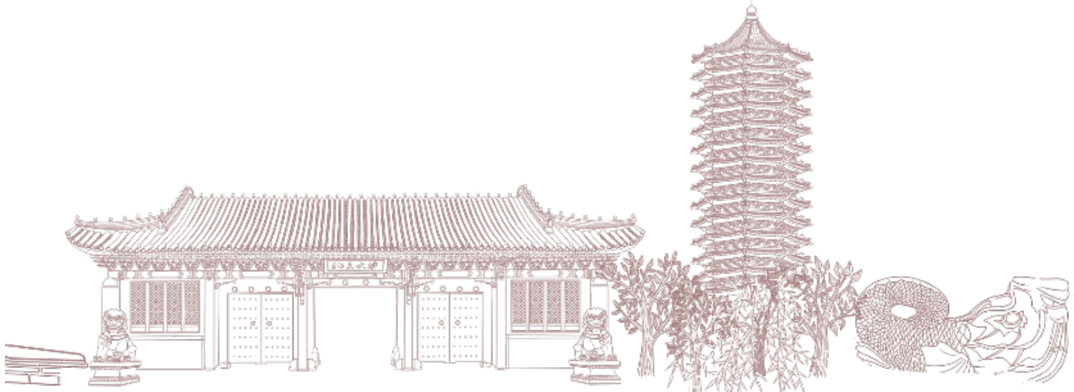
概率论

Probability Theory

主讲人：倪星宇

2024 年 4 月 29 日



从频率到概率

- 频率
  - 在相同条件下，进行了 n 次实验，在这 n 次实验中，事件 A 发生了 n_A 次
    - 比值 n_A/n 称为事件 A 发生的频率，记作 $f_n(A)$
  - 性质：频率落在 $[0,1]$ 上；两两互不相容的事件之并的频率等于各事件频率之和
  - 例：抛硬币的正反面（正面向上为 A、背面向上为 B、有一面向上为 $C = A \cup B$）
- 概率
  - 设 E 是随机试验，S 是其样本空间，对于 E 每一个事件 A 赋予一个实数，记为 $P(A)$
    - $P(A)$ 称为事件 A 的概率
  - 性质
    - 非负性、规范性、可列可加性
    - $P(A) = \lim_{n\to\infty} f_n(A)$

古典概型

- 古典概型（等可能概型）的定义
  - 试验的样本只包含有限个元素
  - 试验中每个基本事件发生的可能性相同
  - 例：抛硬币、掷骰子
- 古典概型的概率计算
  - 若事件 A 包含 k 个基本事件，则 $P(A)$ 取 A 包含的基本事件数除以基本事件的总数
  - 例 1：假设人出生在星期一到星期日的概率是相等的
    - 出生在工作日的概率是多少？出生在周末的概率是多少？
  - 例 2：N 件产品中有 D 件次品
    - 从中任取 n 件，恰有 k 件次品的概率是多少？
  - 例 3：生日悖论

条件概率

- 条件概率与乘法定理
  - $P(B|A) = P(AB)/P(A)$ 为在事件 A 发生的**条件**下事件 B 的概率
  - 乘法定理：$P(AB) = P(B|A)P(A)$
  - 例：抽卡抽中五星的概率为 0.6%，在抽到五星时抽到自己想要的角色的概率是 50%
    - 抽一次卡抽中自己想要的五星角色的概率是多少？
- 全概率公式与贝叶斯公式
  - 设试验 E 的样本空间为 S，A 为 E 的事件，$B_1, B_2, \dots, B_n$ 为 S 的一个划分
    - $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)$ 称为**全概率公式**
  - $P(B_i|A) = P(A|B_i)P(B_i)/\sum_{j=1}^{n} P(A|B_j)P(B_j)$ 称为贝叶斯公式
    - 例：人群中感染率为 q，核酸检测假阳性概率为 0、假阴性概率为 p
      - 核酸阴性者为感染者的概率是多少？

