《概率论与数理统计》
课程手册

- 内容提要
  - 疑难分析
    - 例题解析

目　　录

第一章　随机事件及其概率..... 3

第二章　随机变量及其分布..... 13

第三章　多维随机变量及其分布..... 24

第四章　随机变量的数字特征..... 34

第五章　大数定律和中心极限定理..... 41

第六章　数理统计的基本概念..... 55

第七章　参数估计..... 50

第八章　假设检验..... 55

第九章　方差分析和回归分析..... 60

第一章 随机事件及其概率

内 容 提 要

1、随机试验、样本空间与随机事件

（1）随机试验：具有以下三个特点的试验称为随机试验，记为 E.

1） 试验可在相同的条件下重复进行；

2） 每次试验的结果具有多种可能性，但试验之前可确知试验的所有可能结果；

3） 每次试验前不能确定哪一个结果会出现.

（2）样本空间：随机试验 E 的所有可能结果组成的集合称为 E 的样本空间，记为 Ω；试验的每一个可能结果，即 Ω 中的元素，称为样本点，记为 e.

（3）随机事件：在一次试验中可能出现也可能不出现的事件称为随机事件，简称事件，常用 A、B、C 等大写字母表示；可表述为样本空间中样本点的某个集合，分为复合事件和简单事件，还有必然事件（记为 Ω）和不可能事件（记为 Φ）.

2、事件的关系与运算

（1）包含关系与相等：“事件 A 发生必导致 B 发生”，记为 $A \subset B$ 或 $B \supset A$；

$A = B \Leftrightarrow A \subset B$ 且 $B \subset A$.

（2）和事件（并）：“事件 A 与 B 至少有一个发生”，记为 $A \cup B$.

（3）积事件（交）：“ 事件 A 与 B 同时发生”，记为 $A \cap B$ 或 AB.

（4）差事件、对立事件(余事件)：“事件 A 发生而 B 不发生”，记为 $A-B$ 称为 A 与 B 的差事件；$\Omega - B = \overline{B}$ 称为 B 的对立事件；易知： $A - B = A\overline{B}$.

（5）互不相容性： $AB = \phi$； A、B 互为对立事件 $\Leftrightarrow A \cup B = \Omega$ 且 $AB = \Phi$.

（6）事件的运算法则：1) 交换律： $A \cup B = B \cup A$， $AB = BA$ ；

2) 结合律： $A \cup (B \cup C) = (A \cup B) \cup C$， $(AB)C = A(BC)$；

3) 分配律： $(A \cup B)C = AC \cup BC$， $(AB) \cup C = (A \cup C)(B \cup C)$；

4) 对偶 (De Morgan) 律： $\overline{A \cup B} = \overline{A}\,\overline{B}$ ， $\overline{AB} = \overline{A} \cup \overline{B}$ ， 可推广

$\overline{\bigcup_k A_k}=\bigcap_k \overline{A_k}$，$\overline{\bigcap_k A_k}=\bigcup_k \overline{A_k}$.

3、频率与概率

（1）频率的定义：事件 A 在 n 次重复试验中出现 n_A 次，则比值 $\frac{n_A}{n}$ 称为事件 A 在 n 次重复试验中出现的频率，记为 $f_n(A)$，即 $f_n(A)=\frac{n_A}{n}$.

（2）统计概率：当 $n\to\infty$ 时，频率 $f_n(A)=\frac{n_A}{n}\to P(A)$. 当 n 很大时，$P(A)=P\approx f_n(A)$ 称为事件 A 的统计概率.

（3）古典概率：若试验的基本事件数为有限个，且每个事件发生的可能性相等，则试验对应古典概型（等可能概型），事件 A 发生的概率为：$P(A)=\frac{A\text{中所含样本点数}}{\Omega\text{中样本点总数}}=\frac{k}{n}=\frac{k(A)}{n}$.

（4）几何概率：若试验基本事件数无限，随机点落在某区域 g 的概率与区域 g 的测度(长度、面积、体积等)成正比，而与其位置及形状无关，则试验对应几何概型，“在区域 Ω 中随机地取一点落在区域 g 中”这一事件 A_g 发生的概率为：$P(A_g)=\frac{g\text{的测度}}{\Omega\text{的测度}}$.

（5）概率的公理化定义：设 (Ω,F) 为可测空间，在事件域 F 上定义一个实值函数 $P(A),\ A\in F$，满足：1）非负性：$P(A)\geq 0$，对任意 $A\in F$；2）规范性：$P(\Omega)=1$；3）可列可加性：若有一列 $A_i\in F, i=1,2,\cdots$，$A_iA_j=\Phi$，使得 $P(\bigcup_{j=1}^{\infty}A_j)=\sum_{j=1}^{\infty}P(A_j)$，则称 $P(A),\ A\in F$ 为 σ 域 F 上的概率测度，简称“概率”.

4、概率的基本性质

（1）不可能事件概率零：$P(\Phi)=0$.

（2）有限可加性：设 $A_1,A_2,\cdots,A_n$ 是 n 个两两互不相容的事件，即 $A_iA_j=\Phi$，（$i\neq j$），$i,j=1,2,\cdots n$，则有 $P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n)$.

（3）单调不减性：若事件 $B\supset A$，则 $P(B)\geq P(A)$，且 $P(B-A)=P(B)-P(A)$.（4）互补性：$P(\overline{A})=1-P(A)$，且 $P(A)\leq 1$.（5）加法公式：对任意两事件 A、B，有 $P(A\cup B)=P(A)+P(B)-P(AB)$；此性质可推广到任意 n 个事件

$A_1, A_2, \cdots, A_n$ 的情形.

（6）可分性：对任意两事件 A、B，有 $P(A) = P(AB) + P(A\overline{B})$.

5、条件概率与乘法公式

（1）条件概率：设 A、B 是 Ω 中的两个事件，即 A、$B \in F$，则 $P(B|A) = \dfrac{P(AB)}{P(A)}$ 称为事件 A 发生的条件下事件 B 发生的条件概率.

（2）乘法公式：设 A、$B \subset F$，则 $P(AB) = P(A)P(B|A) = P(B)P(A|B)$ 称为事件 A、B 的概率乘法公式.

6、全概率公式与贝叶斯(Bayes)公式

（1）全概率公式：设 $A_1, A_2, \cdots, A_n$ 是 Ω 的一个划分，且 $P(A_i) > 0$，$(i = 1,2,\cdots,n)$，则对任何事件 $B \in F$，有 $P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$，称为全概率公式.

（2）贝叶斯(Bayes)公式：设 $A_1, A_2, \cdots, A_n$ 是 Ω 的一个划分，且 $P(A_i) > 0$ $(i = 1,2,\cdots,n)$，则对任何事件 $B \in F$，有 $P(A_j|B) = \dfrac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}, (j = 1,\cdots,n)$，称为贝叶斯公式或逆概率公式.

7、事件的独立性

（1）两事件的独立：设 (Ω, F, P) 为一概率空间，事件 A、$B \in F$，且 $P(A) > 0$，若 $P(B) = P(B|A)$，则称事件 A 与 B 相互独立；等价于：$P(AB) = P(A)P(B)$.

（2）多个事件的独立：设 $A_1, A_2, \cdots, A_n$ 是 n 个事件，如果对任意的 $k(1 < k \le n)$，任意的 $1 \le i_1 < i_2 < \cdots < i_k \le n$，具有等式 $P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$，称 n 个事件 $A_1, A_2, \cdots, A_n$ 相互独立.

8、贝努里(Bernoulli)概型

（1）只有两个可能结果的试验称为贝努里试验，常记为 E. E 也叫做“成功—失败”试验，“成功”的概率常用 $p = P(A)$ 表示，其中 A =“成功”.

（2）把 E 重复独立地进行 n 次，所得的试验称为 n 重贝努里试验，记为 E^n.

（3）把 E 重复独立地进行可列多次，所得的试验称为可列重贝努里试验，记为 E^{∞}．以上三种贝努里试验统称为贝努里概型.

（4） E^n 中成功 k 次的概率是： $C_n^k p^k (1-p)^{n-k} = C_n^k p^k q^{n-k}, (0 \le k \le n)$ 其中 $p+q=1$.

疑 难 分 析

1、必然事件与不可能事件

必然事件是在一定条件下必然发生的事件，不可能事件指的是在一定条件下必然不发生的事件. 它们都不具有随机性，是确定性的现象，但为研究的方便，把它们看作特殊的随机事件.

2、互逆事件与互斥事件

如果两个事件 A 与 B 必有一个事件发生，且至多有一个事件发生，则 A 、 B 为互逆事件；如果两个事件 A 与 B 不能同时发生，则 A 、 B 为互斥事件. 因而，互逆必定互斥，互斥未必互逆. 区别两者的关键是：当样本空间只有两个事件时，两事件才可能互逆，而互斥适用与多个事件的情形. 作为互斥事件在一次试验中两者可以都不发生，而互逆事件必发生一个且只发生一个.

3、两事件独立与两事件互斥

两事件 A 、 B 独立，则 A 与 B 中任一个事件的发生与另一个事件的发生无关，这时 $P(AB)=P(A)P(B)$；而两事件互斥，则其中任一个事件的发生必然导致另一个事件不发生，这两事件的发生是有影响的，

这时 $AB=\Phi, P(AB)=0$. 可以用图形作一直观解释. 在图 1.1 左边的正方形中，

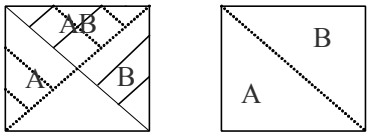

图 1.1

$P(AB)=\frac{1}{4}, P(A)=\frac{1}{2}=P(B)$，表示样本空间中两事件的独立关系，而在右边的正方形中，

$P(AB)=0$，表示样本空间中两事件的互斥关系.

4、条件概率 $P(A|B)$ 与积事件概率 $P(AB)$

$P(AB)$ 是在样本空间 Ω 内，事件 AB 的概率，而 $P(A|B)$ 是在试验 E 增加了新条件 B 发生后的缩减的样本空间 Ω_B 中计算事件 A 的概率. 虽然 A 、 B 都发生，但两者是不同的，一般说来，当 A 、 B 同时发生时，常用 $P(AB)$，而在有包含关系或明确的主从关系时，用 $P(A|B)$. 如袋中有 9 个白球 1 个红球，作不放回抽样，每次任取一球，取 2 次，求：（1）第二次才取到白球的概

率；（2）第一次取到的是白球的条件下，第二次取到白球的概率. 问题（1）求的就是一个积事件概率的问题，而问题（2）求的就是一个条件概率的问题.

5、全概率公式与贝叶斯(Bayes)公式

当所求的事件概率为许多因素引发的某种结果，而该结果又不能简单地看作这诸多事件之和时，可考虑用全概率公式，在对样本空间进行划分时，一定要注意它必须满足的两个条件. 贝叶斯公式用于试验结果已知，追查是何种原因（情况、条件）下引发的概率.

例 题 解 析

【例 1】写出下列随机试验的样本空间及下列事件包含的样本点：

（1）掷一棵骰子，出现奇数点.

（2）投掷一枚均匀硬币两次：

1）第一次出现正面；2）两次出现同一面；3）至少有一次出现正面.

（3）在 1，2，3，4 四个数中可重复地抽取两个数，其中一个数是另一个数的两倍.

（4）将 a,b 两只球随机地放到 3 个盒子中去，第一个盒子中至少有一个球.

分析：可对照集合的概念来理解样本空间和样本点：样本空间可指全集，样本点是元素，事件则是包含在全集中的子集.

解:(1) 掷一棵骰子，有六种可能结果，如果用“1”表示“出现 1 点”这个样本点，其余类似. 则样本空间为：Ω={1，2，3，4，5，6}，出现奇数点的事件为：{1，3，5}.

（2）投掷一枚均匀硬币两次，其结果有四种可能，若用（正，反）表示“第一次出现正面，第二次出现反面”这一样本点，其余类似. 则样本空间为：Ω={（正,正），（正,反），（反,正），（反,反）}，用 A、B、C 分别表示上述事件 1）、2）、3），则事件 A={（正,正），（正,反）}；事件 B={（正,正），（反,反）}；事件 C={（正,正），（正,反），（反,正）}.

