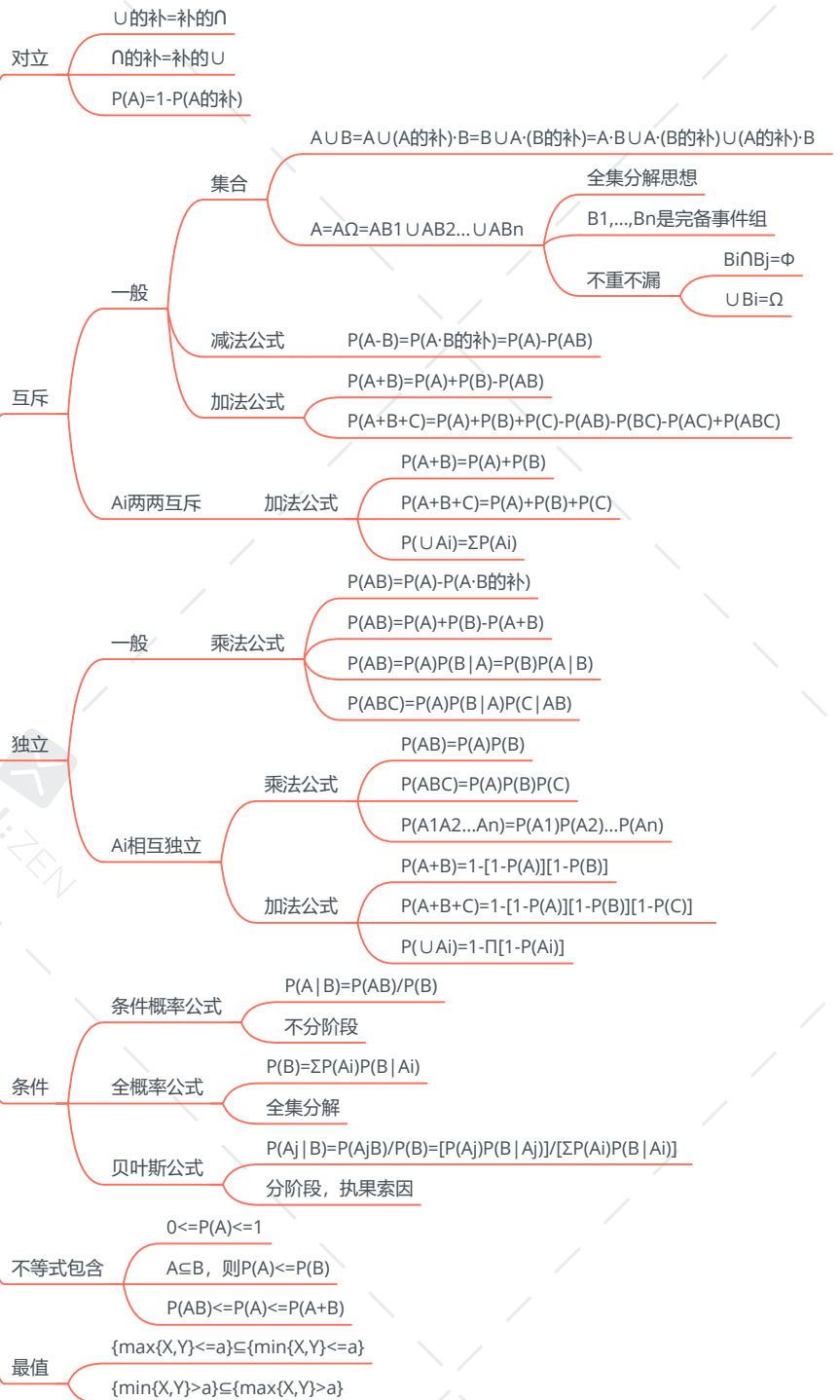


概率第1讲：随机事件与概率

重要公式求概率



古典概型求概率

几何概型求概率

事件独立性

概率第1讲：随机事件与概率

重要公式求概率

古典概型求概率

古典概型的概念

条件 有限个、等可能的样本点

公式 $P = A \text{ 中样本点的个数} / \Omega \text{ 中样本点的个数}$

随机分配问题（随机占位）

核心 放

分配方式 每盒容纳任意多个质点 N^n
每盒容纳至多一个质点 排列 $P(N, n)$

简单随机抽样

核心 取

抽样方式 先后有放回取 n 个 N^n

先后无放回取 n 个 排列 $P(N, n)$

任取 n 个 组合 $C(N, n)$

备注 $P(\text{先后无放回取}) = P(\text{任取})$

抓阄模型 依概率摸球
和第几次无关

几何概型求概率

落入区域A的可能性大小

与A的几何度量成正比

与A的位置、形状无关

公式 $P = A \text{ 的度量} / \Omega \text{ 的度量}$

长度之比

面积之比

事件独立性

定义

两个事件

$P(AB) = P(A)P(B)$

三个事件

$P(AB) = P(A)P(B)$

两两独立 $P(BC) = P(B)P(C)$

$P(AC) = P(A)P(C)$

相互独立 两两独立，且

$P(ABC) = P(A)P(B)P(C)$

判定

$\Leftrightarrow A$ 与 $(B \text{ 的补})$ 独立

$\Leftrightarrow (A \text{ 的补})$ 与 B 独立

$\Leftrightarrow (A \text{ 的补})$ 与 $(B \text{ 的补})$ 独立

独立事件组，不同事件做四则运算，结果仍独立

若 $P(A) > 0$ ，则 A 与 B 独立 $\Leftrightarrow P(B|A) = P(B)$

若 $0 < P(A) < 1$ ，则 A 与 B 独立 $\Leftrightarrow P(B|A) = P(B|(A \text{ 的补}))$

$\Leftrightarrow P(B|A) + P((B \text{ 的补})|(A \text{ 的补})) = 1$

若 $P(A) = 0$ 或 1 ，则 A 与任意事件 B 独立 概率关系

