随机过程

0 课程介绍和概率论预备内容

廖振宇

华中科技大学电子信息与通信学院

2025年2月25日

目录

课程介绍

- 课程内容与方法：在**理解关键概念**的基础上，掌握**分析和计算方法**（习题练习和讨论很重要！）
  1. 第0章：概率论预备知识；
  2. 第1章：随机过程基本概念；
  3. 第2章：泊松过程；
  4. 第3章：马尔可夫链；
  5. 第4章：平稳随机过程；
  6. 第5章：高斯过程。
- 参考书目：
  - 随机过程， 刘澍 主编，华中科技大学出版社，2023
  - 随机过程及应用，陆大絟 编著，清华大学出版社，2006
  - 随机过程，Sheldon M. Ross著，机械工业出版社，2013
  - 随机变量与随机过程，A. Papoulis, S. U. Pillai编著，机械工业出版社，2013
- 在线课程：
  - 随机过程 清华大学张颢
  - MIT 6.262 Discrete Stochastic Processes

课程评分方法

- 课堂讨论与作业（10%）：结合点名、课堂讨论、课堂作业
- 课后作业（20%）：结课答疑之前交即可、交了就有分，但建议随课每周完成
- 期末考试（70%）：12周周三晚（5月7日），之前安排一次答疑，时间待定

课程相关联系方式

- 我的邮箱：zhenyu_liao@hust.edu.cn
- 助教：卢佳龙，邮箱：jialong@hust.edu.cn；电话：15827029400
- 课件和作业在微助教课堂“随机过程：电磁2301-2”（编号：OO478）发布

个人介绍

- 教育和工作经历：
  - 2010-2014，本科，光电信息工程，华中科技大学
  - 2014-2016，硕士，信号与图像处理，法国巴黎萨克雷大学
  - 2016-2019，博士，计算机与数学，法国巴黎萨克雷大学
  - 2019-2020，博士后研究员，统计系 & ICSI，美国加州大学伯克利分校
- 个人荣誉：
  - 法国巴黎萨克雷大学ED STIC优秀博士生论文奖、华中科技大学东湖青年学者、武汉市“武汉英才”优秀青年人才
  - 第十四批湖北省“百人计划”（创新人才）
  - 法国自然科学基金 ANR-LabEx-CIMI 访问教授、加拿大 CRM-Simons 访问教授等
- 学术服务：
  - 欧盟自然科学基金（ERC）、加拿大国家自然科学基金（NSERC）和以色列科学基金会（ISF）评审
  - 机器学习旗舰会议ICML、ICLR、IJCNN领域主席（Area Chair）
  - Springer Statistics and Computing副编辑（Associate Editor）
  - 中国现场统计研究会随机矩阵理论与应用分会副秘书长，中国现场统计研究会大数据统计分会理事
- 个人主页：`https://zhenyu-liao.github.io/`

目录

概率的公理化定义

- 概率：描述、度量和预测现实生活中不确定事件的规律
- 古典概率的直观理解：针对事件 A，如果重复进行 n 次实验，事件 A 发生 n_A 次，则当重复实验次数足够大时，事件 A 发生的频率接近其概率：

$$\frac{n_A}{n} \approx \mathbb{P}(A)$$

- 在数学上是含糊的，不精确、不严谨的
- 概率的**公理化定义**：“概率是满足XXX规则的XXX（数学对象）”

定义 (样本空间Ω)

随机实验所得到的所有可能结果构成的集合Ω。

- 一次掷硬币实验，结果有两种可能 $\Omega = \{0, 1\}$
- 一次掷骰子实验，结果有六种可能 $\Omega = \{1, 2, 3, 4, 5, 6\}$

概率的公理化定义：样本空间与基本事件

定义 (样本空间Ω)

一次随机实验所得到的所有可能结果构成的集合。

- 一次抛硬币实验，结果有两种可能 $\Omega = \{0, 1\}$
- 一次抛骰子实验，结果有六种可能 $\Omega = \{1, 2, 3, 4, 5, 6\}$

定义 (样本点、基本事件)

