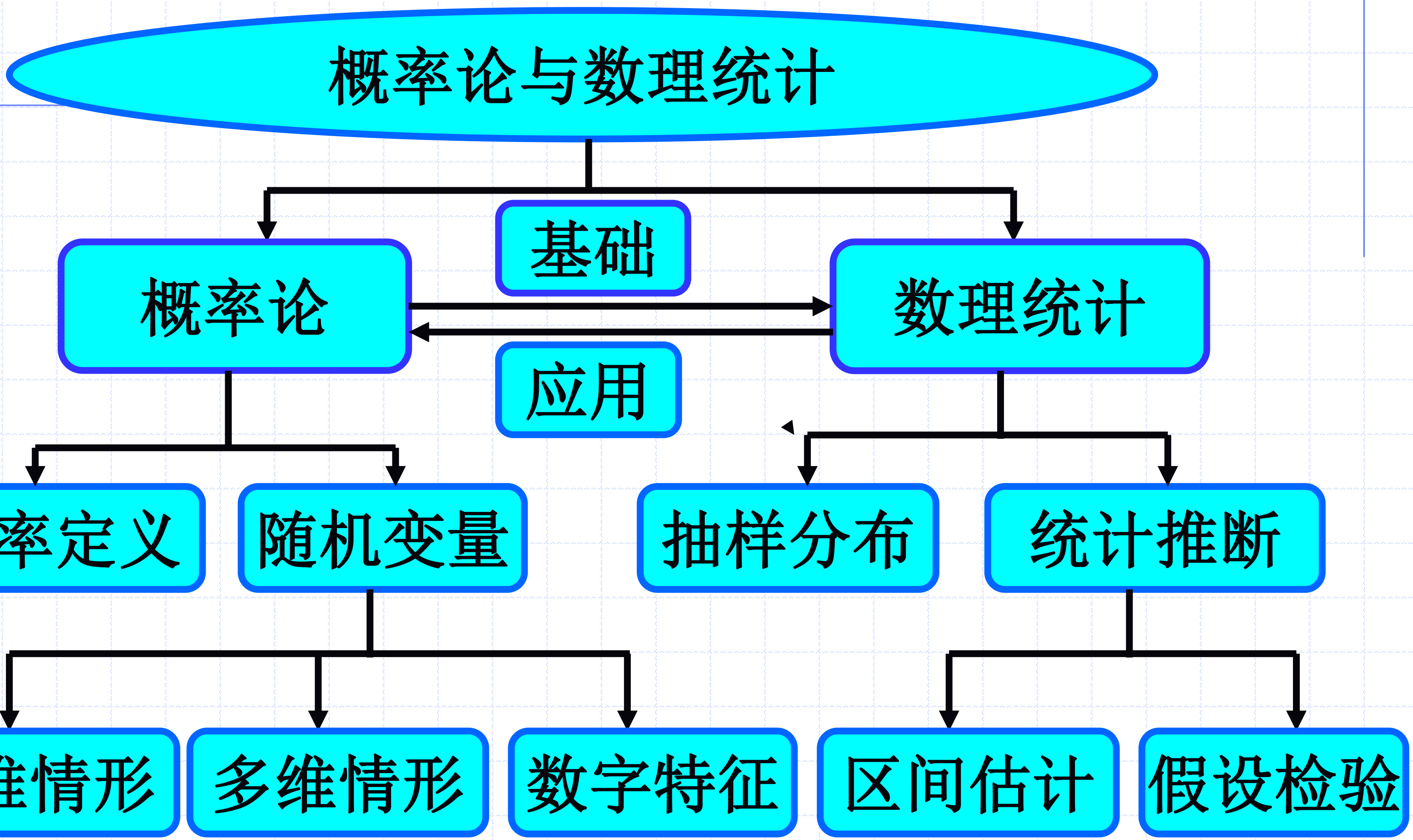
