概率论部分

理论基础

1. 大数定律

① 弱大数定理(辛钦)：
$\lim_{n\to\infty} P(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < \varepsilon) = 1$.

② Bernoulli：
$\lim_{n\to\infty} P(|\frac{f_n}{n} - p| < \varepsilon) = 1$.

（要求同分布.）

③ Chebyshev：
$\lim_{n\to\infty} P(|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E(X_k)| < \varepsilon) = 1$.

2. 中心极限定理 (CLT)

Lemma: Lindeberg's Thm.

① 独立同分布：$(X_1,\cdots,X_n)$ (Lindeberg-Lévy CLT) ($n\to\infty$时)
$Y_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \sim N(0,1)$ (近似服从)
($\sum_{i=1}^{n} X_i = \sqrt{n}\sigma Y_n + n\mu \sim N(n\mu, n\sigma^2)$.)

② De Moivre-Laplace Thm：
$(p+q=1)$ $X_i \sim B(n,p)$. $n\uparrow$, $X \sim N(np, npq)$
$\lim_{n\to\infty} P(\frac{Y_n - np}{\sqrt{npq}} < x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$.

③ 不同分布：$X_1,\cdots,X_n$：$f_{X_k}(x)$. $E(X_k) = \mu_k$.
$D(X_k) = \sigma_k^2$. $Y_n = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n}\mu_k}{\sum_{k=1}^{n}\sigma_k^2}$. (类似 Lindeberg 条件)
$\lim_{n\to\infty} P(\frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n}\mu_k}{\sum_{k=1}^{n}\sigma_k^2} \le x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$.

应用

(1) 基础知识

① 样本空间 Ω, 随机事件
② 频率 f、概率 p
③ 古典概型、几何概型…
④ 条件概率. (全概率公式、Bayes 公式.)
⑤ 独立性.

(2) 一维随机变量

① 连续型
- 均匀分布 ($\frac{1}{b-a}$)
- 指数分布 ($\frac{1}{\theta}e^{-\frac{x}{\theta}}$)
- 欧拉分布 (…)

② 离散型
- 退化分布(两点)
- 几何分布
- ~~[illegible]~~
- 二项分布 ($\binom{n}{k}p^k(1-p)^{n-k}$.)
- 泊松分布 ($\frac{\lambda^k}{k!}e^{-\lambda}$.)

③ $F(x)$、$f(x)$.
④ $f \circ g$ 时 $f(x)$、$F(x)$.

(3) 二维随机变量 (X,Y)

① 联合分布、边际分布. 条件分布.
② 相互独立的分布.
③ $X\pm Y$, XY、$\frac{X}{Y}$. $\max\{X,Y\}$, $\min\{X,Y\}$.

(4) 数字特征

① $E(X)$
- 连续 (绝对收敛) $\int_{-\infty}^{+\infty} x f(x) dx$
- 离散 (绝对收敛) $\sum_{k=1}^{\infty} x_k p_k$
- $f \circ g$.

② $D(X)$
- $E(X^2) - [E(X)]^2$.
- $\sigma(X) = \sqrt{D(X)}$.
- Chebyshev 不等式.
- $Cov(X,Y)$.
- ρ_{XY}
- ($D(X\pm Y) = D(X) + D(Y) \pm 2cov(X,Y)$.)

③ $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$.

① 有限个.
② 可数无穷个. (N)
③ 连续. ($[0,1]$)

样本空间 Ω ← 样本点 ($\forall$一个结果) ← 随机试验 → 随机事件 $A、B、C,\cdots$

① 必然、不可能事件.
(注: 不可能事件: $P=0$, 反之未必成立.)
② 随机事件.

频率 f
$f_n(A):=\frac{m}{n}$.
性质: ① $0\le f_n(A)\le 1$.
② 必然事件 $f=1$, 不可能… $f=0$.
Buffon 投针试验.

① $\Omega=\{\omega_1,\omega_2,\omega_3\}$.
ω_1 出现, 事件发生
② 必然事件 $=\Omega$.
③ 不可能事件 $=\varnothing$.

(elementary event)
随机事件分类
① 基本事件 (simple event)
② 复合事件

概率的统计定义.
$\lim_{n\to\infty} f_n(A) = p = P(A)$.
(实操性较差.)

性质
① $0\le P\le 1$.
② 必然 $P=1$; 不可能 $P=0$.
③ 可加性. (互斥)
(可列可加性)
$\forall A,B$. $P(A+B)=P(A)+P(B)-P(AB)$.

(要求满足P的几条基本性质)

① 条件概率: $P(A|B)=\frac{P(AB)}{P(B)}$.
($\Rightarrow P(AB)=P(A|B)P(B)=P(B|A)P(A)$.)
※ 本质是样本空间的缩小.

