $\{p_{i_1, \dots, i_n}\}$ ，其中

$$p_{i_1, \dots, i_n} = \mathbb{P}(X_1 = x_{i_1}, \dots, X_n = x_{i_n})$$

$x_{i_k}$  表示  $X_k$  的第  $i_k$  个取值

- 注意和联合累计分布之间概念上的差别
- 2-元离散随机变量  $X, Y$  对应的分布可表示为一个矩阵  $[p_{ij}]$ ，其中
$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j)$$

## 多元随机变量：连续型

给定  $n$  个 r.v.  $X_1, \dots, X_n$

- 连续型：此时联合分布可以写为非负函数  $f(y_1, \dots, y_n)$  的多元积分

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(y_1, \dots, y_n) dy_1 \cdots dy_n$$

$f$  称为  $X_1, \dots, X_n$  的联合密度函数 (joint DF)

- 与 1-元情形类似，非精确意义下

$$f(x_1, \dots, x_n) = \frac{\partial^n F(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n}$$

## 多元随机变量：边缘分布

给定  $n$  个 r.v.  $X_1, \dots, X_n$  及联合分布  $F$

- $X_i$  的边缘分布  $F_i(x_i)$  仍然定义为  $\mathbb{P}(X_i \leq x_i)$ ，其与联合分布的联系为

$$\begin{aligned} F_i(x_i) &= \mathbb{P}(\{X_i \leq x_i\} \cap \Omega \cap \dots \cap \Omega) \\ &= \mathbb{P}\left(\{X_i \leq x_i\} \cap \bigcap_{j \neq i} \{X_j \leq \infty\}\right) \\ &= \mathbb{P}(X_1 \leq \infty, \dots, X_i \leq x_i, \dots, X_n \leq \infty) \\ &= F(\infty, \dots, x_i, \dots, \infty) \end{aligned}$$

练习 对 2-元离散 r.v.  $X, Y$  及联合分布  $p_{ij}$ ，将  $X$  的边缘分布  $p_i$  写为  $p_{ij}$  的求和形式

练习 对 2-元连续 r.v.  $X, Y$  及联合密度  $f(x, y)$ ，将  $X$  的边缘分布  $F_X$  写为  $f$  的 2-元积分形式

## 多元随机变量：混合矩

给定  $n$  个 r.v.  $X_1, \dots, X_n$  及联合分布  $F$

- $\forall \alpha_1, \dots, \alpha_n \in \mathbb{R}$ , 令  $\alpha = [\alpha_1, \dots, \alpha_n]$ , 则  $X_1, \dots, X_n$  的混合  $\alpha$ -阶矩定义为

$$\mathbb{E}[X_1^{\alpha_1} \cdots X_n^{\alpha_n}] = \int x_1^{\alpha_1} \cdots x_n^{\alpha_n} dF(x_1, \dots, x_n)$$

此处亦可理解为多元 Riemann-Stieltjes 积分

注 1-元情形下,  $dF(x)$  对应离散和中的项为  $F(x + \Delta x) - F(x) = \mathbb{P}(X \in [x, x + \Delta x])$ , 即  $X$  处于  $x$  附近的概率; 多元情形下,  $dF(x_1, \dots, x_n)$  表示  $X_1, \dots, X_n$  处于  $(x_1, \dots, x_n)$  附近的概率

## 多元随机变量：离散/连续型分布下的混合矩

- 离散情形：混合矩可写为多重求和

$$\mathbb{E}[X_1^{\alpha_1} \cdots X_n^{\alpha_n}] = \sum_{i_1} \cdots \sum_{i_n} x_{i_1}^{\alpha_1} \cdots x_{i_n}^{\alpha_n} p_{i_1, \dots, i_n}$$

- 连续情形：混合矩可写为多重积分

$$\mathbb{E}[X_1^{\alpha_1} \cdots X_n^{\alpha_n}] = \int \cdots \int x_1^{\alpha_1} \cdots x_n^{\alpha_n} f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