若 $A = \Phi$ 或 Ω ，则 A 与任意事件 B 独立 集合关系

若 $0 < P(A) < 1$ ， $0 < P(B) < 1$ ， A 与 B 互斥或包含：集合关系

互斥或包含：集合关系
独立性：概率关系

概率第2讲：一维 随机变量及其分布

求分布

60

用分布

14

求函数分布

17

判分布

随机变量

$$X=X(\omega), \omega \in \Omega$$

统一化, 数量化

单调不减

分布函数 $F(x)$

右连续

$$F(x_i)=F(x_i+0)$$

等号跟着大于号

有界性

$$F(-\infty)=0$$

$$F(+\infty)=1$$

概率分布 p_i

非负性

$$p_i \geq 0$$

归一性

$$\sum p_i = 1$$

概率密度 $f(x)$

非负性

$$f(x) \geq 0$$

归一性

$$\int f(x) dx = 1$$

$F(x)$ 与 $f(x)$

$$F'(x)=f(x)$$

$$dF(x)=f(x)dx$$

反问题: 建方程, 求参数

$$F(-\infty)=0$$

$$F(+\infty)=1$$

$$\sum p_i = 1$$

$$\int f(x) dx = 1$$

结论

若 $F_1(x), F_2(x)$ 是分布函数,
 $a > 0, b > 0, a+b=1$, 则

$F(x)=aF_1(x)+bF_2(x)$ 是分布函数

$$F(-a)=1-F(a)=0.5-\int_{[0,a]} f(x) dx$$

$$F(-a)+F(a)=1$$

$$P\{|X| < a\} = 2F(a)-1$$

$$P\{|X| > a\} = 2[1-F(a)]$$

若 $f(x)$ 是偶函数, 则

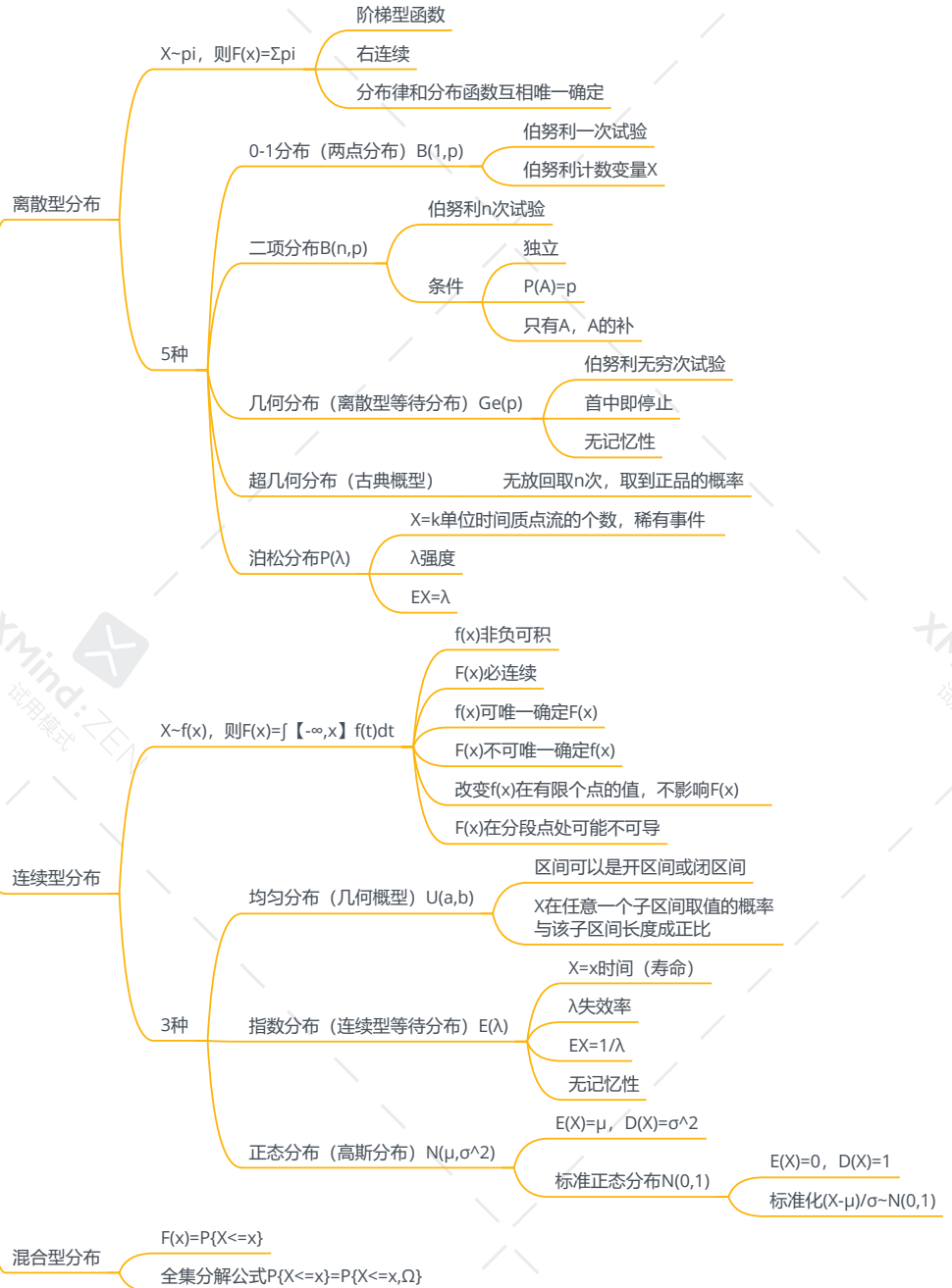
概率第2讲：一维随机变量及其分布

判分布^①

求分布

用分布^④

求函数分布^⑦



概率第2讲：一维随机变量及其分布

判分布⁽⁵⁷⁾

求分布⁽⁵⁰⁾

用分布

- $X \sim F(x)$
 - $P\{X \leq a\} = F(a)$ 带等号取函数值
 - $P\{X < a\} = F(a-0)$ 不带等号取左极限
- $X \sim \pi_i$
 - $P\{X \in I\} = \sum P\{X = x_i\}$
- $X \sim f(x)$
 - $P\{X \in I\} = \int f(x) dx$