② 乘法公式
$P(A_1\cdots A_{n-1})>0$, 则:
$P(A_1\cdots A_n)=P(A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_{n-1})$.

③ 全概率公式: $B_1, B_2, \cdots$ 完备事件组.
$\boxed{P(A)=\sum_i P(B_i)P(A|B_i)}$
(若只有 B、$\bar B$, $P(A)=P(B)P(A|B)+P(\bar B)P(A|\bar B)$.)

④ Bayes公式: 由果溯因.
$$\underbrace{P(B_i|A)}_{\text{后验概率}}=\frac{P(AB_i)}{\sum_j P(B_j)P(A|B_j)}=\frac{\overbrace{P(B_i)}^{\text{先验概率}}P(A|B_i)}{\sum_j P(B_j)P(A|B_j)}.$$

⑤ 独立性 (Independent)
$P(AB)=P(A)P(B)$. ⇒ n个事件两两相互独立
$P(A_1\cdots A_n)=P(A_1)\cdots P(A_n)$.
⋮
$P(A_iA_j)=P(A_i)P(A_j)$.

① 古典概型. (等可能; 有限个.)
$P(A)=\frac{n(A)}{n(\Omega)}$. ⇒ 抽签先后不影响公平.
性质
① $0\le P(A)\le 1$.
② $P(\varnothing)=0$, $P(\Omega)=1$.

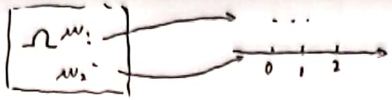
② 几何概型. (无限个; 测度之比)
(可测集A) $P(A)=\frac{m(A)}{m(\Omega)}$ (m: Measure)
(Survivorship Bias.)

运算

① 包含: $A\subset B$. (同集合论)
② 相等: $A\supset B$ 且 $A\subset B \Leftrightarrow A=B$.
③ 交: $A\cap B$ / AB. (乘) $\Rightarrow \bigcap_{i=1}^{n} A_i$; $\bigcap_{i=1}^{\infty} A_i\cdots$
④ 并: $A\cup B$ / $A+B$. (加) $\Rightarrow \bigcup_{i=1}^{n} A_i$; $\bigcup_{i=1}^{\infty} A_i\cdots$
⑤ ~~辨~~差: $A\setminus B$ / $A-B$.
⑥ 补: $\Omega-A$ / $\bar A$ / $C_\Omega A$ / A^C. 对立. (⇒互斥)
$A-B=A\bar B$. $AB=\varnothing$, $A+B=\Omega \Leftrightarrow$ A,B对立.

⑥: $AB=\varnothing$ 互斥

⑦ 交换律、结合律、分配律、De-Morgan律.
($\overline{A+B}=\bar A\bar B$; $\overline{AB}=\bar A+\bar B$.)
⑧ 划分/分割/完备事件组: $B_1, B_2, \cdots$
① B_i 两两互斥. ② $\bigcup_i B_i=\Omega$.



随机变量 (r.v.)

① 离散 (discrete) ② 连续 (continuous).

③ 奇异 (singular) ④ 混合型 (mixed).

连续型

① $F(x)=P(X\le x)$. $x\in\mathbb{R}$.

$\Rightarrow P(a<x\le b)=F(b)-F(a)$.

⇔ 单增；右连续；规范性.

↓ F(x)绝对连续.

② 概率密度 $f(x)=F'(x)$.

$\int_{-\infty}^{+\infty} f(x)dx=1$. (即F(x).)

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

③ 均匀分布 $X\sim U[a,b]$.

$$f(x)=\begin{cases}\frac{1}{b-a}, & x\in[a,b]\\ 0, & \text{else}\end{cases}$$

④ Cauchy分布: $\boxed{f(x)=\frac{1}{\pi(1+x^2)}}$

$F(x)=\boxed{\frac{1}{2}}+\frac{1}{\pi}\arctan x$.

⑤ 指数分布: $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0.\\ 0, & x<0.\end{cases}$

$$F(x)=\begin{cases}1-e^{-\lambda x}, & x\ge 0.\\ 0, & x<0.\end{cases}$$

等待时间. 无记忆性.

⑥ Γ分布: $X\sim\Gamma(\alpha,\frac{1}{\beta})$.

$f(x,\beta,\alpha)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$, $x>0$.

⑦ $F_Y(y)=P(Y\le y)=P(g(X)\le y)$ $\Rightarrow f(y)=F'(y)$.

$=P(X\le x\,|\,g(x)\le y)=\int_I f(x)dx$.

$g(X)=Y$
反函数 $x=h(y)$.
$f_Y(y)=f_X[h(y)]\cdot h'(y)$.

离散型

① 只能取有限个/可列的.