（3）在 1，2，3，4 四个数中可重复地抽取两个数，共有 $4^2=16$ 种可能，若用 (i,j) 表示“第一次取数 i，第二次取数 j”这一样本点，则样本空间为：$\Omega=\{(i,j)\}\ (i,j=1,2,3,4)$；其中一个数是另一个数的两倍的事件为：{（1,2），（2,1），（2,4），（4,2）}.

(4)三个盒子分别记为甲、乙、丙，将 a,b 两只球随机地放到 3 个盒子中去共有九种结果. 若用(甲、乙)表示“a 球放入甲盒,b 球放入乙盒”这一样本点,其余类似. 则样本空间为：Ω={（甲，

甲），（甲，乙），（甲，丙），（乙，乙），（乙，甲），（乙，丙），（丙，甲），（丙，乙），（丙，丙）}；第一个盒子中至少有一个球的事件为：{（甲，甲），（甲，乙），（甲，丙），（乙，甲），（丙，甲）}.

【例 2】设 A、B、C 为三个事件，用 A、B、C 的运算关系表示下列各事件：

（1）仅 A 发生；　　（2）A 与 C 都发生，而 B 不发生；

（3）所有三个事件都不发生；（4）至少有一个事件发生；

（5）至多有两个事件发生；　（6）至少有两个事件发生；

（7）恰有两个事件发生；　　（8）恰有一个事件发生.

分析：利用事件的运算关系及性质来描述事件.

解：（1）$A\bar{B}\bar{C}$；（2）$A\bar{B}C$；（3）$\bar{A}\bar{B}\bar{C}$ 或 $\overline{A\cup B\cup C}$；（4）$A\cup B\cup C$ 或 $A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup ABC$；（5）$\bar{A}\cup\bar{B}\cup\bar{C}$ 或 $A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup\bar{A}\bar{B}\bar{C}$；

（6）$AB\cup AC\cup BC$ 或 $ABC\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$；

（7）$AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$；（8）$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$.

【例 3】把 n 个不同的球随机地放入 $N(N\geq n)$ 个盒子中，求下列事件的概率：

（1）某指定的 n 个盒子中各有一个球；

（2）任意 n 个盒子中各有一个球；

（3）指定的某个盒子中恰有 $m(m<n)$ 个球.

分析：这是古典概率的一个典型问题，许多古典概率的计算问题都可归结为这一类型. 每个球都有 N 种放法，n 个球共有 N^n 种不同的放法.“某指定的 n 个盒子中各有一个球”相当于 n 个球在 n 个盒子中的全排列；与（1）相比，（2）相当于先在 N 个盒子中选 n 个盒子，再放球；（3）相当于先从 n 个球中取 m 个放入某指定的盒中，再把剩下的 $n-m$ 个球放入 $N-1$ 个盒中.

解：样本空间中所含的样本点数为 N^n.

（1）该事件所含的样本点数是 $n!$，故：$p=n!/N^n$；

（2）在 N 个盒子中选 n 个盒子有 C_N^n 种选法，故所求事件的概率为： $p=C_N^n\cdot n!/N^n$；

（3）从 n 个球中取 m 个有 C_n^m 种选法，剩下的 $n-m$ 个球中的每一个球都有 $N-1$ 种放法，故所求事件的概率为： $p=C_N^n\cdot(N-1)^{n-m}/N^n$.

【例 4】随机地向由 $0<y<1,|x|<\frac{1}{2}$ 所围成的正方形内掷一点，点落在该正方形内任何区域的概率与区域面积成正比，求原点和该点的连线与 x 轴正向的夹角小于 $\frac{3}{4}\pi$ 的概率.

分析：这是一个几何概率问题，通常可借助几何上的度量（长度、面积、体积或容积等）来合理地规定其概率.

解：用 S 表示该正方形的面积， S_1 表示图 1.2 阴影部分 面积，则所求的概率为：

$$p=\frac{S_1}{S}=\frac{1-\frac{1}{2}(\frac{1}{2})^2}{1}=\frac{7}{8}.$$

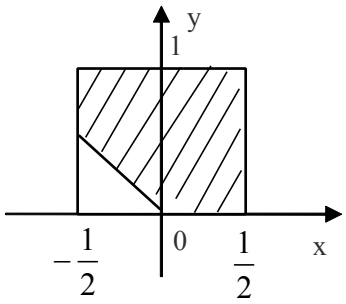

图 1.2

【例 5】设事件 A 与 B 互不相容，且 $P(A)=p,P(B)=q$，求下列事件的概率：

$P(AB),P(A\cup B),P(A\overline{B}),P(\overline{A}\overline{B})$.

分析：按概率的性质进行计算.

解： A 与 B 互不相容，所以 $AB=\Phi$， $P(AB)=P(\Phi)=0$；

$P(A+B)=P(A)+P(B)=p+q$；由于 A 与 B 互不相容，这时 $A\overline{B}=A$，从而

$P(A\overline{B})=P(A)=p$；由于 $\overline{A}\overline{B}=\overline{A\cup B}$，从而

$P(\overline{A}\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-(p+q)$.

【例 6】某住宅楼共有三个孩子，已知其中至少有一个是女孩，求至少有一个是男孩的概率（假设一个小孩为男或为女是等可能的）.

分析：在已知“至少有一个是女孩”的条件下求“至少有一个是男孩”的概率，所以是条件概率问题. 根据公式 $P(B|A)=\frac{P(AB)}{P(A)}$，必须求出 $P(AB),P(A)$.

解：设 A={至少有一个女孩}， B={至少有一个男孩}，则 $\overline{A}$={三个全是男孩}， $\overline{B}$={三个全是女孩}，于是

$P(\overline{A})=\frac{1}{2^3}=\frac{1}{8}=P(\overline{B})$，事件 AB 为“至少有一个女孩且至少有一个男孩”，因为 $\overline{AB}=\overline{A}\cup\overline{B}$，且 $\overline{A}\overline{B}=\Phi$，所以 $P(AB)=1-P(\overline{AB})=1-P(\overline{A}\cup\overline{B})=$ $1-[P(\overline{A})+P(\overline{B})]=1-(\frac{1}{8}+\frac{1}{8})=\frac{3}{4}, P(A)=1-P(\overline{A})=\frac{7}{8}$，从而，在已知至少有一个为女孩的条件下，求至少有一个是男孩的概率为： $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{3}{4}}{\frac{7}{8}}=\frac{6}{7}$.

【例 7】某电子设备制造厂所用的晶体管是由三家元件制造厂提供的. 根据以往的记录有以下的数据(表 1-1).

表 1-1

元件制造厂	次品率	提供晶体管的份额
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05

设这三家工厂的产品在仓库中均匀混合的，且无区别的标志.（1）在仓库中随机地取一只晶体管，求它是次品的概率.（2）在仓库中随机地取一只晶体管，若已知取到的是次品，为分析此次品出自何厂，需求出此次品由三家工厂生产的概率分别是多少. 试求这些概率.

分析：事件“取出的一只晶体管是次品”可分解为下列三个事件的和：“这只次品是一厂提供的”、“这只次品是二厂提供的”、“这只次品是三厂提供的”，这三个事件互不相容，可用全概率公式进行计算. 一般地，当直接计算某一事件 A 的概率 $P(A)$ 比较困难，而 $P(B_i), P(A|B_i)$ 比较容易计算，且 $\sum_i B_i=\Omega$ 时，可考虑用全概率公式计算 $P(A)$.（2）为条件概率，可用贝叶斯公式进行计算.

解：设 A 表示“取到的是一只次品”，$B_i(i=1,2,3)$ 表示“所取到的产品是由第 i 家工厂提供的”. 易知，B_1,B_2,B_3 是样本空间 Ω 的一个划分，且有

$P(B_1)=0.15, P(B_2)=0.80, P(B_3)=0.05, P(A|B_1)=0.02, P(A|B_2)=$

$0.01, P(A|B_3)=0.03$.

（1）由全概率公式：$P(A)=\sum_{i=1}^{3}P(B_i)P(A|B_i)=0.0125$.

（2）由贝叶斯公式：

$P(B_1|A)=\dfrac{P(A|B_1)P(B_1)}{P(A)}=0.24, P(B_2|A)=0.64, P(B_3|A)=0.12$. 以上结果表明，这只次品来自第二家工厂的可能性最大.

【例 8】一名工人照看 A、B、C 三台机床，已知在 1 小时内三台机床各自不需要工人照看的概率为 $P(\overline{A})=0.9, P(\overline{B})=0.8, P(\overline{C})=0.7$. 求 1 小时内三台机床至多有一台需要照看的概率.

分析：每台机床是否需要照看是相互独立的，这样，可根据事件的独立性性质及加法公式进行计算.

解：各台机床需要照看的事件是相互独立的，而三台机床至多有一台需要照看的事件 D 可写成：$D=\overline{A}\overline{B}\overline{C}+A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C$，则由加法公式与独立性性质得：

$P(D)=P(\overline{A}\overline{B}\overline{C}+A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C)=P(\overline{A}\overline{B}\overline{C})+$

$P(A\overline{B}\overline{C})+P(\overline{A}B\overline{C})+P(\overline{A}\overline{B}C)=P(\overline{A})P(\overline{B})P(\overline{C})+P(A)P(\overline{B})P(\overline{C})+$

$P(\overline{A})P(B)P(\overline{C})+P(\overline{A})P(\overline{B})P(C)$=0. 902.

【例 9】某车间有 10 台同类型的设备，每台设备的电动机功率为 10 千瓦. 已知每台设备每小时实际开动 12 分钟，它们的使用是相互独立的. 因某种原因，这天供电部门只能给车间提供 50 千瓦的电力. 问该天这 10 台设备能正常运作的概率是多少？

分析：由题意知，所要求的概率就是求“该天同时开动的设备不超过 5 台”这一事件的概率. 因为每台设备的使用是相互独立的，且在某一时刻，设备只有开动与不开动两种情况，所以本题可视为 10 重贝努里试验，可用二项概率公式进行求解.

解：设 A 表示事件“设备开动”， X 表示“同时开动的设备数”，则由二项概率公式得：

$P\{X=k\}=C_{10}^{k}(\frac{1}{5})^{k}(\frac{4}{5})^{10-k}$，同时开动不超过 5 台的概率：

$P\{X\le 5\}=P\{X=0\}+P\{X=1\}+\cdots+P\{X=5\}\approx 0.994$；

故该天这 10 台设备能正常运作的概率为 0. 994.

第二章　随机变量及其分布

内 容 提 要

1、随机变量

设 Ω 是随机试验的样本空间，如果对于试验的每一个可能结果 $\omega \in \Omega$，都有唯一的实数 $X(\omega)$ 与之对应，则称 $X(\omega)$ 为定义在 Ω 上的随机变量，简记为 X. 随机变量通常用大写字母 X、Y、Z 等表示.

2、分布函数及其性质

设 X 为随机变量，x 为任意实数，函数 $F(x)=P\{X\le x\}(-\infty<x<+\infty)$　称为随机变量 X 的分布函数.

分布函数完整地描述了随机变量取值的统计规律性，具有以下性质：

（1）$0\le F(x)\le 1 \qquad (-\infty<x<+\infty)$；

（2）如果 $x_1<x_2$，则 $F(x_1)\le F(x_2)$；

（3）$F(x)$ 为右连续，即 $F(x+0)=F(x)$；

（4）$\lim\limits_{x\to-\infty}F(x)=0,\ \lim\limits_{x\to+\infty}F(x)=1$；

（5）$P\{x_1<X\le x_2\}=P\{X\le x_2\}-P\{X\le x_1\}=F(x_2)-F(x_1)$.

3、离散型随机变量及其概率分布

如果随机变量 X 只能取有限个或可列个可能值，则称 X 为离散型随机变量. 如果 X 的一切可能值为 $x_1,x_2,\cdots$，并且 X 取 x_k 的概率为 p_k，则称

$p_k=P\{X=x_k\}(k=1,2,3,\cdots)$ 为离散型随机变量 X 的概率函数（概率分布或分布律）. 列成表格形式，也称为分布列(表 2-1)：

表 2-1

X	x_1	x_2	x_3	$\cdots$

P	p_1	p_2	p_3	$\cdots$

其中 $p_i \geq 0, \sum_i p_i = 1$.

