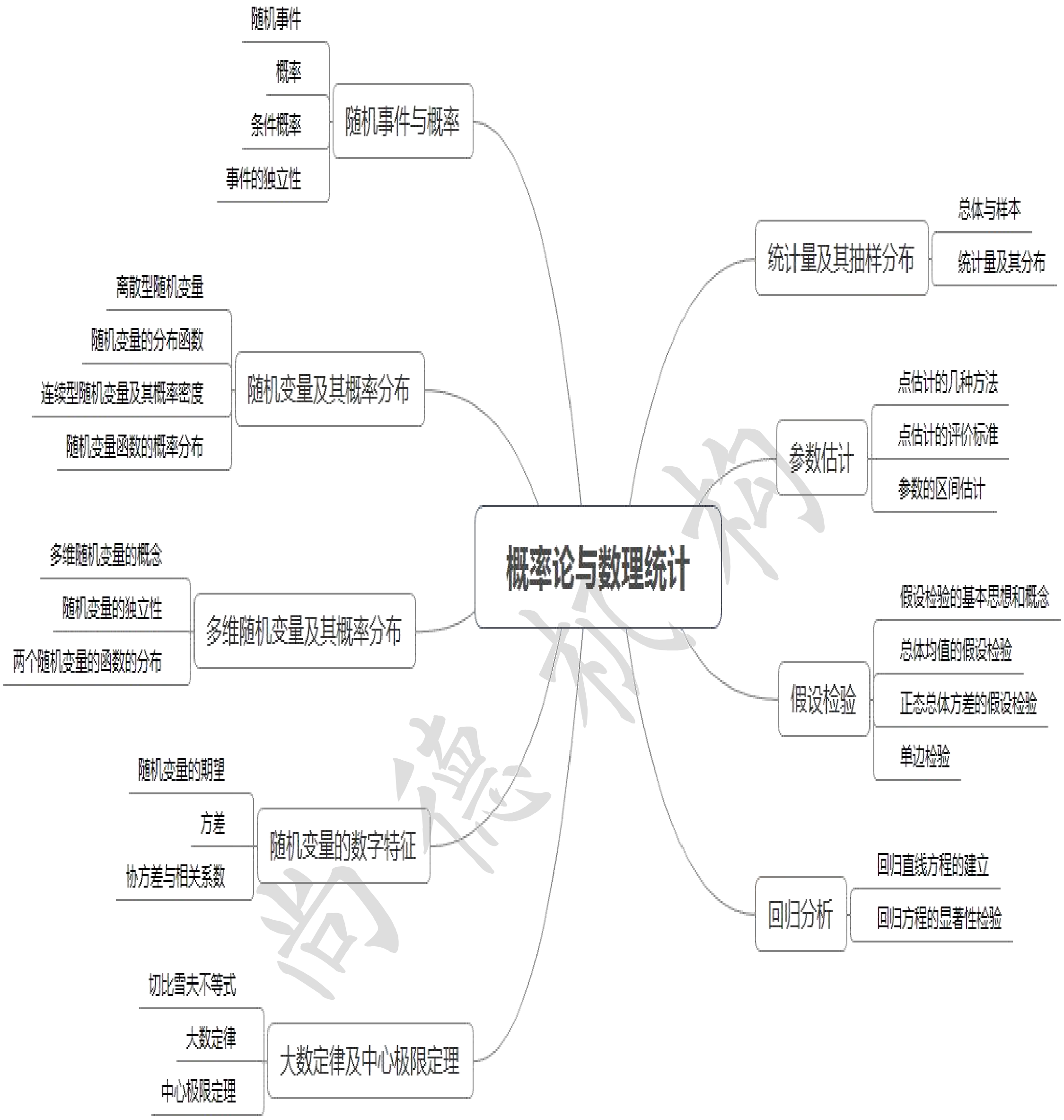
自考工商管理本科专业课
概率论与数理统计（课程代码：04183）
通关宝典（讲义）

概率论与数理统计

GAILVLUNYUSHULITONGJI

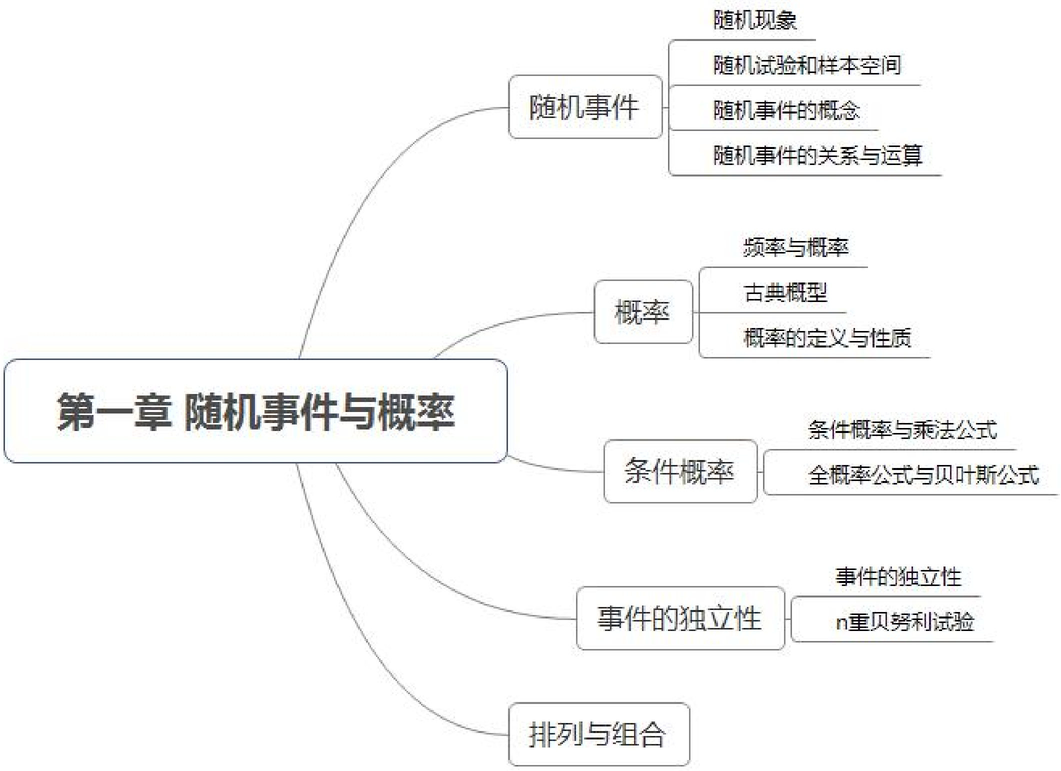
教材框架图


概率论与数理统计
随机事件与概率
随机事件
概率
条件概率
事件的独立性
随机变量及其概率分布
离散型随机变量
随机变量的分布函数
连续型随机变量及其概率密度
随机变量函数的概率分布
多维随机变量及其概率分布
多维随机变量的概念
随机变量的独立性
两个随机变量的函数的分布
随机变量的数字特征
随机变量的期望
方差
协方差与相关系数
大数定律及中心极限定理
切比雪夫不等式
大数定律
中心极限定理
统计量及其抽样分布
总体与样本
统计量及其分布
参数估计
点估计的几种方法
点估计的评价标准
参数的区间估计
假设检验
假设检验的基本思想和概念
总体均值的假设检验
正态总体方差的假设检验
单边检验
回归分析
回归直线方程的建立
回归方程的显著性检验

目 录

第一章 随机事件与概率

框架图



第一节 随机事件

1. 随机现象

随机现象：在一定的条件下，可能出现这样的结果，也可能出现那样的结果，我们预先无法断言，这类现象称为**随机现象**.

2. 随机试验和样本空间

随机试验：若把科学实验或观察都称为试验，则满足下列条件的试验称为**随机试验**（简称试验）：

（1）试验的可重复性——在相同条件下可重复进行.

（2）一次试验结果的随机性——每次试验前不能确定哪个结果会出现.

（3）全部试验结果的可知性——每次试验的可能结果不止一个，且在试验开始前能明确所有可能的结果.

样本空间：随机试验的每一个可能出现的不可分解的结果称为样本点，全体样本点的集合称为**样本空间**，用Ω（或S）来表示. 注：样本空间的元素可以是数，也可以不是数，样本空间所含有的样本点可以是有限多个也可以是无限多个. 样本点是随机试验最基本并且不可再分的结果. 当随机试验的内容确定后，样本空间也将确定.

3. 随机事件的概念

随机事件：样本空间Ω的子集合称为试验E的**随机事件**，简称**事件**，以大写字母$A,B,\cdots$来表示. 俗话说：在一次试验中可能出现也可能不出现的事件，统称为**随机事件**.

随机事件可以分为：

（1）基本事件	只含一个样本点的子集合.
（2）复合事件	含若干个样本点的子集合.
(3)不可能事件	不含样本点的子集合（空集），即它在每次试验中都不会发生，记为Φ.
（4）必然事件	样本空间本身，即它在每次试验中必然发生，记为Ω.

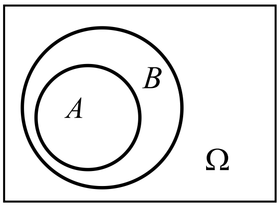
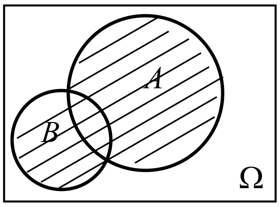
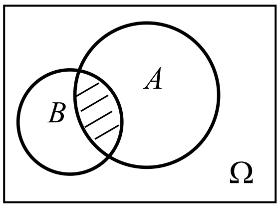
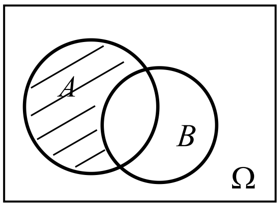
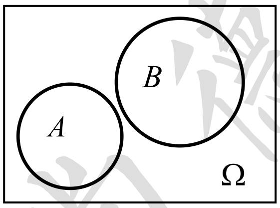
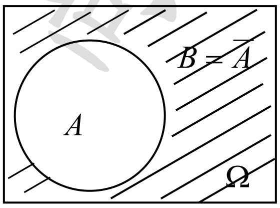
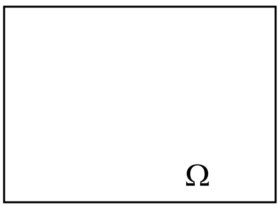
注意：（3）和（4）具有确定性，不是随机事件，但作随机事件处理.

4. 随机事件的关系与运算

设Ω为试验E的样本空间，A,B,C为Ω的子集，事件的关系及其运算可以与集合的关系和运算相对应：

符号	概率论	
Ω	样本空间，必然事件	
Φ	不可能事件	
ω	样本点（基本事件）	
A	事件	
$\overline{A}$	对立（逆或余）事件：A 的对立事件 若事件 A 与事件 B 中至少有一个发生，且 A 与 B 互不相容，即 $A\bigcup B=\Omega$，$AB=\Phi$，则称 A 与 B 互为对立事件.	$\overline{\overline{A}}=A$ $\overline{\Omega}=\Phi,\overline{\Phi}=\Omega$ $A-B=A\overline{B}=A-AB$
$A\subset B$	事件的包含：事件 A 发生必有事件 B 发生	$\Phi\subset A\subset\Omega$
$A=B$	事件的相等：事件 A 与事件 B 等价（同一事件的不同描述）	若 $A\subset B$ 且 $B\subset A$，则 $A=B$
$A\bigcup B$ 或 $A+B$	和事件：事件 A 与事件 B 至少有一个发生. 推广：$\bigcup\limits_{i=1}^{n}A_i$ 表示事件 A_1，A_2，…，A_n 中至少发生一个.	$A\subset A\bigcup B,B\subset A\bigcup B$ 若 $A\subset B$，则 $A\bigcup B=B$
$A\bigcap B$ 或 AB	积事件：事件 A 与事件 B 同时发生. 推广：$\bigcap\limits_{i=1}^{n}A_i$ 表示事件 A_1，A_2，…，A_n 都发生.	$AB\subset A,AB\subset B$ 若 $A\subset B$，则 $AB=A$
$A-B$	差事件：事件 A 发生而事件 B 不发生	$A-B\subset A$ 若 $A\subset B$，则 $A-B=\Phi$
$AB=\Phi$	互不相容（互斥）：事件 A 与事件 B 不能同时发生. 推广：n 个事件 A_1，A_2，…，A_n 两两之间互不相容，即 $A_iA_j=\Phi$ $(i\neq j,i,j=1,2,\cdots,n)$	

设 A,B,C 为事件，事件之间的关系与运算直观图：

事件	文氏图（韦恩图）	运算律
事件的包含 $A\subset B$		交换律：$A\cup B=B\cup A, A\cap B=B\cap A$
和事件 $A\cup B$		结合律：$A\cup(B\cup C)=(A\cup B)\cup C$， $A\cap(B\cap C)=(A\cap B)\cap C$
积事件 $A\cap B$		分配律：$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$， $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$
差事件 $A-B$		对偶律（德摩根律）：$\overline{A\cup B}=\overline{A}\overline{B}$， $\overline{AB}=\overline{A}\cup\overline{B}$
互不相容 $AB=\Phi$		$A\Phi=\Phi$
对立事件 $\overline{A}$		$A+\overline{A}=\Omega$，$A\cap\overline{A}=\Phi$
必然事件 Ω		$A\Omega=A$，$A+\Omega=\Omega$， $A+A=A$

注意：（1）事件的运算顺序一般是：先逆后积，再和差.（2）有包含关系的事件的和事件，“谁大要谁”；有包含关系的事件的积事件，“谁小要谁”.（3）讨论事件的关系时，借助“文氏图”.

第二节 概率

1. 频率与概率

频率：在相同条件下进行 n 次试验，在这 n 次试验中，事件 A 发生的次数 n_A 称为事件 A 发生的频数，而比值 $\frac{n_A}{n}$ 称为事件 A 发生的频率，并记作 $f_n(A)$. 频率具有如下性质：

（1）$0 \le f_n(A) \le 1$.

（2）$f_n(\Phi)=0$，$f_n(\Omega)=1$.

（3）若 A 与 B 互不相容，则 $f_n(A\cup B)=f_n(A)+f_n(B)$.

推广：当 A_1，A_2，…，A_m（或 A_1，A_2，…，A_m，…）互不相容时，

$$f_n\left(\bigcup_{k=1}^{m}A_k\right)=\sum_{k=1}^{m}f_n(A_k)\text{（或 }f_n\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}f_n(A_k)\text{）}.$$

其中 m 是正整数. 随着试验重复次数 n 的大量增加，频率 $f_n(A)$ 逐渐稳定于某一常数，称这个常数为事件 A 的概率 $P(A)$.

2. 古典概型

古典概型：具有下面两个特点的随机试验的概率模型，称为**古典概型**：

（1）基本事件的总数是有限的，换句话说样本空间仅含有有限个样本点；

（2）每个基本事件发生的可能性相同.

设 Ω 为随机试验 E 的样本空间，其中所含样本点总数为 n, A 为一随机事件，其中所含样本点数为 r，则古典概型事件概率计算公式为：

$$P(A)=\frac{r}{n}=\frac{A\text{中样本点数}}{\Omega\text{中样本点总数}}=\frac{A\text{所包含的基本事件数}}{\Omega\text{中基本事件总数}}.$$

古典概型中事件概率具有如下性质：

（1）$0 \le P(A) \le 1$；

（2）$P(\Phi)=0$，$P(\Omega)=1$；

（3）当 A 与 B 互不相容时，有 $P(A\cup B)=P(A)+P(B)$.

推广：当 A_1，A_2，…，A_m（或 A_1，A_2，…，A_m，…）互不相容时，

$$P\left(\bigcup_{k=1}^{m}A_k\right)=\sum_{k=1}^{m}P(A_k)\text{（或}P\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}P(A_k)\text{）}.$$

3. 概率的定义与性质

概率：设 E 是随机试验，Ω 是它的样本空间. 对于 E 的每一事件 A 赋予一个实数 $P(A)$，若满足：

（1）非负性：$P(A)\geq 0$；（2）规范性：$P(\Omega)=1$；（3）可列可加性：设 A_1，A_2，…，A_n 互不相容，则 $P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$，称 $P(A)$ 是事件 A 的概率.