样本空间的元素 $\omega \in \Omega$。

定义 (事件)

样本空间的子集 $A \subseteq \Omega$。

- 一次抛硬币实验得到正面结果
- 一次抛骰子实验得到“偶数面”事件$A = \{2, 4, 6\}$

概率的公理化定义

定义 (概率)

针对事件A，其概率$\mathbb{P}(A)$满足：

- 非负性：$\mathbb{P}(A) \geq 0$
- 归一性：$\mathbb{P}(\Omega) = 1$
- 可加性：如果 $A \cap B = \emptyset$，则 $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

常称$(\Omega, \mathcal{F}, \mathbb{P})$为概率空间，其中$\mathcal{F}$为事件构成的$\sigma$-代数。

- 本质上，概率是一种测度（类似“面积测量”、“体积测量”），是一个函数映射，其值域为$[0, 1]$。
- 在这个定义下，我们可以讨论一次抛硬币实验得到正面概率、一次抛骰子实验得到“偶数面”的概率。

条件概率

定义 (条件概率)

假定事件M发生的情况下，事件A发生的条件概率为

$$\mathbb{P}(A|M) = \frac{\mathbb{P}(A \cap M)}{\mathbb{P}(M)}$$

- 根据定义，显然有：如果$M \subseteq A$，则$\mathbb{P}(A|M) = 1$；如果$A \subseteq M$，则$\mathbb{P}(A|M) \geq \mathbb{P}(A)$。
- 概率理解：$\mathbb{P}(A) \approx \frac{n_A}{n}, \mathbb{P}(M) \approx \frac{n_M}{n}, \mathbb{P}(A \cap M) \approx \frac{n_{AM}}{n}$，有

$$\mathbb{P}(A|M) = \frac{\mathbb{P}(A \cap M)}{\mathbb{P}(M)} \approx \frac{n_{AM}/n}{n_M/n} = \frac{n_{AM}}{n_M}$$

Example (条件概率)

盒子里有三个白球w_1, w_2, w_3和两个红球r_1, r_2，随机相继取出两个球。求第一个是白球，第二个是红色的概率是多少？（用条件概率计算）

独立事件的概率

定义 (独立事件)

如事件A, B 满足$\mathbb{P}(A\cap B)=\mathbb{P}(A)\cdot\mathbb{P}(B)$，则$A, B$独立。

- 独立事件的频率理解：

$$\mathbb{P}(A)\approx\frac{n_A}{n},\quad \mathbb{P}(B)\approx\frac{n_B}{n},\quad \mathbb{P}(A\cap B)\approx\frac{n_{AB}}{n},$$

如果事件A, B独立，则

$$\frac{n_A}{n}\approx\mathbb{P}(A)=\frac{n_B}{n}\quad\Rightarrow\quad\mathbb{P}(A)\cdot\mathbb{P}(B)=\mathbb{P}(A\cap B)=\frac{n_{AB}/n}{n_B/n}\approx\frac{n_{AB}}{n_B}$$

全概率定理与贝叶斯定理

定理 (全概率定理)

考虑Ω可以分为K个互斥事件$A_1, A_2, \ldots, A_K$，且$A_i \cap A_j = \emptyset$，则对于任意事件B，我们有

$$\mathbb{P}(B) = \sum_{i=1}^{K} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$$

由条件概率定义 $\mathbb{P}(B \cap A_i) = \mathbb{P}(A_i|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$，可得$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B\cap A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\cdot\mathbb{P}(A_i)}{\mathbb{P}(B)}$，进一步

定理 (贝叶斯定理)

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}{\sum_{i=1}^{K} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}$$

其中$\mathbb{P}(A_i)$和$\mathbb{P}(A_i|B)$分别分成为先验和后验概率。

目录

随机变量和分布函数

随机变量是概率论的**主要研究对象**，其统计规律用分布函数来描述。

定义 (随机变量和分布函数)