随机变量

- 随机变量
  - 定义：随机试验的样本空间为 $S=\{e\}$，随机变量 $X=X(e)$ 是 S 上的实值单值函数
  - 离散型随机变量
    - 0–1 分布：$P\{X=k\}=p^k(1-p)^{1-k}, k=0,1\ (0<p<1)$
    - 二项分布：$P\{X=k\}=\binom{n}{k}p^k q^{n-k}, k=0,1,2,\dots,n\ (q=1-p)$
      - 0–1 分布重复 n 次求和
    - 泊松分布：$P\{X=k\}=\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, k=0,1,2,\dots\ (\lambda>0)$
      - n 很大时，二项分布近似于泊松分布
  - 连续型随机变量
    - 分布函数 $F(x)=P\{X\le x\}, (-\infty<x<+\infty)$
      - 如何计算 $P\{x_0<X\le x_1\}$?
    - 概率密度函数 $f(x)$：$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$

随机变量

- 连续型随机变量
  - 均匀分布
    - 概率密度 $f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$
  - 指数分布
    - 概率密度 $f(x) = \begin{cases} \frac{1}{\theta} \mathrm{e}^{-\frac{x}{\theta}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$，分布函数 $F(x) = \begin{cases} 1 - \mathrm{e}^{-\frac{x}{\theta}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$
    - 无记忆性：$P\{X > s + t | X > s\} = P\{X > t\}$
  - 正态分布
    - 概率密度 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, (-\infty < x < +\infty)$

随机变量的数字特征

- 数学期望
  - 离散型 $E(X)=\sum_{k=1}^{\infty}x_kp_k=\sum_{k=1}^{\infty}P\{X\geq k\}$
  - 连续型 $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$
  - 重要性质
    - $E(C)=C$ (C 为常数)
    - $E(CX)=CE(X)$ (C 为常数)
    - $E(X+Y)=E(X)+E(Y)$ (X,Y 任意)
    - $E(XY)=E(X)E(Y)$ (X,Y 相互独立)
  - 例
    - 一民航送客车载有 20 位旅客自机场开出，旅客有 10 个车站可以下车。如到达一个车站没有旅客下车就不停车。以 X 表示停车的次数，求 $E(X)$。
      - 设每位旅客在各个车站下车是等可能的，并设各位旅客是否下车相互独立

随机变量的数字特征

- 方差
  - 度量随机变量与其均值（期望）的偏离程度
  - $D(X) = \mathrm{Var}(X) = E\{[X - E(X)]^2\}$
    - $\sigma(X) = \sqrt{D(X)}$ 称为标准差或均方差
  - 方差是随机变量 X 的函数 $g(X) = [X - E(X)]^2$ 的数学期望
    - $D(X) = \sum_{k=1}^{\infty}[x_k - E(X)]^2 p_k$
    - $D(X) = \int_{-\infty}^{+\infty}[x - E(X)]^2 f(x)\mathrm{d}x$
    - $D(X) = E(X^2) - [E(X)]^2$
  - 性质
    - $D(CX) = C^2 D(X), D(X + C) = D(X)$ (C 为常数)
    - $D(X + Y) = D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}$

随机变量

- 泊松分布的推导 ($E(X) = D(X) = \lambda$)
  - 条件
    - X 是在一个区间内发生特定事件的次数，取值范围为自然数
    - 一个事件的发生不影响其它事件的发生，即事件独立发生
    - 事件的发生率是相同的，不能有些子区间内发生率高一些而另一些子区间低一些
    - 两个事件不能在同一个时刻发生
    - 一个区间内一个事件发生的概率与区间的大小成比例
  - 例：已知某医院平均一天里有 λ 名新生儿诞生，满足上述条件，求出生 k 人的概率
    - 设一天分为 n 个时间段，每个时间段内要么有人出生，要么没有
      - 每个时间段有人出生的概率为 $p = \lambda/n$
    - $P\{X = k\} = \binom{k}{n} p^k (1-p)^{n-k} = 1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\lambda^k \frac{1}{k!}\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$

随机变量

- 指数分布的推导
  - 无记忆性：$\frac{1-F(s+t)}{1-F(s)} = 1 - F(t)$
    - $1 - F(s+t) = 1 - F(s) - F(t) + F(s)F(t)$
    - $F(x) + f(0)\Delta x = F(x) + f(x)\Delta x + F(x)f(0)\Delta x$
    - $f(0) = f(x) + F(x)f(0)$
  - 设 $f(0) = 1/\theta$
    - $\theta f(x) = 1 - F(x)$
    - $f(x) = \frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}$
    - $F(x) = 1 - \mathrm{e}^{-\frac{x}{\theta}}$
  - $E(X) = \theta, D(X) = \theta^2$

