概率论与数理统计（前半）——笔记

Adopac

2022 年 9 月 5 日

目录

1 概率论的基本概念

1.1 随机事件

试验 观察测量实验

随机试验：在相同条件下可重复、结果不止一次、无法预测。用 E 表示

事件 每种结果

随机事件

基本事件：相对于实验目的来说，不能再分或不必再分

复合事件：由基本事件复合。

必然事件（Ω ）：一定发生

不可能事件（Φ ）：一定不发生

样本空间：所有基本事件的集合（必然事件）

样本点：样本空间中的元素

事件间的关系 (1) 包含 $A \subset B$ $B \supset A$ A 发生必然导致 B 发生

$$\Phi \subset A \subset \Omega$$

常见相等证明 $A \subset B$ 且 $B \supset A$，则有 $A = B$

(2) 和（并）$A \cup B$ $A + B$

AB 中至少有一个发生

$A + B \supset A$ $A + A = A$ $A + \Phi = A$ $A + \Omega = A$

(3) 交（积）$A \cap B$ AB

AB 同时发生

$AB \subset A$ $AA = A$ $A\Phi = \Phi$ $A\Omega = A$

(4) 差 A-B

A 发生而 B 不发生

(5) 互不相容事件

A,B 不同时发生，AB=Φ

(6) 对立事件

A,B 互不相容，并且 A+B=Ω $A = \overline{B}$ $B = \overline{A}$

有公式 $A - B = A - AB = A\overline{B}$

(7) 完备事件组

$A_1A_2A_3...A_n$ 两两互不相容，且 $\bigcup_{i=1}^{n} A_i = \Omega$。

无限可列个　能按照某种规律排成一个序列

实数排不了，直线排不了 $\rightarrow$ 不可列

运算律	交换律	$A \cup B = B \cup A$ $A \cap B = B \cap A$
	结合律	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
	分配律	$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
	对偶	$\overline{A \cup B} = \overline{A} \cap \overline{B}$ $\overline{A \cap B} = \overline{A} \cup \overline{B}$ $\overline{A_1 \cup A_2 \cup A_3 \cup ... \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap ... \cap \overline{A_n}$

1.1.1　事件的概率

古典概率　：（1）有限个样本点　（2）等可能性

排列　(1) 不重复排列从 n 个不同元素中取出 m 个

$$P_n^m = n(n-1)(n-2)...(n-m+1)$$

(2) 全排列

$$P_n^m = n(n-1)(n-2)... \times 3 \times 2 \times 1 = n!$$

(3) 重复排列从 n 个元素抽 m 个排列

$$n \times n \times n... \times n = n^m$$

组合　从 n 个不同元素中取出 m 个不同元素

$$C_n^m = \frac{P_n^m}{m!}$$

$C_n^m = C_n^{n-m}$　$C_n^n = C_n^0 = 1$

1.2　几何概型

度量

$$P_A = \frac{\mu(G)}{\mu\Omega}$$

具有完全可加性

频率与概率 n 次实验，A 发生 m 次，$\omega_n(A) = \frac{m}{n}$ 频率。

性质：非负、规范、可加

$$P(A+V) = P(A) + P(B) - P(AB)$$

1.3

1.3.1 条件概率

$$P(A|B) = \frac{n_{AB}}{n_B} = \frac{P(AB)}{P(B)}$$

若 $A_1, A_2, ..., A_n$ 不相容，有

$$P(\sum_{i=1}^{\infty} A_i|B) = \sum_{i=1}^{\infty} P(A_i|B)$$

1.3.2 乘法公式

$$P(AB) = P(B) \cdot P(A|B)$$

$$P(AB) = P(A) \cdot P(B|A)$$

要求 P(A)>0 P(B)>0

$$P(A_1A_2...A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)...P(A_n|A_1A_2...A_{n-1})$$