- 反问题：已知概率，求参数
 - 算积分
 - 看模型
 - 用图形

8个分布

求函数分布

离散型→离散型

连续型→连续型（或混合型）

连续型→离散型

分布函数法

公式法

先确定Y的可能取值 a_i

再计算概率 $P(Y=a_i)$

写出Y的概率分布

$$F(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = \int [g(x) \leq y] f(x) dx$$

$$g(X) \leq y: \text{曲线 } Y=g(X) \text{ 在直线 } Y=y \text{ 下方}$$

画图，分段讨论

$$f(y) = F'(y)$$

条件：y=g(x)严格单调，且可导

$$y=g(x) \text{ 的可导反函数是 } x=h(y)$$

$$f(y) = f[h(y)] \cdot |h'(y)|, \alpha < y < \beta$$

$$\alpha = \min\{g(a), g(b)\}$$

$$\beta = \max\{g(a), g(b)\}$$

概率第3讲：多维随机变量及其分布

判分布

- 分布函数 $F(x,y)$
 - 单调不减
 - 右连续
 - $F(x_i+0,y)=F(x_i,y)$
 - $F(x,y_i+0)=F(x,y_i)$
 - 等号跟着大于号
 - 有界性
 - $F(-\infty,-\infty)=F(x,-\infty)=F(-\infty,y)=0$
 - $F(+\infty,+\infty)=1$
 - 非负性
 - $p\{x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2\} \geq 0$
- 概率分布 p_{ij}
 - 非负性
 - $p_{ij} \geq 0$
 - 归一性
 - $\sum \sum p_{ij} = 1$
- 概率密度 $f(x,y)$
 - 非负性
 - $f(x,y) \geq 0$
 - 归一性
 - $\iint f(x,y) dx dy = 1$
- 反问题：建方程，求参数
 - $F(-\infty,-\infty)=F(x,-\infty)=F(-\infty,y)=0$
 - $F(+\infty,+\infty)=1$
 - $\sum \sum p_{ij} = 1$
 - $\iint f(x,y) dx dy = 1$