考虑概率空间$(\Omega, \mathcal{F}, \mathbb{P})$，随机变量$X = X(e)$是定义在$\Omega$上的函数，其对任意实数$x$满足集合$\{e : X(e) \leq x\} \in \mathcal{F}$（可理解为“是一个事件”），简记为$X$，将

$$F_X(x) = \mathbb{P}(e : X(e) \leq x), \quad x \in \mathbb{R}$$

称为随机变量X的分布函数，可以简写为$F_X(x) = \mathbb{P}(X \leq x)$。

具有以下性质：

1. $F(+\infty) = 1, \quad F(-\infty) = 0$
2. $F(x)$是非减函数：$F(x_1) \leq F(x_2), x_1 < x_2$
3. $F(x)$右连续：$F(x^+) = F(x)$

- $\mathbb{P}(X > x) = 1 - F_X(x)$
- $\mathbb{P}(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$

常见的随机变量

常见的随机变量分为**离散型随机变量**和**连续型随机变量**。

- **离散型随机变量**的概率分布用分布列描述：

$$\mathbb{P}(X = x_k) = p_k, \quad k = 1, 2, 3, \ldots$$

其分布函数为：

$$F_X(x) = \sum_{x_k \le x} p_k$$

- **连续型随机变量**的概率分布用概率密度函数 $f_X(x)$ 描述，满足：

$$F_X(x) = \int_{-\infty}^{x} f(t)\, dt$$

其中极限定义为

$$f_X(x) = \lim_{\Delta x \to 0} \frac{\mathbb{P}(x \le X \le x + \Delta x)}{\Delta x}$$

- **频率理解**：对于Δn_x是n次实验中满足$x \le X \le x + \Delta x$的次数，则：

$$f_X(x)\Delta x \approx \frac{\Delta n_x}{n}$$

常见的分布：连续型

- **正态（高斯）分布:** $\mathcal{N}(\mu, \sigma^2)$，包含期望参数$\mu$和方差参数$\sigma^2$，其概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  - 其期望为μ和方差为σ^2
  - 最常见的分布，中心极限定理

- **指数分布:** $\mathrm{Exp}(\lambda)$，包含参数$\lambda > 0$，其概率密度函数为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{其他} \end{cases}$$

  - 其期望为$\frac{1}{\lambda}$，方差为$\frac{1}{\lambda^2}$
  - 常用来描述互不相交区间上事件发生相互独立时的累计时间（如电话呼叫的到达时间，或公共汽车到达这一站点的时间）。
  - 指数分布的无记忆性：$\mathbb{P}(X > t + s \mid X > s) = \mathbb{P}(X > t)$

- **均匀分布:** $U(a, b)$，其概率密度函数在区间(a, b)内，$a < b$为常数，期望为$\mu = \frac{a+b}{2}$，方差为$\frac{(b-a)^2}{12}$。

常见分布：离散型

- **伯努利（0-1）分布**: 取值为$\{0,1\}$，包含参数p；$\mathbb{P}(X=1)=p, \mathbb{P}(X=0)=1-p$; 其期望为p，方差为$p(1-p)$
  - 举例: 掷硬币
- **二项（式）分布**: 在n次伯努利试验形成的实验中，每次成功的概率为p，则n次实验中成功的总次数Y服从二项分布，包含参数n, p

$$\mathbb{P}(Y=k)=\binom{n}{k}p^k(1-p)^{n-k}, \quad k=0,1,\dots,n$$

  - 期望为np，方差为$np(1-p)$
- **几何分布**: 重复伯努利试验中第一次成功所需要的试验次数，$\mathbb{P}(Z=k)=p(1-p)^{k-1}, k=0,1,\dots,n$
  - 期望为$\frac{1}{p}$，方差为$\frac{1-p}{p^2}$，无记忆性：$\mathbb{P}(Z>m+n \mid Z>m)=(1-p)^n$
- **泊松分布**: $\mathbb{P}(X=k)=\frac{e^{-\lambda}\lambda^k}{k!}, k=0,1,\dots,\infty$, 包含参数$\lambda>0$
  - 表示大事件中稀有事件发生的次数：如机在固定时间段内接到电话的次数，或彩票中奖的票数等
  - 期望为λ，方差为λ