概率的性质

有界性	$0\leq P(A)\leq 1$. 特别地，$P(\Phi)=0$，$P(\Omega)=1$. 注意，由 $P(A)=0$ 不能推出 $A=\Phi$. 同样，由 $P(A)=1$ 不能推得 $A=\Omega$
有限可加性	若 A_1，A_2，…，A_n 是两两互不相容事件，则 $P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$ 其中 n 为正整数.
单调性	设 A,B 是两个事件，若 $A\subset B$，则有 $P(B-A)=P(B)-P(A)$；$P(B)\geq P(A)$.
逆事件	对任一事件 A，有 $P(\overline{A})=1-P(A)$，$P(A)+P(\overline{A})=1$.（求逆公式）
加法公式	对任意两个事件 A,B 有 $P(A\cup B)=P(A)+P(B)-P(AB)$ 推广：对任意 n 个事件 A_1，A_2，…，A_n，有 $P(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leq i<j\leq n}P(A_iA_j)+\sum_{1\leq i<j<k\leq n}P(A_iA_jA_k)+\cdots$ $+(-1)^{n-1}P(A_1A_2\cdots A_n)$
减法公式	$P(A-B)=P(A\overline{B})=P(A)-P(AB)$. 特别地，当 $B\subset A$ 时，$P(A-B)=P(A)-P(B)$.

第三节 条件概率

1. 条件概率与乘法公式

条件概率：设 A,B 是两个事件且 $P(A)>0$，称 $P(B\mid A)=\dfrac{P(AB)}{P(A)}$ 为在事件 A 发生的条件下事件 B 发生的概率.

定理（乘法公式） 设 $P(A)>0$，则 $P(AB)=P(A)P(B\mid A)$. 同样地，若 $P(B)>0$，则 $P(AB)=P(B)P(A\mid B)$.

推广到 n 个事件的情况：

（1）设 $P(AB)>0$，则 $P(ABC)=P(A)P(B\mid A)P(C\mid AB)$.

（2）设 $P(A_1A_2\cdots A_{n-1})>0$，则

$$P(A_1A_2\cdots A_{n-1}A_n)=P(A_1)P(A_2\mid A_1)\cdots P(A_n\mid A_1A_2\cdots A_{n-1}).$$

2. 全概率公式与贝叶斯（Bayes）公式

完备事件组：设事件 A_1，A_2，…，A_n 满足如下两个条件：

（1）A_1，A_2，…，A_n 互不相容（即 $A_iA_j=\Phi(i\neq j)$），且 $P(A_i)>0$，$i=1,2,\cdots,n$；

（2）$A_1\cup A_2\cup\cdots\cup A_n=\Omega$，即 A_1，A_2，…，A_n 至少有一个发生，则称 A_1，A_2，…，A_n 为样本空间 Ω 的一个**划分（完备事件组）**. 当 A_1，A_2，…，A_n 是 Ω 的一个划分时，每次试验有且只有其中的一个事件发生.

全概率公式：设随机试验对应的样本空间为 Ω，设 A_1，A_2，…，A_n 是样本空间 Ω 的一个划分，B 是任意一个事件，则 $P(B)=\sum\limits_{i=1}^{n}P(A_i)P(B\mid A_i)$.

贝叶斯公式：设 A_1，A_2，…，A_n 是样本空间的一个划分，B 是任一事件，且 $P(B)>0$，则

$$P(A_i\mid B)=\frac{P(A_i)P(B\mid A_i)}{P(B)}=\frac{P(A_i)P(B\mid A_i)}{\sum\limits_{k=1}^{n}P(A_k)P(B\mid A_k)},\quad i=1,2,\cdots,n.$$

第四节 事件的独立性

1. 事件的独立性

事件的独立性：设 A,B 是两事件，若满足等式 $P(AB)=P(A)P(B)$，则称事件 A,B 相互独立，简称 A,B 独立.

注：若 $P(A)>0,P(B)>0$，则 A,B 相互独立与 A,B 互不相容不能同时成立.

事件独立性有下列性质：

（1）设 A,B 是两事件，且 $P(A)>0$，则 A,B 相互独立的充要条件是 $P(B)=P(B\mid A)$. 设 $P(B)>0$，则 A,B 相互独立的充要条件是 $P(A)=P(A\mid B)$.

（2）若事件 A 与 B 相互独立，则下列各对事件也相互独立：A 与 $\overline{B}$，$\overline{A}$ 与 B，$\overline{A}$ 与 $\overline{B}$.

三事件相互独立：设 A,B,C 是三个事件，若满足等式

$$P(AB)=P(A)P(B)，P(BC)=P(B)P(C)，P(AC)=P(A)P(C)，$$

$$P(ABC)=P(A)P(B)P(C)，$$

则称事件 A,B,C 相互独立.

注：事件 A,B,C 相互独立的充要条件是 A,B,C 两两独立，且 $P(ABC)=P(A)P(B)P(C)$.

推广——n 个事件 $A_1,A_2,\cdots,A_n$ 的独立性：

定义 3 设 $A_1,A_2,\cdots,A_n$ 为 n 个事件，若对于任意整数 $k(1\le k\le n)$ 和任意 k 个整数 $1\le i_1<i_2<\cdots<i_k\le n$，有 $P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$，则称 $A_1,A_2,\cdots,A_n$ 相互独立，简称 $A_1,A_2,\cdots,A_n$ 独立. 注：n 个事件相互独立必两两独立，但反之不然.

2. n 重贝努利事件

贝努利概型：具有下面两个特点的随机试验的概率模型，称为**贝努利概型**：

（1）n 次独立重复试验；

（2）每次试验只有两个可能结果 A 和 $\overline{A}$，且 $P(A)=p$，$0<p<1$.

注：只有两种可能结果的独立重复试验称为贝努利试验.

定理 1 在 n 重贝努利试验中，设每次试验中事件 A 的概率为 p（$0<p<1$），则事件 A 恰好发生 k 次的概率

$$P_n(k)=C_n^k p^k(1-p)^{n-k}\text{，}\ k=0,1,2,\cdots,n.$$

A 在指定的 k 次试验中发生，而在其余 $n-k$ 次试验中不发生的概率为 $p^k(1-p)^k$.

常见 n 重贝努利试验：抛掷硬币时注意的是正面是否朝上；产品抽样检验时，注意抽出的产生是否是次品；射手向目标射击时，注意目标是否被命中.

本章基本概型及其概率计算

概型	随机试验特征	概率计算公式
古典概型	（1）基本事件等可能； （2）样本空间有限.	$P(A)=\dfrac{A\text{中基本事件数}}{\Omega\text{中基本事件总数}}$
全概模型	试验分两次进行，第一次所有可能结果 $A_1,A_2,\cdots,A_n$ 为样本空间 Ω 的一个划分，B 为第二次试验后的某随机事件.	全概率公式：$P(B)=\sum\limits_{i=1}^{n}P(A_i)P(B\mid A_i)$ 贝叶斯公式（逆概率公式）： $P(A_i\mid B)=\dfrac{P(A_i)P(B\mid A_i)}{P(B)}=\dfrac{P(A_i)P(B\mid A_i)}{\sum\limits_{k=1}^{n}P(A_k)P(B\mid A_k)}$ $i=1,2,\cdots,n$
贝努利概型	n 次独立重复试验，且每次试验只有两个可能结果：A 和 $\overline{A}$，$P(A)=p$ 不变.	n 次试验中 A 恰好发生 k 次的概率为： $P_n(k)=C_n^k p^k(1-p)^{n-k}$，$k=0,1,2,\cdots,n$.

注意：利用基本公式计算概率的一般步骤：

① 写出欲求概率的事件 A；

② 分析 B 与哪些事件有关，用字母（如 $A_1,A_2,\cdots,A_n$）将这些事件表示出来；

③ 分析 B 与 $A_1,A_2,\cdots,A_n$ 是什么关系，写出事件的关系式；

④ 选择适当的公式计算事件 A 的概率.

附录 排列与组合

1. 两个基本原理

（1）乘法原理：若某件事需经k步才能完成，做第一步有m_1种方法，做第二步有m_2种方法…做第k步有m_k种方法，则完成这件事共有$m_1 \times m_2 \times \cdots \times m_k$种方法.

（2）加法原理：若某件事可由k类不同途径之一去完成，在第一类途径中有m_1种完成方法，在第二类途径中有m_2种完成方法…在第k类途径中有m_k种完成方法，那么完成这件事共有$m_1 + m_2 + \cdots + m_k$种方法.

2. 排列

（1）排列——从n个不同元素中任取$r(r \le n)$个元素排成一列（考虑元素次序）称此为一个排列，此种排列的总数记为A_n^r.

（2）可重复排列——从n个不同元素中每次取出一个，放回后再取下一个，如此连续取r次所得的排列称为可重复排列，此种排列总数共有n^r个.注意这里的r允许大于n.

3. 组合

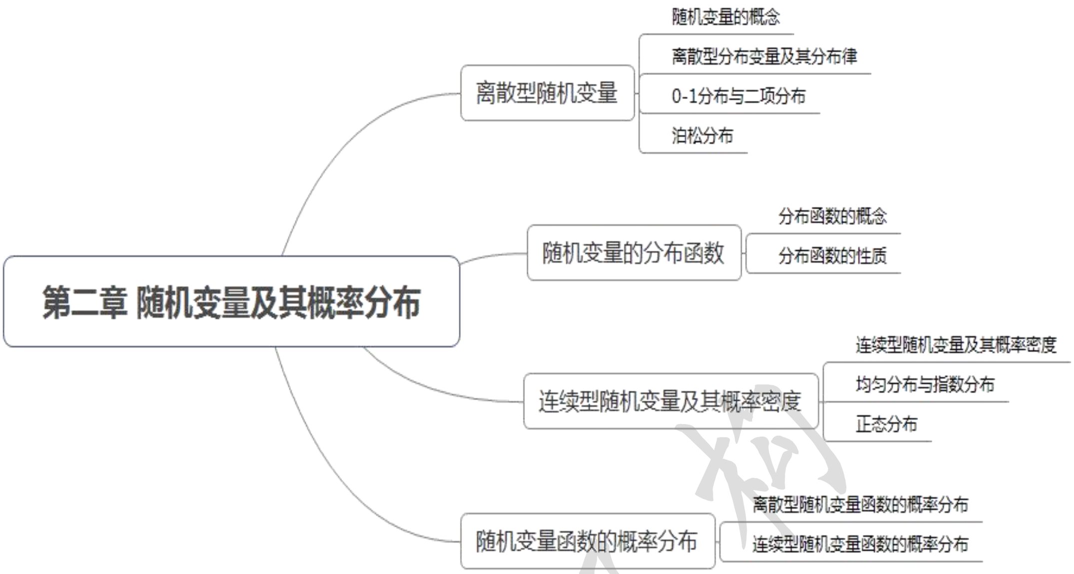
（1）组合——从n个不同的元素中任取r（$r \le n$）个元素并成一组（不考虑元素间的次序），称此为一个组合，此种组合的总数为C_n^r或$\binom{n}{r}$，计算公式为$C_n^r = \dfrac{n!}{r!(n-r)!}$.

注意：$0! = 1$，$C_n^0 = \binom{n}{0} = 1$.

（2）性质——$C_n^r = C_n^{n-r}$；$C_n^n = C_n^0 = 1$.

第二章 随机变量及其概率分布

框架图



第一节 离散型随机变量

1. 随机变量的概念

随机变量：设 Ω 是随机试验 E 的样本空间，若对每一个样本点 $\omega \in \Omega$，有一个实数 $X = X(\omega)$ 与之对应，这样得到的一个定义在样本空间 Ω 上的实值函数 $X = X(\omega)$，称为**随机变量**. 随机变量一般用大写字母 $X, Y, Z, \cdots$ 或 $X_1, X_2, X_3, \cdots$ 来表示. 随机变量可分为：离散型、连续型、非离散且非连续型随机变量.

2. 离散型分布变量及其分布律

离散型随机变量：若随机变量 X 只能取有限多个或可列无限多个值，则称 X 为离散型随机变量. 对应的随机试验样本空间的样本点为有限多个或可列无限多个.

分布律：设 X 为离散型随机变量，可能取值为 $x_1, x_2, \cdots, x_k, \cdots$，且 $P\{X = x_k\} = p_k$，$k = 1,2,3,\cdots$，则称 $\{p_k\}$ 为 X 的分布律（或分布列，或概率分布）.

分布律用表格表示如下：

X	x_1	x_2	$\cdots$	x_k	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

分布律的性质：

（1）非负性： $p_k \geq 0, k = 1,2,3,\cdots$；

（2）规范性： $\sum\limits_{k=1}^{\infty} p_k = 1$. 注意：正数列 p_i 为离散型随机变量 X 的分布律 $\Leftrightarrow$（1）（2）同时成立.

（3）利用分布律计算概率： $P\{X \in I\} = \sum\limits_{x_i \in I} P\{X = x_i\}$，其中 I 为任一实数集合.

（4）利用分布律求分布函数 $F(x) = \sum\limits_{x_i \leq x} p_i$， $x \in R$，其图形是右连续的阶梯函数.

分布律或分布函数均可完整描述离散型随机变量.

三种重要的离散随机变量及其分布—— 0-1 分布、二项分布与泊松分布.

3. $0-1$ 分布与二项分布

4. 泊松分布

0-1分布：若随机变量只取两个可能值 0，1，且 $P\{X=1\}=p$， $P\{X=0\}=q$，其中 $0<p<1$， $q=1-p$，则称 X 服从 **$0-1$ 分布**. X 的分布律为

X	0	1
P	q	p

二项分布：若随机变量 X 的可能取值为 $0,1,2,\cdots,n$，而 X 的分布律为

$$p_k=P\{X=k\}=C_n^k p^k q^{n-k}，\ n=0,1,2,\cdots,n，$$

其中 $0<p<1$， $p+q=1$，则称 X 服从参数为 n,p **二项分布**，简记为 $X\sim B(n,p)$. 显然，当 $n=1$ 时， X 服从 $0-1$ 分布，即 $0-1$ 分布实际上是二项分布的特例.

常见二项分布：如一批产品的不合格率为 p，检查 n 件产品，n 件产品中不合格品数 X 服从二项分布；调查 n 个人， n 个人中色盲人数 Y 服从参数为 n,p 的二项分布，其中 p 为色盲率；n 部机器独立运转，每台机器出故障的概率为 p，则 n 部机器中出故障的机器 Z 服从二项分布.

泊松定理 设 $\lambda>0$ 是常数，n 是任意正整数，且 $np_n=\lambda$，则对于任意取定的非负整数 k，有 $\lim\limits_{n\to\infty}C_n^k p_n^k(1-p_n)^{n-k}=\dfrac{\lambda^k}{k!}e^{-\lambda}$.

注意：当 n 很大 p 很小时，有近似公式 $C_n^k p^k q^{n-k}\approx\dfrac{\lambda^k}{k!}e^{-\lambda}$，其中 $\lambda=np$.

泊松分布：设随机变量 X 的可能取值为 $0,1,2,\cdots,n,\cdots$，而 X 的分布律为

$$p_k=P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}，\ k=0,1,2,\cdots，$$

其中 $\lambda>0$，则称 X 服从参数为 λ 的**泊松分布**，简记为 $X\sim P(\lambda)$.

常见泊松分布：如某一时段进入某商店的顾客数，某一地区一个时间间隔内发生交通事故的次数，一天内 110 报警台接到报警的次数，在一个时间间隔内某种放射性物质发生的粒子数.

第二节 随机变量的分布函数

1. 分布函数的概念

分布函数：设 X 为随机变量，x 是任意实数，则称函数 $F(x)=P\{X\le x\}$，$x\in(-\infty,+\infty)$，为 X 的**分布函数**（也称为累积概率函数）.

注意：（1）分布函数 $F(x)$ 是普通的一元函数，它在任意一点 x_0 处的值 $F(x_0)$ 表示随机变量 X 在区间 $(-\infty,x_0]$ 内取值的概率.（2）分布函数的定义域为 $(-\infty,+\infty)$，值域为 $[0,1]$.

2. 分布函数的性质

分布函数的性质

非负性	$0\le F(x)\le 1$
规范性	$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$，$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$
不减性	$F(x)$ 单调不减，即对任意 $x_1,x_2\in(-\infty,+\infty)$，当 $x_1<x_2$ 时，$F(x_1)\le F(x_2)$
右连续性	$F(x)$ 右连续，即 $F(x+0)=\lim\limits_{\Delta x\to0^+}F(x+\Delta x)=F(x)$.
利用分布函数计算概率	$P\{x_1<X\le x_2\}=F(x_2)-F(x_1)$，$P\{X=x\}=F(x)-F(x-0)$

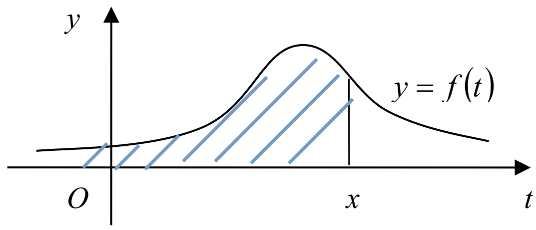
注意：$F(x)$ 为某个随机变量 X 的分布函数 $\Leftrightarrow$ 非负性、规范性、不减性、右连续性同时成立.