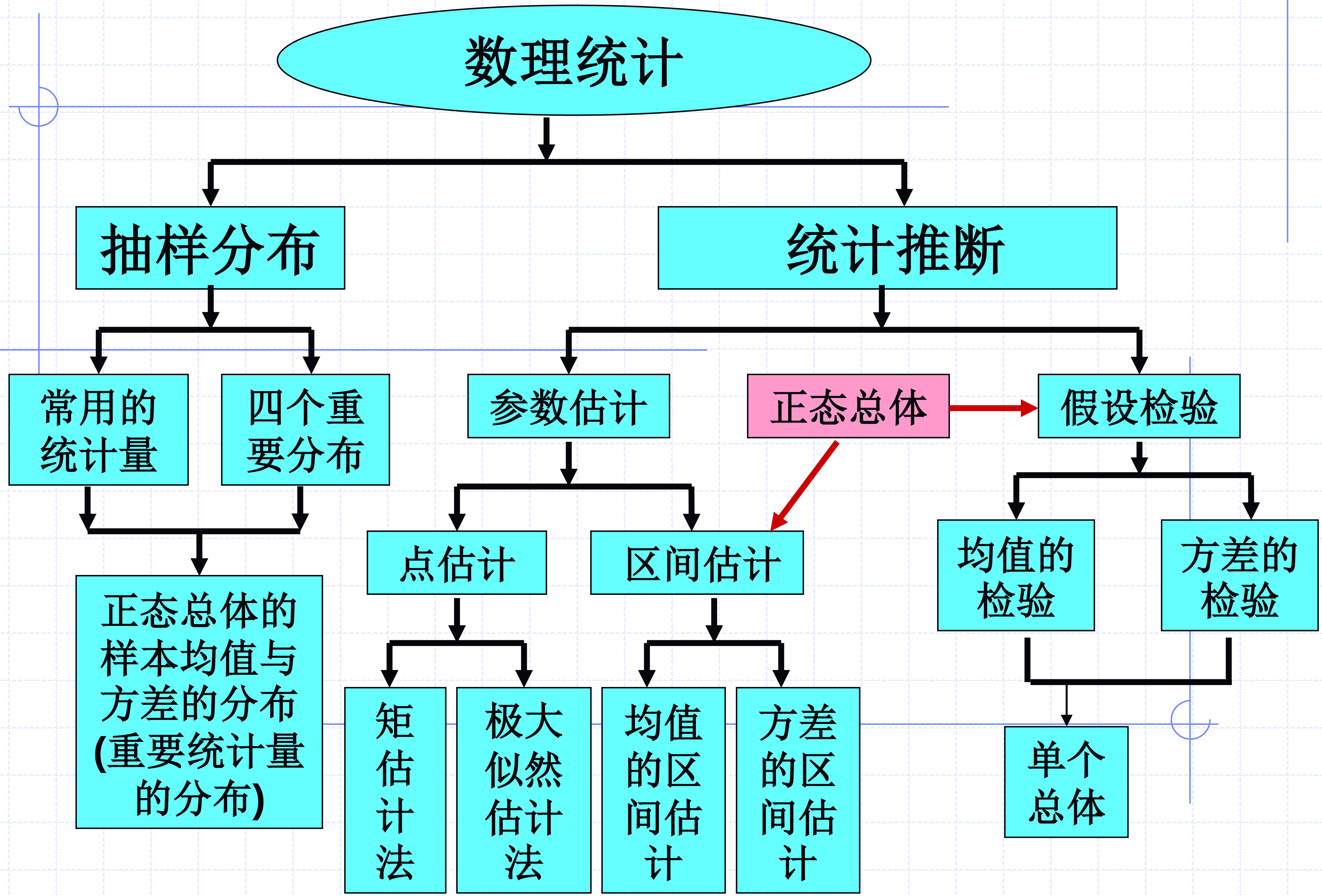
课程总知识结构图