求分布

用分布

- $(X,Y) \sim p_{ij}$ $P\{(X,Y) \in D\} = \sum \{p_{ij} | (x_i, y_j) \in D\}$
- $(X,Y) \sim f(x,y)$ $P\{(X,Y) \in D\} = \iint_D f(x,y) dx dy$
- $(X,Y) \sim$ 混合型 全集分解思想，全概率公式
- 反问题：已知概率，求参数

求函数分布

- 多维 \rightarrow 一维
 - (离散型, 离散型) \rightarrow 离散型
 - $(X,Y) \sim p_{ij}, Z = g(X,Y) \sim q_{ij}$
 - $X \sim p_i, Y \sim q_j, X$ 与 Y 独立且取值在某个范围
 - $X+Y$
 - XY
 - $\max\{X,Y\}$
 - (连续型, 连续型) \rightarrow 连续型
 - 分布函数法
 - 卷积公式法
 - $X+Y$
 - $X-Y$
 - XY
 - X/Y
 - 最值函数
 - $\max\{X,Y\}$
 - $\min\{X,Y\}$
 - 独立
 - 独立同分布
 - $F_{\max} = [F(x)]^n$
 - $F_{\min} = 1 - [1 - F(x)]^n$
 - (离散型, 连续型) \rightarrow 连续型
 - X 与 Y 独立 分布函数法+全集分解
 - X 与 Y 不独立 分布函数法
- 一维 \rightarrow 多维
 - 离散型 \rightarrow (离散型, 离散型)
 - 连续型 \rightarrow (离散型, 离散型)
- 多维 \rightarrow 多维
 - (离散型, 离散型) \rightarrow (离散型, 离散型)
 - (离散型, 连续型) \rightarrow (离散型, 离散型)
 - (连续型, 连续型) \rightarrow (离散型, 离散型)

概率第3讲：多维随机变量及其分布

判分布

求联合分布

求 $F(x,y)$

$(X,Y) \sim pij$, 则 $F(x,y) = P\{X \leq x, Y \leq y\} = \sum_{xi \leq x, yj \leq y} pij$

$(X,Y) \sim f(x,y)$, 则 $F(x,y) = P\{X \leq x, Y \leq y\} = \int_{-\infty, x} \int_{-\infty, y} f(u,v) du dv$

求 p_{ij}

求 $f(x,y)$

二维均匀分布

$f(x,y) = 1/S, (x,y) \in D$

$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 注意顺序

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

$Y \sim N(\mu_2, \sigma_2^2)$

二维正态分布

$X \sim N(\mu_1, \sigma_1^2)$

$Y \sim N(\mu_2, \sigma_2^2)$

且 X 与 Y 独立

$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, 0)$

$(X,Y) \sim N$, 则 $aX + bY \sim N$ (a, b 不全为0)