第三节 连续型随机变量及其概率密度

1. 连续型随机变量及其概率密度

连续型随机变量及概率密度函数：设 $F(x)$ 是随机变量 X 的分布函数，若存在非负可积函数 $f(x)$，使其对任意实数 $x\in(-\infty,+\infty)$，都有 $F(x)=\int_{-\infty}^{x}f(t)dt$，则称 X 为**连续型随机变量**，并称 $f(x)$ 为 X 的**概率密度函数**，简称概率密度（或密度函数）.

其几何意义如图所示，其中图中阴影部分的面积为 $F(x)$（随机变量落在区间 $(-\infty,x]$ 上的概率）.



连续型随机变量的概率分布的性质：

（1）概率密度 $f(x)\geq 0$（$x\in R$）.

（2）$\int_{-\infty}^{+\infty} f(x)dx=1$. 注意：非负函数 $f(x)$ 为连续型随机变量 X 的概率密度.

（3）设 x 为概率密度 $f(x)$ 的连续点，则 $F'(x)$ 存在，且 $F'(x)=f(x)$.

（4）连续型随机变量的分布函数 $F(x)$ 是连续函数.

（5）对于连续型随机变量，任意 $x_0\in(-\infty,+\infty)$，有 $P\{X=x_0\}=0$.

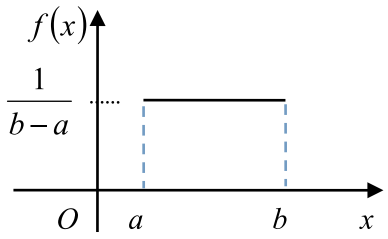
（6）$P\{a<X\leq b\}=P\{a\leq X<b\}=P\{a<X<b\}=\int_a^b f(x)dx$.

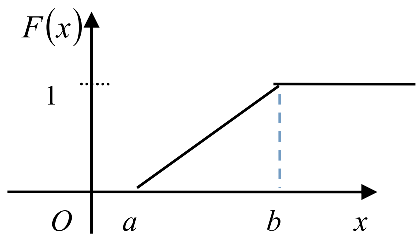
三种重要的连续型概率分布——均匀分布、指数分布和正态分布.

2. 均匀分布与指数分布

均匀分布：若随机变量 X 的概率密度为 $f(x)=\begin{cases}\dfrac{1}{b-a}, & x\leq a\leq b,\\ 0, & 其他.\end{cases}$ 则称 X 服从区间 $[a,b]$ 上的**均匀分布**，简记为 $X\sim U(a,b)$.

均匀分布的概率密度 $f(x)$ 与分布函数 $F(x)$ 的图形为





均匀分布的均匀性是指随机变量 X 落在区间 $[a,b]$ 内长度相等的子区间上的概率都是相等同的.

均匀分布的概率计算——概率公式：

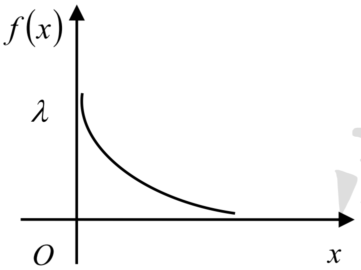
设 $X\sim U(a,b)$，$a\le c<d\le b$，即 $[a,b]\supset[c,d]$，则 $P\{c\le X\le d\}=\dfrac{d-c}{b-a}$.

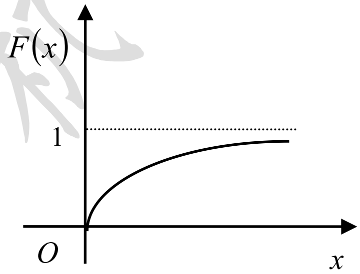
常用均匀分布：计算结果保留到小数点后第 n 位，则舍入误差 X 通常假定服从 $\left(-0.5\times10^{-n},0.5\times10^{-n}\right)$ 上的均匀分布；从刻度器上读数时把零头数化为最靠近的整分度时发生的误差也服从均匀分布.

指数分布：若随机变量 X 的概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\le 0,\end{cases}$ 其中 $\lambda>0$ 为常数，则称 X 服从参数为 λ 的**指数分布**，简记为 $X\sim E(\lambda)$，其中分布函数为

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

指数分布的概率密度 $f(x)$ 与分布函数 $F(x)$ 的图形为





指数分布常被用作各种“寿命”分布：如电子元件的使用寿命、动物的寿命、电话的通话时间、顾客在某一服务系统接受服务的时间等都可假定服从指数分布.

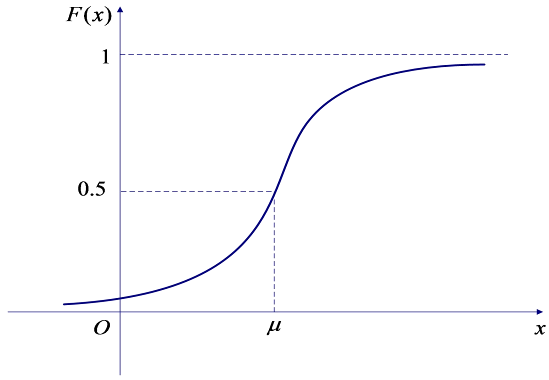
3. 正态分布

正态分布：若随机变量 X 的概率密度为 $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$-\infty<x<+\infty$，其中 μ,σ^2 为常数，$-\infty<\mu<+\infty$，$\sigma>0$，则称 X 服从参数为 μ,σ^2 的**正态分布**，简记为 $X\sim N(\mu,\sigma^2)$. 习惯上称服从正态分布的随机变量为正态随机变量，又称正态分布的概率密度曲线为正态分布曲线.

正态分布曲线的性质：

（1）曲线关于直线 $x=\mu$ 对称，对任意的 $h>0$，有 $P\{\mu-h<X\le\mu\}=P\{\mu<X\le\mu+h\}$.

（2）当 $x=\mu$ 时取得最大值 $f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$，

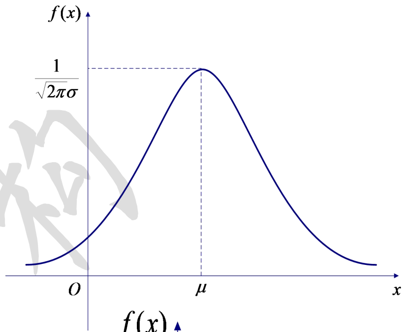


在 $x=\mu\pm\sigma$ 处曲线有拐点，曲线以 x 轴为渐近线.

（3）当 σ 取定，$\mu_1<\mu_2$ 时，

$$f_1(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu_1)^2}{2\sigma^2}},$$

$$f_2(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu_2)^2}{2\sigma^2}},$$



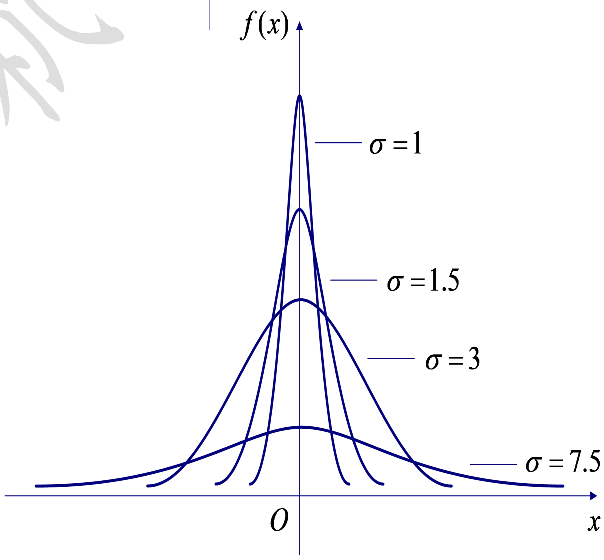
$f_1(x)$ 与 $f_2(x)$ 的图形可互相沿着 x 轴平行移动而得.

注：正态分布曲线的位置完全由 μ 决定，μ 是正态分布的中心.（固定 σ，改变 μ，图形沿 x 轴平移而不改变其形状.）

（4）当 μ 给定，且 $\sigma_1<\sigma_2$ 时，

$$f_1(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu)^2}{2\sigma_1^2}},$$

$$f_2(x)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(x-\mu)^2}{2\sigma_2^2}},$$



注意：正态分布曲线中 σ 的值刻画了正态随机变量取值的分散程度. σ 越小，取值分散程度越小；σ 越大，取值分散程度越大.（固定 μ，改变 σ，当 σ 很小时，形状与尖塔相似；当 σ 增大时，曲线趋于平坦）.

正态分布 $X\sim N(\mu,\sigma^2)$ 的分布函数 $F(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$. 特别地，当 $\mu=0,\sigma=1$ 的正态分布 $N(0,1)$ 称为标准正态分布. 标准正态分布的概率密度和分布函数分别记为 $\varphi(x)$ 和 $\Phi(x)$，即

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}，\ \Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt，\ -\infty<x<+\infty.$$

标准正态分布函数$\Phi(x)$的性质：

（1）$\Phi(-x)=1-\Phi(x)$；

（2）$\Phi(0)=\frac{1}{2}$；

（3）对一般正态分布$X\sim N(\mu,\sigma^2)$，其分布函数$F(x)=P\{X\le x\}=\Phi\left(\frac{x-\mu}{\sigma}\right)$；

（4）$P\{a<x\le b\}=P\{a\le x<b\}=P\{a\le x\le b\}=P\{a<x<b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$

即$F(b)-F(a)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$.

（5）$P\{X>a\}=P\{X\ge a\}=1-\Phi\left(\frac{a-\mu}{\sigma}\right)$.

标准正态分布的上侧α分位数：设$X\sim N(0,1)$，若u_a满足条件$P\{X>u_a\}=\alpha$，$0<\alpha<1$，则称点u_a为**标准正态分布的上侧α分位数**.

第四节 随机变量函数的概率分布

1. 离散型随机变量函数的概率分布

设离散型随机变量X的分布律为

X	x_1	x_2	$\cdots$	x_k	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

则X的函数$Y=g(X)$（$g(x)$一般是连续函数或分段连续函数）也是离散型随机变量，且Y的分布律为

Y	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_k)$	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

注意：若某些$g(x_k)$的值相等，则对应Y的分布律中只写一项$g(x_k)$，并将相应的概率求和作为随机变量Y取$g(x_k)$值的概率.

2. 连续型随机变量函数的概率分布

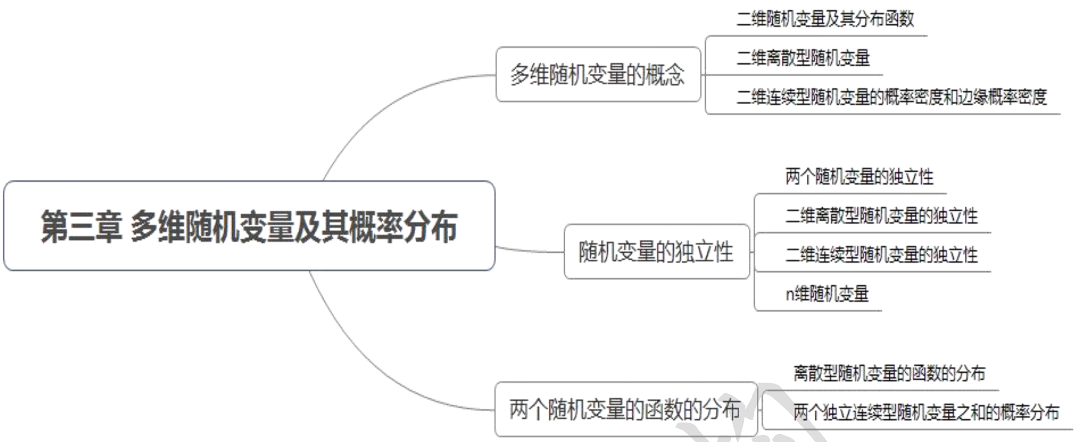
连续型随机变量函数的概率分布：设X为连续型随机变量，其概率密度为$f_X(x)$. 设$g(x)$是一严格单调的可导函数，其值域为$[\alpha, \beta]$，且$g'(x) \neq 0$. 记$x = h(y)$为$y = g(x)$的反函数，则$Y = g(X)$的概率密度

$$f_Y(y) = \begin{cases} f_X(h(y))|h'(y)|, & \alpha < y < \beta, \\ 0, & \text{其他}. \end{cases}$$

特别地，当$\alpha = -\infty$，$\beta = +\infty$时，$f_Y(y) = f_X(h(y))|h'(y)|$，$-\infty < y < +\infty$.

第三章 多维随机变量及其概率分布

框架图



第一节 多维随机变量的概念

1. 二维随机变量及其分布函数

二维随机变量：设 $X_1, X_2, \cdots, X_n$ 是定义在同一样本空间 Ω 上的随机变量，构成向量 $X=(X_1, X_2, \cdots, X_n)$ 称为一个 n 维随机变量或 n 维随机向量，当 $n=2$ 时，称为二维随机变量，记为 (X,Y) 或 (ξ,η). 当 $n=1$ 时，称为一维随机变量，即第二章中的随机变量.

分布函数：设 (X,Y) 为一个二维随机变量，记

$$F(x,y)=P\{X\le x, Y\le y\}，-\infty<x<+\infty，-\infty<y<+\infty，$$

称二元函数 $F(x,y)$ 为 X 与 Y 的**联合分布函数**，或称为 (X,Y) 的**分布函数**.

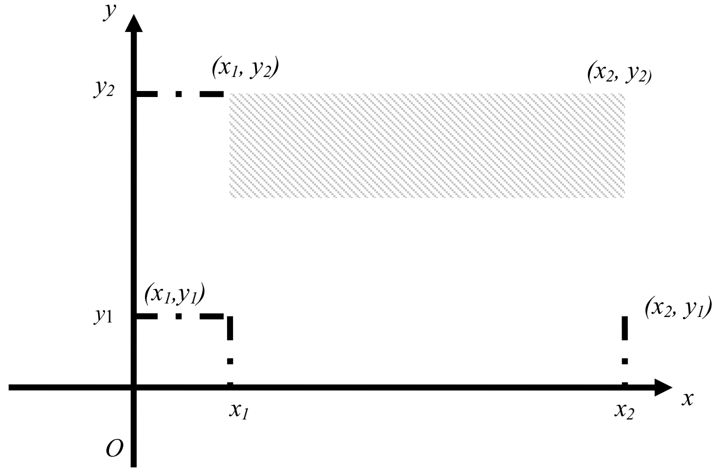
边缘分布函数：(X,Y) 的两个分量 X 与 Y 各自的分布函数分别称为二维随机变量 (X,Y) 关于 X 与关于 Y 的**边缘分布函数**，记为 $F_X(x)$ 与 $F_Y(y)$.

边缘分布函数可由联合分布函数确定：

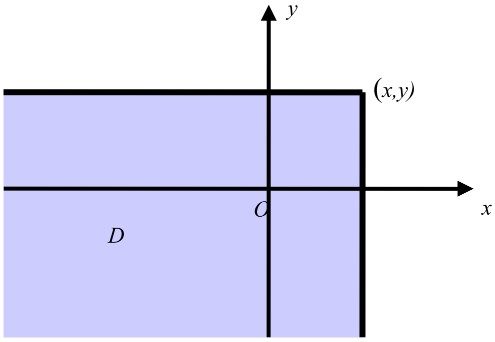
$$F_X(x)=P\{X\le x\}=P\{X\le x, Y<+\infty\}=F(x,+\infty)=\lim_{y\to+\infty}F(x,y),$$

$$F_Y(y)=P\{Y\le y\}=P\{X<+\infty, Y\le y\}=F(+\infty,y)=\lim_{x\to+\infty}F(x,y).$$

几何意义：分布函数 $F(x,y)$ 在 (x,y) 处的函数值就是随机点 (X,Y) 落在以 (x,y) 为顶点、位于该点左下方的无穷矩形 D 内的概率, 如图.