联合分布与联合密度函数

考虑两个随机变量(X, Y)，研究其联合统计特性，也即二元随机变量(X, Y)落在$x - y$二维平面定义上的指定区域D内的概率

定义 (联合分布函数)

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

- $F(x, y)$对于x, y均为非减函数
- 满足$F(-\infty, y) = 0, F(x, -\infty) = 0, F(+\infty, +\infty) = 1$
- 且$\mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2, y_2) - F(x_1, y_1)$

定义 (联合密度函数)

$f_{(X,Y)}(x, y) = \frac{\partial^2 F_{(X,Y)}(x,y)}{\partial x \partial y}$

针对xy二维平面区域D有

$$\mathbb{P}((X, Y) \in D) = \int\int_D f_{(X,Y)}(x, y)dxdy$$

边际分布函数与边际密度函数

定义 (边际分布/密度函数)

针对二元随机变量(X, Y)，仅关心**其中一个随机变量**的统计特性，研究其边际分布函数$F_X(x), F_Y(y)$和边际密度函数$f_X(x), f_Y(y)$

$$F_X(x) = F_{(X,Y)}(x, \infty), \quad F_Y(y) = F_{(X,Y)}(\infty, y),$$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y)dy, \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y)dx$$

定义 (独立随机变量)

如果对于二维平面上任意(x, y)有

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

则随机变量X, Y独立，并且

$$F_{(X,Y)}(x, y) = F_X(x) \cdot F_Y(y), \quad f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y)$$

多元随机变量和独立性

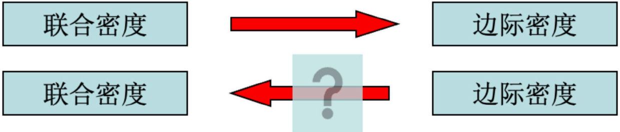
定理

如随机变量X, Y独立，则其函数$g(X), h(Y)$独立。

定理

由独立实验产生的随机变量独立。

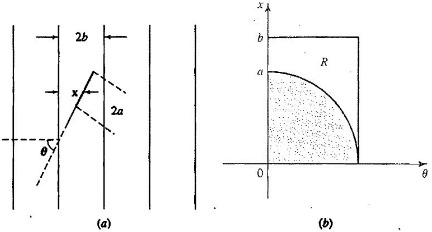
- 针对二维随机变量(X, Y)，其联合密度$f_{(X,Y)}$包含更多信息，总可以得到边际/边缘密度f_X, f_Y。
- 而通常情况下，无法从边缘密度得到联合密度（缺少相互关联、影响的信息），除非两者独立！
- 相关概念和关系均可推广到多(> 2)元分布



布丰投针实验

Example (布丰投针实验)

一个长度为2a的针随机抛在由一组间距为2b的平行线分割的平板上，如下图所示，其中$b > a$。证明：针与平行线相交的概率为$\frac{2a}{\pi b}$。

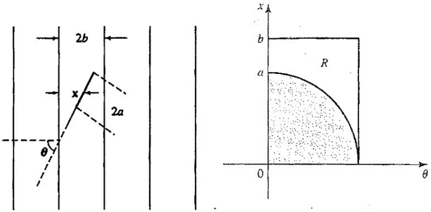


- 用概率论概念描述：用随机变量X表示投针中心到最近（平行）线的距离，Θ表示投针和平行线垂直方向之间的夹角；考虑X, Θ两个随机变量相互独立，$X \sim U(0, b), \Theta \sim U(0, \pi/2)$
- 联合概率密度函数$f_{(X,\Theta)}(x, \theta) = f_X(x) \cdot f_\Theta(\theta) = \frac{1}{b} \times \frac{2}{\pi}, x \in [0, b], \theta \in [0, \pi/2]$

布丰投针实验

Example (布丰投针实验)

一个长度为2a的针随机抛在由一组间距为2b的平行线分割的平板上，如下图所示，其中$b > a$。证明：针与平行线相交的概率为$\frac{2a}{\pi b}$。



(a) (b)