$(X,Y) \sim N$, 则 X 与 Y 独立 \Leftrightarrow 不相关 $\rho = 0$

求分布

求边缘分布

求 $F_X(x), F_Y(y)$

$F_X(x) = F(x, +\infty)$

$F_Y(y) = F(+\infty, y)$

求 $p\{i\}, p\{j\}$

$p\{i\} = \sum_j pij$

$p\{j\} = \sum_i pij$

求 $f(x), f(y)$

$f(x) = \int f(x,y) dy$

$f(y) = \int f(x,y) dx$

求谁不积谁, 不积先定限

求条件分布

条件=联合/边缘

联合=边缘 \times 条件

判独立

X, Y 独立

$\Leftrightarrow F(x,y) = F(x) \cdot F(y)$

$\Leftrightarrow p\{ij\} = p\{i\} \cdot p\{j\}$

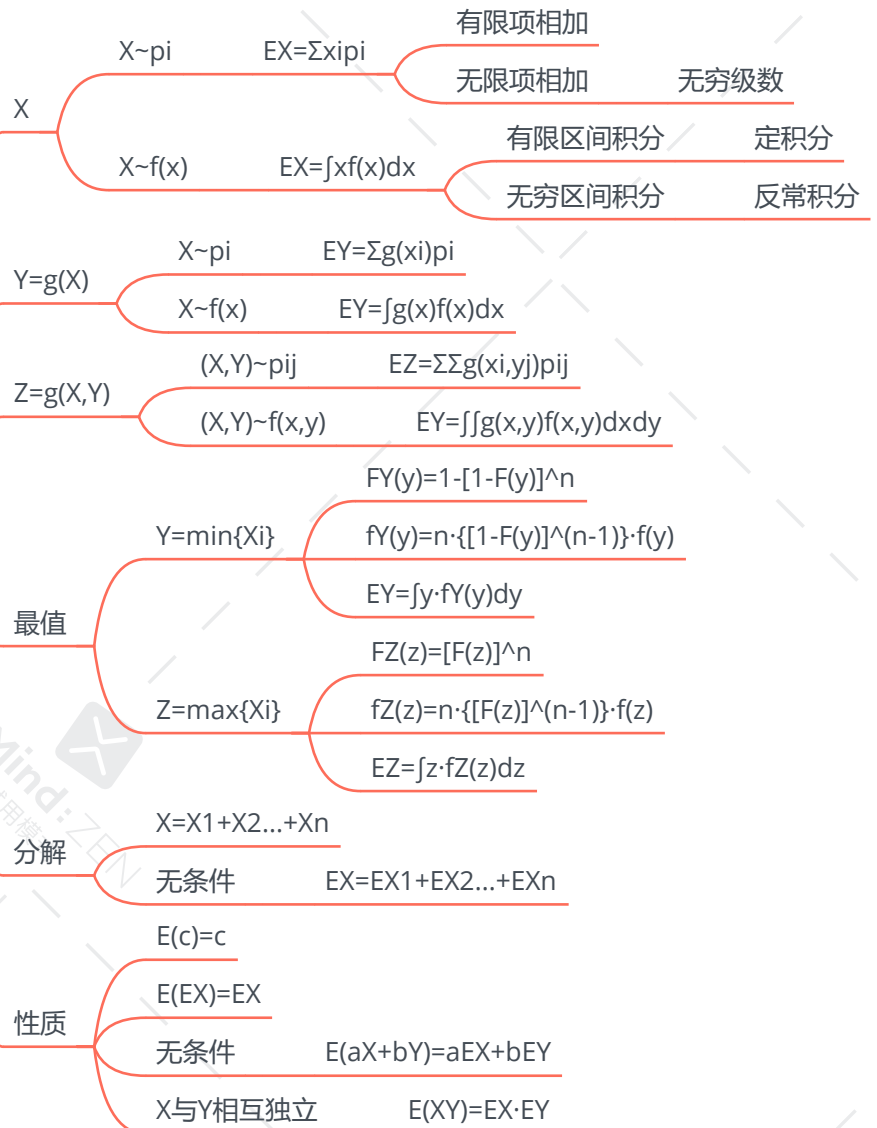
$\Leftrightarrow f(x,y) = f(x) \cdot f(y)$

用分布

求函数分布

概率第4讲： 数字特征

期望



方差

常用分布的期望和方差

协方差

相关系数

独立性和相关性的判定

切比雪夫不等式

概率第4讲： 数字特征

期望⁽⁴²⁾

定义 $DX = E[(X-EX)^2]$ $DX \geq 0$
 $E[(X-c)^2] \geq DX$

定义法 $X \sim \pi$ $DX = \sum [(x_i - EX)^2] \cdot \pi$
 $X \sim f(x)$ $DX = \int [(x - EX)^2] \cdot f(x) dx$

公式法 $DX = E(X^2) - (EX)^2$
 $E(X^2) = DX + (EX)^2 \geq (EX)^2$

方差

最值

$Y = \min\{X_i\}$ $EY = \int y \cdot f_Y(y) dy$
 $E(Y^2) = \int (y^2) \cdot f_Y(y) dy$
 $DY = E(Y^2) - (EY)^2$

$Z = \max\{X_i\}$ $EZ = \int z \cdot f_Z(z) dz$
 $E(Z^2) = \int (z^2) \cdot f_Z(z) dz$
 $DZ = E(Z^2) - (EZ)^2$

分解

$X = X_1 + X_2 + \dots + X_n$
无条件 $DX = DX_1 + DX_2 + \dots + DX_n + 2 \sum \text{Cov}(X_i, X_j)$
 X_i 与 X_j 相互独立 $DX = DX_1 + DX_2 + \dots + DX_n$

性质

$D(c) = 0$
 $DX = 0 \Leftrightarrow X$ 几乎处处为某常数, $P\{X=a\}=1$
 $D(aX+b) = (a^2) \cdot DX$
无条件 $D(X \pm Y) = DX + DY \pm 2 \text{Cov}(X, Y)$
 X 与 Y 相互独立 $D(aX \pm bY) = DX + DY$
 $D(XY) = DX \cdot DY + DX \cdot (EY)^2 + DY \cdot (EX)^2 \geq DX \cdot DY$