1.4

1.4.1 全概率公式

$A_1A_2...A_n$ 是 E 的完备事件组

$P(A_i) > 0$

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

1.4.2 贝叶斯公式

$A_1A_2...A_n$ 是 E 的完备事件组
$P(A_i) > 0 \ P(B) > 0$

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} = \frac{P(A_kB)}{P(B)}$$

$P(A_i)$ 先验概率　$P(A_i|B)$ 后验概率

1.5

1.5.1 事件的独立性

A 的概率不受 B 发生与否的影响 $P(A|B) = P(B)$
定理：$P(A) > 0P(B) > 0$
AB 独立 $\Leftrightarrow P(AB) = P(A)P(B)$
AB 独立 也有 A 与 $\overline{B}\,\overline{A}$ 与 B $\overline{A}$ 与 $\overline{B}$ 独立

1.5.2 伯努利模型

定理：A 的概率 P (0<P<1) $\overline{A} = 1 - P$ 在 n 重伯努利中 A 发生 k 次

$$P_n(k) = C_n^k p^k (1-p)^{n-k}$$

2 随机变量及其分布

2.1 随机变量的概念

定义 $\Omega X = X(\omega)$ 为实值函数，$\{\omega|X(\omega)\}, \{x = a\}$

离散型：有限个、无线可列

非离散型：连续型

2.2

2.2.1 离散型随机变量及其概率分布

X 的所有取值 $X_k(k = 1, 2, ...)$ 可列个

概率函数 $P = \{X = x_k\} = P_k$

2.2.2 连续型随机变量及其概率密度函数

非负可积 f(x) $f(x) \geqslant 0$ $a \leqslant b$

$P\{a < x \leqslant b\} = \int_a^b f(x)dx$

x 连续，则称 f(x) 为概率分布密度函数

性质：（1）$f(x) \geqslant 0$ （2）$\int_{-\infty}^{+\infty} f(x) = 1$ （3）连续变量取个别值的概率为 0

概率为 0 的事件未必是不可能事件，概率为 1 的事件未必是必然事件。

分布函数

离散型：$F(x) = P(X \geqslant x)$ 普通的实函数，X 取值不超过 x 的概率

连续型：$F(x) = P\{X \leqslant x\} = \int_{-\infty}^{x} f(t)dt$

2.2.3 常见的分布

0-1 分布

$$P\{x = k\} = p^k(1-p)^{1-k} \ k = 0, 1$$

几何分布 $X \sim G(P)$P(A)=P，第 k 此首次发生，前 k-1 次未发生。

$$P\{x = k\} = (1-p)^{k-1}p$$

二项分布 P(A)=P n 次实验内发生了 k 次

$$P\{x=k\}=C_n^k p^k(1-p)^{n-k}\ k=0,1,2,...n$$

泊松分布

$$P\{x=k\}=\frac{\lambda^k}{k!}e^{-\lambda}\ k=0,1,2,...$$

应用类型：电话台收到的呼叫次数、公共设施服务的人数（等车、收银台等）

将二项分布近似成泊松分布 条件：$n\geqslant 100\ np\leqslant 10$，将 $\lambda=np$ 代入泊松公式

超几何分布 N 个元素，N_1 个属于第一类，N_2 个属于第二类。x，取 n 个属于第一类

$$P\{x=k\}=\frac{C_{N_1}^k C_{N_2}^{n-k}}{C_N^n}\ k=0,1,2,...,min(n,N_1)$$

均匀分布

$$f(x)=\begin{cases}\frac{1}{b-a} & a\leqslant x\leqslant b\\ 0 & else\end{cases} \tag{1}$$

指数分布

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\leqslant 0\end{cases} \tag{2}$$

$\lambda>0$

正态分布 $X\sim N(\mu,\sigma^2)$

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

性质：σ 固定，μ 变化，左右移动；μ 固定，σ 变小，最高点上移、σ 变大，最高点下移

标准正态分布 $\mu=0\ \sigma=1$

2.3 随机变量函数分布

连续型模板：设 x 的 $f_x(x)$ $y = g(x)$ $Y = g(X)$ 应 $F_Y(X) \to F_X(x)$，再 $f_X(x) \to f_Y(x)$