- 联合概率密度函数$f_{(X,\Theta)}(x,\theta) = f_X(x) \cdot f_\Theta(\theta) = \frac{1}{b} \times \frac{2}{\pi}, x \in [0,b], \theta \in [0,\pi/2]$
- 当$x < a\cos\theta$时，投针于平行线相交，对应概率为

$$\mathbb{P}(x < a\cos\theta) = \boxed{\frac{2}{\pi b}} \int_0^{2/\pi} a\cos\theta d\theta = \frac{2a}{\pi b} \tag{1}$$

目录

数学期望或均值

定义 (数学期望、均值)

对于随机变量X，其概率密度函数为$f_X(x)$，则数学期望或均值定义为

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

- 如果积分不收敛，则该随机变量的均值不存在
- 对于离散随机变量，由$f(x) = \sum_i p_i \delta(x - x_i)$，数学期望可写成求和

$$\mathbb{E}[X] = \int x f_X(x) dx = \int \sum_i p_i x \delta(x - x_i) dx = \sum_i p_i x_i = \sum_{i=1} \mathbb{P}(X = x_i) x_i.$$

数学期望或均值

- 要注意数学期望/均值和样本均值区分开来！！！
  - 通过n次独立重复实验，产生/观察到n个独立同分布随机变量（样本）$X_1, X_2, \ldots, X_n$，其均值为$\mathbb{E}[X_i] = \mu$，通过样本平均$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 估计数学期望
  - 由大数定理（Law of Large Numbers）：当$n \to \infty$时，$\bar{X} \to \mu$依概率/几乎处处收敛
- 数学期望的频率理解
  - 用Δn_k表示随机变量X落在区间$[z_k, z_{k+1}], z_{k+1} = z_k + \Delta x$之间的次数，有

$$\sum_{i=1}^{n} X_i \approx \sum_{k} z_k \Delta n_k, \quad f(z_k)\Delta x \approx \frac{\Delta n_k}{n} \quad (\text{“密度函数的频率理解”})$$

  - 可以得到$\bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i \approx \frac{1}{n}\sum_k z_k \Delta n_k \approx \sum_k z_k f(z_k)\Delta x \approx \int x f(x) dx$

Example (数学期望计算)

一次抛掷骰子实验，得到点数结果的数学期望是多少？`https://probability.visualized.fun/law-of-large-numbers/`

随机变量函数期望

- 随机变量函数的期望计算：$\mathbb{E}[g(X)] = \int g(x)f_X(x)dx$ （比如，方差）
- 线性性：针对随机变量 X_1, X_2
  - 无论独立与否，对于 $Y = aX_1 + bX_2$ 有

$$\mathbb{E}[Y] = a\mathbb{E}[X_1] + b\mathbb{E}[X_2]$$

  - 证明：考虑积分

$$\begin{aligned}
\int yf_Y(y)dy &= \int\int (ax_1 + bx_2)f_{(X_1,X_2)}(x_1, x_2)dx_1dx_2 \\
&= a\int\int x_1f_{(X_1,X_2)}(x_1, x_2)dx_1dx_2 + b\int\int x_2f_{(X_1,X_2)}(x_1, x_2)dx_1dx_2 \\
&= a\int x_1f_{X_1}(x_1)dx_1 + b\int x_2f_{X_2}(x_2)dx_2 \\
&= a\mathbb{E}[X_1] + b\mathbb{E}[X_2]
\end{aligned}$$

  - 用到：联合密度积分得到边缘密度 $f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dy$
- 对于 $Z = ag(X_1) + bg(X_2)$ 是不是有相似结论？$Z = X_1X_2$ 呢？

方差

定义 (方差)

描述随机变量的“集中/分散程度”： 对于随机变量 X，$X-\mathbb{E}[X]$ 表示其离均值的偏差，考虑 $(X-\mathbb{E}[X])^2 \geq 0$，则X的方差为
$\sigma_X^2 = D(X) = \mathrm{Var}[X] = \mathbb{E}[(X-\mathbb{E}[X])^2]$，$\sigma_X = \sqrt{\mathrm{Var}[X]} = \sqrt{D(X)}$ 为标准差。