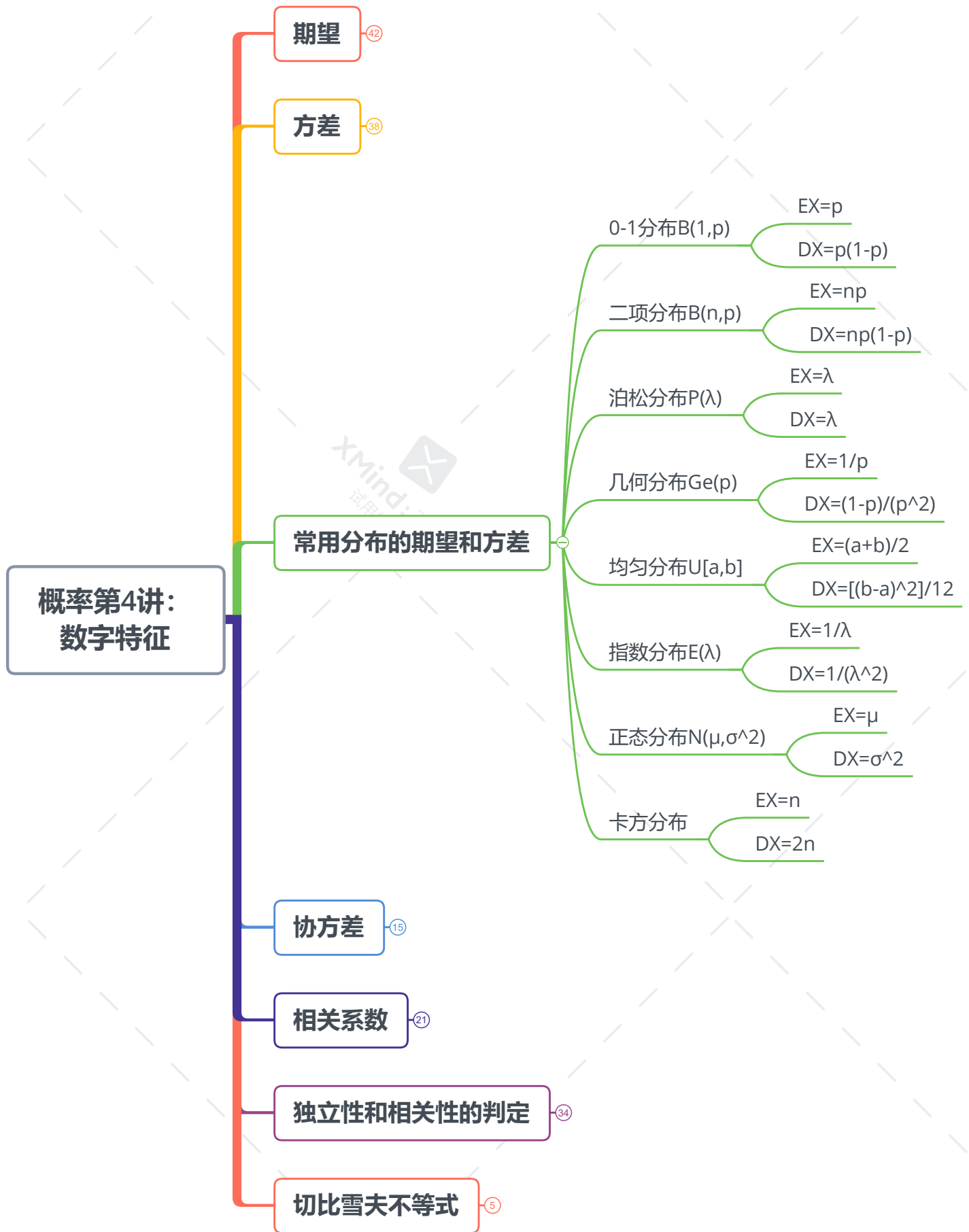
常用分布的期望和方差⁽²⁴⁾

协方差⁽¹⁵⁾

相关系数⁽²¹⁾

独立性和相关性的判定⁽³⁴⁾

切比雪夫不等式⁽⁵⁾



概率第4讲： 数字特征

期望⁽⁴²⁾

方差⁽³⁸⁾

常用分布的期望和方差⁽²⁴⁾

协方差

定义

$$\text{Cov}(X, Y) = E[(X - EX)(Y - EY)]$$

$$\text{Cov}(X, X) = DX$$

定义法

$$(X, Y) \sim p_{ij} \quad \text{Cov}(X, Y) = \sum \sum [(x_i - EX)(y_j - EY)] \cdot p_{ij}$$