概率论与数理统计
基础
概率论
数理统计
应用
概率定义
随机变量
抽样分布
统计推断
一维情形
多维情形
数字特征
区间估计
假设检验

第六章----第八章知识结构图



一 常用的统计量

名称	统计量	观察值
样本均值	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
样本方差	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$
样本标准差	$S = \sqrt{S^2}$	$s = \sqrt{s^2}$
样本 k 阶原点矩	$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$	$a_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$
样本 k 阶中心矩	$B_k = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^k$	$b_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^k$

二、 常用统计量的分布

1) χ^2 – 分布

设$(X_1,\cdots X_n)$为来自于正态总体$N(0,1)$的样本，

则称统计量： $\chi^2 = X_1^2 + \cdots + X_n^2$

记为 $\chi^2 \sim \chi^2(n)$

2) $X \sim N(0,1), Y \sim \chi^2(n), X,Y$独立，则称随机变量

$$t = \frac{X}{\sqrt{Y/n}} \qquad t \sim t(n).$$

3) F – 分布 若 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，$X,Y$ 独立，

$$F = \frac{X/n_1}{Y/n_2} \qquad F \sim F(n_1,n_2).$$

返回主目录

4） 正态总体的样本均值与样本方差的分布：

设$X_1,\cdots,X_n$是总体$N(\mu,\sigma^2)$的样本，$\bar{X},S^2$分别是样本均值与样本方差，则有：

(1) $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$; $\quad \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$,

(2) $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$;

(3) $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$

练习 设 $X_1, X_2, \cdots, X_{11}$ 为总体 $X \sim N(\mu, \sigma^2)$ 的一样本，

S^2 是样本方差，求 $P\left\{\dfrac{S^2}{\sigma^2} \le 1.61\right\}$，$D(S^2)$

解: $$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{10S^2}{\sigma^2} \sim \chi^2(10)$$

$$P\left\{\frac{S^2}{\sigma^2} \le 1.61\right\} = P\left\{\frac{10S^2}{\sigma^2} \le 16.1\right\} = P\left\{\chi^2 \le 16.1\right\}$$

$$= 1 - P\left\{\chi^2 > 16.1\right\} \triangleq 1 - \alpha$$

即 $P\left\{\chi^2 > 16.1\right\} = \alpha$ 即 $\chi^2_\alpha(10) = 16.1$

查表得 $\alpha \approx 0.10 \longrightarrow P\left\{\dfrac{S^2}{\sigma^2} \le 1.61\right\} \approx 0.90$

练习 设 $X_1, X_2, \cdots, X_{11}$ 为总体 $X \sim N\left(\mu, \sigma^2\right)$ 的一样本，

S^2 是样本方差，求 $P\left\{\frac{S^2}{\sigma^2} \le 1.61\right\}$，$D\left(S^2\right)$

解： $$\frac{10S^2}{\sigma^2} \sim \chi^2(10)$$

$$\therefore D\left(\frac{10S^2}{\sigma^2}\right) = 20$$

$$\rightarrow \left(\frac{10}{\sigma^2}\right)^2 D\left(S^2\right) = 20$$

$$\rightarrow D\left(S^2\right) = \frac{20}{10^2}\sigma^4 = \frac{1}{5}\sigma^4$$

练习

设 $X_1, X_2, \cdots X_9$是来自正态总体X的简单随机样本

$$Y_1 = \frac{1}{6}(X_1 + X_2 + \cdots X_6) \qquad Y_2 = \frac{1}{3}(X_7 + X_8 + \cdots X_9)$$

$$S_2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2 \qquad Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S}$$

证明：统计量 $\boldsymbol{Z} \sim t\,(2)$

- 统计推断
  - 估计问题
    - 参数估计问题
      - 点估计问题
        - 矩估计法
        - 极大似然估计
      - 区间估计问题
    - 非参数估计问题
  - 假设检验问题
    - 参数假设检验问题
      - 一个正态总体的假设检验
    - 非参数假设检验问题

矩估计法 令 $A_k = \mu_k, \quad k = 1, \cdots, l, \cdots$

$$= f_k(\theta_1, \theta_2 \cdots \theta_l),$$

得到包含l个未知参数 $\theta_1, \theta_2 \cdots \theta_l$ 的方程组

$$\begin{cases} A_1 = f_1(\theta_1, \theta_2, \cdots, \theta_l) \\ A_2 = f_2(\theta_1, \theta_2, \cdots, \theta_l) \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ A_k = f_k(\theta_1, \theta_2, \cdots, \theta_l) \end{cases}$$

从中解出 $\hat{\theta}_1, \hat{\theta}_2 \cdots \hat{\theta}_l$

极大似然法求估计量的步骤：（一般情况下）

1) 构造似然函数$L(\theta)$：

$L(\theta)=\prod_{i=1}^{n}P(x_i)$ (离散型)，$L(\theta)=\prod_{i=1}^{n}f(x_i)$ (连续型)；

2) 取对数：$\ln L(\theta)$；

3) 令 $\dfrac{d\ln L}{d\theta}=0$；

4) 解似然方程得θ的极大似然估计量$\hat{\theta}$.