利用分布函数及其几何意义，随机点(X,Y)落在矩形域$\{x_1 < X \le x_2, y_1 < Y \le y_2\}$内（如图）的概率为：



$$P\{x_1 < X \le x_2, y_1 < Y \le y_2\} = F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1).$$

分布函数$F(x,y)$的性质：

规范性：$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$， $F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$	$0\le F(x,y)\le 1$，且 对任意固定的y，$F(-\infty,y)=0$； 对任意固定的x，$F(x,-\infty)=0$； $F(-\infty,-\infty)=0$，$F(+\infty,+\infty)=1$.
不减性：$F(x)$单调不减，即对任意$x_1,x_2\in(-\infty,+\infty)$，当$x_1<x_2$时，$F(x_1)\le F(x_2)$	$F(x,y)$是变量x（或y）的不减函数，即对于任意固定的y，当$x_2>x_1$时$F(x_2,y)\ge F(x_1,y)$； 对于任意固定的x，当$y_2>y_1$时， $F(x,y_2)\ge F(x,y_1)$.
右连续性：$F(x)$右连续，即 $F(x+0)=\lim\limits_{\Delta x\to0^+}F(x+\Delta x)=F(x)$	$F(x,y)$关于x和关于y均右连续，即 $F(x,y)=F(x+0,y)$；$F(x,y)=F(x,y+0)$.
利用分布函数计算概率： $P\{x_1<X\le x_2\}=F(x_2)-F(x_1)$， $P\{X=x\}=F(x)-F(x-0)$.	随机点(X,Y)落在矩形域$\{x_1<X\le x_2,y_1<Y\le y_2\}$内的概率为：$P\{x_1<X\le x_2,y_1<Y\le y_2\}$ $=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)$.

2. 二维离散型随机变量

二维离散型随机变量：若二维随机变量只能取有限多对或可列无穷多对$\left(x_i, y_j\right)$，$\left(i, j=1,2,\cdots\right)$，则称$\left(X, Y\right)$为**二维离散型随机变量.** 设二维随机变量$\left(X, Y\right)$的所有可能取值为$\left(x_i, y_j\right)$，$\left(i, j=1,2,\cdots\right)$，$\left(X, Y\right)$在各个可能取值的概率为：$P\left\{X=x_i, Y=y_j\right\}=p_{ij}$ $\left(i, j=1,2,\cdots\right)$，则称$P\left\{X=x_i, Y=y_j\right\}=p_{ij}$ $\left(i, j=1,2,\cdots\right)$为$\left(X, Y\right)$的**分布律**，其表格形式为：

Y / X	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}		$\cdots$		$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	1

满足：① $p_{ij} \geq 0$ $\left(i, j=1,2,\cdots\right)$；② $\sum\limits_i \sum\limits_j p_{ij}=1$（①，②称为$\left(X, Y\right)$的分布律的性质）.

注意：若数集$\left\{p_{ij}\right\}$ $\left(i, j=1,2,\cdots\right)$具有以上两条性质，则它必可作为二维离散型随机变量的分布律.

边缘分布律：对于离散型随机变量$\left(X, Y\right)$，分量X（或Y）的分布律称为$\left(X, Y\right)$关于X（或Y）的边缘分布律，记为$p_{i\cdot}$ $\left(i=1,2,\cdots\right)$（或$p_{\cdot j}$ $\left(j=1,2,\cdots\right)$），它可由$\left(X, Y\right)$的分布律求出.

$$p_{i\cdot}=P\left\{X=x_i\right\}=P\left\{X=x_i, Y=y_1\right\}+P\left\{X=x_i, Y=y_2\right\}+\cdots+P\left\{X=x_i, Y=y_j\right\}+\cdots$$

$$=\sum_j P\left\{X=x_i, Y=y_j\right\}=\sum_j p_{ij}$$，（类似地，可计算$p_{\cdot j}$）.

(X,Y)的边缘分布律为：

(X,Y)关于 X 的边缘分布律为： $p_{i\cdot}=P\{X=x_i\}=\sum_j p_{ij}$， $i=1,2,\cdots$，或记为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
p	$p_{1\cdot}$	$p_{2\cdot}$	$\cdots$	$p_{i\cdot}$	$\cdots$

(X,Y)关于 Y 的边缘分布律为： $p_{\cdot j}=P\{Y=y_j\}=\sum_i p_{ij}$， $j=1,2,\cdots$，或记为

Y	y_1	y_2	$\cdots$	y_i	$\cdots$
p	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$

注意：① $p_{i\cdot}\geq 0$， $p_{\cdot j}\geq 0$， $(i,j=1,2,\cdots)$ ② $\sum_i p_{i\cdot}=1$， $\sum_j p_{\cdot j}=1$（①，②称为(X,Y)的边缘分布律的性质）.

3. 二维连续型随机变量的概率密度和边缘概率密度

二维连续型随机变量：设二维随机变量(X,Y)的分布函数为$F(x,y)$，若存在非负可积函数 $f(x,y)$，使得对任意的实数 x,y，有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv,$$

则称(X,Y)为**二维连续型随机变量**，并称 $f(x,y)$为(X,Y)的概率密度或 X 与 Y 的联合密度函数.

概率密度 $f(x,y)$的性质：

（1） $f(x,y)\geq 0$；

（2） $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$. 注：该性质为判定一个二元函数是否为概率密度函数的依据.

（3）若 $f(x,y)$在(x,y)处连续，则有$\dfrac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)$.

（4）若(X,Y)的概率密度为 $f(x,y)$，则(X,Y)在平面区域D（即$\{(X,Y)\in D\}$）内取值的概率为

$$P\{(X,Y)\in D\}=\iint\limits_{D} f(x,y)dxdy.$$

注：随机点(X,Y)落在平面区域D上的概率等于以平面区域D为底、以曲面$z=f(x,y)$为顶的曲顶柱体的体积.

连续型随机变量(X,Y)的边缘分布

边缘概率密度：设连续型随机变量(X,Y)的概率密度为$f(x,y)$，分量X（或Y）的概率密度称为(X,Y)关于X（或Y）的**边缘概率密度**，简称边缘密度，记为$f_X(x)$（或$f_Y(y)$），即

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy,\quad -\infty<x<+\infty;\quad f_Y(x)=\int_{-\infty}^{+\infty}f(x,y)dx,\quad -\infty<y<+\infty.$$

下面介绍两种重要的二维连续型随机变量的分布：均匀分布与二维正态分布.

均匀分布：设D为平面上的有界区域，其面积为S且$S>0$，若二维随机变量(X,Y)的概率密度为$f(x,y)=\begin{cases}\dfrac{1}{S}, & (x,y)\in D,\\ 0, & 其他,\end{cases}$则称$(X,Y)$服从区域$D$上的均匀分布（或称$(X,Y)$在$D$上服从均匀分布），记作$(X,Y)\sim U_D$.

特殊情形：

（1）D为矩形区域$a\le x\le b$，$c\le y\le d$. 此时，

$$f(x,y)=\begin{cases}\dfrac{1}{(b-a)(d-c)}, & a\le x\le b, c\le y\le d,\\ 0, & 其他.\end{cases}.$$

（2）D为圆形区域，若(X,Y)在以原点为圆心、R为半径的圆域上服从均匀分布，则(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi R^2}, & x^2+y^2\le R^2,\\ 0, & 其他.\end{cases}.$$

正态分布：若二维随机变量(X,Y)概率密度为

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

$$(-\infty<x<+\infty,-\infty<y<+\infty),$$

其中 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 都是常数，且 $\sigma_1>0$，$\sigma_2>0$，$|\rho|<1$. 记二维正态分布为 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

第二节 随机变量的独立性

1. 两个随机变量的独立性

二维随机变量的独立性：设 $F(x,y)$，$F_X(x)$ 和 $F_Y(x)$ 分别是二维随机变量 (X,Y) 的分布函数和两个边缘分布函数. 若对任意实数 x,y，有 $F(x,y)=F_X(x)F_Y(x)$，则称 X 与 Y 相对独立.

注：（1）$F(x,y)=F_X(x)F_Y(x)$ 等价于对任意实数 x,y，有

$$P\{X\le x,Y\le y\}=P\{X\le x\}P\{Y\le y\}.$$

（2）随机变量 X 与 Y 相互独立，即对任意实数 x,y，事件 $\{X\le x\}$ 与 $\{Y\le y\}$ 相互独立.

2. 二维离散型随机变量的独立性

设 (X,Y) 为离散型随机变量，则 X 与 Y 独立

$\Leftrightarrow$ 对一切 i,j，有 $P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$，$P_{ij}=P_{i\cdot}P_{\cdot j}$.

注意：只要有一对 (i,j) 值使得上式不成立，则 X,Y 不独立.

3. 二维连续型随机变量的独立性

设二维连续型随机变量 (X,Y) 的概率密度为 $f(x,y)$，$f_X(x)$，$f_Y(y)$ 分别为 (X,Y) 关于 X 和 Y 的边缘概率密度，则 X 与 Y 相互独立的充要条件是：等式 $f(x,y)=f_X(x)f_Y(y)$，几乎处处成立（即，在平面上除去“面积为零”的集合外，处处成立）.

联合分布与边缘分布的关系——联合分布可确定边缘分布，但一般情况下，边缘分布不

能确定联合分布. 只有当 X 与 Y 相互独立时，(X,Y) 的分布可由它的两个边缘分布完全确定.

4. n 维随机变量

边缘概率分布函数和边缘概率密度： 设 $(X_1,X_2,\cdots,X_n)$ 的分布函数为

$$F(X_1,x_2,\cdots,x_n)=P\{X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n\},$$

其概率密度为 $f(x_1,x_2,\cdots,x_n)$，则函数

$$F_{X_i}(x_i)=P\{X_1<+\infty,X_2<+\infty,\cdots,X_i\le x_i,\cdots,X_n<+\infty\}$$

和 $f_{X_i}(x_i)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_n)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n$ 分别称为 $(X_1,X_2,\cdots,X_n)$ 关于 X_i 的**边缘概率分布函数和边缘概率密度**，$i=1,2,\cdots,n$.

相互独立： 若对一切 $x_1,x_2,\cdots,x_n$ 有 $P\{X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n\}=\prod_{i=1}^{n}P\{X_i\le x_i\}$

即，$F(x_1,x_2,\cdots,x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$，则称 $X_1,X_2,\cdots,X_n$ 是**相互独立**的.

第三节 两个随机变量的函数的分布

二维随机变量 $Z=g(X,Y)$ 的分布

1. 离散型随机变量的函数的分布

设 (X,Y) 是二维离散型随机变量，其分布律为 $P\{X=x_i,Y=y_j\}=p_{ij}$，$i,j=1,2,\cdots$，则 $Z=g(X,Y)$ 也是离散型随机变量，其分布律为 $P\{Z=g(x_i,y_j)\}=p_{ij}$，$i,j=1,2,\cdots$. 若不同的 (x_i,y_j) 有相同的 $g(x_i,y_j)$，则相应的概率应相加合并.

2. 两个独立连续型随机变量之和的概率分布

设 (X,Y) 是二维连续型随机变量，其概率密度为 $f(x,y)$，当 $g(x,y)$ 是二维连续可微函数时，$Z=g(X,Y)$ 的分布函数为

$$F_Z(z)=P\{Z\le z\}=P\{g(X,Y)\le z\}=\iint_{g(x,y)\le z}f(x,y)dxdy,$$

概率密度为 $f_Z(z)=F_Z'(z)$.

（1）$Z=X+Y$ 的分布：$Z=X+Y$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx=\int_{-\infty}^{+\infty} f(z-y,y)dy\,.$$

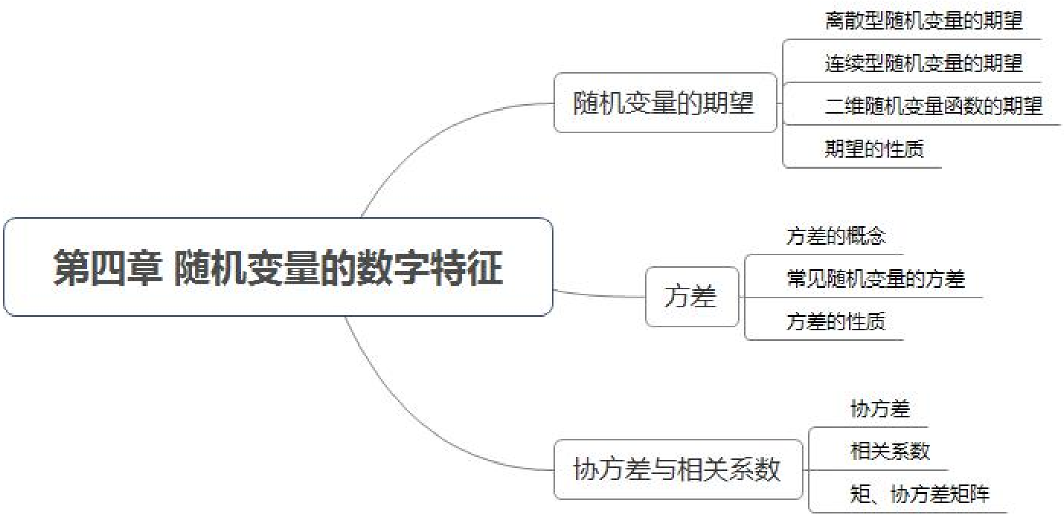
特别地，当 X 与 Y 独立时，则 $f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx$.

（2）若 $X_1,X_2,\cdots,X_n$ 相互独立，且服从正态分布 $X_i\sim N(\mu_i,\sigma_i^2)$ $(i=1,2,\cdots,n)$，则 n 个独立正态随机变量的线性组合仍服从正态分布，即

$$X=a_1X_1+a_2X_2+\cdots+a_nX_n\sim N\left(\sum_{i=1}^n a_i\mu_i,\sum_{i=1}^n a_i^2\sigma_i^2\right).$$

第四章 随机变量的数字特征

框架图



第一节 随机变量的期望

1. 离散型随机变量的期望

离散型随机变量的期望：设离散型随机变量 X 的分布律为

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots.$$

若 $\sum_i x_i p_i$ 绝对收敛（即 $\sum_i |x_i| p_i$ 收敛），则定义 X 的数学期望（简称均值或期望）为

$$E(X)=\sum_i x_i p_i .$$

注意：（1）当 X 的可能取值为有限多个 $x_1,x_2,\cdots,x_n$ 时，$E(X)=\sum_{i=1}^{n} x_i p_i$，

（2）当 X 的可能取值为可列多个 $x_1,x_2,\cdots,x_n,\cdots$ 时，$E(X)=\sum_{i=1}^{\infty} x_i p_i$.

（3）$\sum_i x_i p_i$ 绝对收敛，数学期望才存在，否则数学期望不存在.

定理 1 设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k$，$k=1,2,\cdots$. 令 $Y=g(X)$，

若 $\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛，则随机变量 Y 的数学期望为 $E(Y)=E[g(X)]=\sum_{k=1}^{\infty} g(x_k)p_k$.

2. 连续型随机变量的期望

连续型随机变量的数学期望：设连续型随机变量 X 的概率密度为 $f(x)$，若反常积分

$\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，则称该积分为随机变量 X 的**数学期望**（简称期望或均值），记为

$E(X)$，即 $E(X)=\int_{-\infty}^{+\infty} xf(x)dx$.

注：反常积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，数学期望才存在，否则数学期望不存在.

定理 2 设 X 为连续型随机变量，其概率密度为 $f_X(x)$，又随机变量 $Y=g(X)$，则当

$\int_{-\infty}^{+\infty} |g(x)| f_X(x)dx$ 收敛时，有 $E(Y)=E[g(X)]=\int_{-\infty}^{+\infty} g(x)f_X(x)dx$.

注意：随机变量的数学期望反映了随机变量取值的集中位置.

3. 二维随机变量函数的期望

定理 3 （1）若 (X,Y) 为离散型随机变量，若其分布律为 $p_{ij}=P\{X=x_i,Y=y_j\}$，边

缘分布律为

$$p_{i\cdot}=\sum_j p_{ij}, p_{\cdot j}=\sum_i p_{ij},$$

则 $E(X)=\sum_i x_i p_{i\cdot}=\sum_i\sum_j x_i p_{ij}$， $E(Y)=\sum_j y_j p_{\cdot j}=\sum_i\sum_j y_j p_{ij}$.

（2）若 (X,Y) 为二维连续型随机变量， $f(x,y)$， $f_X(x)$， $f_Y(y)$ 分别为 (X,Y) 的概率密度与边缘概率密度，则

$$E(X)=\int_{-\infty}^{+\infty}xf_X(x)dx=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy，\ E(Y)=\int_{-\infty}^{+\infty}yf_Y(y)dy=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dxdy.$$

定理 4 设 $g(X,Y)$ 为连续函数，对于二维随机变量 (X,Y) 的函数 $g(X,Y)$，

（1）若 (X,Y) 为离散型随机变量， $\sum_i\sum_j\left|g(x_i,y_j)\right|p_{ij}$ 收敛，则

$$E[g(X,Y)]=\sum_i\sum_j g(x_i,y_j)p_{ij}.$$

（2）若 (X,Y) 为连续型随机变量，且积分 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|g(x,y)|f(x,y)dxdy$ 收敛，则

$$E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy.$$

4. 期望的性质

（1）常数的期望等于这个常数，即 $E(C)=C$，其中 C 为常数.