X 服从 [a,b] 均匀分布，$Y = kX + C\,(k \neq 0)$ 服从相应区间上的均匀分布

X 服从正态分布，Y 是线性函数，则 Y 也服从正态分布

3 多维随机变量及其分布

3.1

3.1.1 二维随机变量及其分布

E 是随机试验，Ω 为样本空间，X,Y 是 Ω 的两个变量。

分布函数 $F(x,y)=P\{X\leqslant x,Y\leqslant y\}$ 联合分布

性质：1、$0\leqslant F(x,y)\leqslant 1$；2、$F(x,y)$ 不减；3、$F(x,y)$ 分别关于 x,y 右连续；4、$x_1\leqslant x_2$ $y_1\leqslant y_2$ $P\{x_1<x\leqslant x_2\,,\,y_1<y\leqslant y_2\}=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)$

3.1.2 二维离散型的联合分布及边缘分布

$P\{X=x_i\,,\,Y=y_i\}=P_{ij}$

性质：1、$P_{ij}\geqslant 0$；2、$\sigma_i\sigma_jP_{ij}=1$

联合分布可唯一确定边缘分布，边缘分布不能确定联合分布。

3.1.3 二维连续型的联合分布与边缘分布

$F(x,y)=P\{X\leqslant x\,,\,Y\leqslant y\}=\int_{-\infty}^{x}\int_{-\infty}^{y}f(s,t)dsdt$

性质：1、$f(x,y)>0$2、$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$3、$\frac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y)$

边缘密度函数

$F_X(x)=F_X(x,+\infty)=\int_{-\infty}^{x}[\int_{-\infty}^{+\infty}f(s,t)dt]ds$

$f_X(x)=F_X'(x)=\int_{-\infty}^{+\infty}f(x,y)dy$

$f_Y(y)=F_Y'(y)=\int_{-\infty}^{+\infty}f(x,y)dx$

二维正态分布的边缘也是正态分布；两边缘分布是正态，二维并非是正态

3.2

3.2.1 条件分布

$F(x)=P\{X\leqslant x\}$

$F(x|A)=P\{X\leqslant x|A\}$

连续型的条件分布 $f(x|y)=\frac{f(x,y)}{f_Y(y)}$

$f(y|x) = \frac{f(x,y)}{f_X(x)}$

3.2.2 随机变量的独立性

$f(x,y) = f_X(x) \cdot f_Y(y)$
$F(x,y) = F_X(x) \cdot F_Y(y)$

二维离散的独立性

$P\{x = x_i\,, y = y_i\} = P\{x = x_i\}P\{y = y_i\}$

变量独立，构造函数也独立

3.3 二维随机变量函数的分布

二维离散 Z=XY

XY 独立，$\lambda_1\lambda_2$ 的泊松分布 $Z = X_1 + X_2$

$\{z = k\} = \int_{i=0}^{k}\{X = i\,, Y = k - i\}$

$Pz = k = \Sigma_{i=0}^{k}P\{X = i\}\,P\{Y = k - i\} = \Sigma_{i=0}^{k}\frac{\lambda_1^i}{i!}e^{-\lambda_1}\frac{\lambda_2^i}{(k-i)!}e^{-\lambda_2} = \frac{(x_1+x_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}$

连续变量函数

$F_Z(\zeta) = P\{Z \leqslant \zeta\} = P\{g(x,y) \leqslant \zeta\} = \int_{D_\zeta} f(x,y)dxdy$ 其中 $D_\zeta = \{(x,y)|g(x,y) \leqslant \zeta\}$

卷积公式 Z=X+Y，X,Y 独立

$$f_z(\zeta) = \int_{-\infty}^{+\infty} f(x, \zeta - x)dx = \int_{-\infty}^{+\infty} f(\zeta - y, y)dy$$