估计量的评选标准

1）、无偏性

若$\hat{\theta}=\hat{\theta}(X_1,\cdots,X_n)$的数学期望存在，且$E\hat{\theta}=\theta$,

则称$\hat{\theta}$是θ的无偏估计量。

2）、有效性

若$\hat{\theta}_1=\hat{\theta}_1(X_1,\cdots,X_n)$，$\hat{\theta}_2=\hat{\theta}_2(X_1,\cdots,X_n)$都是$\theta$的无偏估计量

3）、一致性

当$n\to\infty$时$\hat{\theta}$依概率收敛于θ,

则称$\hat{\theta}$为θ的一致估计量(相合估计量)

1、$X_1,\cdots X_n$ 为总体$X \sim N(\mu,\sigma^2)$ 的样本，试求常数C,

使得 $C\sum_{i=1}^{n-1}(X_{i+1}-X_n)^2$ 为 σ^2 的无偏估计

2、$X_1,\cdots X_n$ 为总体$X \sim N(\mu,\sigma^2)$ 的样本，

$\hat{\sigma}=\frac{1}{k}\sum_{i=1}^{n-1}\left|X_{i+1}-X_i\right|$ 为总体参数 σ 的无偏估计量

求 k

区间估计　$X\sim N(\mu,\sigma^2)$，对 μ,σ^2 进行区间估计

未知参数	统计量	置信区间
μ (σ^2已知)	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$	$\left(\bar{X}\pm\dfrac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$
μ (σ^2未知)	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$	$\left(\bar{X}\pm\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$
σ^2	$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$	$\left(\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\ \dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$
σ	$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$	$\left(\sqrt{\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}},\ \sqrt{\dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}\right)$

假设检验

$X \sim N(\mu,\sigma^2)$, 对 μ,σ^2进行假设检验

$X_1 X_2,\cdots,X_n$
$x_1 x_2,\cdots,x_n$
显著性水平α,

	原假设 H_0	备择假设H_1	检验统计量	★ 拒绝域
1) μ的检验 σ^2为已知	$\mu=\mu_0$	$\mu\neq\mu_0$	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	$\lvert U\rvert>z_{\alpha/2}$
	$\mu\le\mu_0$	$\mu>\mu_0$		$U>z_{\alpha}$
	$\mu\ge\mu_0$	$\mu<\mu_0$	$\sim N(0,1)$	$U<-z_{\alpha}$
2) μ的检验 σ^2为未知	$\mu=\mu_0$	$\mu\neq\mu_0$	$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$	$\lvert t\rvert>t_{\alpha/2}(n-1)$
	$\mu\le\mu_0$	$\mu>\mu_0$		$t>t_{\alpha}(n-1)$
	$\mu\ge\mu_0$	$\mu<\mu_0$	$\sim t(n-1)$	$t<-t_{\alpha}(n-1)$
3) σ^2的检验	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2>\chi^2_{\alpha/2}(n-1)$或 $\chi^2<\chi^2_{1-\alpha/2}(n-1)$
	$\sigma^2\le\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2>\chi^2_{\alpha}(n-1)$
	$\sigma^2\ge\sigma_0^2$	$\sigma^2<\sigma_0^2$	$\sim\chi^2(n-1)$	$\chi^2<\chi^2_{1-\alpha}(n-1)$

1、某批零件，其重量应服从正态分布 $N\left(\mu,\sigma^2\right)$，其中 μ,σ^2未知从中抽取容量为**9**的一个样本，样本值为（单位：公斤）

5.5　5.4　5.8　5.3　5.3　5.6　5.7　5.2　5.7。

（**1**）求零件重量 μ 的置信区间，置信度为**0.95.**

（**2**）是否可以认为这批零件的重量 $\mu = 5.1$?　$\alpha = 0.05$

（**3**）是否可以认为这批零件的重量 $\mu \le 5.1$?　$\alpha = 0.05$

2、设总体X的概率密度函数为$f(x)=(\theta+1)x^{\theta}$，$0<x<1$

其中 $\theta>-1$ 是未知参数

$x_1,x_2,\cdots,x_n$是来自这个总体的一组观测值。求：

（1）求未知参数 θ 的矩估计值

（2）求未知参数 θ 的极大似然估计值

$$EX=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^1 x\cdot(\theta+1)x^{\theta}dx=\frac{\theta+1}{\theta+2}$$