（2）常数 C 与随机变量 X 乘积的期望等于该常数与随机变量 X 的期望的乘积，即

$$E(CX)=C\cdot E(X).$$

（3）随机变量和的期望等于随机变量期望之和，即 $E(X+Y)=E(X)+E(Y)$.

一般地，设 $X_1,X_2,\cdots,X_n$ 为 n 个随机变量，则有 $E\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n E(X_i)$；

$$E\left(\sum_{i=1}^n C_iX_i\right)=\sum_{i=1}^n C_iE(X_i)，其中 C_i(i=1,2,\cdots) 是常数.$$

（4）两个相互独立的随机变量乘积的期望等于期望的乘积，即若 X,Y 是相互独立的随机变量，则 $E(XY)=E(X)E(Y)$. 一般地，若 $X_1,X_2,\cdots,X_n$ 相互独立，则

$$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n).$$

（5） $[E(XY)]^2\le E(X^2)E(Y^2)$.

第二节 方差

1. 方差的概念

离差：对任一随机变量 X，设期望为 $E(X)$，记 $Y=X-E(X)$，称为随机变量 X 的**离差**. 离差 Y 代表随机变量 X 与期望之间的随机误差，其值可正可负.

方差：设随机变量 $(X-E(X))^2$ 的期望存在，则称 $E(X-E(X))^2$ 为随机变量 X 的**方差**，记作 $D(X)$，即

$$D(X)=E\left(X-E(X)^2\right).$$

称 $\sigma=\sqrt{D(X)}$ 为 X 的**标准差**（或均方差）.

注意：方差反映了随机变量偏离其中心——期望的平均偏离程度.

离散型：设 X 的分布律为 $P\{X=x_k\}=p_k, k=1,2,\cdots$，则 $D(X)=\sum_{i=1}^{n}(x_i-E(X))^2 p_i$.

连续型：设 X 的概率密度为 $f(x)$，则 $D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2 f(x)dx$.

2. 方差的性质

（1）常数的方差等于零，随机变量与常数之和的方差等于随机变量的方差，即

$$D(C)=0，\ D(X+C)=D(X).$$

（2）常数与随机变量乘积的方差等于这个常数的平方与随机变量方差的乘积. 即

$$D(CX)=C^2D(X)，\text{其中 } C \text{ 为常数}.$$

（3）设 X,Y 是两个随机变量，则有

$$D(X+Y)=D(X)+D(Y)+2E[(X-E(X))(Y-E(Y))].$$

特别地，若 X,Y 相互独立，则有 $D(X\pm Y)=D(X)+D(Y)$.

推广：n 个相互独立的随机变量情况——若 $X_1,X_2,\cdots,X_n$ 相互独立，则

$$D(k_1X_1+k_2X_2+\cdots+k_nX_n)=k_1D(X_1)+k_2D(X_2)+\cdots+k_nD(X_n)$$

3. 常见随机变量的方差

几种重要的随机变量的分布及其数字特征.

随机变量的分布及其数字特征

	分布	分布律或概率密度	期望 $E(X)$	方差 $D(X)$
离散型	X 服从参数为 p 的 0-1 分布	$P\{X=0\}=q$，$q=1-p$， $P\{X=1\}=p$，$0<p<1$	p	$p(1-p)$
	X 服从二项分布 $X\sim B(n,p)$	设 $X\sim B(n,p)$，即 $p_i=P\{X=i\}=C_n^i p^i q^{n-i}$ （$i=0,1,2,\cdots,n$），$q=1-p$	np	npq
	X 服从泊松分布 $X\sim P(\lambda)$	设 $X\sim P(\lambda)$ 其分布律为 $P\{X=i\}=\dfrac{\lambda^i e^{-\lambda}}{i!}$（$i=0,1,2,\cdots$）	λ	λ
连续型	X 服从均匀分布 $X\sim U(a,b)$	设随机变量 X 在 $[a,b]$ 上服从均匀分布，概率密度为 $f(x)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b,\\ 0, & 其他,\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
	X 服从指数分布 $X\sim E(\lambda)$	设随机变量 X 服从参数为 $\lambda>0$ 的指数分布，概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\le 0,\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
	X 服从正态分布 $X\sim N(\mu,\sigma^2)$	设 $X\sim N(\mu,\sigma^2)$ 概率密度为 $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$， $-\infty<x<+\infty$	μ	σ^2

第三节 协方差与相关系数

1. 协方差

协方差：设有二维随机变量(X,Y)，且$E(X)$，$E(Y)$存在，若$E[(X-E(X))(Y-E(Y))]$存在，则称此值为X与Y的**协方差**，记为$Cov(X,Y)$，即

$$Cov(X,Y)=E[(X-E(X))(Y-E(Y))].$$

当(X,Y)为二维离散型随机变量时，其分布律为$p_{ij}=P\{X=x_i,Y=y_i\}$（$i=1,2,\cdots$，$j=1,2,\cdots$），则$Cov(X,Y)=\sum_i\sum_j(x_i-E(X))(y_i-E(Y))p_{ij}$.

当(X,Y)为二维连续型随机变量时，$f(x,y)$为(X,Y)的概率密度，则

$$Cov(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-E(X))(y-E(Y))f(x,y)dxdy\cdot$$

协方差有计算公式：$Cov(X,Y)=E(XY)-E(X)E(Y)$. 特别地，取$X=Y$时，有$Cov(X,Y)=E[(X-E(X))(X-E(X))]=D(X)$.

协方差性质：

（1）$Cov(X,Y)=Cov(Y,X)$.

（2）$Cov(aX,bY)=abCov(Y,X)$，其中a，b为任意常数.

（3）$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$.

（4）若X与Y相互独立，则有$Cov(X,Y)=0$.

注：$Cov(X,Y)=0$是X与Y相互独立的必要不充分条件.

2. 相关系数

相关系数：若$D(X)>0$，$D(Y)>0$，称$\dfrac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$为$X$与$Y$的相关系数，记为$\rho_{XY}$，即$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$.

相关系数的性质：

（1）$|\rho_{XY}|\le 1$，$\rho_{xx}=1$

（2）$|\rho_{XY}|=1 \Leftrightarrow$ 存在常数 a,b 使 $P\{X=aX+b\}=1$ 且 $a\ne 0$.

不相关：若相关系数 $\rho_{XY}=0$ 则称 X 与 Y 不相关.

注意:(1)协方差与相关系数是描述两个随机变量之间线性相关的数字特征. $|\rho_{XY}|$ 越接近 1，X 与 Y 之间的线性关系越密切. 当 $|\rho_{XY}|=1$ 时，Y 与 X 存在完全线性关系，即 $Y=aX+b$；当 $|\rho_{XY}|=0$ 时，X 与 Y 之间无线性关系.

（2）当 $D(X)>0$，$D(Y)>0$ 时，随机变量 X 与 Y 不相关 $\Leftrightarrow Cov(X,Y)=0$.

若随机变量相互独立，则 $Cov(X,Y)=0$，从而，知 X 与 Y 不相关. 随机变量 X 与 Y 不相关，但 X 与 Y 不一定相关独立.

（3）若二维随机变量 (X,Y) 服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，$X$ 与 Y 的相关系数 $\rho_{XY}=\rho$，X 与 Y 不相关 $\Leftrightarrow$ X 与 Y 相互独立（X 与 Y 不相关和 X 与 Y 相互独立都等价于 $\rho=0$）.

3. 矩、协方差矩阵

K 阶原点距及中心矩：设 X 是随机变量，k 为正整数，若 $E(X^k)$ 存在，则称 $E(X^k)$ 为 X 的 **k 阶原点矩**，记为 v_k，即 $v_k=E(X^k)$. 若 $E[(X-E(X))^k]$ 存在，则称 $E[(X-E(X))^k]$ 为 X 的 **k 阶中心矩**，记为 μ_k，即 $\mu_k=E[(X-E(X))^k]$.

当 X 为离散随机变量时，设其概率分布律为 $P\{X=x_i\}=p_i$，$i=1,2,\cdots$，则

$$v_k=\sum_i x_i^k p_i,\quad \mu_k=\sum_i (x_i-E(X))^k p_i.$$

若 X 为连续型随机变量，其概率密度为 $f(x)$，则 $v_k=\int_{-\infty}^{+\infty}x^k f(x)dx$，

$$\mu_k=\int_{-\infty}^{+\infty}(x-E(X))^k f(x)dx.$$

一阶原点矩为数学期望：$v_1=E(X)$；二阶中心矩为方差：$\mu_2=D(X)$.

K+1 阶混合原点矩及中心矩：设 X，Y 为随机变量，若 $E(X^kY^l)\,(k,l=1,2,\cdots)$ 存在，

则称它为 X 和 Y 的 $k+l$ **阶混合原点矩**，若 $E\left\{[X-E(X)]^k[Y-E(Y)]^l\right\}$存在，则称它为 X 和 Y 的 $k+l$ **阶混合中心矩**. 协方差$Cov(X,Y)$是 X ，Y 的**二阶混合中心矩**.

四个二阶中心矩：将二维随机变量(X_1,X_2)的 4 个二阶中心矩

$$c_{11}=E[X_1-E(X_1)]^2=D(X_1)=Cov(X_1,X_1),$$

$$c_{12}=E[(X_1-E(X_1))(X_2-E(X_2))]=Cov(X_1,X_2),$$

$$c_{21}=E[(X_2-E(X_2))(X_1-E(X_1))]=Cov(X_2,X_1),$$

$$c_{22}=E[X_2-E(X_2)]^2=D(X_2)=Cov(X_2,X_2)$$

排成矩阵的形式，称此矩阵$C=(c_{ij})_{2\times2}$为随机变量(X_1,X_2)的协方差矩阵.

n 维随机变量协方差矩阵：设 n 维随机变量 $X_1,X_2,\cdots,X_n$ 的二阶混合中心矩

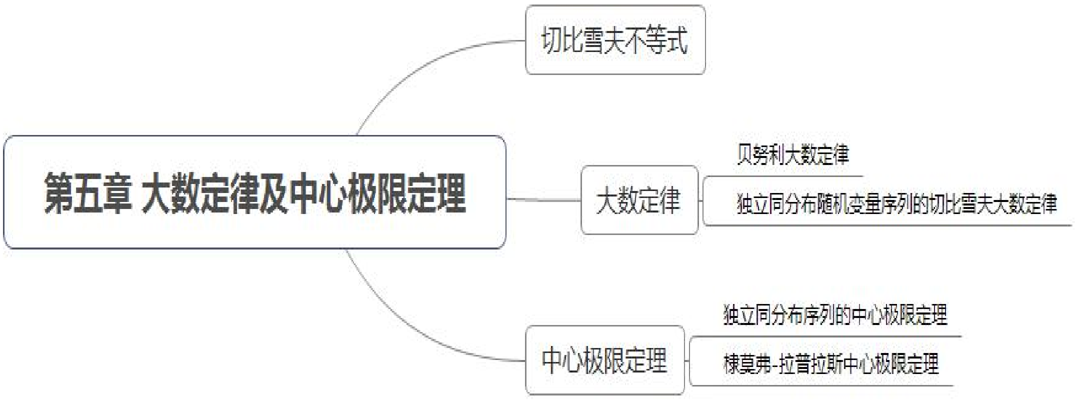
$$c_{ij}=E[(X_i-E(X_i))(X_j-E(X_j))]=Cov(X_i,X_j)\quad(i,j=1,2,\cdots,n)$$

存在，则称矩阵$C=(c_{ij})_{n\times n}$为 n 维随机变量 $X_1,X_2,\cdots,X_n$ 的协方差矩阵.

注：由于 $c_{ij}=c_{ji}$，$i,j=1,2,\cdots,n$，故协方差矩阵为对称矩阵；由于 $Cov(X_i,X_i)=D(X_i)$ $(i,j=1,2,\cdots,n)$，故协方差矩阵的对角线元素即为 $X_i(i=1,2,\cdots,n)$的方差.

第五章 大数定律及中心极限定理

框架图



第一节 切比雪夫 Chebyshev 不等式

1. 切比雪夫不等式

定理（切比雪夫不等式）设随机变量 X 的期望 $E(X)$ 及方差 $D(X)$ 存在，则对任意小正数 $\varepsilon>0$，有

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

说明：切比雪夫不等式仅用数学期望及方差就对随机变量在某范围取值的概率作出估计.

注意：（1）切比雪夫不等式在随机变量 X 分布未知的情况下给出了概率 $P\{|X-E(X)|<\varepsilon\}$ 的估计值.

（2）由切比雪夫不等式知，若方差越小，则概率 $P\{|X-E(X)|<\varepsilon\}$ 越大，表明方差是描述随机变量取值偏离数学期望程度的数字特征.

（3）当 $D(X)$ 很小时，X 落入区间 $(E(X)-\varepsilon,E(X)+\varepsilon)$ 几乎一定会发生；X 落入区间 $(E(X)-\varepsilon,E(X)+\varepsilon)$ 之外几乎一定不发生.

第二节 大数定律

1. 贝努利大数定律

定理 1 设 m 是 n 次独立重复试验中事件 A 发生的次数，p 是事件 A 的概率，则对任意正数 ε，有 $\lim\limits_{n\to\infty}P\left\{\left|\frac{m}{n}-p\right|<\varepsilon\right\}=1$.

贝努利大数定律表明：当 n 充分大时，“事件 A 发生的频率 $\frac{m}{A}$ 与概率 p 的绝对偏差小于任意给定的正数 ε”这一事件的概率可以任意接近于 1，即当 n 充分大时“频率与概率的绝对偏差小于任意给定的正数 ε”几乎必然发生,（即，为“概率是频率稳定值”的确切含义）.

2. 独立同分布随机变量序列的切比雪夫大数定律

定理 2 设 $X_1,X_2,\cdots,X_n,\cdots$ 是独立同分布随机变量序列，$E(X_i)=\mu$，$D(X_i)=\sigma^2$

$(i=1,2,\cdots)$均存在，则对于任意$\varepsilon>0$有$\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum\limits_{i=1}^{n}X_i-\mu\right|<\varepsilon\right\}=1$.

切比雪夫大数定律表明：经过算术平均后得到的随机变量$\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$在统计上具有一种稳定性，它的取值将比较紧密地聚集在它的期望附近. 贝努利大数定律是切比雪夫大数定律特殊情况.

第三节 中心极限定理

1. 独立同分布序列的中心极限定理

定理 1 设$X_1,X_2,\cdots,X_n,\cdots$是独立同分布的随机变量序列，且具有相同数学期望和方差$E(X_i)=\mu$，$D(X_i)=\sigma^2(i=1,2,\cdots)$. 记随机变量$Y_n=\dfrac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$的分布函数为$F_n(x)$，则对于任意实数$x$，有

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}P\{Y_n\le x\}=\lim_{n\to\infty}P\left\{\frac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x),$$

其中$\Phi(x)$为标准正态分布函数.

说明：

（1）当n充分大时，独立同分布的随机变量之和$Z_n=\sum\limits_{i=1}^{n}X_i$的分布近似于正态分布$N(n\mu,n\sigma^2)$. n个独立同分布的正态随机变量之和服从正态分布. 中心极限定理表明，不论$X_1,X_2,\cdots,X_n,\cdots$同服从什么分布，当$n$充分大时，其和$Z_n$近似服从正态分布.

（2）若$X_1,X_2,\cdots,X_n,\cdots$的平均值$\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$，有

$$E(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu，\quad D(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}D(X_i)=\frac{\sigma^2}{n},$$

它的标准化随机变量为$Y_n=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$. 从而有$Y_n$的分布函数为$F_n(x)$，有

$$\lim_{n\to\infty}F_n(x)=\int_{-\infty}^{t}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x).$$

当n充分大时，独立同分布随机变量的平均值$\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$的分布近似于正态分布

$N\left(\mu,\dfrac{\sigma^2}{n}\right)$.

2. 棣莫弗—拉普拉斯中心极限定理

定理 2 （棣莫弗—拉普拉斯中心极限定理）设随机变量Z_n是n次独立重复试验中事件A发生的次数，p是事件A发生的概率，则对于任意实数x，

$$\lim_{n\to\infty}P\left\{\frac{Z_n-np}{\sqrt{npq}}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x),$$

其中$q=1-p$, $\Phi(x)$为标准正态分布函数.