$\mathbf{x} \sim N(0,1)\,, Y \sim N(0,1)\,,$ ，X,Y 独立，Z=X+Y

$\Phi\zeta = \int_{-\infty}^{+\infty}\varphi_X(x)\varphi_Y(\zeta - X)dx = Z \sim N(0,2)$

结论：有 $X \sim N(\mu_1, \sigma_1^2)\,, Y \sim N(\mu_2, \sigma_2^2)$，有

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

4 随机变量的数字特征

4.1 数学期望

离散型的期望 $P(X=X_k)=P_k$，$E_x=\Sigma_{k=1}^{\infty}X_kP_k$

连续型的期望 $E_x=\Sigma{-\infty}^{+\infty}xf(x)dx$

数学期望的性质：

$EC=C$

$E(kX+b)=k\cdot EX+b$

$E(x\pm Y)=EX\pm EY$

若 X,Y 独立，有 $E(XY)=E(X)\cdot E(Y)$

条件期望：一个变量取某值，另一个变量的期望

4.2

4.2.1 方差

表示偏离程度 $(X-EX)^2$；方差 $DX=E(X-EX)^2$；标准差 $\sqrt{DX}$

离散型：$DX=\Sigma_k(X_k-EX)^2P_k$

连续型：$DX=\int_{-\infty}^{+\infty}(X-EX)^2f(x)dx$

$$DX=E(X-EX)^2=EX^2-(EX)^2$$

方差的性质：

$DC=0$

$D(kX+b)=k^2DX$

X,Y 独立，有 $D(X\pm Y)=DX\pm DY$

$DX=0\leftrightarrow P(X=EX)=1$

标准化：

$$X^*=\frac{X_EX}{\sqrt{DX}}$$

4.3 常见期望与方差

分布	EX	dx
0-1	P	$P\cdot(1-P)$
二项	np	$np\cdot(1-P)$
几何	$\frac{1}{p}$	$\frac{1-p}{p^2}$
泊松	λ	λ
均匀	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
正态	μ	σ^2

4.4

4.4.1 协方差

$$\begin{aligned}CoV(X,Y) &= E[(X-EX)(Y-EY)]\\ &= E(XY-XEY-YEX+EXEY)\\ &= E(XY)-EXEY\end{aligned}$$

4.4.2 相关系数

$$\rho=\frac{CoV(x,y)}{\sqrt{DX}\sqrt{DY}}=\frac{E(XY)-EXEY}{\sqrt{DX}\sqrt{DY}}$$

ρ 与 CoV 同正，同负，同零

定理 4.4，$|\rho|=1 \Longleftrightarrow$X 与 Y 与 ρ=1 成线性关系。$P(Y=ax+b)=1$

$\rho=1$，XY 完全正相关

$\rho=-1$，XY 完全负相关

$|\rho|\to 0$，X 线性相关弱

$\rho=0$，XY 不存在线性关系

XY 独立，则 XY 不相关；XY 不相关，则 XY 不一定独立

5 大数定律及中心极限定理

5.1 大数定律

大量重复实验的稳定性

5.1.1 切比雪夫不等式

EX 和 DX 存在 $\forall \xi > 0$，有 $P(|X - EX| \geqslant \xi) \leqslant \frac{DX}{\xi^2}$

5.1.2 切比雪夫大数定律

伯努利大数定律 做 n 重伯努利试验，A 发生了 m_n 次，概率 P，频率 $\frac{m_n}{n}$，有 $\lim_{n\to\infty}\{|\frac{mn}{n} - p| < \xi\} = 1$

切比雪夫大数定律

$x_1, x_2, ...x_n, ...$ 不相关的变量，EX_i 和 DX_i 都存在方差有界，$DX_i \leqslant m$

对于 $\forall \xi > 0$

$$\lim_{n\to\infty} P\{|\frac{1}{n}\int_{i=1}^{n} X_i - \frac{1}{n}\int_{i=1}^{n} EX_i| < \xi\} = 1$$