令 $$\bar{X}=\frac{\theta+1}{\theta+2}$$

$$\hat{\theta}=\frac{2\bar{X}-1}{1-\bar{X}}.$$

设总体X密度 $f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$

$\theta>-1$未知， 求 θ 的极大似然估计。

解：

$x_1,x_2,\cdots,x_n$ 为样本值。

$$L(\theta)=\begin{cases}\prod_{i=1}^{n}(\theta+1)x_i^{\theta}=(\theta+1)^n(x_1x_2\cdots x_n)^{\theta}, & 0<x_i<1,\ i=1,2,\cdots,n.\\ 0, & \text{其他}\end{cases}$$

$0<x_i<1,\ i=1,2,\cdots,n$ 时，

$$\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i$$

设总体X 密度 $f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$

$\theta>-1$未知， 求 θ 的极大似然估计。

$$\ln L(\theta)= n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i$$

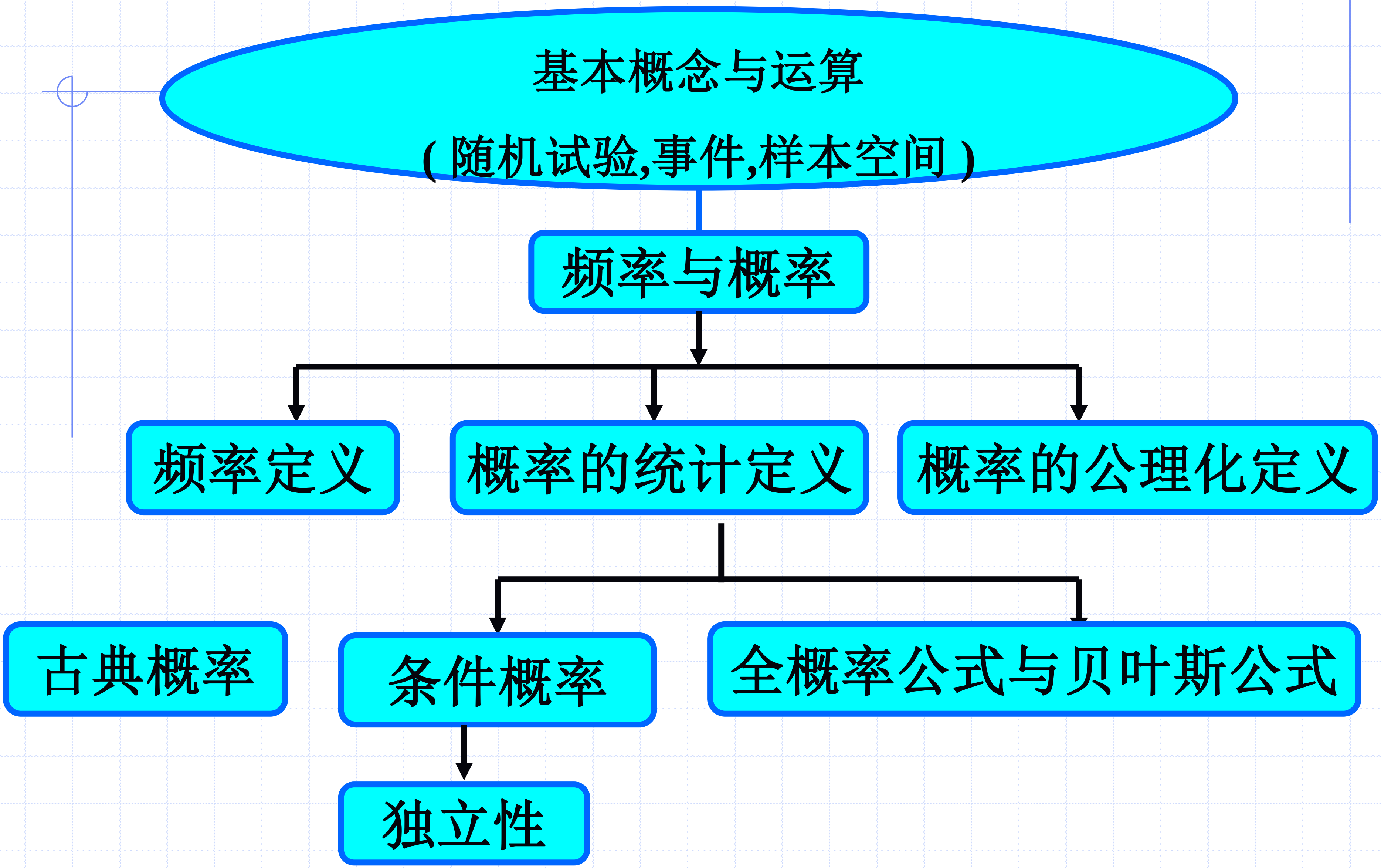
$$\frac{d}{d\theta}\ln L(\theta)=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i=0$$