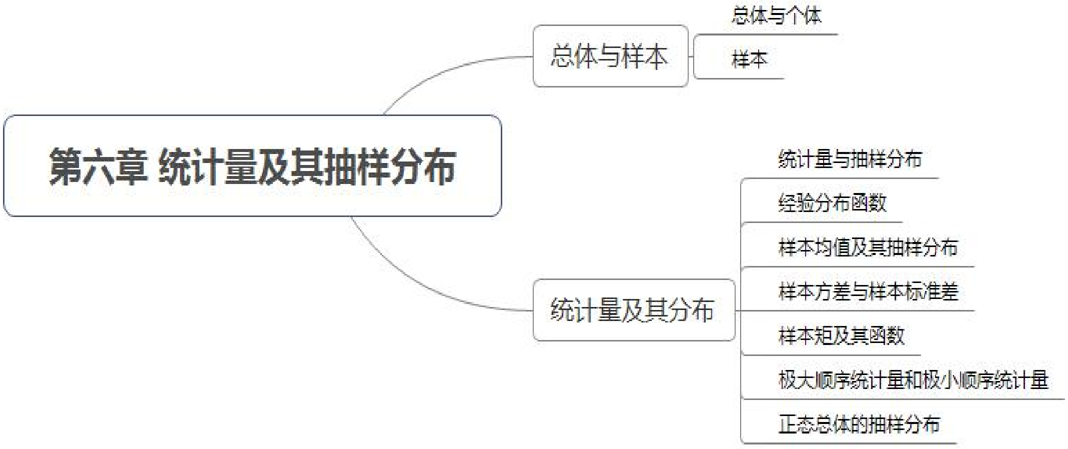
说明：

（1）在贝努利试验中，若事件A发生的概率为p. 又设Z_n为n次独立重复试验中事件A发生的频数，则当n充分大时，Z_n近似服从正态分布$N(np,npq)$.

（2）在贝努利试验中，若事件A发生的概率为p，$\dfrac{Z_n}{n}$为n次独立重复试验中事件A发生的频率，则当n充分大时，$\dfrac{Z_n}{n}$近似服从正态分布$N\left(p,\dfrac{pq}{n}\right)$.

第六章 统计量及其抽样分布

框架图



第一节 总体与样本

1. 总体与个体

总体：数理统计中，把研究对象的某个数量指标的全体称为**总体**（如研究某型号灯泡的寿命），它是一个随机变量，一般用 X,Y,Z 表示.

个体：组成总体的每一个元素称为**个体**（如灯泡的寿命），以 $X_1,X_2,\cdots,X_n$ 等表示，其中 X_n 表示第 n 个元素的取值（第 n 个灯泡的寿命）. 容量为有限的称为有限总体，容量为无限的称为无限总体.

2. 样本

样本：从总体 X 中按某种方式抽取部分个体 $X_1,X_2,\cdots,X_n$ 称为**样本**. 样本中包含个体的数量称为**样本容量**（或简称样本量）. 一般情况下，总是规定样本 $X_1,X_2,\cdots,X_n$ 是相互独立且与总体 X 同分布的随机变量，称之为**简单随机样本**（简称为样本）, 样本中的个体称为样品. 简单随机样本的一次抽样结果的 n 个具体数值 $(x_1,x_2,\cdots,x_n)$ 称为**样本观察值**. 因为样本是一组观测值，此时用小写字母 $x_1,x_2,\cdots,x_n$ 表示是恰当的.

简单随机样本：一个随机样本 $x_1,x_2,\cdots,x_n$，若满足：

（1）样本具有随机性，即 $x_1,x_2,\cdots,x_n$ 服从同一总体分布（说明：要求总体中每个个体都有同等机会被选入样本，也就是每一样品 x_i 与总体 X 有相同的分布）；

（2）样本要有独立性，即 $x_1,x_2,\cdots,x_n$ 相互独立（说明：要求样本中每一样品的取值不影响其他样品的取值，即 $x_1,x_2,\cdots,x_n$ 相互独立）.

则称为**简单随机样本**，简称样本. 若 $x_1,x_2,\cdots,x_n$ 是一个简单随机样本，则 $x_1,x_2,\cdots,x_n$ 是独立同分布的随机变量.

联合分布函数：设总体 X 具有分布函数 $F(x)$，$x_1,x_2,\cdots,x_n$ 为取自该总体容量为 n 的样本，则样本联合分布函数为 $F(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}F(x_i)$.

（1）若总体具有密度函数 $f(x)$，则样本的联合密度函数为 $f(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}f(x_i)$.

（2）若总体为离散型随机变量，则样本的（联合）概率函数为

$p(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}P\{X=x_i\}$，通常，样本分布指多维随机变量 $(x_1,x_2,\cdots,x_n)$ 的联合分布.

第二节 统计量及其分布

1. 统计量与抽样分布

统计量： 设 $x_1,x_2,\cdots,x_n$ 为取自某总体的样本，若样本函数 $T=T(x_1,x_2,\cdots,x_n)$ 中不含有任何未知参数，则称 T 为统计量. 统计量的分布称为抽样分布.

2. 经验分布函数

经验分布函数： 义 2 设 $x_1,x_2,\cdots,x_n$ 是取自总体分布函数为 $F(x)$ 的样本，若将样本观测值由小到大进行排列为 $x_{(1)},x_{(2)},\cdots,x_{(n)}$，则 $x_{(1)},x_{(2)},\cdots,x_{(n)}$ 称为有序样本，用有序样本定义如下函数

$$F_n(x)=\begin{cases}0, & x<x_{(1)}\\ \dfrac{k}{n}, & x_{(k)}\le x<x_{(k+1)},k=1,2,\cdots,n-1,\\ 1, & x>x_{(n)},\end{cases}$$

则 $F_n(x)$ 是一非减右连续函数，且满足 $F_n(-\infty)=0$ 和 $F_n(+\infty)=1$. 易知，$F_n(x)$ 是一分布函数，并称 $F_n(x)$ 为经验分布函数.

3. 样本均值及其抽样分布

样本均值： 设 $x_1,x_2,\cdots,x_n$ 为取自某总体的样本，其算术平均值称为样本均值，一般用 $\bar{x}$ 表示，即 $\bar{x}=\dfrac{x_1+x_2+\cdots+x_n}{n}=\dfrac{1}{n}\sum_{i=1}^{n}x_i$.

在分组样本场合，样本均值的近似公式为 $\bar{x}=\dfrac{x_1f_1+\cdots+x_kf_k}{n}\quad\left(n=\sum_{i=1}^{n}f_i\right)$. 其中 k 为

组数，x_i为第i组的组中值，f_i为第i组的频数.

样本均值的性质：

（1）若把样本中的数据与样本均值之差称为偏差，则样本所有偏差之和为 0，即

$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)=0$.

（2）数据观察值与均值的偏差平方和最小，即在形如$\sum\left(x_i-c\right)^2$的函数中，$\sum\left(x_i-\bar{x}\right)^2$最小，其中$c$为任意给定常数.

定理 1 设$x_1,x_2,\cdots,x_n$是来自某个总体X的样本，$\bar{x}$为样本均值.

（1）若总体分布为$N\left(\mu,\sigma^2\right)$，则$\bar{x}$的精确分布为$N\left(\mu,\dfrac{\sigma^2}{n}\right)$；

（2）若总体X分布未知（或不是正态分布），且$E\left(X\right)=\mu$，$D\left(X\right)=\sigma^2$，则当样本容量n较大时，$\bar{x}=\dfrac{1}{n}\sum_{i=1}^{n}x_i$的渐近分布（指$n$较大时的近似分布）为$N\left(\mu,\dfrac{\sigma^2}{n}\right)$.

4. 样本方差与样本标准差

样本方差：设$x_1,x_2,\cdots,x_n$为取自某总体的样本，则它关于样本均值$\bar{x}$的平均偏差平方和

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2,$$

称为**样本方差**，其算术根$s=\sqrt{s^2}$称为样本标准差. 样本标准差与样本均值具有相同的度量单位.

定理 2 设总体X具有二阶矩，即$E\left(X\right)=\mu$，$D\left(X\right)=\sigma^2<+\infty$，$x_1,x_2,\cdots,x_n$，为从该总体得到的样本，$\bar{x}$和$s^2$分别是样本均值和样本方差，则

$$E\left(\bar{x}\right)=\mu,\quad D\left(\bar{x}\right)=\frac{\sigma^2}{n},\quad E\left(s^2\right)=\sigma^2.$$

定理表明：样本均值的均值与总体均值相同，而样本均值的方差是总体方差的$\dfrac{1}{n}$.

5. 样本矩及其函数

样本矩：设$x_1,x_2,\cdots,x_n$是样本，则统计量$a_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k$，称为样本$k$阶原点矩.特别地，样本一阶原点矩就是样本均值.统计量$b_k=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^k$，称为样本$k$阶中心矩.

6. 极大顺序统计量和极小顺序统计量

顺序统计量：设总体X具有分布函数$F(x)$，分布密度$f(x)$，$x_1,x_2,\cdots,x_n$为其样本，我们分别称$x_{(1)}\triangleq\min\{x_1,x_2,\cdots,x_n\}$，$x_{(n)}\triangleq\max\{x_1,x_2,\cdots,x_n\}$为极小顺序统计量和极大顺序统计量.

定理3 若$x_{(1)}$，$x_{(n)}$分别为极小、极大顺序统计量，则$x_{(1)}$具有分布密度$f_1(x)=n(1-F(x))^{n-1}f(x)$，$x_{(n)}$具有分布密度$f_n(x)=nF^{n-1}(x)f(x)$.

7. 正态总体的抽样分布

三大抽样分布：χ^2分布（卡方分布）、t分布、F分布.

名称	定义	性质
χ^2分布	设$X_1,X_2,\cdots,X_n$独立同分布于标准正态分布$N(0,1)$，则$\chi^2=X_1^2+\cdots+X_n^2$的分布称为自由度为$n$的$\chi^2$分布，记为$\chi^2\sim\chi^2(n)$.	1. $E(\chi^2(n))=n$，$D(\chi^2(n))=2n$. 2. 设$X\sim\chi^2(m)$，$Y\sim\chi^2(n)$，且X与Y独立，则$X+Y\sim\chi^2(m+n)$.
F分布	设$X_1\sim\chi^2(m)$，$X_2\sim\chi^2(n)$，X_1与X_2独立，则称$F=\frac{X_1/m}{X_2/n}$的分布是自由度为m与n的F分布，记为$F\sim F(m,n)$，其中m称为分子自由度，n称为分母自由度.	1. 若$F\sim F(m,n)$，则$\frac{1}{F}\sim F(n,m)$. 2. $F_{1-\alpha}(m,n)=\frac{1}{F_\alpha(n,m)}$.
t分布	设随机变量X_1与X_2独立且$X_1\sim N(0,1)$，$X_2\sim\chi^2(n)$，则称$t=\frac{X_1}{\sqrt{X_2/n}}$的分布为自由度为$n$的$t$分布，记为$t\sim t(n)$	1. 自由度为1的t分布就是标准柯西分布，它的均值不存在； 2. $E(T)=0\ (n>1)$，$D(T)=\frac{n}{n-2}\quad(n>2)$；

		3. $t_{\frac{\alpha}{2}}(n)=-t_{1-\frac{\alpha}{2}}(n)$； 4. $\lim\limits_{n\to\infty}f(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

正态总体的样本均值与样本方差的分布

定理 4（单个正态总体的抽样分布）设 $x_1,x_2,\cdots,x_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

其样本均值和样本方差分别为 $\bar{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$ 和 $s^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2$，则有

（1）$\bar{x}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$，进一步有 $\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$；

（2）$\dfrac{\bar{x}-\mu}{s/\sqrt{n}}\sim t(n-1)$（（1）中 σ 未知时，以 s 替代 σ）；

（3）$\dfrac{(n-1)\cdot s^2}{\sigma^2}\sim\chi^2(n-1)$；

（4）$\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2=\sum\limits_{i=1}^{n}\left(\dfrac{x_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$；

（5）$\bar{x}$ 与 s^2 相互独立，且 $E(\bar{x})=\mu$，$E(s^2)=\sigma^2$.

定理 5（两个正态总体的抽样分布）设 $X_1,X_2,\cdots,X_n$ 与 $Y_1,Y_2,\cdots,Y_n$ 分别是来自正态总体 $X\sim N(\mu_1,\sigma_1^2)$ 和 $Y\sim N(\mu_2,\sigma_2^2)$ 的两个相互独立的样本，记

$$S_X^2=\frac{1}{m-1}\sum_{i=1}^{m}(X_i-\bar{X})^2,\quad S_Y^2=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2,$$

其中 $\bar{X}=\dfrac{1}{m}\sum\limits_{i=1}^{m}X_i$，$\bar{Y}=\dfrac{1}{n}\sum\limits_{i=1}^{n}Y_i$，则有

（1）$\bar{X}-\bar{Y}\sim N\left(\mu_1-\mu_2,\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}\right)$，$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}\sim N(0,1)$；

（2）$\dfrac{\sum\limits_{i=1}^{m}\left(X_i-\mu_1\right)^2/m\sigma^2}{\sum\limits_{i=1}^{n}\left(Y_i-\mu_2\right)^2/n\sigma^2}\sim F(m,n)$；

（3）$F=\dfrac{S_X^2/\sigma_1^2}{S_Y^2/\sigma_2^2}\sim F(m-1,n-1)$，若 $\sigma_1^2=\sigma_2^2$，则 $F=\dfrac{S_X^2}{S_Y^2}\sim F(m-1,n-1)$；

（4）$\dfrac{(m-1)S_X^2}{\sigma_1^2}+\dfrac{(n-1)S_Y^2}{\sigma_2^2}\sim\chi^2(m+n-2)$；

（5）当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时，记 $S_W^2=\dfrac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}$，则

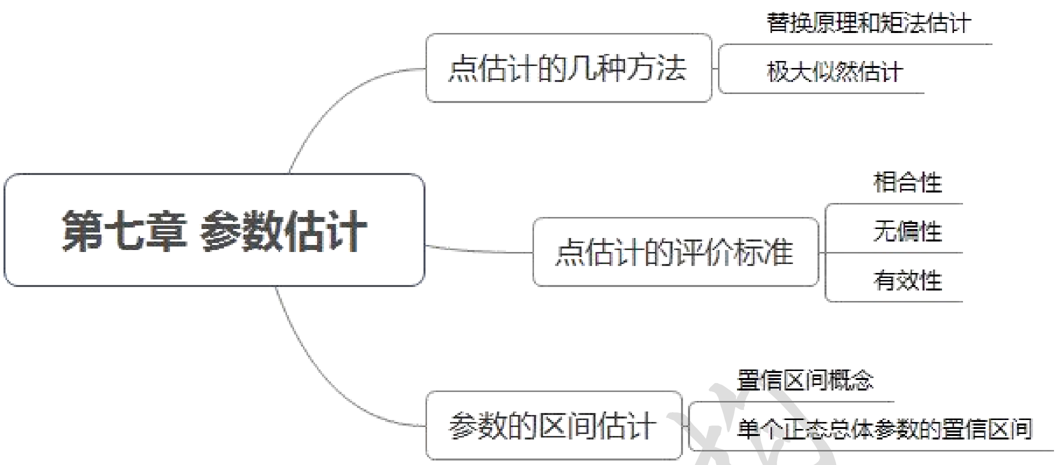
$$(m+n-2)S_W^2/\sigma^2\sim\chi^2(m+n-2),\quad \frac{\left(\overline{X}-\overline{Y}\right)-\left(\mu_1-\mu_2\right)}{S_W\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}\sim t(m+n-2).$$

注：若 $x_1,x_2,\cdots,x_n$ 为来自 $N\left(\mu,\sigma^2\right)$ 的样本，则 $\dfrac{\overline{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\sim\chi^2(n-1),\quad \frac{\overline{x}-\mu}{s/\sqrt{n}}\sim t(n-1).$$

第七章 参数估计

框架图


第七章 参数估计
点估计的几种方法
替换原理和矩法估计
极大似然估计
点估计的评价标准
相合性
无偏性
有效性
参数的区间估计
置信区间概念
单个正态总体参数的置信区间

第一节 点估计的几种方法

估计量：设总体 X 的分布函数为 $F(x;\theta)$，其中 θ 是未知参数（未知参数可以不止一个），$X_1,X_2,\cdots,X_n$ 是取自总体 X 的一个样本，用样本构造一个统计量 $T(X_1,X_2,\cdots X_n)$ 作为参数 θ 的估计，则称统计量 $T(X_1,X_2,\cdots,X_n)$ 为 θ 的**估计量**，记为 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$. 若 $(x_1,x_2,\cdots,x_n)$ 是样本的一个观察值，将其代入估计量 $\hat{\theta}$ 中得 $\hat{\theta}(X_1,X_2,\cdots,X_n)$，称之为 θ 的**估计值**. 估计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 与估计值 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 均称为 θ 的**点估计**.

参数估计的形式有两种：点估计和区间估计. 构造估计量的两种常用方法：矩估计法和极大似然估计法

1. 替换原理和矩法估计

矩估计法——替换原理（后称此法为矩法）：

（1）用样本矩替换总体矩（矩——可以是原点矩也可以是中心矩）；

（2）用样本矩的函数替换相应的总体矩的函数.