常见的离散型随机变量的分布有：

（1）0-1 分布，记为 $X \sim (0-1)$，概率函数

$$P\{X=k\} = p^k (1-p)^{1-k}, k = 0,1, \quad 0 < p < 1;$$

（2）二项分布，记为 $X \sim B(n,p)$，概率函数

$$P\{X=k\} = C_n^k p^k (1-p)^{n-k}, k = 0,1,\cdots,n, \quad 0 < p < 1;$$

（3）泊松分布，记为 $X \sim P(\lambda)$，概率函数

$$P\{X=k\} = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0,1,\cdots, \quad \lambda > 0;$$

泊松定理　设 $\lambda > 0$ 是一常数，n 是任意正整数，设 $np_n = \lambda$，则对于任一固定的非负整数 k，

有 $\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$.

当 n 很大且 p 很小时，二项分布可以用泊松分布近似代替，即

$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!}$，其中 $\lambda = np$.

（4）超几何分布，记为 $X \sim H(n,M,N)$，概率函数

$P\{X=k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k = 0,1,\cdots,\min(n,M)$，其中 n、N、M 为正整数，且 $M \leq N, n \leq N$.

当 N 很大，且 $p = \frac{n}{N}$ 较小时，有 $\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \approx C_n^k p^k (1-p)^{n-k}$.］

（5）几何分布，记为 $X \sim G(p)$，概率函数

$$P\{X=k\} = p(1-p)^{k-1}, k = 0,1,\cdots, \quad 0 < p < 1.$$

4、连续型随机变量及其概率分布

如果对于随机变量 X 的分布函数 $F(x)$，存在非负函数 $f(x)$，使对于任一实数 x，有 $F(x)=\int_{-\infty}^{x} f(t)dt$，则称 X 为连续型随机变量. 函数 $f(x)$ 称为 X 的概率密度函数.

概率密度函数具有以下性质：

（1）$f(x)\geq 0$；　　（2）$\int_{-\infty}^{+\infty} f(t)dt=1$；

（3）$P\{x_1<X\leq x_2\}=\int_{x_1}^{x_2} f(t)dt$；　（4）$P\{X=x_1\}=0$；

（5）如果 $f(x)$ 在 x 处连续，则 $F'(x)=f(x)$.

常见的连续型随机变量的分布有：

（1）均匀分布，记为 $X\sim U(a,b)$，概率密度为

$f(x)=\begin{cases}\dfrac{1}{b-a}, a\leq x\leq b,\\ 0, \text{其它}\end{cases}$. 相应的分布函数为 $F(x)=\begin{cases}0, x<a\\ \dfrac{x-a}{b-a}, a\leq x\leq b\\ 1, x>b\end{cases}$；

（2）指数分布，记为 $X\sim E(\lambda)$，概率密度为

$f(x)=\begin{cases}\lambda e^{-\lambda}, x\geq 0,\\ 0, \text{其它}\end{cases}$. 相应的分布函数为 $F(x)=\begin{cases}1-e^{-\lambda x}, x\geq 0\\ 0, x<0\end{cases}$；

（3）正态分布，记为 $X\sim N(\mu,\sigma^2)$，概率密度为

$f(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty<X<+\infty$，相应的分布函数为 $F(x)=\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dt$；

当 $\mu=0,\sigma=1$ 时，即 $X\sim N(0,1)$ 时，称 X 服从标准正态分布. 这时分别用 $\varphi(x)$ 和 $\Phi(x)$ 表示 X 的密度函数和分布函数，即 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \Phi(x)=\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$. 具有性质：

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

一般正态分布 $X\sim N(\mu,\sigma^2)$ 的分布函数 $F(x)$ 与标准正态分布的分布函数 $\Phi(x)$ 有关系：

$F(x)=\Phi(\dfrac{x-\mu}{\sigma})$.

5、随机变量函数的分布

（1）离散型随机变量函数的分布

设 X 为离散型随机变量，其分布列为(表 2-2)：

表 2-2

X	x_1	x_2	x_3	…	x_n	…
P	p_1	p_2	p_3	…	p_n	…

则 $Y=g(X)$ 任为离散型随机变量，其分布列为(表 2-3)：

表 2-3

Y	$y_1=g(x_1)$	$y_2=g(x_2)$	$y_3=g(x_3)$	…	$y_n=g(x_n)$	…
P	p_1	p_2	p_3	…	p_n	…

y_i 有相同值时，要合并为一项，对应的概率相加.

（2）连续型随机变量函数的分布

设 X 为离散型随机变量，概率密度为 $f_X(x)$，则 $Y=g(X)$ 的概率密度有两种方法可求.

1）定理法：若 $y=g(x)$ 在 X 的取值区间内有连续导数 $g'(x)$，且 $g(x)$ 单调时，$Y=g(X)$ 是连续型随机变量，其概率密度为

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

其中 $\alpha=\min\{g(-\infty),g(+\infty)\}, \beta=\max\{g(-\infty),g(+\infty)\}$. $h(y)$ 是 $g(x)$ 的反函数.

2）分布函数法：先求 $Y=g(X)$ 的分布函数

$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\sum\limits_k \int\limits_{\Delta_k(y)} f_x(x)dx$，然后求 $f_Y(y)=[F_Y(y)]'$.

疑 难 分 析

1、随机变量与普通函数

随机变量是定义在随机试验的样本空间 Ω 上，对试验的每一个可能结果 $\omega \in \Omega$，都有唯一的实数 $X(\omega)$ 与之对应. 从定义可知：普通函数的取值是按一定法则给定的，而随机变量的取值是由统计规律性给出的，具有随机性；又普通函数的定义域是一个区间，而随机变量的定义域是样本空间.

2、分布函数 $F(x)$ 的连续性

定义左连续或右连续只是一种习惯. 有的书籍定义分布函数 $F(x)$ 左连续，但大多数书籍定义分布函数 $F(x)$ 为右连续. 左连续与右连续的区别在于计算 $F(x)$ 时，$X=x$ 点的概率是否计算在内. 对于连续型随机变量，由于 $P\{X=x_1\}=0$，故定义左连续或右连续没有什么区别；对于离散型随机变量，由于 $P\{X=x_1\}\neq 0$，则定义左连续或右连续时 $F(x)$ 值就不相同，这时，就要注意对 $F(x)$ 定义左连续还是右连续.

例 题 解 析

【例 1】分析下列函数是否是分布函数. 若是分布函数，判断是哪类随机变量的分布函数.

（1） $F(x)=\begin{cases}0, & x<-2,\\ \dfrac{1}{2}, & -2\le x<0,\\ 1, & x\ge 0.\end{cases}$ （2） $F(x)=\begin{cases}0, & x<0,\\ \sin x, & 0\le x<\pi,\\ 1, & x\ge \pi.\end{cases}$

（3） $F(x)=\begin{cases}0, & x<0,\\ x+\dfrac{1}{2}, & 0\le x<\dfrac{1}{2},\\ 1, & x\ge \dfrac{1}{2}.\end{cases}$

分析：可根据分布函数的定义及性质进行判断.

解：（1） $F(x)$ 在 $(-\infty,+\infty)$ 上单调不减且右连续. 同时，$\lim\limits_{x\to-\infty}F(x)=0,\lim\limits_{x\to+\infty}F(x)=1$. 故 $F(x)$ 是随机变量的分布函数. 有 $F(x)$ 的图形可知是阶梯形曲线，故 $F(x)$ 是离散型随机变量的分布函数；

（2）由于 $F(x)$ 在 $[\dfrac{\pi}{2},\pi]$ 上单调下降，故 $F(x)$ 不是随机变量的分布函数. 但只要将 $F(x)$ 中的 π 改为 $\dfrac{\pi}{2}$，$F(x)$ 就满足单调不减右连续，且 $\lim\limits_{x\to-\infty}F(x)=0,\lim\limits_{x\to+\infty}F(x)=1$，这时 $F(x)$ 就是随

机变量的分布函数. 由 $F(x)$ 可求得 $f(x)=F'(x)=\begin{cases}0, & \text{其它},\\ \cos x, 0<x\le\dfrac{\pi}{2}.\end{cases}$ 显然， $F(x)$ 是连续型随机变量的分布函数；

（3） $F(x)$ 在（$-\infty,+\infty$）上单调不减且右连续，且 $F(-\infty)=0, F(+\infty)=1$，是随机变量的分布函数. 但 $F(x)$ 在 $x=0$ 和 $x=\dfrac{1}{2}$ 处不可导，故不存在密度函数 $f(x)$，使得 $\int_{-\infty}^{x}f(x)dx=F(x)$. 同时， $F(x)$ 的图形也不是阶梯形曲线，因而 $F(x)$ 既非连续型也非离散型随机变量的分布函数.

【例 2】盒中装有大小相等的球 10 个，编号分别为 0、1、2、…、9. 从中任取 1 个，观察号码是“小于 5”、“等于 5”、“大于 5”的情况. 试定义一个随机变量，求其分布律和分布函数.

分 析 :“任取 1 球的号码”是随机变量，它随着试验的不同结果而取不同的值. 根据号码是“小于 5”、“等于 5”、“大于 5”的三种情况，可定义该随机变量的取值. 进一步，可由随机变量的分布律与分布函数的定义，求出其分布律与分布函数.

解 : 分别用 ω_1、ω_2、ω_3 表示试验的三种结果“小于 5”、“等于 5”、“大于 5”，这时试验的样本空间为 $\Phi=\{\omega_1,\omega_2,\omega_3\}$，定义随机变量 X 为： $X=X(\omega)=\begin{cases}0,\omega=\omega_1\\ 1,\omega=\omega_2\\ 2,\omega=\omega_3\end{cases}$， X 取每个值的概率为： $P\{X=0\}=\dfrac{5}{10}$，

$P\{X=1\}=\dfrac{1}{10}$， $P\{X=2\}=\dfrac{4}{10}$；故 X 的分布律为(表 2-4)：

表 2-4

X	0	1	2
P_k	$\dfrac{5}{10}$	$\dfrac{1}{10}$	$\dfrac{4}{10}$

当 $x<0$ 时， $F(x)=P\{X\le x\}=0$；

当 $0\le x<1$ 时， $F(x)=P\{X\le x\}=P\{X=0\}=\dfrac{5}{10}$；

当 $1\le x<2$ 时， $F(x)=P\{X\le x\}=P\{X=0\}+P\{X=1\}=\dfrac{6}{10}$；

当 $2\le x$ 时， $F(x)=P\{X\le x\}=P\{X=0\}+P\{X=1\}+P\{X=2\}=1$；

由此求得分布函数为：$F(x)=P\{X\le x\}=\begin{cases}0, x<0\\ \frac{5}{10}, 0\le x<1\\ \frac{6}{10}, 1\le x<2\\ 1, x\ge 2\end{cases}$.

【例 3】设 1 小时内进入某图书馆的读者人数服从泊松分布. 已知 1 小时内无人进入图书馆的概率为 0.01. 求 1 小时内至少有 2 个读者进入图书馆的概率.

分析：1 小时内进入图书馆的人数是一个随机变量 X，且 $X\sim P(\lambda)$. 这样，$\{X=0\}$ 表示在 1 小时内无人进入图书馆，$\{X\ge 2\}$ 表示在 1 小时内至少有 2 人进入图书馆. 通过求参数 λ，进一步，求 $P\{X\ge 2\}$.

解：设 X 为在 1 小时内进入图书馆的人数，则 $X\sim P(\lambda)$，这时：$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}, k=0,1,\cdots$ 已知 $P\{X=0\}=e^{-\lambda}=0.01$，故 $\lambda=2\ln 10$. 所求概率为：$P\{X\ge 2\}=1-e^{-\lambda}-\lambda e^{-\lambda}=1-0.01(1+2\ln 10)=0.944$.