$$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln x_i}-1$$ 极大似然估计值；

$$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln X_i}-1$$ 极大似然估计量；

矩估计量： $$\hat{\theta}=\frac{2\bar{X}-1}{1-\bar{X}}$$

第一章知识结构图


基本概念与运算
(随机试验,事件,样本空间)
频率与概率
频率定义
概率的统计定义
概率的公理化定义
古典概率
条件概率
全概率公式与贝叶斯公式
独立性

1、从以往的资料分析得知，在出口罐头导致索赔的事件中，有**50%**是质量问题；有**30%**是数量短缺问题；有**20%**是产品包装问题．又知在质量问题的争议中，经过协商解决的占**40%**；在数量短缺问题的争议中，经过协商解决的占**60%**；在产品包装问题的争议中，经过协商解决的占**75%**．如果在发生的索赔事件中，经过协商解决了，问这一事件不属于质量问题的概率是多少？

2、一个袋中有六只球，但是不知道是四只白球，两只红球（情况**I**），还是两只白球，四只红球（情况**II**）。情况**I**出现的可能性是情况**II**出现的两倍。

（**1**）如果出现情况**I**，那么从袋中随机抽取三只球，恰好是两只白球的概率是多少？

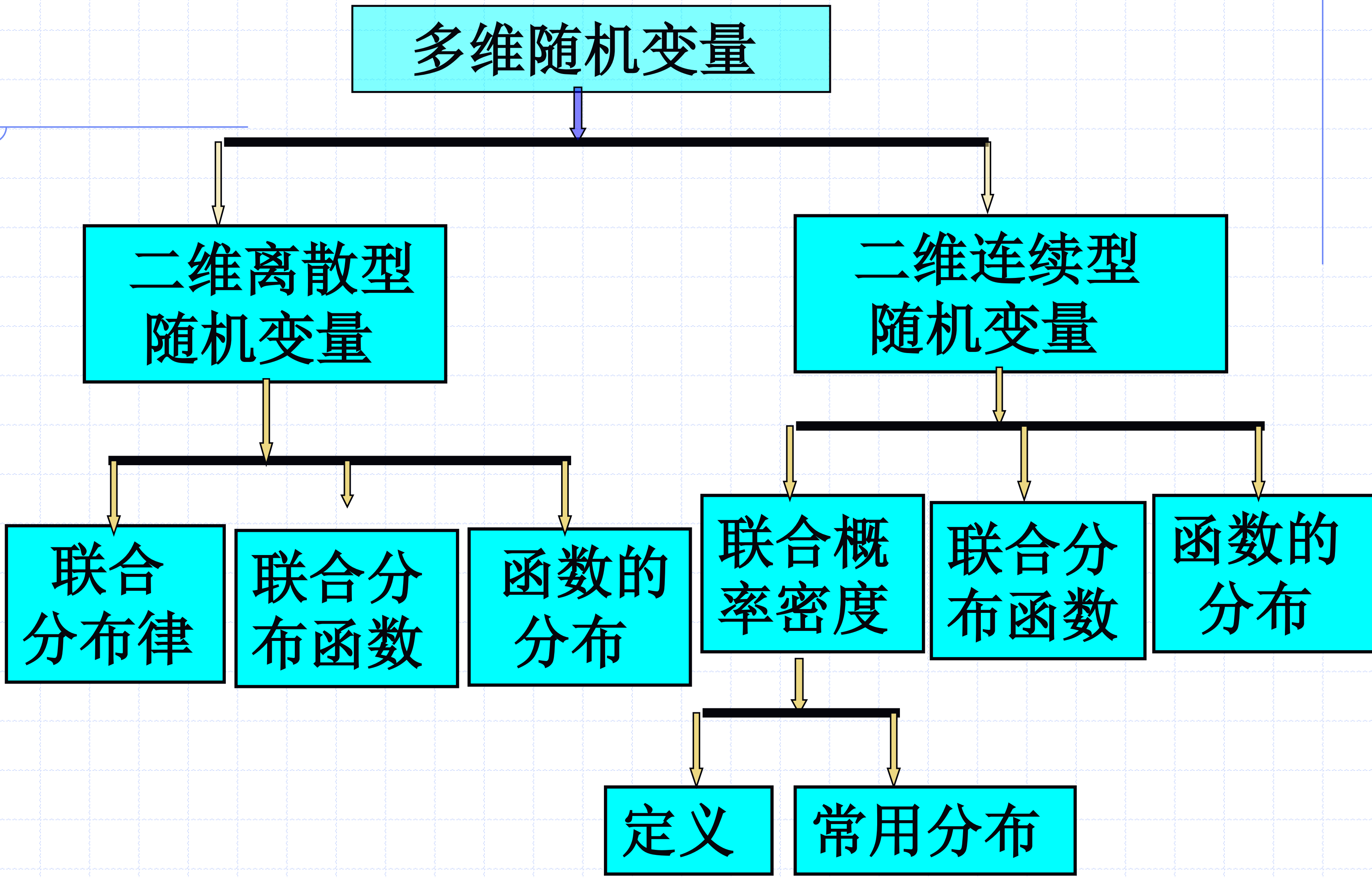
（**2**）随机从袋中抽取一只球，发现是红色的。问袋中是情况**I**的概率是多大？