- $\mathrm{Var}[X] = \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- $\mathbb{E}[X^2]$ 称作二阶矩 (second order moment)，有 $\mathbb{E}[X^2] - (\mathbb{E}[X])^2 \geq 0$
- 针对常数 c 有 $\mathrm{Var}[c] = 0, \mathrm{Var}[cX] = c^2\mathrm{Var}[X]$
- $\mathrm{Var}[c] = 0 \iff \mathbb{P}(X = \mathbb{E}[X]) = 1$
- Cauchy-Schwarz 不等式：$(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$当右边有界时成立
  - 本质来源于内积
  - 证明：注意到，$\mathbb{E}[(aX-Y)^2] = a^2\mathbb{E}[X^2] - 2a\mathbb{E}[XY] + \mathbb{E}[Y^2] \geq 0$对于任意实数a总成立

协方差和相关系数

定义 (协方差和相关系数)

考虑随机变量 X, Y 满足 $\mathbb{E}[X^2] < \infty, \mathbb{E}[Y^2] < \infty$，称

$$B_{XY} = \mathrm{Cov}[X, Y] = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$$

为 X, Y 的协方差；称

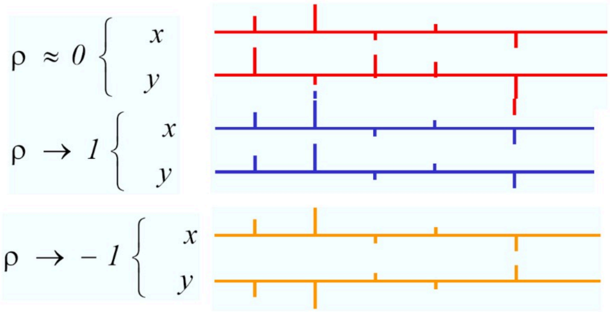
$$\rho_{XY} = \frac{\mathrm{Cov}[X, Y]}{\sqrt{\mathrm{Var}[X] \cdot \mathrm{Var}[Y]}}$$

为 X, Y 的相关系数。

- $\mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$
- “自己的”协方差就是方差：$\mathrm{Cov}[X, X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathrm{Var}[X]$
- $|\rho_{XY}| \leq 1$（无量纲）反映随机变量线性相关性，$\rho_{XY} = 0$表示X, Y不（线性）相关，注意和“独立”区分开！！！

相关系数

- $|\rho_{XY}| \leq 1$（无量纲）反映随机变量线性相关性，$\rho_{XY} = 0$表示X, Y不（线性）相关，注意和“独立”区分开！！！



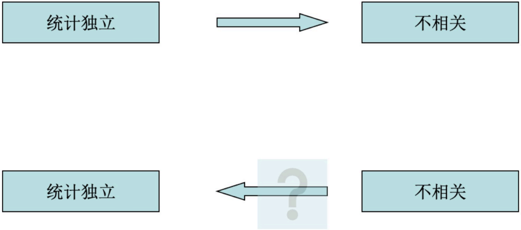
相关性和独立性

Example (相关性和独立性)

考虑随机变量 Z，具有概率密度

$$f_Z(z) = \begin{cases} \frac{1}{2\pi}, & z \in [0, 2\pi] \\ 0, & \text{其他} \end{cases}$$

问 $X = \sin(Z)$, $Y = \cos(Z)$ 是否相关，是否独立？ 提示：$\sin(t)\cos(t) = \frac{1}{2}\sin(2t)$



随机变量的矩

定义 (矩)

考虑随机变量 $X, k = 1, 2, \ldots$

- 其 k 阶原点矩 $m_k = \mathbb{E}[X^k]$
- 其 k 阶中心矩 $\mathbb{E}[(X - \mathbb{E}[X])^k]$
  - 可能不存在
  - 二阶矩 $m_2 = \mathbb{E}[X^2] = (\mathbb{E}[X])^2 + \mathrm{Var}[X]$
  - 可由矩生成函数/矩母函数得到 $M_X(t) = \mathbb{E}[e^{tX}] = \int e^{tx} f_X(x) dx$

定理 (矩定理)