$$(X, Y) \sim f(x, y) \quad EY = \iint [(x - EX)(y - EY)] \cdot f(x, y) dx dy$$

公式法

$$\text{Cov}(X, Y) = E(XY) - EX \cdot EY$$

性质

$$\text{Cov}(X, Y) = \text{Cov}(Y, X)$$

$$\text{Cov}(aX, bY) = ab \cdot \text{Cov}(X, Y)$$

$$\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y) \quad \text{单个可拆}$$

相关系数

定义

$$\rho\{XY\} = \text{Cov}(X, Y) / [(\sqrt{DX}) \cdot (\sqrt{DY})]$$

线性相依性

量纲为1，无单位

$$\rho = 0 \quad \Leftrightarrow X \text{与} Y \text{不相关}$$

$$\rho \neq 0 \quad \Leftrightarrow X \text{与} Y \text{相关}$$

$$-1 \leq \rho\{XY\} \leq 1$$

性质

$$\rho\{XY\} = 1 \quad \Leftrightarrow P\{Y = aX + b\} = 1, a > 0$$

$$\Leftrightarrow a > 0, Y = aX + b \text{几乎处处成立}$$

$$\rho\{XY\} = -1 \quad \Leftrightarrow P\{Y = aX + b\} = 1, a < 0$$

$$\Leftrightarrow a < 0, Y = aX + b \text{几乎处处成立}$$

变体

$$Y = aX + b, a > 0 \quad \Rightarrow \rho\{XY\} = 1$$

$$Y = aX + b, a < 0 \quad \Rightarrow \rho\{XY\} = -1$$

独立性和相关性的判定⁽³⁴⁾

切比雪夫不等式⁽⁵⁾

概率第4讲：
数字特征

期望⁽⁴²⁾

方差⁽³⁸⁾

常用分布的期望和方差⁽²⁴⁾

协方差⁽¹⁵⁾

相关系数⁽²¹⁾

独立性和相关性的判定

用分布判独立性

$$F(x,y)=F(x) \cdot F(y)$$

$$p_{ij}=\{p_{i.}\} \cdot \{p_{.j}\}$$

$$f(x,y)=f(x) \cdot f(y)$$

用数字特征判相关性

$$\rho=0$$

$$\Leftrightarrow \text{Cov}(X,Y)=0$$

$$\Leftrightarrow E(XY)=E X \cdot E Y$$

$$\Leftrightarrow D(X+Y)=D X+D Y$$

$$\Leftrightarrow D(X-Y)=D X+D Y$$

先计算 $\text{Cov}(X,Y)=E(XY)-E X \cdot E Y$

程序

再判断相关性

$$\text{Cov}(X,Y) \neq 0$$

$$\Leftrightarrow X \text{与} Y \text{相关}$$

$$\text{Cov}(X,Y)=0$$

$$\Leftrightarrow X \text{与} Y \text{不相关}$$

最后判断独立性

$$X \text{与} Y \text{相关}$$

$$\Rightarrow X \text{与} Y \text{不独立}$$

$$X \text{与} Y \text{不相关}$$

用分布

反证法

重要结论

单向

$$X \text{与} Y \text{独立}$$

$$\Rightarrow X \text{与} Y \text{不相关} \rho=0$$

$$X \text{与} Y \text{相关} \rho \neq 0$$

$$\Rightarrow X \text{与} Y \text{不独立}$$

双向

$$\text{若} (X,Y) \sim N$$

$$X \text{与} Y \text{独立} \Leftrightarrow X \text{与} Y \text{不相关} \rho=0$$

$$\text{若} X \sim B(1,p_1), Y \sim B(1,p_2)$$

$$X \text{与} Y \text{独立} \Leftrightarrow X \text{与} Y \text{不相关} \rho=0$$

切比雪夫不等式

条件

$E(X)$ 和 $D(X)$ 存在且有限

公式

$$P\{|X-E X| \geq \varepsilon\} \leq [D X / (\varepsilon^2)]$$

$$P\{|X-E X| < \varepsilon\} \geq 1 - [D X / (\varepsilon^2)]$$

概率第5讲：大数定律和中心极限定理

依概率收敛

定义

$$n \rightarrow \infty, \lim P\{|X_n - X| < \varepsilon\} = 1$$

应用

$n \rightarrow \infty$, 随机变量序列 $\{X_n\}$ 依概率收敛于随机变量 X

讨论参数估计量的一致性（相合性）

大数定律

伯努利大数定律

条件