（**3**）为判断袋中情况，随机抽取两只球。请根据取球的结果判断袋中球的情况。

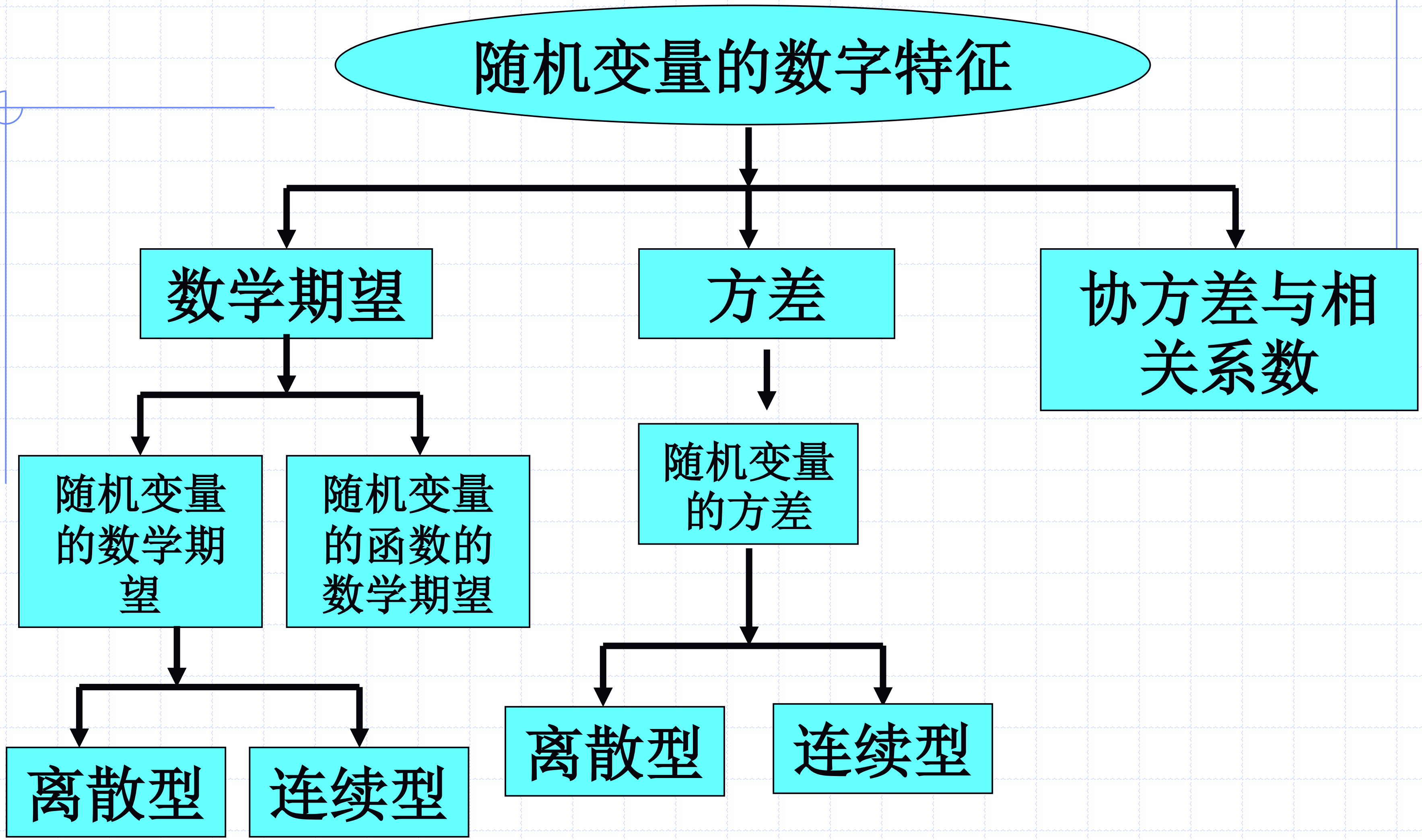
第二章知识结构图

- 随机变量
  - 离散型随机变量
    - 分布律
      - 定义
      - 常用分布
    - 分布函数
    - 函数的分布
  - 连续型随机变量
    - 概率密度
      - 定义
      - 常用分布
    - 分布函数
    - 函数的分布

第三章知识结构图


多维随机变量
二维离散型随机变量
二维连续型随机变量
联合分布律
联合分布函数
函数的分布
联合概率密度
联合分布函数
函数的分布
定义
常用分布

第四章知识结构图



随机变量的数字特征
数学期望
方差
协方差与相关系数
随机变量的数学期望
随机变量的函数的数学期望
随机变量的方差
离散型
连续型
离散型
连续型

设随机向量（*X*,*Y*）的联合概率密度函数为

$$f(x,y)=3x,\quad 0\le x\le 1,\ 0\le y\le x$$

求：**1**）（*X*,*Y*）的边缘密度函数，并判断*X*与*Y*是否相互独立

2) $P\{X+2Y>1\}$

3) $Z=X+Y$ 的概率密度

4）求 $D(X)$

第五章知识结构图

- 概率论的理论结果
  - 大数定律
    - 切比雪夫大数定理
    - 贝努利大数定理
    - 辛钦(大数)定理
  - 中心极限定理
    - 独立同分布的中心极限定理
    - 李雅普诺夫中心极限定理
    - 棣莫弗--拉普拉斯中心极限定理