- 一般的多元 r.v. 混合矩用处有限；重要的是 2-元随机变量的混合 2-阶矩

## 2-元随机变量：协方差及相关系数

给定 r.v.  $X, Y$  及其联合分布  $F(x, y)$

- $X, Y$  的混合 2-阶矩 (cross moment) 定义为

$$\mathbb{E}[XY] = \int xy dF(x, y)$$

- $X, Y$  的协方差 (covariance) 定义为

$$\text{cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$$

练习 证明  $\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y$

- $X, Y$  的相关系数 (correlation) 定义为  $\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$

## 随机向量：均值

将 r.v.  $X_1, \dots, X_n$  列为一个列向量  $\mathbf{X} = [X_1, \dots, X_n]^\top$

- 该向量称为随机向量 (random vector)
- 以  $\mathbb{E}\mathbf{X}$  表示各分量的期望组成的列向量

$$\boldsymbol{\mu}_{\mathbf{X}} \equiv \mathbb{E}\mathbf{X} = \begin{bmatrix} \mathbb{E}X_1 \\ \vdots \\ \mathbb{E}X_n \end{bmatrix}$$

## 随机向量：协方差矩阵

- 矩阵

$$\Sigma_X = \begin{bmatrix} \text{cov}(X_1, X_1) & \cdots & \text{cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \cdots & \text{cov}(X_n, X_n) \end{bmatrix}$$

称为  $X$  的协方差矩阵 (covariance matrix)

- 注意区分矩阵  $\Sigma$  与求和号  $\sum$
- 多元正态分布:  $X = [X_1, \dots, X_n] \sim N(\mu_X, \Sigma_X)$ , 令  $x = [x_1, \dots, x_n]^T$ , 其密度函数为

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma_X}} \exp \left\{ -\frac{1}{2}(x - \mu_X)^T \Sigma_X^{-1} (x - \mu_X) \right\}$$

## 随机向量：协方差矩阵的简化表示

由定义  $\text{cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)]$ , 可将  $\Sigma_X$  改写为

$$\begin{aligned}
 \Sigma_X &= \mathbb{E} \begin{bmatrix} (X_1 - \mathbb{E}X_1)(X_1 - \mathbb{E}X_1) & \cdots & (X_1 - \mathbb{E}X_1)(X_n - \mathbb{E}X_n) \\ \vdots & \ddots & \vdots \\ (X_n - \mathbb{E}X_n)(X_1 - \mathbb{E}X_1) & \cdots & (X_n - \mathbb{E}X_n)(X_n - \mathbb{E}X_n) \end{bmatrix} \\
 &= \mathbb{E} \begin{bmatrix} X_1 - \mathbb{E}X_1 \\ \vdots \\ X_n - \mathbb{E}X_n \end{bmatrix} \begin{bmatrix} X_1 - \mathbb{E}X_1 & \cdots & X_n - \mathbb{E}X_n \end{bmatrix} \\
 &= \mathbb{E}[(X - \mu_X)(X - \mu_X)^\top] \\
 &= \mathbb{E}[XX^\top] - \mu_X \mu_X^\top
 \end{aligned}$$

## 随机向量：线性变换

给定  $n$ -维随机向量  $\mathbf{X} = [X_1, \dots, X_n]^\top$  和  $m \times n$  常数矩阵  $A$

- $\mathbf{Y} = A\mathbf{X}$  是一个  $m$ -维随机向量
- $\mathbf{Y}$  的期望是  $\mathbf{X}$  的线性变换

$$\mu_Y = \mathbb{E}Y = \mathbb{E}AX = A\mathbb{E}X = A\mu_X$$

- $\mathbf{Y}$  的协方差矩阵  $\Sigma_Y$  可写为

$$\begin{aligned}\Sigma_Y &= \mathbb{E}[\mathbf{Y}\mathbf{Y}^\top] - \mu_Y\mu_Y^\top \\ &= \mathbb{E}[A\mathbf{X}\mathbf{X}^\top A^\top] - A\mu_X\mu_X^\top A^\top \\ &= A\left(\mathbb{E}[\mathbf{X}\mathbf{X}^\top] - \mu_X\mu_X^\top\right)A^\top \\ &= A\Sigma_X A^\top\end{aligned}$$