5.2 中心极限定理

现象由大量相互独立的因素影响，大量独立同分布的变量和的极限分布是正态分布。

定理 $x_1, x_2, ..., x_n, ...$ 独立同分布，$E_x i = \mu$ $D_x i = \sigma^2$ $0 < \sigma^2 < +\infty$

$$\lim_{n\to\infty} P(\frac{\Sigma_{i=1}^{n} - nu}{\sqrt{n}\sigma} \leqslant x) = \varPhi_0(x)$$

6 样本及抽样分布

6.1 总体与样本

简单随机抽样：1、同分布；2、独立同分布

6.1.1 统计量定义

不含任何未知参数的样本的函数

6.1.2 常用统计量

样本均值、样本方差...

设总体 X 的均值为 $EX = \mu$，方差为 $DX = \sigma^2$，样本 $(x_1, x_2, ..., x_n)$ 来自总体 X，则

$E\overline{X} = \mu$ $D\overline{X} = \frac{1}{n}\sigma^2$ $ES^2 = \sigma^2$

6.2 抽样分布

正态分布，卡方分布 $\mathcal{X}^2(n)$

性质：1、$\mathcal{X}^2(2)$，为 $\lambda = \frac{1}{2}$ 的指数分布。2、单峰曲线 n=2 取最大值，不对称；n 增大，峰向右，越对称，n 很大时可用正态分布。

定理：$X_1, X_2, ..., X_n$ 独立 $N(0,1)$ $\Sigma_{i=1}^n X_i^2 \sim \mathcal{X}^2(n)$

有卡方分布 $EX = n$ $DX = 2n$

定理：$X \sim \mathcal{X}^2(n)$ $Y \sim \mathcal{X}^2(m)$，XY 独立，有 $X + Y \sim \mathcal{X}^2(n+m)$

上 α 分位数：$P(X^2 > \mathcal{X}^2_\alpha(n)) = \alpha$

t 分布 $X \sim N(0,1)$ $Y \sim \mathcal{X}^2(n)$ XY 独立，有 $\frac{X}{\sqrt{\frac{Y}{n}}} \sim t(n)$

上 α 分位数 $P(T > t_\alpha(n)) = \alpha$

F 分布 $X \sim \mathcal{X}^2(n_1)$ $Y \sim \mathcal{X}^2(n_2)$ 独立，有 $\frac{\frac{X}{n_1}}{\frac{Y}{n_2}} \sim F(n_1,)$

若 $F \sim F(n_1, n_2)$，则 $\frac{1}{F} \sim F(n_2, n_1)$

上 α 分位数 $P(F > F_\alpha(n_1, n_2)) = \alpha$

7 参数估计

7.1 矩估计法

用样本的矩估计总体的矩

一阶矩 $\overline{X} = \frac{1}{n}\Sigma X_i$ 二阶矩 $\overline{X} = \frac{1}{n}\Sigma X_i^2$

7.1.1 极大似然估计

模板：1、写出总体概率函数或者密度函数；2、写出似然函数；3、两边取 ln；4、对 λ 求导，令导数 =0。

7.2 点估计的优良性原则

无偏性：$E\widehat{\theta} = \theta$

有效性：$D(\widehat{\theta}_1) \leqslant D(\widehat{\theta}_2)$

7.3 置信区间和枢纽变量

区间估计 $P(\widehat{\theta}_1 \leqslant \theta \leqslant \widehat{\theta}_2) = 1 - \alpha$

$1 - \alpha$ 为置信度

枢纽变量 $I = I(T, \theta)$

给定 $1 - \alpha$，确定 F 的上 α 分位数，上 $(1 - \frac{\alpha}{2})$ 分位数有

$$P(U_{1-\frac{\alpha}{2}} \leqslant I(T, \theta) \leqslant U_{\frac{\alpha}{2}}) = 1 - \alpha$$