$P(X=x_k)=p_k$. $k=1,2,\cdots$

非负性；$\sum_k p_k=1$.

② 几何分布: $P(X=k)=pq^{k-1}$, $k=1,2,\cdots$

$(p+q=1)$

③ n重Bernoulli (二项分布): $C_n^k p^k q^{n-k}$ $X\sim B(n,p)$.

④ 两点分布/退化分布: ← (n=1时)

$P(X=x_1)=p$, $P(X=x_2)=1-p$.

⇓

0-1分布: $P(X=0)=p$, $P(X=1)=1-p$.
$X\sim B(1,p)$.

⑤ 超几何分布: $X\sim H(n,M,N)$.

$$P(X=k)=\frac{C_M^k\cdot C_{N-M}^{n-k}}{C_N^n}.$$

$\Rightarrow n$一, $\lim \frac{M}{N}=p\in(0,1)$ (即M,N↑, n小.)

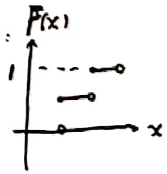
$\Rightarrow P(X=k)=\frac{C_M^k\cdot C_{N-M}^{n-k}}{C_N^n}=C_n^k p^k(1-p)^{n-k}$.

⑥ 泊松分布 (Poisson) $X\sim P(\lambda)$.

$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$. $(\lambda>0, k\in\mathbb{N}^*)$

$n\ge 100$, $p\le 0.01$, $np\le 20$. 用二项分布的近似.

⑦ 分布函数:



⑧ $Y=g(X)$. 算对应的Y即可.

③ 零测集. 奇异型.

$P(X=x_i)=0$ F(X)连续(不绝对连续)

eg. Cantor f.

多维随机变量 $\vec{X}=(X_1,\cdots,X_n)$.

$P(X=x_i, Y=y_j)=P_{ij}$. $(i,j\in\mathbb{N}^*)$.

(1) $\boxed{\text{离散}}$

X \ Y	y_1	···	y_n
x_1			
⋮	··	···	···
x_n			
		···	

边缘分布 (行/列和)

⇑⇓

联合分布.

矩形面积差 $(x_1,x_2]\times(y_1,y_2]\to F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)$.

(2) $\boxed{\text{连续}}$: $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(s,t)\,ds\,dt$ 联合 $f(x,y)$.

边缘: $F_X(x)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(s,t)ds\right]dt$ ($F_Y(y)$同理)

(3) $\boxed{\text{条件分布}}$: $P(X=x_i\,|\,Y=y_j)=\frac{P_{ij}}{P_j}$. (对X同理)

$F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{f(s,y)}{f_Y(y)}ds$.

(4) $\boxed{\text{独立性}}$: $f(x,y)=f(x)\cdot f(y)$.

数字特征

(1) Chebyshev 不等式：

$P(|X-E(X)|\geq\varepsilon)\leq\frac{D(X)}{\varepsilon^2}$.

同理 $P(|X-E(X)|<\varepsilon)\geq 1-\frac{D(X)}{\varepsilon^2}$.

(2) Chebyshev's law：相互独立. $D(X_n)<M$.

$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<\varepsilon\right)=1$.

(3) 辛钦大数定律：$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\mu$. 逼近期望.

(4) Bernoulli 大⋯：$\frac{f_n(A)}{n}\xrightarrow{P}p$. f 近似 p.

(5) 小概率原理.

(6) 强大数定律：$\lim_{n\to\infty}P\left(\bigcup_{k=n}^{\infty}(|\xi_k-\xi|\geq\varepsilon)\right)=0$.

Borel：$\frac{f_n(A)}{n}\xrightarrow{a.s.}p$.

方差、标准差（集中程度）

① 离差：$X-E(X)$.

② $D(X)/Var(X)=E[X-E(X)]^2$

$=E(X^2)-[E(X)]^2$.

③ $\sigma(X)=\sqrt{D(X)}$.

④ Thm：$D(X_1+X_2)=D(X_1)+D(X_2)$.（推广n）

相互独立：$D(X_1X_2)=D(X_1)D(X_2)$.

$D(c)=0$. $D(kX)=k^2D(X)$.

数学期望 $E(X)$

离散：$\sum_k x_kp_k$. 条件：绝对收敛.

连续：$\int_{-\infty}^{+\infty}xf(x)dx$ 定绝对收敛.

（注意：$f(x)=\frac{1}{\pi(1+x^2)}$ 不存在.）

二维
- 离散：$\sum_i\sum_j x_ip_{ij}$.
- 连续：~~⋯~~ $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy$.

$X\to F(X)$. $Y=g(X)$. 离散同理.

$E(Y)=E[g(X)]=\int_{\mathbb{R}}g(x)dF(x)$.

$=\int_{\mathbb{R}}g(x)\underset{(f(x))}{F(x)}dx$.