基本思想：在总体存在所需要的各阶矩的条件下，用样本的各阶矩去估计总体的相应的各阶矩（统计思想（替换思想），实质是用经验分布函数去替换总体分布）.

方法步骤：求两矩作方程，解方程得估计.

在总体分布形式未知场合可对各种参数作出估计，如：

用样本均值 $\bar{x}$ 估计总体均值 $E(X)$，即 $\hat{E}(X)=\overline{X}$；

用样本二阶中心矩 S_n^2 估计总体方差 $D(X)$，即 $D(X)=S_n^2$，其中 $S_n^2=\frac{1}{2}\sum_{i=1}^{n}(X_i-\overline{X})$；

用事件 A 出现的频率估计事件 A 发生的概率.

2. 极大似然估计法

极大似然估计：设总体的概率函数为 $p(x;\theta)$，$\theta\in\Theta$，其中 θ 是一个未知参数或几个未知参数组成的参数向量，Θ 是参数 θ 可能取值的参数空间，$x_1,\cdots,x_n$ 是来自总体的样本，将样本的联合概率函数看成 θ 的函数，用 $L(\theta;x_1,\cdots,x_n)$ 表示，简记为 $L(\theta)$，

$$L(\theta)=L(\theta;x_1,\cdots,x_n)=p(x_1;\theta)p(x_2;\theta)\cdots p(x_n;\theta),$$

$L(\theta)$ 称为样本的似然函数. 若某统计量 $\hat{\theta}=\hat{\theta}(x_1,\cdots,x_n)$ 满足 $\hat{L}(\theta)=\max\limits_{\theta\in\Theta}(\theta)$，则称 $\hat{\theta}$ 是 θ 的**极大似然估计**，简记为 MLE.

基本思想：根据样本的具体情况，选择参数 θ 的估计 $\hat{\theta}$ 使得该样本发生的概率最大. 用上述基本思想得到未知参数估计的方法称为**极大似然估计法.**

方法步骤："造似然"求导数，找驻点得估计.

（1）构造似然函数

设总体 X 分布中有 k 个未知参数 $\theta_1,\theta_2,\cdots,\theta_k$， $X_1,X_2,\cdots,X_n$ 是取自总体 X 的样本. 若总体是离散型，其分布律为 $P\{X=x_i\}=p_i(\theta_1,\theta_2,\cdots,\theta_k)$， $i=1,2,\cdots$，则似然函数为

$$L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)=P\{X_1=x_1,\cdots,X_n=x_n\}$$

$$=P\{X_1=x_1\}\cdots P\{X_n=x_n\}=\prod_{i=1}^{n}p_i(\theta_1,\cdots,\theta_k).$$

若总体是连续型，其概率密度为 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$，则似然函数为

$$L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)=f(x_1,x_2,\cdots,x_n)=f(x_1;\theta_1,\cdots,\theta_k)(x_2;\theta_1,\cdots,\theta_k)\cdots(x_n;\theta_1,\cdots,\theta_k)$$

$$=\prod_{i=1}^{n}f(x_i;\theta_1,\cdots,\theta_k).$$

（2）取对数（因 $L(x)$ 与 $\ln L(x)$ 的驻点相同）

$$\ln L(x)=\sum_{i=1}^{n}\ln p_i(\theta_1,\cdots,\theta_k)\ \text{或}\ \sum_{i=1}^{n}\ln f(x_i;\theta_1,\cdots,\theta_k).$$

（3）求导数找驻点得估计

$$\frac{\partial \ln L}{\partial \theta_i}=0，\ i=1,2,\cdots,k，$$

若方程有解，则可求得 θ_i 的最大似然估计值 $\hat{\theta}_i=\hat{\theta}_i(X_1,X_2,\cdots,X_n)$, $i=1,2,\cdots,k$.

第二节 点估计的评价标准

1. 相合性

相合估计：设 $\theta\in\Theta$ 为未知参数，$\hat{\theta}_n=\hat{\theta}_n(x_1,\cdots,x_n)$ 是 θ 的一个估计量，n 是样本容量，

若对任何一个 $\varepsilon > 0$，有 $\lim\limits_{n\to\infty} P\left\{\left|\hat{\theta}_n - \theta\right| > \varepsilon\right\} = 0$，则称 $\hat{\theta}_n$ 为参数 θ 的**相合估计**（或一致估计）.

定理 设 $\hat{\theta}_n = \hat{\theta}_n(x_1,\cdots,x_n)$ 是 θ 的一个估计量，若 $\lim\limits_{n\to\infty} E(\hat{\theta}_n) = \theta$，$\lim\limits_{n\to\infty} D(\hat{\theta}_n) = 0$，则 $\hat{\theta}_n$ 是 θ 是**相合估计**.

2. 无偏性

无偏估计：设 $\hat{\theta} = \hat{\theta}(x_1,\cdots,x_n)$ 是 θ 的一个估计，θ 的参数空间为 Θ，若对任意的 $\theta \in \Theta$，有 $E(\hat{\theta}) = \theta$，则称 $\hat{\theta}$ 是 θ 的**无偏估计**，否则称为有偏估计.

注意：设 $x_1, x_2, \cdots, x_n$ 是取自总体 X 的样本，$E(X) = \mu$，$D(X) = \sigma^2$，由于 $E(\overline{X}) = E(X) = \mu$，$E(S^2) = \sigma^2$，故样本均值 $\overline{X}$ 是总体均值的无偏估计量，样本方差 S^2 是总体方差 σ^2 的无偏估计量. 无偏性不具有不变性，从而 S 不是 σ 的无偏估计.

3. 有效性

有效：设 $\hat{\theta}_1, \hat{\theta}_2$ 是 θ 的两个无偏估计，若 $D(\hat{\theta}_1) \le D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 较 $\hat{\theta}_2$ 有效.

第三节 参数的区间估计

1. 置信区间概念

置信区间：设 θ 为总体的未知参数. $\hat{\theta}_1 = \hat{\theta}_1(x_1,\cdots,x_n)$，$\hat{\theta}_2 = \hat{\theta}_2(x_1,\cdots,x_n)$ 是由样本 $x_1, x_2, \cdots, x_n$ 定出的两个统计量，若对于给定的概率 $1-\alpha\ (0<\alpha<1)$，有 $P\{\hat{\theta}_1 \le \theta \le \hat{\theta}_2\} = 1-\alpha$，则随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 称为参数 θ 的置信度为 $1-\alpha$ 的**置信区间**，$\hat{\theta}_1$ 称为**置信下限**，$\hat{\theta}_2$ 称为**置信上限**.

置信区间的意义：θ 包含在随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 中的概率为 $100(1-\alpha)\%$；或者说，随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 以 $100(1-\alpha)\%$ 的概率包含 θ. α 常取的数值为 0.05，0.01，此时置信度 $1-\alpha$ 分别为 0.95，0.99.

2. 单个正态总体参数的置信区间

（1）σ 已知时 μ 的置信区间；

（2）σ 未知时 μ 的置信区间；

（3）σ^2 的置信区间；

所估参数	条件	估计函数	置信区间
μ	σ^2 已知	$u=\dfrac{(\bar{x}-\mu)}{\sigma}\sqrt{n}$	$\left[\bar{x}-u_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}},\bar{x}+u_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}\right]$
	σ^2 未知	$t=\dfrac{(\bar{x}-\mu)}{s}\sqrt{n}$	$\left[\bar{x}-t_{\frac{\alpha}{2}}(n-1)\dfrac{s}{\sqrt{n}},\bar{x}+t_{\frac{\alpha}{2}}(n-1)\dfrac{s}{\sqrt{n}}\right]$
σ^2	μ 未知	$\chi^2=\dfrac{(n-1)s^2}{\sigma^2}$	$\left[\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]$

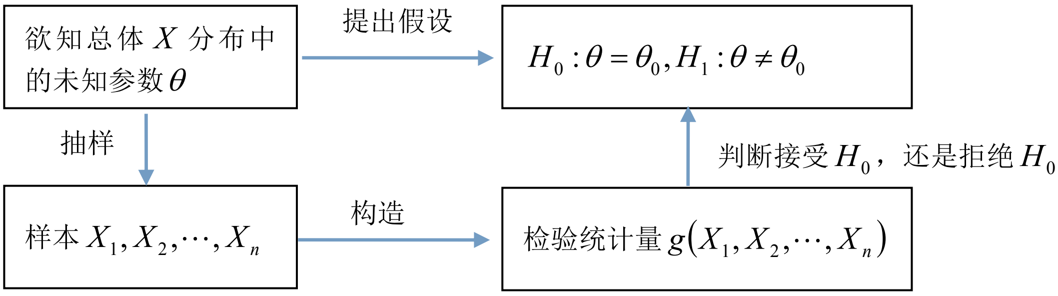
第八章 假设检验

框架图


第八章 假设检验
假设检验的基本思想和概念
基本思想
统计假设的概念
两类错误
假设检验的基本步骤
总体均值的假设检验
u检验
t检验
正态总体方差的假设检验
卡方检验
F检验
单边检验

第一节 假设检验的基本思想和概念

1. 假设检验的基本思想



判定的理论依据是小概率原理：小概率事件在一次试验中几乎不可能发生. 若在一次试验中小概率事件居然发生了，完全有理由拒绝 H_0 的正确性，否则只能接受 H_0，认同 H_0 的正确性.

2. 统计假设的概念

原假设和备择假设：当总体类型已知时，对分布的一个或几个未知参数的值作出假设，或者对总体分布的类型或某些特征提出某种假设，这种假设称为**原假设**或零假设，通常用 H_0 表示，在抛弃原假设后可供选择的假设称为**备择假设**，记为 H_1. H_0 与 H_1 是互不相容的.

注：原假设与备择假设的建立主要根据具体问题来解决. 一般把没有充分理由不能轻易否定的命题作为原假设，只有理由充分时才能拒绝它，而把其他容许的命题作为备择假设.

参数假设：在总体类型已知，仅有某个或某几个参数未知的情况下，对未知参数作出假设称为**参数假设**. 若对总体的某些特征提出假设，这称为非参数假设.

拒绝域与接受域：检验统计量把样本空间分成两个区域，使原假设 H_0 被拒绝的样本观察值 $x_1,x_2,\cdots,x_n$ 所组成的区域称为**拒绝域**. 而保留原假设 H_0 的样本观察值所组成的区域称为**接受域**. 拒绝域与接受域都是样本空间的子集，并且是互不相容的，而其区域之和为样本空间.

拒绝域与接受域的分界线处于一个特殊的地位，当样本越过这个分界线时，结论就发生了根本性的改变，因此把分界线的值称为**临界值**.

例：设总体 X 的分布是 $N(\mu,\sigma^2)$，且 σ^2 已知. 作假设

$$H_0:\mu=\mu_0\text{（}\mu_0\text{是已知数）}.$$

给定α（α是小概率），可得$u_{\frac{\alpha}{2}}$. 进行一次抽样得样本均值$\overline{x}$. 若H_0为真时，则

$u=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$，且$\overline{x}$应在$\mu_0$的两侧附近取值. 否则，若$\overline{x}$较$\mu_0$的偏离度较大，应视

为小概率事件发生，即$P\left\{|u|\geq u_{\frac{\alpha}{2}}\right\}=P\left\{\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq u_{\frac{\alpha}{2}}\right\}=\alpha$. 故认为原假设$H_0$有问题，故

应拒绝H_0，而接受H_1. 若$|u|\geq u_{\frac{\alpha}{2}}$，我们拒绝$H_0$；若$|u|<u_{\frac{\alpha}{2}}$则我们不拒绝$H_0$. 称

$u=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$为检验统计量，而称区域$\left\{(x_1,x_2,\cdots,x_n):|u|\geq u_{\frac{\alpha}{2}}\right\}$为拒绝域，简记为

$$W=\left\{|u|\geq u_{\frac{\alpha}{2}}\right\}.$$

3. 两类错误

第一类错误：原假设H_0为真，但由于样本的随机性，使样本观察值落入拒绝域，这时所作的判断便是拒绝H_0，这类错误称为第一类错误，简称弃真. 它发生的概率就是显著性水平α：

$$\alpha=P\{拒绝H_0\mid H_0为真\}.$$

第二类错误：原假设H_0为假，但由于样本的随机性，使样本观察值落入接受域，这时所作的判断便是保留H_0，这类错误称为第二类错误，简称采伪，它发生的概率记为β：

$$\beta=P\{接受H_0\mid H_1为真\}.$$

现列表说明两类错误

判断 真实情况	接受H_0 （$(x_1,x_2,\cdots,x_n)\notin W$）	拒绝H_0 （$(x_1,x_2,\cdots,x_n)\notin W$）
H_0成立	正确	第一类错误
H_1成立	第二类错误	正确

4. 假设检验的基本步骤

步骤：（1）看问题提假设——根据题意合理地建立原假设H_0和备择假设H_1；

（2）选统计量定分布——根据H_0的内容，选取适当的检验统计量，要求在H_0成立的条件下，能确定检验统计量的分布；

（3）给定显著性水平α的拒绝域——按问题具体要求选取α，构造小概率事件的拒绝域W（α是事先给定的，一般不超过0.1）

（4）代样本值作判断——由样本值计算出统计量的值，若该值落在拒绝域内，则拒绝原假设H_0，否则接受H_0. 以下例说明步骤：

设$X \sim N(\mu,\sigma^2)$，σ^2已知时，检验均值μ，

$$H_0:\mu=\mu_0，\ H_1:\mu\neq\mu_0\text{（看问题提假设）}$$

因$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$，（选统计量定分布）故由$P\left\{\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq z_{\frac{\alpha}{2}}\right\}=\alpha$（由给定的$\alpha$构造小概率事件得拒绝域），得拒绝域为$|Z|\geq z_{\frac{\alpha}{2}}$，若$\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|>z_{\frac{\alpha}{2}}$，则拒绝$H_0$，接受$H_1$；否则接受$H_0$.（代样本值作判断）.

第二节 总体均值的假设检验

1. u 检验

方差已知时，单个正态总体均值检验	方差已知时，两个正态总体均值检验
设 $x_1,\cdots,x_n$ 是从正态总体 $N\left(\mu,\sigma_0^2\right)$ 中抽取的一个样本，σ_0^2 是已知常数，欲检验假设：$H_0:\mu=\mu_0$，$H_1:\mu\neq\mu_0$，其中 μ_0 为已知数.	设 $X\sim N\left(\mu_1,\sigma_1^2\right)$，$X\sim N\left(\mu_2,\sigma_2^2\right)$，其中 σ_1^2，σ_2^2 为已知常数. $x_1,\cdots,x_m$ 和 $y_1,\cdots,y_n$ 分别是取自 X 和 Y 的样本且相互独立. 欲检验假设：$H_0:\mu_1=\mu_2$，$H_1:\mu_1\neq\mu_2$.
已求出检验统计量 $$u=\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}$$ （在假设 H_0 成立时，它服从标准正态分布），拒绝域 $$W=\left(-\infty,-u_{\frac{\alpha}{2}}\right)\bigcup\left(u_{\frac{\alpha}{2}},+\infty\right).$$ 若由样本检测值计算出 u 的值落在 W 内，则作出拒绝 H_0 的判定，否则认为与 H_0 相容.	检验假设 $\mu_1=\mu_2$，等价于检验假设 $$\mu_1-\mu_2=0.$$ 已知 $\bar{x}-\bar{y}$ 是 $\mu_1-\mu_2$ 的一个估计量，且当 H_0 为真时，有 $u=\dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}\sim N(0,1)$. 于是对给定的水平 α，可得临界值 $u_{\frac{\alpha}{2}}$，使 $$P\left\{\lvert u\rvert>u_{\frac{\alpha}{2}}\right\}=\alpha,$$ 从而得拒绝域 $W=\left(-\infty,-u_{\frac{\alpha}{2}}\right)\bigcup\left(u_{\frac{\alpha}{2}},+\infty\right)$. 再由样本值计算 u 的观测值 $$u=\frac{\bar{x}-\bar{y}}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}},$$ 若 $u\in W$，则拒绝 H_0；否则认为相容.