【例 4】设随机变量 X 的密度函数为 $f(x)=\begin{cases}\frac{c}{\sqrt{1-x^2}}, |x|<1\\ 0, \quad \text{其它}\end{cases}$，试求：

（1）常数 c；（2）$P\{0\le X\le \frac{1}{2}\}$；（3）$X$ 的分布函数.

分析：由密度函数的性质 $\int_{-\infty}^{+\infty}f(x)dx=1$ 可求得常数 c；对密度函数在 $[0,\frac{1}{2}]$ 上积分，即得 $P\{0\le X\le \frac{1}{2}\}$；根据连续型随机变量分布函数的定义可求 X 的分布函数.

解：（1）由 $1=\int_{-\infty}^{+\infty}f(x)dx=\int_{-1}^{+1}\frac{c}{\sqrt{1-x^2}}dx=c\cdot\arcsin x|_{-1}^{+1}=c\pi$ 得：$c=\frac{1}{\pi}$；（2）$P\{0\le X\le \frac{1}{2}\}=\int_0^{\frac{1}{2}}\frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}dx=\frac{1}{\pi}\arcsin x|_0^{\frac{1}{2}}=\frac{1}{6}$；

（3）当$x\le -1$时，$\{X\le x\}$是不可能事件，所以$F(x)=P\{X\le x\}=0$；当$|x|<1$时，

$$F(x)=\int_{-\infty}^{x}f(x)dx=\int_{-1}^{x}\frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}dx=\frac{1}{\pi}\arcsin x\Big|_{-1}^{x}=\frac{1}{\pi}\arcsin x+\frac{1}{2};$$

当$x\ge 1$时，$F(x)=\int_{-\infty}^{x}f(x)dx=\int_{-1}^{1}\frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}dx=1$；所以，$X$的分布函数为：

$$F(x)=\begin{cases}0, & x\le -1\\ \frac{1}{\pi}\arcsin x+\frac{1}{2}, & |x|<1\\ 1, & x\ge 1\end{cases}.$$

【例 5】设顾客在某银行窗口等待服务的时间X（以分计）服从指数分布，其概率密度为$f_X(x)=\begin{cases}\frac{1}{5}e^{-\frac{x}{5}}, & x>0\\ 0, & \text{其它}\end{cases}$，某顾客在窗口等待服务，若超过 10 分钟，他就离开. 他一个月要到银行 5 次，以Y表示一个月内他未等到服务而离开窗口的次数，写出Y的分布律，并求$P\{Y\ge 1\}$.

分析：显然，Y为随机变量，取值为 0、1、2、3、4、5，且$Y\sim B(5,p)$. 由$p=P\{X>10\}$及分布律的定义，可求得Y的分布律，进而求$P\{Y\ge 1\}$.

解：Y的取值为 0、1、2、3、4、5，$Y\sim B(5,p)$. 由题意得：

$p=P\{X>10\}=\int_{10}^{+\infty}f_X(x)dx=\int_{10}^{+\infty}\frac{1}{5}e^{-\frac{x}{5}}dx=e^{-2}$，故$Y$的分布律为：

$P\{X=k\}=C_5^k e^{-2k}(1-e^{-2})^{5-k},k=0,1,2,3,4,5$，即(表 2-5)：

表 2-5

Y	0	1	2	3	… 5
P_k	$(1-e^{-2})^5$	$5e^{-2}(1-e^{-2})^4$	$10e^{-4}(1-e^{-2})^3$	$10e^{-6}(1-e^{-2})^2$	$\cdots e^{-10}$

所以，$P\{Y\ge 1\}=1-P\{Y<1\}=1-P\{X=0\}=0.5167$.

【例 6】某单位招聘 2500 人，按考试成绩从高分到低分依次录用，共有 10000 人报名，假设报名者的成绩 $X\sim N(\mu,\sigma^2)$，已知 90 分以上有 359 人，60 分以下有 1151 人，问被录用者中最低分为多少？

分析：已知成绩 $X\sim N(\mu,\sigma^2)$，但不知 μ、σ 的值，所以，本题的关键是求 μ、σ，再进一步根据正态分布标准化方法进行求解.

解：根据题意：$P\{X>90\}=\dfrac{359}{10000}=0.0359$，故 $P\{X\le 90\}=1-P\{X>90\}=0.9641$，而

$P\{X\le 90\}=P\{\dfrac{X-\mu}{\sigma}\le\dfrac{90-\mu}{\sigma}\}=\Phi(\dfrac{90-\mu}{\sigma})=0.9641$，反查标准正态分布表，得：

$\dfrac{90-\mu}{\sigma}=1.8$ （1）

同样，$P\{X<60\}=\dfrac{1151}{10000}=0.1151$，而

$P\{X<60\}=P\{X\le 60\}=P\{\dfrac{X-\mu}{\sigma}\le\dfrac{60-\mu}{\sigma}\}=\Phi(\dfrac{60-\mu}{\sigma})=0.1151$，通过反查标准正态分布表，得：$\dfrac{60-\mu}{\sigma}=1.2$ （2）

由（1）、（2）两式解得：$\mu=72,\sigma=10$，所以 $X\sim N(72,10^2)$；

已知录用率为 $\dfrac{2500}{10000}=0.25$，设被录用者中最低分为 x_0，则

$P\{X\le x_0\}=1-P\{X\ge x_0\}=0.75$，而

$P\{X\le x_0\}=P\{\dfrac{X-72}{10}\le\dfrac{x_0-72}{10}\}=\Phi(\dfrac{x_0-72}{10})=0.75$，反查标准正态分布表，得：

$\dfrac{x_0-72}{10}\approx 0.675$，解得：$x_0\approx 78.75$

故：被录用者中最低分为 79 分.

【例 7】设 X 的分布律为(表 2-6)：

表 2-6

X	1	2	3	4	5	6
P	$\dfrac{1}{4}$	$\dfrac{1}{6}$	$\dfrac{1}{12}$	$\dfrac{1}{8}$	$\dfrac{5}{24}$	$\dfrac{1}{6}$

求 $Y=\cos\frac{\pi}{2}X$ 的分布律.

分析：X 是离散型随机变量，　Y 也是离散型随机变量. 当 X 取不同值时，将 Y 那些取相等的值分别合并，并把相应的概率相加. 从而得到 Y 的分布律.

解：X 与 Y 的对应关系如下表 2-7：

表 2-7

X	1	2	3	4	5	6
Y	0	-1	0	1	0	-1
P	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{5}{24}$	$\frac{1}{6}$

由上表可知，Y 的取值只有-1，0，1 三种可能，由于

$P\{Y=-1\}=P\{X=2\}+P\{X=6\}=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$，

$P\{Y=0\}=P\{X=1\}+P\{X=3\}+P\{X=5\}=\frac{1}{4}+\frac{1}{12}+\frac{5}{24}=\frac{13}{24}$，

$P\{Y=1\}=P\{X=4\}=\frac{1}{8}$，所以，$Y=\cos\frac{\pi}{2}X$ 的分布律为(表 2-8)：

表 2-8

Y	-1	0	1
P	$\frac{1}{3}$	$\frac{13}{24}$	$\frac{1}{8}$

【例 8】设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，求随机变量函数 $Y=e^X$ 的概率密度.

分析：由于函数 $y=e^x$ 在 $(-\infty,+\infty)$ 上单调增加，且可导，故可按公式法求 Y 的概率密度.

解：由 $f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$-\infty<x<+\infty$ 知 $y=e^x>0$，所以 Y 的取值区间为 $(0,+\infty)$.

当 $y\le 0$ 时，$f_Y(y)=0$；当 $y>0$ 时，有反函数 $x=\ln y$，从而

$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}\cdot\frac{1}{y}=\frac{1}{\sqrt{2\pi}\sigma y}e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}$，由此得随机变量 Y 的概率密度为：

$$f_Y(y)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma y}e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}, & y>0\\ 0, & y\le 0\end{cases}.$$

【例 9】已知 $X\sim N(0,1)$，求 $Y=X^2$ 的概率密度.

分析：根据分布函数的定义，先求 $Y=X^2$ 的分布函数，然后对其求导，即可得到 Y 的概率密度.

解：若 $y\le 0$，则 $\{Y\le y\}$ 是不可能事件，因而 $F_Y(y)=P\{Y\le y\}=0$，

若 $y>0$，则有 $F_Y(y)=P\{Y\le y\}=P\{X^2\le y\}=P\{-\sqrt{y}\le X\le\sqrt{y}\}$

$=\Phi(\sqrt{y})-1$，$f_Y(y)=F_Y'(y)=2[\Phi(\sqrt{y})]'=2\varphi(\sqrt{y})\cdot\frac{1}{2\sqrt{y}}=\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}$，从而，$Y$ 的概率密度为：$f_Y(y)=\begin{cases}\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}, & y>0\\ 0, & y\le 0\end{cases}$.

第三章　多维随机变量及其分布

内 容 提 要

1、二维随机变量及其联合分布函数

设 X， Y 为随机变量，则称它们的有序数组（ X,Y ）为二维随机变量.

设（ X,Y ）为二维随机变量，对于任意实数 x、 y，称二元函数

$F(x,y)=P\{X\le x,Y\le y\}$ 为（ X,Y ）的联合分布函数.

联合分布函数具有以下基本性质：

（1） $F(x,y)$ 是变量 x 或 y 的非减函数；

（2） $0\le F(x,y)\le 1$ 且

$F(-\infty,y)=0$， $F(x,-\infty)=0$， $F(-\infty,-\infty)=0$， $F(+\infty,+\infty)=1$；

（3） $F(x,y)$ 关于 x 右连续，关于 y 也右连续；

（4）对任意点 $(x_1,y_1),(x_2,y_2)$，若 $x_1<x_2,y_1<y_2$，则

$$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\ge 0.$$

上式表示随机点 (X,Y) 落在区域 $[x_1<X\le x_2,y_1<Y\le y_2]$ 内的概率为：

$P\{x_1<X\le x_2,y_1<Y\le y_2\}$.

2、二维离散型随机变量及其联合分布律

如果二维随机变量 (X,Y) 所有可能取值是有限对或可列对，则称 (X,Y) 为二维离散型随机变量.

设 (X,Y) 为二维离散型随机变量，它的所有可能取值为 $(x_i,y_j),i,j=1,2,\cdots$ 将 $P\{X=x_i,Y=y_j\}=p_{ij}\quad(i,j=1,2,\cdots)$ 或表 3.1 称为 (X,Y) 的联合分布律.

表 3.1

Y / X	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$
$\vdots$					

联合分布律具有下列性质：（1） $p_{ij} \geq 0$；（2） $\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} p_{ij} = 1$.

3、二维连续型随机变量及其概率密度函数

如果存在一个非负函数 $p(x,y)$，使得二维随机变量 (X,Y) 的分布函数 $F(x,y)$ 对任意实数 x,y 有 $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(x,y)dxdy$，则称 (X,Y) 是二维连续型随机变量，称 $p(x,y)$ 为 (X,Y) 的联合密度函数（或概率密度函数）.

联合密度函数具有下列性质：

（1）对一切实数 x,y，有 $p(x,y) \geq 0$；

（2） $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = 1$；

（3）在任意平面域 D 上，(X,Y) 取值的概率

$$P\{(X,Y) \in D\} = \iint_D p(x,y)dxdy;$$

（4）如果 $p(x,y)$ 在 (x,y) 处连续，则 $\dfrac{\partial^2 F(x,y)}{\partial x \partial y} = p(x,y)$.

4、二维随机变量的边缘分布

设 (X,Y) 为二维随机变量，则称

$F_X(x) = P\{X \leq x, -\infty < Y < +\infty\}$，$F_Y(y) = P\{-\infty < X < +\infty, Y \leq y\}$ 分别为 (X,Y) 关于 X 和关于 Y 的边缘分布函数.

当(X,Y)为离散型随机变量，则称

$p_{i\cdot}=\sum_{j=1}^{\infty}p_{ij}\quad(i=1,2,\cdots)$　　$p_{\cdot j}=\sum_{i=1}^{\infty}p_{ij}\quad(j=1,2,\cdots)$分别为$(X,Y)$关于$X$和关于$Y$的边缘分布律.