Thm：$E(C)=C$;（可推广）

$E(kX)=kE(X)$; $E(X_1+X_2)=E(X_1)+E(X_2)$.

若相互独立：$E(X_1X_2)=E(X_1)E(X_2)$.

原点矩、中心矩

(raw moments) (central moments)

raw：$\nu_k=E(X^k)$. k阶原点矩.

central：$\mu_k=E[(X-E(X))^k]$ k阶中心矩.

相关程度：

(1) $cov(X,Y)=E(X-E(X))(Y-E(Y))$.

① $cov(X,X)=D(X)$.

② $cov(X,Y)=cov(Y,X)$.

③ $cov(aX+b,cY+d)=ac\cdot cov(X,Y)$.

④ $cov(X_1+X_2,Y)=cov(X_1,Y)+cov(X_2,Y)$.

⑤ $cov(X,Y)=E(XY)-E(X)\cdot E(Y)$.

⑥ 独立(X,Y)：$cov(X,Y)=0$. 反之,未必.

⑦ $D(X\pm Y)=D(X)+D(Y)\pm 2cov(X,Y)$.

⑧ C-S 不等式：$[cov(X,Y)]^2\leq D(X)\cdot D(Y)$.

（等号成立：$Y=aX+b$.）

(2) $R(X,Y)=\rho_{XY}=corr(X,Y)=\frac{cov(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}$.

- $R(X,Y)=0$. 线性无关
- $R(X,Y)\in[0,1]$. 正相关.
- $R(X,Y)\in[-1,0)$. 负相关.

⇓ 协方差矩阵.（半正定.）

CLT（pro+f 用到特征函数.）依概率均匀地小.

① Lindeberg 定理：$\lim\limits_{n\to\infty} P(Z_n \le x) = \Phi(x)$.（满足 Lindeberg 条件）

② 同分布：列维定理：（μ. σ^2 同.）$n\to\infty$，逐点收敛于 $N(0,1)$.

$$\lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\right) = \Phi(x).$$

③ De Moivre – Laplace 定理：A发生概率 p. 前 n 次 A 发生 Y_n 次.

$$\lim_{n\to\infty} P\left(\frac{Y_n - np}{\sqrt{npq}} < x\right) = \Phi(x).\quad (q = 1-p.) \Rightarrow X \sim N(np, npq).$$

线性函数：$X \sim N(\mu, \sigma^2)$

$\Rightarrow Y = aX + b \sim N(a\mu+b, a^2\sigma^2)$

$(a \ne 0)$

$\Rightarrow X \sim N(\mu, \sigma^2)$，则

$\frac{X-\mu}{\sigma} \overset{\text{标}}{\sim} N(0,1)$ 标准化.

线性组合. X、Y独立且

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

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

$\Rightarrow X_1, \cdots, X_n$ 相互独立且 ⋯

则 $\sum_{k=1}^{n} C_k X_k \sim N\left(\sum_{k=1}^{n} C_k\mu_k, \sum_{k=1}^{n} C_k^2\sigma_k^2\right)$.

正态分布：（Gauss 分布）

$X \sim N(\mu, \sigma^2)$.

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

$X \sim N(0,1)$ 标准⋯

$\varphi(x)$ / $\varphi_{0,1}(x)$.

$\Downarrow$

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

$\Phi(x)$. 标准.

性质：① $\Phi(-x) = 1 - \Phi(x)$.

② $P(-x < X < x) = 2\Phi(x) - 1$.

③ ~~[illegible]~~ $\Phi(0) = 0.5$.

④ 3σ 原则作估算.（<3‰）

二维正态分布：$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\cdot\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$

（显然 $|\rho| < 1$.）

$\Downarrow$ n 维正态分布

$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$，$\vec{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$，$B$ n 阶正定矩阵.

$$\varphi(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |B|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}[(\vec{x}-\vec{\mu})^T B^{-1} (\vec{x}-\vec{\mu})]}$$

$\vec{X} \sim N(\vec{\mu}, B)$.

数字特征：$X \sim N(\mu, \sigma^2)$

$E(X) = \mu$，$D(X) = \sigma^2$.

X 的 k 阶中心矩：$E[(X-E(X))^k] = \mu_k(X)$.

$$= \begin{cases} 0, & k = \text{奇}. \\ (k-1)!!\,\sigma^k, & k = \text{偶}. \end{cases}$$

二维数字特征：

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

$R(X,Y) = \rho$.

$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，X、Y 相互独立

$\Longleftrightarrow R(X,Y) = \rho = 0$.

---

n 维：$\vec{X} \sim N(\vec{\mu}, B)$.　$E(\vec{X}) = \vec{\mu}$.

$\mathrm{cov}\,\vec{X} = B$.

（二维 $B = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$.）