矩母函数 n 次导满足 $M_X^{(n)}(t) = \mathbb{E}[X^n e^{tX}]$, 注意到

$$M_X^{(n)}(0) = \mathbb{E}[X^n] = m_n$$

特征函数

定义 (特征函数)

随机变量 X 的特征函数定义为

$$\Phi_X(t) = \mathbb{E}[e^{\imath tX}] = \int e^{\imath tx} f_X(x)dx$$

- 是实数变量的复值函数，在 0 点有最大值 $|\Phi_X(t)| \leq \Phi_X(0) = 1$
- 和矩母函数 $M_X(t)$ 满足 $M_X(\imath t) = \Phi_X(t)$，若 n 阶原点矩存在，注意到

$$M_X^{(k)}(0) = \imath^k \mathbb{E}[X^k], \quad k \leq n$$

- 对于独立随机变量 $X_1, X_2, \ldots, X_n$，考虑$X \equiv \sum_{i=1}^n X_i$，则有

$$\Phi_X(t) = \prod_{i=1}^n \Phi_{X_i}(t)$$

- 唯一性：随机变量的分布函数和特征函数相互唯一确定（因为 Laplace 变换关系）

高斯分布特征函数

Example (高斯分布特征函数)

考虑标准高斯随机变量 $X \sim \mathcal{N}(0,1)$ 和 $Y = \mu + \sigma X$，问 X, Y 的概率密度函数和特征函数。

- 概率密度函数：

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad 和 \quad f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

- 矩母函数：

$$M_X(t) = \int e^{tx} f_X(x)dx = e^{t^2/2} \cdot \frac{1}{\sqrt{2\pi}} \int e^{-\frac{(x-t)^2}{2}} dx = e^{t^2/2}$$

- 则特征函数为：

$$\Phi_X(t) = M_X(\imath t) = e^{-t^2/2}$$

- 则有 $\Phi_Y(t) = \mathbb{E}[e^{\imath t(\mu+\sigma X)}] = e^{\imath t\mu} \cdot \mathbb{E}[e^{\imath t\sigma X}] = e^{\imath t\mu} \cdot \Phi_X(\sigma t) = e^{\imath t\mu - \sigma^2 t^2/2}$

条件分布与条件期望

定义 (条件分布与条件期望)

类似之前条件概率，针对随机变量X, Y，给定（随机事件） $Y = y$ 时

- $X|Y = y$的条件概率密度函数：$f_{X|Y}(x|y) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)}$
- 在 X 和 Y 独立时，$f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y)$，则 $f_{X|Y}(x|y) = f_X(x)$
- $X|Y = y$的条件期望： $\mathbb{E}[X|Y = y] = \int x f_{X|Y}(x|y)dx$，是 y 的函数!

- 回顾全概率公式:

$$\mathbb{P}(B) = \sum_{i=1}^{K} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i) \Rightarrow f_Y(y) = \int f_{Y|X}(y|x) f_X(x)dx$$

定理 (全期望定理)

$$\mathbb{E}[\mathbb{E}[X|Y]] = \int \mathbb{E}[X|Y = y] f_Y(y)dy \tag{2}$$

例子：全期望定理

$$\mathbb{E}[\mathbb{E}[X|Y]] = \int \mathbb{E}[X|Y = y]f_Y(y)dy \tag{3}$$

Example (全期望定理)

假设市场上的所有电灯泡都来自于 A, B 两个工厂，市场占比为 $6:4$，A 工厂生产的灯泡寿命数学期望是 5000 小时，B 工厂生产的灯泡寿命数学期望是 4000 小时，请问市场上购买到的灯泡寿命的数学期望是多少小时?

- 【概率的概率，第一部分 - 二项分布】https://www.bilibili.com/video/BV1Bz411b7Jy/?share_source=copy_web&vd_source=405ba20a8a5d0e3e737677cdbaa02da6&t=0
- 【概率的概率，第二部分 - 为什么”概率为0”不等同”不可能”】https://www.bilibili.com/video/BV1ga4y147sC/?share_source=copy_web&vd_source=405ba20a8a5d0e3e737677cdbaa02da6