当(X,Y)为连续型随机变量，则称

$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy$,　　$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx$分别为$(X,Y)$关于$X$和关于$Y$的边缘密度函数.

5、二维随机变量的条件分布

（1）离散型随机变量的条件分布

设(X,Y)为二维离散型随机变量，其联合分布律和边缘分布律分别为

$P\{X=x_i,Y=y_j\}=p_{ij},P\{X=x_i\}=p_{i\cdot},P\{Y=y_j\}=p_{\cdot j}\quad(i,j=1,2,\cdots)$，则当$j$固定，且$P\{Y=y_j\}=p_{\cdot j}>0$时，称

$P\{X=x_i\mid Y=y_j\}=\dfrac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\dfrac{p_{ij}}{p_{\cdot j}},i=1,2,\cdots$为$Y=y_j$条件下随机变量$X$的条件分布律. 同理，有$P\{Y=y_j\mid X=x_i\}=\dfrac{p_{ij}}{p_{i\cdot}},j=1,2,\cdots$

（2）连续型随机变量的条件分布

设(X,Y)为二维连续型随机变量，其联合密度函数和边缘密度函数分别为：$p(x,y),p_X(x),p_Y(y)$. 则当$p_Y(y)>0$时，在$p(x,y)$和$p_X(x)$的连续点处，(X,Y)在条件$Y=y$下，X的条件概率密度函数为：$p_{X|Y}(x\mid y)=\dfrac{p(x,y)}{p_Y(y)}$.

同理，有$p_{Y|X}(x\mid y)=\dfrac{p(x,y)}{p_X(x)}$.

6、随机变量的独立性

设$F(x,y)$及$F_X(x)$、$F_Y(y)$分别是(X,Y)的联合分布函数及边缘分布函数. 如果对任何实数x,y有$F(x,y)=F_X(x)\cdot F_Y(y)$则称随机变量$X$与$Y$相互独立.

设(X,Y)为二维离散型随机变量，X与Y相互独立的充要条件是

$p_{ij}=p_{i\cdot}p_{\cdot j}(i,j=1,2,\cdots)$.

设(X,Y)为二维连续型随机变量，X与Y相互独立的充要条件是对任何实数x,y，有

$p(x,y)=p_X(x)p_Y(y)$.

7、两个随机变量函数的分布

设二维随机变量(X,Y)的联合概率密度函数为$p(x,y)$，$Z=\varphi(X,Y)$是X,Y的函数，则Z的分布函数为$F_Z(z)=\iint\limits_{\varphi(x,y)\le z}p(x,y)dxdy$.

（1）$Z=X+Y$的分布

若(X,Y)为离散型随机变量，联合分布律为p_{ij}，则Z的概率函数为：

$P_Z(z_k)=\sum\limits_i p(x_i,z_k-x_i)$或$P_Z(z_k)=\sum\limits_j p(y_j,z_k-y_j)$.

若(X,Y)为连续型随机变量，概率密度函数为$p(x,y)$，则Z的概率函数为：

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

（2）$Z=\dfrac{X}{Y}$的分布

若(X,Y)为连续型随机变量，概率密度函数为$p(x,y)$，则Z的概率函数为：

$p_Z(z)=\int_{-\infty}^{+\infty}|y|p(yz,y)dy$.

疑 难 分 析

1、事件$\{X\le x,Y\le y\}$表示事件$\{X\le x\}$与$\{Y\le y\}$的积事件，为什么$P\{X\le x,Y\le y\}$不一定等于$P\{X\le x\}\cdot P\{Y\le y\}$？

如同仅当事件A、B相互独立时，才有$P(AB)=P(A)\cdot P(B)$一样，这里

$P\{X\le x,Y\le y\}$依乘法原理$P\{X\le x,Y\le y\}=P\{X\le x\}\cdot P\{Y\le y\,|\,X\le x\}$.只有事件

$P\{X\le x\}$与$P\{Y\le y\}$相互独立时，才有

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

2、二维随机变量(X,Y)的联合分布、边缘分布及条件分布之间存在什么样的关系？

由边缘分布与条件分布的定义与公式知，联合分布唯一确定边缘分布，因而也唯一确定条件分布. 反之，边缘分布与条件分布都不能唯一确定联合分布. 但由 $p(x,y)=p_X(x)\cdot p_{Y|X}(y\mid x)$ 知，一个条件分布和它对应的边缘分布，能唯一确定联合分布.

但是，如果 X、Y 相互独立，则 $P\{X\le x,Y\le y\}=P\{X\le x\}\cdot P\{Y\le y\}$，即 $F(x,y)=F_X(x)\cdot F_Y(y)$. 说明当 X、Y 独立时，边缘分布也唯一确定联合分布，从而条件分布也唯一确定联合分布.

3、两个随机变量相互独立的概念与两个事件相互独立是否相同？为什么？

两个随机变量 X、Y 相互独立，是指组成二维随机变量 (X,Y) 的两个分量 X、Y 中一个分量的取值不受另一个分量取值的影响，满足 $P\{X\le x,Y\le y\}=P\{X\le x\}\cdot P\{Y\le y\}$. 而两个事件的独立性，是指一个事件的发生不受另一个事件发生的影响，故有 $P(AB)=P(A)\cdot P(B)$. 两者可以说不是一个问题.

但是，组成二维随机变量 (X,Y) 的两个分量 X、Y 是同一试验 E 的样本空间上的两个一维随机变量，而 A、B 也是一个试验 E_1 的样本空间的两个事件. 因此，若把“$X\le x$”、“$Y\le y$”看作两个事件，那么两者的意义近乎一致，从而独立性的定义几乎是相同的.

例 题 解 析

【例 1】设一盒内有 2 件次品，3 件正品，进行有放回的抽取和无放回的抽取. 设 X 为第一次抽取所得次品个数，Y 为第二次抽取所取得次品个数. 试分别求出两种抽取下：（1）(X,Y) 的联合分布律；

（2）二维随机变量 (X,Y) 的边缘分布律；

（3）X 与 Y 是否相互独立.

分析：求二维随机变量 (X,Y) 的边缘分布律，仅需求出概率 $P\{X=i,Y=j\}$. 由二维随机变量 (X,Y) 的边缘分布律的定义，$p_{i\cdot}=\sum_j p_{ij}, p_{\cdot j}=\sum_i p_{ij}$；将联合分布律表中各列的概率相加，

即得关于 X 的边缘分布律；将联合分布律表中各行的概率相加，即得关于 Y 的边缘分布律. 关于 X 与 Y 是否相互独立问题可由二维离散型随机变量 X 与 Y 相互独立的充要条件来验证.

解： X、 Y 都服从 0-1 分布，分别记

$$X=\begin{cases}0, & \text{第一次取得正品},\\ 1, & \text{第一次取得次品}.\end{cases} \quad Y=\begin{cases}0, & \text{第二次取得正品},\\ 1, & \text{第二次取得次品}.\end{cases}$$

（1）在有放回抽样时，联合分布律为：

$$P\{X=0,Y=0\}=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25}, P\{X=0,Y=1\}=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25},$$

$$P\{X=1,Y=0\}=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25}, P\{X=1,Y=1\}=\frac{2}{5}\cdot\frac{2}{5}=\frac{4}{25},$$ 可列成表，如表 3-1 所示.

在不放回抽样时，联合分布律为：

$$P\{X=0,Y=0\}=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10}, P\{X=0,Y=1\}=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10},$$

$$P\{X=1,Y=0\}=\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{10}, P\{X=1,Y=1\}=\frac{2}{5}\cdot\frac{1}{4}=\frac{1}{10},$$ 可列成表，如表 3-2 所示.

表 3-1

Y \ X	0	1
0	9/25	6/25
1	6/25	4/25

表 3-2

Y \ X	0	1
0	3/10	3/10
1	3/10	1/10

（2）在有放回抽样时，对表 3-1，按各列、各行相加，得关于 X、 Y 的边缘分布律为表 3-3、表 3-4. 在不放回抽样时，对表 3-2，按各列、各行相加，得关于 X、 Y 的边缘分布律为表 3-5、表 3-6.

表 3-3

X	0	1
$p_{i\cdot}$	3/5	2/5

表 3-4

Y	0	1
$p_{\cdot j}$	3/5	2/5

表 3-5

表 3-6

X	0	1
$p_{i.}$	3/5	2/5

Y	0	1
$p_{.j}$	3/5	2/5

（3）在有放回抽样时，因为 $p_{ij}=p_{i.}p_{.j}(i、j=0,1)$，所以 X 与 Y 相互独立；在不放回抽样时，因为 $p_{1.}p_{.1}=\frac{2}{5}\cdot\frac{2}{5}=\frac{4}{25}\neq p_{11}=\frac{1}{10}$，所以 X 与 Y 不相互独立.

【例 2】设 (X,Y) 的联合密度函数为 $p(x,y)=\begin{cases}Cxy,0<x<1,0<y<1\\0,\quad\quad 其它\end{cases}$ 试求：

（1）常数 C；（2）$p_X(x),p_Y(y)$；（3）X 与 Y 是否相互独立.

分析：由联合密度函数 $p(x,y)$ 的性质 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)dxdy=1$ 确定常数 C，由边缘密度函数的定义：$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy,p_Y(x)=\int_{-\infty}^{+\infty}p(x,y)dx$，计算广义积分得 $p_X(x),p_Y(y)$. 关于 X 与 Y 是否相互独立的问题，可用二维连续型随机变量 X 与 Y 相互独立的充要条件来验证.

解：（1）因为 $1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)dxdy=\int_0^1\int_0^1Cxydxdy=\frac{C}{4}$，因此 $C=4$；

（2）因为 $p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy$，

当 $0<y<1,0<x<1$ 时，$p_X(x)=\int_0^1 4xydy=2x$，当 x,y 为其它情况时，$p_X(x)=0$，所以 $p_X(x)=\begin{cases}2x,0<x<1\\0,\quad 其它\end{cases}$；同理 $p_Y(y)=\begin{cases}2y,0<y<1\\0,\quad 其它\end{cases}$；

（3）$p_X(x)p_Y(y)=\begin{cases}4xy,\quad 0<x<1,0<y<1\\0,\quad\quad\quad 其它\end{cases}$ 则有

$p(x,y)=p_X(x)p_Y(y)$，因此，X 与 Y 相互独立.

【例 3】设二维随机变量 (X,Y) 的密度函数为

$p(x,y)=\begin{cases}[\sin(x+y)]/2,0\le x<\pi/2,0\le y<\pi/2\\0,\quad\quad\quad\quad\quad 其它\end{cases}$，求 (X,Y) 的分布函数 $F(x,y)$.

分析：根据密度函数的定义可以看出分布函数 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(x,y)dxdy$ 与 (x,y) 所在的区域有关，可分区域分别进行讨论.

解：当 $x<0,y<0$ 时， $p(x,y)=0$，于是 $F(x,y)=0$；

当 $0\le x<\pi/2,0\le y<\pi/2$ 时， $p(x,y)=[\sin(x+y)]/2$，

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(x,y)dx\mathrm{d}y=\frac{1}{2}\int_{0}^{x}\int_{0}^{y}\sin(x+y)dxdy$$

$$=[\sin x+\sin y-\sin(x+y)]/2;$$

当 $x\ge\pi/2,0\le y<\pi/2$ 时，

$$F(x,y)=\frac{1}{2}\int_{0}^{\pi/2}\int_{0}^{y}\sin(x+y)dxdy=(1+\sin y-\cos y)/2;$$

当 $0\le x<\pi/2,y\ge\pi/2$ 时，

$$F(x,y)=\frac{1}{2}\int_{0}^{x}\int_{0}^{\pi/2}\sin(x+y)dxdy=(1+\sin x-\cos x)/2;$$

当 $x\ge\pi/2,y\ge\pi/2$ 时，

$F(x,y)=\frac{1}{2}\int_{0}^{\pi/2}\int_{0}^{\pi/2}\sin(x+y)dxdy=1$；所以

$$F(x,y)=\begin{cases}0, & x<0,y<0,\\ [\sin x+\sin y-\sin(x+y)]/2, & 0\le x<\pi/2,0\le y<\pi/2,\\ (1+\sin y-\cos y)/2, & x\ge\pi/2,0\le y<\pi/2,\\ (1+\sin x-\cos x)/2, & 0\le x<\pi/2,y\ge\pi/2,\\ 1, & x\ge\pi/2,y\ge\pi/2.\end{cases}$$

【例 4】随机变量 (X,Y) 的密度函数为 $p(x,y)=\begin{cases}2/(1+x+y)^3,x>0,y>0\\ 0, \quad \text{其它}\end{cases}$，求 $X=1$ 条件下 Y 的条件分布密度.