例1：

$\{x_k\} \quad k=1,2,\cdots$ 是相互独立的随机变量序列

$$x_k \sim \begin{pmatrix} -3^k & 0 & 3^k \\ \frac{1}{3^{2k+2}} & 1-\frac{2}{3^{2k+2}} & \frac{1}{3^{2k+2}} \end{pmatrix}$$

问 $\{x_k\} \quad k=1,2,\cdots$ 是否服从大数定律？

2、 设$\{X_n\}$为独立同分布的随机变量序列，

且 $$P(X_n = \frac{2^k}{k^2}) = \frac{1}{2^k}, k = 1, 2, \cdots$$

证明$\{X_n\}$服从大数定律。

证明： 因为 1) $\{X_n\}$为独立同分布随机序列，

2) $\forall n,\ E(X_n) = \sum_{k=1}^{+\infty} \frac{2^k}{k^2} \cdot \frac{1}{2^k} = \sum_{k=1}^{+\infty} \frac{1}{k^2} < +\infty$ 存在，

故满足辛钦大数定律的条件，

所以，$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu,$

即，$\{X_n\}$服从大数定律。

3、一生产线生产的产品成箱包装，假设每箱平均重**50kg**，标准差为**5kg**．若用最大载重量为**5000kg**的汽车来承运，试用中心极限定理计算每辆车最多装多少箱，才能保证汽车不超载的概率大于**0.977**？