随机变量的数字特征

- 协方差
  - $E\{[X-E(X)][Y-E(Y)]\}$ 称为随机变量 X 与 Y 的协方差，记为 $\mathrm{Cov}(X,Y)$
    - $\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$ 称为随机变量 X 与 Y 的相关系数
  - 协方差的性质
    - $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X), \mathrm{Cov}(X,X)=D(X)$
    - $D(X+Y)=D(X)+D(Y)+2\,\mathrm{Cov}(X,Y)$
    - $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$
    - $\mathrm{Cov}(aX,bY)=ab\,\mathrm{Cov}(X,Y)$ (a,b 是常数)
    - $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$
  - 相关系数为 0 称为两个随机变量**不相关**
    - 相互独立一定不相关，不相关不一定相互独立！
    - 相关系数只是对于两个随机变量**线性**相关度的描述

随机变量的数字特征

- 矩
  - 定义（设 $k = 1,2,\ldots$）
    - $E(X^k)$ 称为 X 的 k 阶原点矩，简称 k 阶矩
    - $E\{[X - E(X)]^k\}$ 称为 X 的 k 阶中心距
    - $E(X^k Y^l)$ 称为 X 和 Y 的 $k + l$ 阶混合矩
    - $E\{[X - E(X)]^k [Y - E(Y)]^l\}$ 称为 X 和 Y 的 $k + l$ 阶混合中心矩
  - 矩与期望、方差
    - X 的数学期望 $E(X)$ 是 X 的一阶原点矩、方差 $D(X)$ 是 X 的二阶中心矩
    - 协方差 $\mathrm{Cov}(X, Y)$ 是 X 和 Y 的二阶混合中心矩

随机变量的数字特征

- 协方差矩阵
  - $c_{ij} = \mathrm{Cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}$
    - n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的二阶混合中心距
  - $$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$
    - n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的协方差矩阵（为对称阵）

大数定律

- 弱大数定律（辛钦大数定律）
  - $\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1$
    - 设 $X_1, X_2, \dots$ 是相互独立、服从同一分布的随机变量序列，满足 $E(X_k) = \mu\ (\forall k)$
    - ε 是任意大于 0 的小量
  - 序列 $\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ **依概率**收敛于 μ
    - $\bar{X} \xrightarrow{P} \mu$
- 伯努利大数定律
  - $\lim_{n\to\infty} P\left\{\left|\frac{f_A}{n} - p\right| < \varepsilon\right\} = 1$
    - f_A 是 n 次独立重复试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率

中心极限定理

- 中心极限定理（独立同分布）
  - $\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t^2/2} \mathrm{d}t = \Phi(x)$
    - 随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立且服从同一分布
    - $E(X_k) = \mu, D(X_k) = \sigma^2\ (\forall k)$
    - $Y_n = \frac{\sum_{k=1}^{n} X_k - E\left(\sum_{k=1}^{n} X_k\right)}{\sqrt{D\left(\sum_{k=1}^{n} X_k\right)}} = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$
  - 高斯分布（正态分布）
    - $N(\mu, \sigma^2)$ 表示均值为 μ、标准差为 σ 的高斯分布
    - $\bar{Y} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
    - $\bar{X} \sim N(\mu, \sigma^2/n)$

中心极限定理

- 李雅普诺夫定理
  - $\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_k} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t^2/2} \mathrm{d}t = \Phi(x)$
    - 随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立，满足 $E(X_k) = \mu_k$ 和 $D(X_k) = \sigma_k^2$
    - $B_n^2 = \sum_{k=1}^{n} \sigma_k^2$
    - 存在正数 δ 使得 $n \to \infty$ 时 $\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} E\{|X_k - \mu_k|^{2+\delta}\} \to 0$
- 棣莫弗—拉普拉斯定理
  - $\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t^2/2} \mathrm{d}t = \Phi(x)$
    - η_n 服从参数为 n, p $(0 < p < 1)$ 的二项分布
    - 证明：二项分布是 0–1 分布之和