分析：通过 (X,Y) 的联合密度和边缘密度函数，来求在 $X=1$ 条件下 Y 条件分布密度.

解：当 $x>0$ 时，有 $p_X(x)=\int_0^{\infty}2/(1+x+y)^3dy=1/(1+x)^2$；

故 $p_{Y|X}(y|x=1)=p(1,y)/p_X(1)=\begin{cases}8/(2+y)^3,y>0\\ 0, \quad y\le 0.\end{cases}$

【例 5】随机变量 (X,Y) 的密度函数为 $p(x,y)=\begin{cases}e^{-y},x>0,y>x\\ 0, \quad \text{其它}\end{cases}$，求 $P\{X>2\,|\,Y<4\}$.

分析：先求得边缘密度函数，再根据条件概率的定义进行求解.

解：因为

$$p_X(x)=\begin{cases}\int_x^{+\infty}e^{-y}dy=e^{-x}, x>0;\\ 0, \quad x\le 0.\end{cases}\quad p_Y(y)=\begin{cases}\int_0^{y}e^{-y}dx=ye^{-y}, y>0;\\ 0, \quad y\le 0.\end{cases}$$ 故

$$P\{X>2,Y<4\}=\iint_G p(x,y)dxdy=\int_2^4\int_2^y e^{-y}dx=\int_2^4(y-2)e^{-y}dy=e^{-2}-3e^{-4}$$

又 $P(Y<4)=\int_{y<4}p_Y(y)dy=\int_0^4 ye^{-y}dy=1-5e^{-4}$

所以 $P\{X>2\mid Y<4\}=(e^{-2}-3e^{-4})/(1-5e^{-4})$.

【例 6】设随机变量 X 和 Y 相互独立，有 $p_X(x)=\begin{cases}1, 0\le x\le 1;\\ 0, \quad \text{其它}.\end{cases}\quad p_Y(y)=\begin{cases}2y, 0\le y\le 1;\\ 0, \quad \text{其它}.\end{cases}$ 求随机变量 $Z=X+Y$ 的概率密度函数 $p_Z(z)$.

分析：可按分布函数的定义先求得 $F_Z(z)=P\{Z\le z\}$，再进一步求得概率密度函数 $p_Z(z)$；在计算累次积分时要分各种情况进行讨论.

解：$F_Z(z)=P\{X+Y\le z\}=\iint_{x+y\le z}p(x,y)dxdy$，积分仅当 $p(x,y)>0$ 时才不为 0，考虑 $p(x,y)>0$ 的区域与 $x+y\le z$ 的取值，分四种情况计算（如图 3-1）.

当 $z<0$ 时，$F_Z(z)=0$；

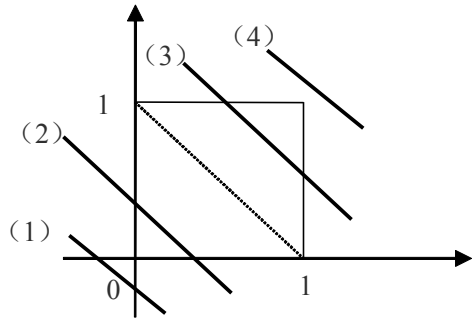

图 3-1

当 $0\le z\le 1$ 时，$F_Z(z)=\int_0^z\int_0^{z-x}2ydy=z^3/3$；

当 $1<z\le 2$ 时，

$$F_Z(z)=\int_0^{z-1}\int_0^1 2ydy+\int_{z-1}^1\int_0^{z-x}2ydy$$

$$=z^2-z^3/3-1/3;$$

当 $z>2$ 时，$F_Z(z)=1$；所以

$$F_Z(z)=\begin{cases}0, & z<0,\\ z^3/3, & 0\le z\le 1,\\ z^2-z^3/3-1/3, & 1<z\le 2,\\ 1, & z>2.\end{cases}\quad p_Z(z)=F_Z(z)=\begin{cases}z^2, & 0\le z\le 1,\\ 2z-z^2, & 1\le z\le 2,\\ 0, & \text{其它}.\end{cases}$$

第四章　随机变量的数字特征

内 容 提 要

1、随机变量的数学期望

设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k, k=1,2,\cdots$，如果级数 $\sum_{k=1}^{\infty}x_k p_k$ 绝对收敛，则称级数的和为随机变量 X 的数学期望.

设连续型随机变量 X 的密度函数为 $p(x)$，如果广义积分 $\int_{-\infty}^{+\infty}xp(x)dx$ 绝对收敛，则称此积分值 $E(X)=\int_{-\infty}^{+\infty}xp(x)dx$ 为随机变量 X 的数学期望.

数学期望有如下性质：

（1）设 C 是常数，则 $E(C)=C$；

（2）设 C 是常数，则 $E(CX)=CE(X)$；

（3）若 X_1、X_2 是随机变量，则 $E(X_1+X_2)=E(X_1)+E(X_2)$；

对任意 n 个随机变量 $X_1,X_2,\cdots,X_n$，有

$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)$$；

（4）若 X_1、X_2 相互独立，则 $E(X_1X_2)=E(X_1)E(X_2)$；

对任意 n 个相互独立的随机变量 $X_1,X_2,\cdots,X_n$，有

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

2、随机变量函数的数学期望

设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k, k=1,2,\cdots$，则 X 的函数 $Y=g(X)$ 的数学期望为 $E[g(x)]=\sum_{k=1}^{\infty}g(x_k)p_k, k=1,2\cdots$，式中级数绝对收敛.

设连续型随机变量 X 的密度函数为 $p(x)$，则 X 的函数 $Y=g(X)$ 的数学期望为 $E[g(x)]=\int_{-\infty}^{+\infty}g(x)p(x)dx$，式中积分绝对收敛.

3、随机变量的方差

设 X 是一个随机变量，则 $D(X)=Var(X)=E\{[X-E(X)]^2\}$ 称为 X 的方差. $\sqrt{D(X)}=\sigma(X)$ 称为 X 的标准差或均方差.

计算方差也常用公式 $D(X)=E(X^2)-[E(X)]^2$.

方差具有如下性质：

（1）设 C 是常数，则 $D(C)=0$；

（2）设 C 是常数，则 $D(CX)=C^2D(X)$；

（3）若 X_1、X_2 相互独立，则 $D(X_1+X_2)=D(X_1)+D(X_2)$；

对任意 n 个相互独立的随机变量 $X_1,X_2,\cdots,X_n$，有

$D(X_1+X_2+\cdots+X_n)=D(X_1)+D(X_2)+\cdots+D(X_n)$；

（4）$D(X)=0$ 的充要条件是：存在常数 C，使 $P\{X=C\}=1(C=E(X))$.

4、几种常见分布的数学期望与方差

（1）$X\sim(0-1).E(X)=p,D(X)=p(1-p)$；

（2）$X\sim B(n,p).E(X)=np,D(X)=np(1-p)$；

（3）$X\sim H(n,M,N).E(X)=\frac{nM}{N},D(X)=\frac{nM(N-M)(N-n)}{N^2(N-1)}$；

（4）$X\sim\pi(\lambda).E(X)=\lambda,D(X)=\lambda$；

（5）$X\sim G(p).E(X)=1/p,D(X)=(1-p)/p^2$；

（6）$X\sim U(a,b).E(X)=(a+b)/2,D(X)=(b-a)^2/12$；

（7）$X\sim e(\lambda).E(X)=1/\lambda,D(X)=1/\lambda^2$；

（8）$X\sim N(\mu,\sigma^2).E(X)=\mu,D(X)=\sigma^2$.

5、矩

设 X 是随机变量，则 $\alpha_k = E(X^k), k = 1,2,\cdots$ 称为 X 的 k 阶原点矩.

如果 $E(X)$ 存在，则 $\mu_k = E\{[X - E(X)]^k\}, k = 1,2,\cdots$ 称为 X 的 k 阶中心矩.

设 (X,Y) 是二维随机变量，则 $\alpha_{kl} = E(X^k Y^l), k,l = 1,2,\cdots$ 称为 (X,Y) 的 $k+l$ 阶混合原点矩； $\mu_{kl} = E\{[X - E(X)]^k \cdot [Y - E(Y)]^l\}, k,l = 1,2,\cdots$ 称为 (X,Y) 的 $k+l$ 阶混合中心矩.

5、二维随机变量的数字特征

(1) (X,Y) 的数学期望 $E(X,Y) = [E(X), E(Y)]$；

若 (X,Y) 是离散型随机变量，则 $E(X) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i p_{ij}$， $E(Y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} y_i p_{ij}$.

若 (X,Y) 是连续型随机变量，则

$E(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xp(x,y)dxdy$， $E(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yp(x,y)dxdy$. 这里，级数与积分都是绝对收敛的.

（2）(X,Y) 的方差 $D(X,Y) = [D(X), D(Y)]$

若 (X,Y) 是离散型随机变量，则 $D(X) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}[x_i - E(X)]^2 p_{ij}$，

$D(Y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}[y_i - E(Y)]^2 p_{ij}$.

若 (X,Y) 是连续型随机变量，则 $D(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x - E(X)]^2 p(x,y)dxdy$，

$D(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[y - E(Y)]^2 p(x,y)dxdy$. 这里，级数与积分都是绝对收敛的.

6、协方差与相关系数

随机变量 (X,Y) 的协方差为 $\mathrm{cov}(X,Y) = E\{[X - E(X)][Y - E(Y)]\}$. 它是 1+1 阶混合中心矩，有计算公式： $\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y)$.

随机变量 (X,Y) 的相关系数为 $\rho_{XY} = \dfrac{\mathrm{cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}$.

相关系数具有如下性质：

（1）$|\rho_{XY}| \le 1$；

（2）$|\rho_{XY}| = 1 \Leftrightarrow$ 存在常数 a,b，使 $P\{Y = aX + b\}$=1，即 X 与 Y 以概率 1 线性相关；

（3）若 X,Y 独立，则 $\rho_{XY}=0$，即 X,Y 不相关. 反之，不一定成立.

疑 难 分 析

1、随机变量的数字特征在概率论中有什么意义？

知道一个随机变量的分布函数，就掌握了这个随机变量的统计规律性. 但求得一个随机变量的分布函数是不容易的，而且往往也没有这个必要. 随机变量的数字特征则比较简单易求，也能满足我们研究分析具体问题的需要，所以在概率论中很多的应用，同时也刻画了随机变量的某些特征，有重要的实际意义.

例如，数学期望反映了随机变量取值的平均值，表现为具体问题中的平均长度、平均时间、平均成绩、期望利润、期望成本等；方差反映了随机变量取值的波动程度；偏态系数、峰态系数则反映了随机变量取值的对称性和集中性. 因此，在不同的问题上考察不同的数字特征，可以简单而切实地解决我们面临的实际问题.

2、在数学期望定义中为什么要求级数和广义积分绝对收敛？

首先，数学期望是一个有限值；其次，数学期望反映随机变量取值的平均值. 因此，对级数和广义积分来说，绝对收敛保证了值的存在，且对级数来说，又与项的次序无关，从而更便于运算求值. 而由于连续型随机变量可以离散化，从而广义积分与无穷级数有同样的意义. 要求级数和广义积分绝对收敛是为了保证数学期望的存在与求出.

3、相关系数 ρ_{XY} 反映了随机变量 X 和 Y 之间的什么关系？

相关系数 ρ_{XY} 是用随机变量 X 和 Y 的协方差和标准差来定义的，它反映了随机变量 X 和 Y 之间的相关程度. 当 $|\rho_{XY}|=1$ 时，称 X 与 Y 依概率 1 线性相关；当 $\rho_{XY}=0$ 时，称 X 与 Y 不相关；当 $0<\rho_{XY}<1$ 时，又分为强相关与弱相关.