U 检验法：由服从标准正态分布的检验统计量作检验的方法称为 u **检验法.**

2. t 检验

<table>
<tr><th>方差未知时，单个正态总体均值检验</th><th>方差未知时，两个正态总体均值检验</th></tr>
<tr><td>设 $x_1,\cdots,x_n$ 是从正态总体 $N\left(\mu,\sigma^2\right)$ 中抽取的一个样本，其中 σ^2 未知，欲检验：
$H_0:\mu=\mu_0$, $H_1:\mu\neq\mu_0$, 其中 μ_0 为已知数.</td><td>设 $X\sim N\left(\mu_1,\sigma_1^2\right)$，$Y\sim N\left(\mu_2,\sigma_2^2\right)$，$x_1,\cdots,x_m$ 和 $y_1,\cdots,y_n$ 分别是取自 X 和 Y 的样本且相互独立. 欲检验假设：
$H_0:\mu_1=\mu_2$， $H_1:\mu_1\neq\mu_2$</td></tr>
<tr><td rowspan="2">由于 σ^2 未知，考虑用样本方差 s^2 代替总体方差 σ^2，因而构造检验统计量
$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}.$$
当 H_0 为真时，$t\sim t(n-1)$. 于是，对给定的显著性水平 α ，查表（t 分布临界值表）可得 $t_{\frac{\alpha}{2}}(n-1)$，使得
$$P\left\{|t|>t_{\frac{\alpha}{2}}(n-1)\right\}=\alpha,$$
即得拒绝域
$$W=\left(-\infty,-t_{\frac{\alpha}{2}}(n-1)\right)\cup\left(t_{\frac{\alpha}{2}}(n-1),+\infty\right).$$
通过样本观测值计算出观测值 t，若 $t\in W$ ，则拒绝 H_0，否则认为与 H_0 相容.</td><td>（1） $\sigma_1^2=\sigma_2^2=\sigma^2$ （σ^2 未知）
当 H_0 为真时，构造检验统计量：
$$t=\frac{\bar{x}-\bar{y}}{s_W\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim t(m+n-2).$$
对给定的水平 α ，得临界值 $t_{\frac{\alpha}{2}}(m+n-2)$ ，使得
$$P\left\{|t|>t_{\frac{\alpha}{2}}(m+n-2)\right\}=\alpha,$$
即得拒绝域
$$W=\left(-\infty,-t_{\frac{\alpha}{2}}(m+n-2)\right)\cup\left(t_{\frac{\alpha}{2}}(m+n-2),+\infty\right)$$
.</td></tr>
<tr><td>（2） $\sigma_1^2\neq\sigma_2^2$ 且未知，但 $m=n$ （配对问题）.
$$Z_i=X_i-Y_i\quad(i=1,\cdots,n),$$
即视两个正态总体样本之差来自一个正态总体的样本. 记 $E(Z_i)=\mu_1-\mu_2=d$ ，$D(Z_i)=\sigma_1^2+\sigma_2^2=\sigma^2$（未知），此时，$\mu_1$ 与 μ_2 是否相等的检验等价于假设检验：</td></tr>
</table>

	$H_0:d=0$，$H_1:d\neq 0$. 构造检验统计量 $t=\frac{\overline{Z}}{s}\sqrt{n}$，其中，$\overline{Z}=\frac{1}{n}\sum_{i=1}^{n}Z_i$，$s^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(Z_i-\overline{Z}\right)^2$. 在假设 H_0 为真时，$t\sim t(n-1)$. 于是可得拒绝域为 $W=\left(-\infty,-t_{\frac{\alpha}{2}}(n-1)\right)\cup\left(t_{\frac{\alpha}{2}}(n-1),+\infty\right)$.

第三节 正态总体方差的假设检验

1. χ^2 检验

设总体 $X\sim N\left(\mu,\sigma^2\right)$，$\sigma^2$ 未知，$x_1,\cdots,x_n$ 为取自 X 的样本，欲检验假设

$$H_0:\sigma^2=\sigma_0^2，\ H_1:\sigma^2\neq\sigma_0^2.$$

其中 σ_0^2 为已知数. 比较 σ^2 的无偏估计 s^2 的大小，当 H_0 为真时，s^2 应在 σ_0^2 周围波动，若 $\frac{s^2}{\sigma_0^2}$ 很大或很小，则否定 H_0，构造检验统计量

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}.$$

在假设 H_0 成立时，$\chi^2\sim\chi^2(n-1)$，于是对给定的水平 α，得 $\chi^2_{\frac{\alpha}{2}}(n-1)$ 与 $\chi^2_{1-\frac{\alpha}{2}}(n-1)$，使

$$P\left\{\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n-1)\right\}=P\left\{\chi^2>\chi^2_{\frac{\alpha}{2}}(n-1)\right\}=\frac{\alpha}{2}.$$

得拒绝域

$$W=\left(0,\chi^2_{1-\frac{\alpha}{2}}(n-1)\right)\cup\left(\chi^2_{\frac{\alpha}{2}}(n-1),+\infty\right),$$

若由样本观测值计算出 χ^2 的值 $\chi^2\in W$，则拒绝 H_0，否则认为 H_0 相容.

2. F 检验

设有两个正态总体 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，$x_1,\cdots,x_m$ 和 $y_1,\cdots,y_n$ 分别是取自 X 和 Y 的样本且相互独立. 欲检验统计假设 $H_0:\sigma_1^2=\sigma_2^2$，$H_1:\sigma_1^2\neq\sigma_2^2$. 由于 s_1^2 是 σ_1^2 的无偏估计，s_2^2 是 σ_2^2 的无偏估计，当 H_0 为真时，要求 s_1^2 与 s_2^2 不能相差太多，其比值 $\frac{s_1^2}{s_2^2}$ 不能太大也不能太小. 当 H_0 为真时，

$$F=\frac{s_1^2}{s_2^2}\sim F(m-1,n-1).$$

取 F 为检验统计量，对给定的水平 α，确定临界值

$$F_{\frac{\alpha}{2}}(m-1,n-1),\quad F_{1-\frac{\alpha}{2}}(m-1,n-1),$$

使

$$P\left\{F>F_{\frac{\alpha}{2}}(m-1,n-1)\right\}=P\left\{F\le F_{1-\frac{\alpha}{2}}(m-1,n-1)\right\}=\frac{\alpha}{2}.$$

即得拒绝域

$$W=\left(0,F_{1-\frac{\alpha}{2}}(m-1,n-1)\right)\cup\left(F_{\frac{\alpha}{2}}(m-1,n-1),+\infty\right).$$

若由样本观测值算得 F 值，当 $F\in W$ 时，拒绝 H_0，即认为两总体之方差有显著差异. 否则认为与 H_0 相容，即两总体之方差无明显差异.

第四节 单边检验

1. 单边检验

以单个正态总体方差已知的情况，讨论均值 μ 的单边检验的拒绝域.

设总体 $X\sim N(\mu,\sigma_0^2)$，σ_0^2 为已知，$x_1,\cdots,x_n$ 是取自 X 的一个样本，给定检验水平 σ_0 考虑单边假设检验问题：

$$H_0:\mu\le\mu_0,\quad H_1:\mu>\mu_0,$$

由于 $\bar{x}$ 是 μ 的无偏估计，故当 H_0 为真时，$u=\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}$ 不能太大，当 u 偏大时，拒绝 H_0，

故拒绝域的形式为：$\dfrac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}>c$，$c$ 待定.

由 $\dfrac{\overline{x}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1)$，故可找临界值 u_α，使 $P\left\{\dfrac{\overline{x}-\mu}{\sigma_0/\sqrt{n}}>u_\alpha\right\}=\alpha$. 当 H_0 成立时，

$$\frac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}\le\frac{\overline{x}-\mu}{\sigma_0/\sqrt{n}},$$

因此，

$$P\left\{\frac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}>\mu_\alpha\right\}\le P\left\{\frac{\overline{x}-\mu}{\sigma_0/\sqrt{n}}>\mu_\alpha\right\}=\alpha.$$

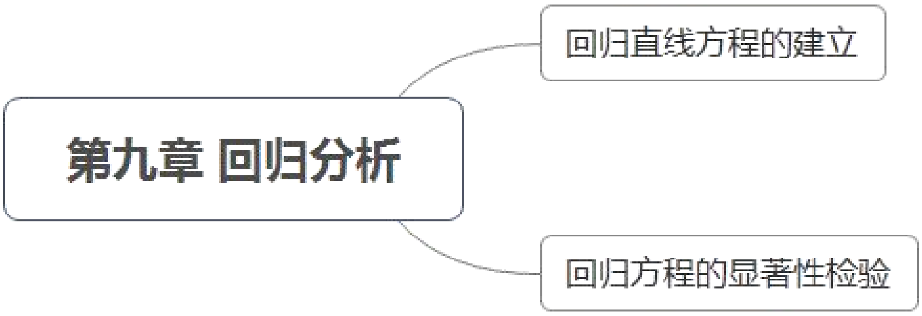
由事件 $\left\{\dfrac{\overline{x}-\mu}{\sigma_0/\sqrt{n}}>\mu_\alpha\right\}$ 是一个小概率事件知，事件 $\left\{\dfrac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}>\mu_\alpha\right\}$ 更为小概率事件.

若根据所给的样本观测值 $x_1,\cdots,x_n$，算出 $\dfrac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}>u_\alpha$，则应该否定原假设 H_0，即拒绝域为 $W=(\mu_\alpha,+\infty)$. 当 $\dfrac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}\le u_\alpha$ 时，承认原假设 $H_0:\mu\le\mu_0$.

类似地，对单边假设检验问题：$H_0:\mu\ge\mu_0$，$H_1:\mu<\mu_0$，仍取 $u=\dfrac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}$ 为检验统计量，拒绝域为 $W=(-\infty,-\mu_\alpha)$，即当由样本观察值算出，$\dfrac{\overline{x}-\mu_0}{\sigma_0/\sqrt{n}}<-u_\alpha$，则应拒绝原假设 H_0.

第九章 回归分析

框架图



第一节 回归直线方程的建立

1. 一元线性回归

一元线性回归：设随机变量 y 与自变量 x（它可精确测定或严格控制的变量）之间成立下列关系

$$y=\beta_0+\beta_1 x+\varepsilon,$$

其中 ε 为随机变量（随机误差），$\varepsilon \sim N\left(0,\sigma^2\right)$，称 $y=\beta_0+\beta_1 x$ 为 y 关于 x 的**一元线性回归**，称 β_2 为**回归系数**，β_0 为**回归常数**,它是回归直线的截距.

回归分析的基本问题是依据样本 $\left(x_i,y_i\right)$，$i=1,2,\cdots,n$ 解决问题：

（1）未知参数 β_0,β_1 及 σ^2 的点估计,求 Y 与 x 之间关系——回归方程.

（2）回归方差的显著性检验，验证 Y 与 x 之间关系是否存在关系式 $Y=\beta_0+\beta_1 x+\varepsilon$.

（3）利用回归方程进行预测和控制.

未知参数 β_0,β_1 及 σ^2 的点估计问题——最小二乘法：

已知样本 $\left(x_i,y_i\right)$，$i=1,2,\cdots n$，画出散点图，适当选取 β_0,β_1 作直线 $\hat{y}=\beta_0+\beta_1 x$，使直线与回归直线最接近，这时离差平方和 $\sum\limits_{i=1}^{n}\left[y_i-\left(\beta_0+\beta_1 x_i\right)\right]^2$ 达到最小. 对函数

$$Q\left(\beta_0,\beta_1\right)=\sum_{i=1}^{n}\left[y_i-\left(\beta_0+\beta_1 x_i\right)\right]^2$$

取关于 β_0 和 β_1 的偏导数，令其等于零，即 $\hat{\beta}_0,\hat{\beta}_1$ 为下列方程组的解：

$$\begin{cases} n\hat{\beta}_0+\left(\sum\limits_{i=1}^{n}x_i\right)\hat{\beta}_1=\sum\limits_{i=1}^{n}y_i, \\ \left(\sum\limits_{i=1}^{n}x_i\right)\hat{\beta}_0+\left(\sum\limits_{i=1}^{n}x_i^2\right)\hat{\beta}_1=\sum\limits_{i=1}^{n}x_i y_i. \end{cases}$$

解得

$$\begin{cases} \hat{\beta}_1=\dfrac{L_{xy}}{L_{xx}}, \\ \hat{\beta}_0=\overline{y}-\hat{\beta}_1\overline{x}, \end{cases}$$

则 $\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x$ 为所求线性回归方程，其中

$$\begin{cases} L_{xx}=\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2=\sum_{i=1}^{n}x_i^2-n\overline{x}^2, \\ L_{xy}=\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)=\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}, \\ L_{yy}=\sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2=\sum_{i=1}^{n}y_i^2-n\overline{y}^2. \end{cases}$$

β_0,β_1 的最小二乘估计 $\hat{\beta}_0,\hat{\beta}_1$ 具有如下性质：

（1） $E\left(\hat{\beta}_0\right)=\beta_0$， $E\left(\hat{\beta}_1\right)=\beta_1$；

（2） $D\left(\hat{\beta}_0\right)=\left(\frac{1}{n}+\frac{\overline{x}^2}{L_{xx}}\right)\sigma^2$， $D\left(\hat{\beta}_1\right)=\frac{1}{L_{xx}}\sigma^2$；

（3） $\hat{\beta}_0\sim N\left(\beta_0,\left(\frac{1}{n}+\frac{\overline{x}^2}{L_{xx}}\right)\sigma^2\right)$, $\hat{\beta}_1\sim N\left(\beta_1,\frac{1}{L_{xx}}\sigma^2\right)$.

第二节 回归方程的显著性检验

1. 显著性检验

对 y 和 x 是否具有线性关系作统计检验.

（1） F 检验法

若 y 与 x 之间不存在线性关系，则一次项系数 $\beta_1=0$，反之，$\beta_1\neq 0$. 检验 y 与 x 之间是否具有线性关系，归纳为检验假设：$H_0:\beta_1=0$， $H_1:\beta_1\neq 0$. 为了检验 H_0 是否为真，必须把由这个原因所引起的 y_i 的波动大小从 y_i 的总波动中分解开，记 $s_T=L_{yy}=\sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2$，称其为总的偏差平方和，它反映各 y_i 的波动大小.

$$s_T=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2+\sum_{i=1}^{n}\left(\hat{y}_i-\overline{y}\right)^2=s_{剩}+s_{回}.\text{（平方和分解式）}$$

其中

$$s_{回}=\sum_{i}\left(\hat{y}_i-\overline{y}\right)^2=\sum_{i}\hat{\beta}_1^2\left(x_i-\overline{x}\right)^2=\hat{\beta}_1^2L_{xx}=\hat{\beta}_1L_{xy}$$

反映了由于 x 的变化所引起的波动大小，称为**回归平方和**，而 $s_{剩}=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$，反映了观测值与回归直线间的偏差，这是由其他一切因素所引起的，称为**剩余平方和.**

于是，当 H_0 为真时，$F=\dfrac{s_{回}}{s_{剩}/(n-2)}\sim F(1,n-2)$.

对给定的显著水平 α，得临界值 $F_\alpha(1,n-2)$，则 H_0 的拒绝域为 $W=[F_\alpha(1,n-2),+\infty)$.

当 H_0 被拒绝时，认为回归效果是显著的，反之称回归方程不显著. 这种用 F 来检验回归方差显著与否的方法称为 F **检验法.**

注意：$\hat{\sigma}^2=\dfrac{s_{剩}}{n-2}=\dfrac{1}{n-2}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 为 σ^2 的无偏估计，且 $\hat{\beta}_1$ 与 $s_{剩}$ 相互独立.

（2）t 检验法

欲检验假设 $H_0:\beta_1=0$， $H_1:\beta_1\neq 0$. 由 $\hat{\beta}_1$ 与 $s_{剩}$ 相互独立，及 t 分布定义，知

$$t=\frac{\dfrac{\hat{\beta}_1-\beta_1}{\sqrt{\dfrac{\sigma^2}{L_{xx}}}}}{\sqrt{\dfrac{s_{剩}}{(n-2)\sigma^2}}}\sim t(n-2),$$

即，$t=\dfrac{\hat{\beta}_1-\beta_1}{\hat{\sigma}/\sqrt{L_{xx}}}\sim t(n-2)$. 当假设 $H_0:\beta_1=0$ 为真时，$t=\dfrac{\hat{\beta}_1}{\hat{\sigma}/\sqrt{L_{xx}}}\sim t(n-2)$.

其中，$\hat{\sigma}=\sqrt{\dfrac{s_{剩}}{n-2}}=\sqrt{\dfrac{1}{n-2}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$，对给定的显著性水平 α，得临界值 $t_{\frac{\alpha}{2}}(n-2)$. 从而得拒绝域

$$W=\left(-\infty,-t_{\frac{\alpha}{2}}(n-2)\right)\cup\left(t_{\frac{\alpha}{2}}(n-2),+\infty\right).$$

当由抽样得到的 n 对观测值 (x_i,y_i)，$i=1,2,\cdots,n$，计算出 $t\in W$ 时，拒绝 H_0，认为一元线性回归显著；否则，认为回归效果不显著.

用 $t=\dfrac{\hat{\beta}_1}{\hat{\sigma}/\sqrt{L_{xx}}}\sim t(n-2)$ 给出的检验统计量作检验的方法称为 t **检验法.**