n 重伯努利试验

事件A发生的次数 μ

每次试验中事件A发生的概率 p

结论

μ/n 依概率收敛于 p

切比雪夫大数定律

条件

X_i 相互独立

方差存在且一致有上界 $DX_i \leq C$

结论

$(\sum X_i)/n$ 依概率收敛于 $E[(\sum X_i)/n]$

辛钦大数定律

条件

X_i 相互独立

X_i 同分布

期望存在 $EX_i = \mu$

结论

$(\sum X_i)/n$ 依概率收敛于 μ

中心极限定理

列维-林德伯格定理

条件

X_i 相互独立

X_i 同分布（无论什么分布）

期望存在 $EX_i = \mu$

方差存在 $DX_i = \sigma^2$

n 很大

结论

$(\sum X_i)$ 近似服从 $N(n \cdot \mu, n \cdot \sigma^2)$

棣莫弗-拉普拉斯定理

条件

X_i 相互独立

$X_i \sim B(1, p)$

$(\sum X_i) \sim B(n, p)$

n 很大

结论

$(\sum X_i)$ 近似服从 $N(n \cdot p, n \cdot p \cdot (1-p))$

二项分布的概率计算

n 不太大 ($n \leq 10$)

直接计算

n 较大, p 较小 ($n > 100, np < 10$)

用泊松分布近似

n 较大, p 不太大

用正态分布近似

概率第6讲：数理统计初步

点估计⁽³⁾

区间估计[数一]⁽²⁾

假设检验[数一]⁽⁴⁾

统计量

统计量

样本均值
样本方差
样本标准差
样本k阶原点矩
样本k阶中心矩
顺序统计量

四大分布

正态分布

$N(\mu, \sigma^2)$

$N(0, 1)$

$\varphi(x)$ 偶函数

$\Phi(-x) = 1 - \Phi(x)$

卡方分布

$EX = n$

$DX = 2n$

$X_1 \sim \chi^2(n_1), X_2 \sim \chi^2(n_2)$,
且 X_1 与 X_2 独立

则 $X_1 + X_2 \sim \chi^2(n_1 + n_2)$

t分布

$Et = 0 (n \geq 2)$

$t(n)\{1-\alpha\text{上分位数}\} = -t(n)\{\alpha\text{上分位数}\}$

F分布

$F(n_1, n_2)$

$F \sim F(n_1, n_2)$, 则 $1/F \sim F(n_2, n_1)$

$F(n_1, n_2)\{1-\alpha\text{上分位数}\} = 1/F(n_2, n_1)\{\alpha\text{上分位数}\}$

正态总体 $N(\mu, \sigma^2)$ 下的常用结论

总体均值 μ 已知 $\{\sum[(X_i - \mu)^2]\}/[\sigma^2] \sim \chi^2(n)$

总体均值 μ 未知, 用 \bar{X} 替代 μ $[(n-1)(S^2)]/[\sigma^2] \sim \chi^2(n-1)$

$\{\sum[(X_i - \bar{X})^2]\}/[\sigma^2] \sim \chi^2(n-1)$

总体方差 σ^2 已知 $\bar{X} \text{ 拔} \sim N(\mu, (\sigma^2)/n)$

$(\sqrt{n})(\bar{X} \text{ 拔} - \mu)/\sigma \sim N(0, 1)$

总体方差 σ^2 未知, 用 S 替代 σ $(\sqrt{n})(\bar{X} \text{ 拔} - \mu)/S \sim t(n-1)$

$[n \cdot (\bar{X} \text{ 拔} - \mu)^2]/[S^2] \sim F(1, n-1)$

样本均值 \bar{X} 拔和样本方差 S^2 相互独立

概率第6讲：数理统计初步

统计量

点估计

矩估计

一个参数

用一阶

求样本均值 \bar{X} 拔

令 $\bar{X}拔 = EX$

若一阶不能用，则用二阶

求 $[\sum(X_i^2)]/n$

令 $[\sum(X_i^2)]/n = E(X^2)$

两个参数

用一阶

求样本均值 \bar{X} 拔

令 $\bar{X}拔 = EX$

同时用二阶

求 $[\sum(X_i^2)]/n$

令 $[\sum(X_i^2)]/n = E(X^2)$

最大似然估计

写似然函数

求参数

有驻点

令导函数=0

无驻点（单调）

用定义

不变性

估计量的评价

无偏性

$E(\theta\text{估计}) = \theta$ ，则 θ 估计是 θ 的无偏估计量

有效性

$E(\theta\text{估计}1) = E(\theta\text{估计}2) = \theta$

$D(\theta\text{估计}1) < D(\theta\text{估计}2)$

则 θ 估计1比 θ 估计2更有效

一致性（相合性）

只针对大样本 $n \rightarrow \infty$

θ 估计依概率收敛于 θ

区间估计[数一]

思路方法

结论

假设检验[数一]

显著性水平

正态总体下的六大检验

两类错误

第一类错误

弃真

$\alpha = P\{\text{拒绝}H_0 | H_0\text{为真}\}$

第二类错误

取伪

$\beta = P\{\text{接受}H_0 | H_0\text{为假}\}$