4、两个随机变量 X 与 Y 相互独立和不相关是一种什么样的关系？

（1）若 X、Y 相互独立，则 X、Y 不相关. 因为 X、Y 独立，则 $E(XY)=E(X)E(Y)$，故 $\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=0$，从而 $\rho_{XY}=0$，所以 X、Y 不相关.

（2）若 X、Y 不相关，则 X、Y 不一定独立. 如：

$p(x,y)=\begin{cases}1/\pi, & x^2+y^2\le 1,\\ 0, & \text{其它.}\end{cases}$ 因为 $E(X)=E(Y)=0$， $D(X)=D(Y)=1/4$

$\operatorname{cov}(X,Y)=0,\rho_{XY}=0$，知 X、Y 不相关. 但 $p_X(x)=2\sqrt{1-x^2}/\pi$，

$p_Y(y)=2\sqrt{1-y^2}/\pi$，$p(x,y)\neq p_X(x)p_y(Y)$，知 X、Y 不独立.

（3）若 X、Y 相关，则 X、Y 一定不独立. 可由反证法说明.

（4）若 X、Y 不相关，则 X、Y 不一定不相关. 因为 X、Y 不独立，

$E(XY)\neq E(X)E(Y)$，但若 $E(X)=E(Y)=E(XY)=0$ 时，可以有 $\rho_{XY}=0$，从而可以有 X、Y 不相关.

但是，也有特殊情况，如 (X,Y) 服从二维正态分布时，X、Y 不相关与 X、Y 独立是等价的.

例 题 解 析

【例 1】设随机变量 X 的分布律为 $P\{X=k\}=\alpha^k/(1+\alpha)^{k+1},\alpha>0,k=0,1,\cdots$ 求 $E(X)$ 和 $D(X)$.

分析：可直接按离散型随机变量的期望和方差的定义进行计算.

解： $E(X)=\sum_{k=0}^{\infty}k\cdot\alpha^k/(1+\alpha)^{k+1}=\alpha/(1+\alpha)^2\cdot\sum_{k=1}^{\infty}k\left(\frac{\alpha}{1+\alpha}\right)^{k-1}=\alpha$；同理

$E(X^2)=\sum_{k=1}^{\infty}k^2\cdot\alpha^k/(1+\alpha)^{k+1}=\alpha/(1+\alpha)^2\cdot\sum_{k=1}^{\infty}k^2\left(\frac{\alpha}{1+\alpha}\right)^{k-1}=\alpha(1+2\alpha)$，

所以 $D(X)=E(X^2)-[E(X)]^2=\alpha(1+\alpha)$.

【例 2】设 (X,Y) 的概率密度函数为 $p(x,y)=\begin{cases}3xy/16, & 0\le x\le 2,0\le y\le x^2,\\ 0, & \text{其它.}\end{cases}$ 求

（1）$E(X),E(Y)$；（2）$D(X),D(Y)$；（3）$\operatorname{cov}(X,Y),\rho_{XY}$.

分析：由数学期望的定义及方差、协方差、相关系数的计算公式，首先须求出关于 X、Y 的边缘密度函数 $p_X(x)$、$p_Y(y)$，然后在分别求数学期望、方差、协方差、相关系数等.

解：（1） $p_X(x)=\int_0^{x^2} 3xy/16\cdot dy=3x^5/32, 0\le x\le 2$，

$p_Y(y)=\int_0^2 3xy/16\cdot dx=3y(4-y)/32, 0\le y\le 4$，所以

$E(X)=\int_0^2 x\cdot 3x^5/32\cdot dx=12/7, E(Y)=\int_0^4 y\cdot 3y(4-y)/32\cdot dy=2$；

（2） $E(X^2)=\int_0^2 x^2\cdot 3x^5/32\cdot dx=3, E(Y^2)=\int_0^4 y^2\cdot 3y(4-y)/32\cdot dy=24/5$

所以 $D(X)=3-(12/7)^2=3/49$， $D(Y)=24/5-2^2=4/5$；

（3） $E(XY)=\int_0^2\int_0^{x^2} xy\cdot 3xy/16\cdot dxdy=32/9$，所以

$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=8/63$；

$\rho_{XY}=\mathrm{cov}(X,Y)/[\sqrt{DX}\sqrt{DY}]=4\sqrt{15}/27\approx 0.574$.

【例 3】设事件 A 在第 i 次试验中出现的概率为 $p_i(i=1,2,\cdots,n)$， X 表示在 n 次独立试验中 A 出现的次数，求 $E(X)$ 和 $D(X)$.

分析：可先求出随机变量的分布，再依公式计算数字特征.

解：设 $X_i=\begin{cases}0, \text{第}i\text{次试验}A\text{不出现},\\ 1, \text{第}i\text{次试验}A\text{出现}.\end{cases}$ 于是： $X=X_1+X_2+\cdots+X_n$.

$P\{X_i=1\}=p_i, P\{X_i=0\}=q_i(i=1,2,\cdots,n)$，故 $E(X_i)=p_i$，

$E(X)=\sum_{i=1}^n E(X_i)=\sum_{i=1}^n p_i$； $D(X_i)=E(X_i^2)-[E(X_i)]^2=p_iq_i$，

由于各 X_i 相互独立，所以 $D(X)=\sum_{i=1}^n D(X_i)=\sum_{i=1}^n p_iq_i$.（式中 $p_i+q_i=1$）

【例 4】设 $X\sim N(\mu,\sigma^2)$， $Y\sim N(\mu,\sigma^2)$，且 X、Y 相互独立，试求

$Z_1=\alpha X+\beta Y$ 和 $Z_2=\alpha X-\beta Y$ 的相关系数. α、β 为不等于零的常数.

分析：求函数的数字特征，可有以下三种方法：（1）先求函数的概率分布，再依公式计算数字特征；（2）直接依随机变量函数数字特征的公式计算；（3）利用数字特征的有关定理计算.

解： $\mathrm{cov}(Z_1,Z_2)=\mathrm{cov}(\alpha X+\beta Y,\alpha X-\beta Y)=\alpha^2\,\mathrm{cov}(X,Y)-\alpha\beta\,\mathrm{cov}(X,Y)$

$+\alpha\beta\,\mathrm{cov}(X,Y)-\beta^2\,\mathrm{cov}(X,Y)=\alpha^2D(X)-\beta^2D(Y)=(\alpha^2-\beta^2)\sigma^2$；

而 $D(Z_1)=D(\alpha X+\beta Y)=\alpha^2\sigma^2+\beta^2\sigma^2=D(Z_2)$，所以

$$\rho_{Z_1Z_2}=\frac{(\alpha^2-\beta^2)\sigma^2}{(\alpha^2+\beta^2)\sigma^2}=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

【例 5】设 $X_1,X_2,\cdots,X_n$ 是相互独立的随机变量，且 $E(X_i)=\mu,D(X_i)=\sigma^2$，

$i=1,2,\cdots,n$. 记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$. 证明

（1）$E(\overline{X})=\mu,D(\overline{X})=\frac{\sigma^2}{n}$；（2）$E(S^2)=\sigma^2$.

分析：运用随机变量数字特征的某些性质及一定的技巧进行证明

证明：（1）$E(\overline{X})=E\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu$，

$$D(\overline{X})=D\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n};$$

（2）$E(S^2)=\frac{1}{n-1}E\{\sum_{i=1}^{n}[(X_i-\mu)-(\overline{X}-\mu)]^2\}$

$$=\frac{1}{n-1}\sum_{i=1}^{n}E[(X_i-\mu)^2]-nE(\overline{X}-\mu)^2=\frac{1}{n-1}\left[n\sigma^2-n\cdot\frac{\sigma^2}{n}\right]=\sigma^2.$$

第五章　大数定律和中心极限定理

内 容 提 要

1、切贝雪夫不等式

设随机变量 X 的数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$，则对任意正数 ε，有不等式

$P\{|X-\mu|\geq\varepsilon\}\leq\dfrac{\sigma^2}{\varepsilon^2}$ 或 $P\{|X-\mu|<\varepsilon\}>1-\dfrac{\sigma^2}{\varepsilon^2}$ 成立.

2、大数定律

（1）切贝雪夫大数定理:设 $X_1,X_2,\cdots,X_n,\cdots$ 是相互独立的随机变量序列，数学期望 $E(X_i)$ 和方差 $D(X_i)$ 都存在，且 $D(X_i)<C\ (i=1,2,\cdots)$，则对任意给定的 $\varepsilon>0$，有

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

（2）贝努利大数定理:设 n_A 是 n 次重复独立试验中事件 A 发生的次数，p 是事件 A 在一次试验中发生的概率，则对于任意给定的 $\varepsilon>0$，有 $\lim\limits_{n\to\infty}P\{|\dfrac{n_A}{n}-p|<\varepsilon\}=1$.

贝努利大数定理给出了当 n 很大时，A 发生的频率 n_A/A 依概率收敛于 A 的概率，证明了频率的稳定性.

3、中心极限定律

（1）独立同分布中心极限定理:设 $X_1,X_2,\cdots,X_n,\cdots$ 是独立同分布的随机变量序列，有有限的数学期望和方差，$E(X_i)=\mu$，$D(X_i)=\sigma^2\neq0(i=1,2,\cdots)$. 则对任意实数 x，随机变量

$Y_n=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)}{\sqrt{n}\sigma}=\dfrac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$ 的分布函数 $F_n(x)$ 满足

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

（2）李雅普诺夫定理:设 $X_1, X_2, \cdots, X_n, \cdots$ 是不同分布且相互独立的随机变量，它们分别有数学期望和方差： $E(X_i)=\mu_i$, $D(X_i)=\sigma_i^{\ 2}\neq 0(i=1,2,\cdots)$. 记 $B_n^2=\sum_{i=1}^{n}\sigma_i^{\ 2}$, 若存在正数 δ ，，使得当 $n\to\infty$ 时，有 $\frac{1}{B_n^{2+\delta}}\sum_{i=1}^{n}E\{|X_i-\mu_i|^{2+\delta}\}\to 0$，则随机变量

$$Z_n=\frac{\sum_{i=1}^{n}X_i-E(\sum_{i=1}^{n}X_i)}{\sqrt{D(\sum_{i=1}^{n}X_i)}}=\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu_i}{B_n}$$ 的分布函数 $F_n(x)$ 对于任意的 x，满足

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}\left\{\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu_i}{B_n}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt.$$

当 n 很大时, $Z_n\sim N(0,1), \sum_{i=1}^{n}X_i\sim N(\sum_{i=1}^{n}\mu_i, B_n^2)$.

（3）德莫佛—拉普拉斯定理:设随机变量 $\eta_n(n=1,2,\cdots)$ 服从参数为 $n, p(0<p<1)$ 的二项分布，则对于任意的 x，恒有 $\lim_{n\to\infty}P\left\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt$.

疑 难 分 析

1、依概率收敛的意义是什么？

依概率收敛即依概率 1 收敛. 随机变量序列 $\{x_n\}$ 依概率收敛于 a，说明对于任给的 $\varepsilon>0$，当 n 很大时，事件“ $|x_n-a|<\varepsilon$ ”的概率接近于 1. 但正因为是概率，所以不排除小概率事件“ $|x_n-a|<\varepsilon$ ”发生. 依概率收敛是不确定现象中关于收敛的一种说法.

2、大数定律在概率论中有何意义？

大数定律给出了在试验次数很大时频率和平均值的稳定性. 从理论上肯定了用算术平均值代替均值，用频率代替概率的合理性，它既验证了概率论中一些假设的合理性，又为数理统计中用样本推断总体提供了理论依据. 所以说，大数定律是概率论中最重要的基本定律.

3、中心极限定理有何实际意义？

许多随机变量本身并不属于正态分布，但它们的极限分布是正态分布. 中心极限定理阐明了在什么条件下，原来不属于正态分布的一些随机变量其总和分布渐进地服从正态分布. 为我们利用正态分布来解决这类随机变量的问题提供了理论依据.