## 本节内容

- 1 概率空间
- 2 1-元随机变量
- 3 多元随机变量
- 4 独立性**

## 随机变量独立性：定义

- 给定  $(\Omega, \mathcal{F}, \mathbb{P})$ , 两个事件  $A, B \in \mathcal{F}$  若满足

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

则称为独立 (independent) 事件

- 给定  $(\Omega, \mathcal{F}, \mathbb{P})$  上的两个 r.v.  $X, Y$ , 若  $\forall x, y \in \mathbb{R}$  满足

$$\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y)$$

则称两个 r.v. 相互独立

- 若  $X, Y$  相互独立, 则其联合分布  $F$  为两个边缘分布  $F_X, F_Y$  的乘积

$$F(x, y) = F_X(x)F_Y(y)$$

## 随机变量独立性：性质

给定两个独立的 r.v.  $X, Y$

- $f, g: \mathbb{R} \rightarrow \mathbb{R}$  为两个函数, 则  $f(X), g(Y)$  相互独立
  - 严格说, 需要  $f, g$  满足可测性条件
- 乘积的期望等于期望的乘积

$$\mathbb{E}[XY] = \int xy dF(x, y) = \int xy dF_X(x) dF_Y(y) = \mathbb{E}X \mathbb{E}Y$$

- 协方差为 0,

$$\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}X \mathbb{E}Y = 0$$

故相关系数为 0

## 随机变量独立性：多元一般情形

给定 r.v.  $X_1, \dots, X_n$

- 若  $\forall x_1, \dots, x_n \in \mathbb{R}$ , 满足

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^n \mathbb{P}(X_i \leq x_i)$$

则称  $X_1, \dots, X_n$  相互独立

- 若  $X_1, \dots, X_n$  相互独立, 则  $f_1(X_1), \dots, f_n(X_n)$  相互独立,

$$\begin{aligned}\mathbb{E}[X_1 \cdots X_n] &= \int x_1 \cdots x_n dF(x_1, \dots, x_n) \\ &= \int x_1 \cdots x_n dF_1(x_1) \cdots dF_n(x_n) = \mathbb{E}X_1 \cdots \mathbb{E}X_n\end{aligned}$$

## 大数定律 (law of large number, LLN)

### 定理 1

当  $\{X_t\}$  为独立同分布 (independent and identical distribution, iid) 随机变量序列时, 样本均值以概率 1 收敛到总体均值  $\mu \equiv \mathbb{E}X_t$ :

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T X_t(\omega) = \mu\right\}\right) = 1,$$

记做

$$\hat{\mu}_T \equiv \frac{1}{T} \sum_{t=1}^T X_t \xrightarrow{\text{a.s.}} \mu,$$

其中, a.s. 表示 almost surely, 意为几乎一定; 以概率 1 收敛, 又称为几乎处处 (almost everywhere) 收敛

## 中心极限定理 (central limit theorem, CLT)

- 考虑 iid 样本  $\{X_t\}$ , 期望为  $\mu$ , 方差为  $\sigma^2$
- 由大数定律,  $\hat{\mu}_T - \mu = T^{-1} \sum_t (X_t - \mu) \xrightarrow{\text{a.s.}} 0$
- 计算可得,  $\text{var}(\hat{\mu}_T) = \sigma^2/T$ , 故  $\hat{\mu}_T$  的标准化

$$\xi_T = \frac{\hat{\mu}_T - \mu}{\sqrt{\text{var}(\hat{\mu}_T)}} = \frac{T^{-1} \sum_t (X_t - \mu)}{\sqrt{T}^{-1} \sigma} = \sqrt{T} \frac{1}{T} \sum_{t=1}^T \frac{X_t - \mu}{\sigma}$$

方差为 1

### 定理 2

假设有 iid 序列  $\{X_t\}$ , 则其标准化样本均值  $\xi_T \xrightarrow{d} N(0, 1)$ , 或等价的,  $\sqrt{T}(\hat{\mu}_T - \mu) \xrightarrow{d} N(0, \sigma^2)$ , 其中  $\xrightarrow{d}$  表示依分布收敛

## CLT: Monte Carlo 模拟示例