4、大数定律与中心极限定理有何异同？

相同点：都是通过极限理论来研究概率问题，研究对象都是随机变量序列，解决的都是概率论中的基本问题，因而在概率论中有重要意义. 不同点：大数定律研究当 时，概率或平均值的极限，而中心极限定理则研究随机变量总和的分布的极限.

例 题 解 析

【例 1】设每次试验中某事件 A 发生的概率为 0.8，请用切贝雪夫不等式估计：n 需要多大，才能使得在 n 次重复独立试验中事件 A 发生的频率在 0.79～0.81 之间的概率至少为 0.95？

分析：根据切贝雪夫不等式进行估计，须记住不等式.

解：设 X 表示 n 次重复独立试验中事件 A 出现的次数，则 $X \sim B(n,0.8)$，A 出现的频率为 $\frac{X}{n}$, $E(X)=0.8n, D(X)=0.8\times 0.2n=0.16n$，

$$P\left\{0.79<\frac{X}{n}<0.81\right\}=P\{|X-0.8n|<0.01n\}\geq 1-\frac{D(X)}{(0.01n)^2}=1-\frac{0.16n}{0.0001n^2}$$

$$=1-\frac{1600}{n}$$

由题意得 $1-\frac{1600}{n}\geq 0.95$，$n\geq 32000$. 可见

做 32000 次重复独立试验中可使事件 A 发生的频率在 0.79～0.81 之间的概率至少为 0.95.

【例 2】证明：（马尔柯夫定理）如果随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$，满足

$\lim\limits_{n\to\infty}\frac{1}{n^2}D(\sum\limits_{k=1}^{n}X_k)=0$，则对任给 $\varepsilon>0$，有 $\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\frac{1}{n}E(\sum\limits_{k=1}^{n}X_k)\right|<\varepsilon\right\}=1$.

证明：$E(\frac{1}{n}\sum\limits_{k=1}^{n}X_k)=\frac{1}{n}\sum\limits_{k=1}^{n}E(X_k), D(\frac{1}{n}\sum\limits_{k=1}^{n}X_k)=\frac{1}{n^2}D(\sum\limits_{k=1}^{n}X_k)$，由切贝雪夫不等式，得

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}E(\sum_{k=1}^{n}X_k)\right|<\varepsilon\right\}\geq 1-\frac{D(\sum\limits_{k=1}^{n}X_k)}{n^2\varepsilon^2},$$

根据题设条件，当 $n\to\infty$ 时，$\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\frac{1}{n}E(\sum\limits_{k=1}^{n}X_k)\right|<\varepsilon\right\}\geq 1$，

但概率小于等于 1，故马尔柯夫定理成立.

【例 3】一本书共有 100 万个印刷符号. 排版时每个符号被排错的概率为 0.0001，校对时每个排版错误被改正的概率为 0.9，求校对后错误不多于 15 个的概率.

分析：根据题意构造一个独立同分布的随机变量序列，具有有限的数学期望和方差，然后建立一个标准化的随机变量，应用中心极限定理求得结果.

解：设随机变量 $X_n=\begin{cases}1, & \text{第}n\text{个印刷符号校对后仍印错}\\ 0, & \text{其它.}\end{cases}$ 则 $X_n(n\geq 1)$ 是独立同分布随机变量序列，有 $p=P\{X_n=1\}=0.0001\times 0.1=10^{-5}$. 作 $Y_n=\sum\limits_{k=1}^{n}X_K,(n=10^6)$，$Y_n$ 为校对后错误总数. 按中心极限定理（德—拉定理），有

$$P\{Y_n\leq 15\}=P\left\{\frac{Y_n-np}{\sqrt{npq}}\leq\frac{15-np}{\sqrt{npq}}\right\}=\Phi(5/[10^3\sqrt{10^{-5}(1-10^{-5})}])\approx\Phi(1.58)$$

$=0.9495$.

第六章　数理统计的基本概念

内 容 提 要

1、总体与样本

在数理统计中，将研究对象的全体称为总体；组成总体的每个元素称为个体.
从总体中抽取的一部分个体，称为总体的一个样本；样本中个体的个数称为样本的容量.

从分布函数为 $F(x)$ 的随机变量 X 中随机地抽取的相互独立的 n 个随机变量，具有与总体相同的分布，则 $X_1, X_2, \cdots, X_n$ 称为从总体 X 得到的容量为 n 的随机样本. 一次具体的抽取记录 $x_1, x_2, \cdots, x_n$ 是随机变量 $X_1, X_2, \cdots, X_n$ 的一个观察值，也用来表示这些随机变量.

2、统计量

设 $X_1, X_2, \cdots, X_n$ 是总体 X 的一个样本，则不含未知参数的样本的连续函数 $f(X_1, X_2, \cdots, X_n)$ 称为统计量. 统计量也是一个随机变量，常见的统计量有

（1）样本均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$；

（2）样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}[\sum_{i=1}^{n} X_i^2 - n\overline{X}^2]$；

（3）样本标准差 $S = \sqrt{S^2}$；

（4）样本 k 阶原点矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, k = 1,2,\cdots$；

（5）样本 k 阶中心矩 $B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k, k = 1,2,\cdots.$

2、经验分布函数

设 $x_1, x_2, \cdots, x_n$ 是总体 X 的一组观察值将它们按大小顺序排列为：

$x_1^* \le x_2^* \le \cdots \le x_n^*$，称它为顺序统计量. 则称

$$F_n(x)=\begin{cases}0, & x<x_1^* \\ \frac{1}{n}, & x_1^*\le x<x_2^* \\ \cdots & \\ \frac{k}{n}, & x_k^*\le x<x_{k+1}^* \\ \cdots & \\ 1, & x\ge x_n^*\end{cases}$$ 为经验分布函数（或样本分布函数）.

3、一些常用统计量的分布

（1） χ^2 分布

设 $X\sim N(0,1)$， $X_1,X_2,\cdots,X_n$ 是 X 的一个样本，则统计量 $\chi^2=\sum_{i=1}^{n}X_i^2$ 服从自由度为 n 的 χ^2 分布，记作 $\chi^2\sim\chi^2(n)$.

（2） t 分布

设 $X\sim N(0,1)$， $Y\sim\chi^2(n)$，且 X,Y 相互独立，则随机变量 $t=\frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，记作 $t\sim t(n)$. t 分布又称为学生分布.

（3） F 分布

设 $X\sim\chi^2(n_1)$， $Y\sim\chi^2(n_2)$，且 X,Y 相互独立，则随机变量 $F=\frac{X/n_1}{Y/n_2}$ 服从自由度为 (n_1,n_2) 的 F 分布，记作 $F\sim F(n_1,n_2)$.

4、正态总体统计量的分布

设 $X\sim N(\mu,\sigma^2)$， $X_1,X_2,\cdots,X_n$ 是 X 的一个样本，则

（1）样本均值 $\overline{X}$ 服从正态分布，有

$\overline{X}\sim N(\mu,\frac{\sigma^2}{n})$ 或 $U=\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\sim N(0,1)$；

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

（3）统计量 $\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$.

设 $X\sim N(\mu_1,\sigma_1^{\ 2}),Y\sim N(\mu_2,\sigma_2^2)$，$X_1,X_2,\cdots,X_{n_1}$ 是 X 的一个样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 是 Y 的一个样本，两者相互独立. 则

（1）统计量 $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$；

（2）当 $\sigma_1=\sigma_2$ 时，统计量 $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{1/n_1+2/n_2}\cdot S_w}\sim t(n_1+n_2-2)$，其中

$S_w=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$；

（3）统计量 $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$；

（4）统计量 $\dfrac{\sum\limits_{i=1}^{n_1}(x_i-\mu_1)^2/\sigma_1^2}{\sum\limits_{j=1}^{n_2}(y_j-\mu_2)^2/\sigma_2^2}\cdot\dfrac{n_2}{n_1}\sim F(n_1,n_2)$.

疑 难 分 析

1、为什么要引进统计量？为什么统计量中不能含有未知参数？

引进统计量的目的是为了将杂乱无序的样本值归结为一个便于进行统计推断和研究分析的形式，集中样本所含信息，使之更易揭示问题实质.

如果统计量中仍含有未知参数，就无法依靠样本观测值求出未知参数的估计值，因而就失去利用统计量估计未知参数的意义.

2、什么是自由度？

所谓自由度，通常是指不受任何约束，可以自由变动的变量的个数. 在数理统计中，自由度是对随机变量的二次型（或称为二次统计量）而言的. 因为一个含有 n 个变量的二次型 $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_{ij}X_iX_j(a_{ij}=a_{ji},i,j=1,2,\cdots,n)$ 的秩是指对称矩阵 $A=(a_{ij})_{n\times n}$ 的秩，它的大小反映 n 个变量中能自由变动的无约束变量的多少. 我们所说的自由度，就是二次型的秩.

例 题 解 析

【例 1】 设 $X_i\sim N(\mu_i,\sigma^2)(i=1,2,\cdots 5)$，（1）$\mu_1,\mu_2,\cdots,\mu_5$ 不全等；（2）

$\mu_1=\mu_2=\cdots=\mu_5$. 问： $X_1,X_2,\cdots,X_5$ 是否为简单随机样本？

分析：相互独立且与总体同分布的样本是简单随机样本，由此进行验证.

解:（1） 由于 $X_i\sim N(\mu_i,\sigma^2)(i=1,2,\cdots 5)$，且 $\mu_1,\mu_2,\cdots,\mu_5$ 不全等，所以 $X_1,X_2,\cdots,X_5$ 不是同分布，因此 $X_1,X_2,\cdots,X_5$ 不是简单随机样本.

（2）由于 $\mu_1=\mu_2=\cdots=\mu_5$，那么 $X_1,X_2,\cdots,X_5$ 服从相同的分布，但不知道 $X_1,X_2,\cdots,X_5$ 是否相互独立，因此 $X_1,X_2,\cdots,X_5$ 不一定是简单随机样本.

【例 2】设 $X\sim N(\mu,\sigma^2)$， $X_1,X_2,\cdots,X_n$ 是取自总体的简单随机样本， $\overline{X}$ 为样本均值， S_n^2 为样本二阶中心矩， S^2 为样本方差，问下列统计量

（1） $\dfrac{nS_n^2}{\sigma^2}$， （2） $\dfrac{\overline{X}-\mu}{S_n/\sqrt{n-1}}$， （3） $\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}$ 各服从什么分布？

分析：利用已知统计量的分布进行分析.

解：（1）由于 $\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，又有 $S_n^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2=\dfrac{n-1}{n}S^2$

$nS_n^2=(n-1)S^2$，因此 $\dfrac{nS_n^2}{\sigma^2}\sim\chi^2(n-1)$；

（2）由于 $\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$，又有 $\dfrac{S}{\sqrt{n}}=\dfrac{S_n}{\sqrt{n-1}}$，因此

$\dfrac{\overline{X}-\mu}{S_n/\sqrt{n-1}}\sim t(n-1)$；

（3）由 $X_i\sim N(\mu,\sigma^2)(i=1,2,\cdots,n)$ 得： $\dfrac{X_i-\mu}{\sigma}\sim N(0,1)(i=1,2,\cdots,n)$，由 χ^2 分布的定义得： $\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$.

【例 3】设总体服从参数为 λ 的指数分布，分布密度为 $p(x;\lambda)=\begin{cases}\lambda e^{-\lambda x},x>0\\0,\quad x\le 0\end{cases}$

求 $E\overline{X}, D\overline{X}$ 和 ES^2.

分析：利用已知指数分布的期望、方差和它们的性质进行计算.

解：由于 $EX_i = 1/\lambda, DX_i = 1/\lambda^2 (i = 1,2,\cdots,n)$，所以

$$E\overline{X} = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{\lambda};$$

$$D\overline{X} = D(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{1}{n\lambda^2};$$

$$ES^2 = E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2] = \frac{1}{n-1}\sum_{i=1}^{n} D(X_i) = \frac{n}{n-1}\cdot\frac{1}{n\lambda^2} = \frac{1}{(n-1)\lambda^2}.$$

【例 4】设总体 $X \sim N(\mu,4)$，$X_1, X_2, \cdots, X_n$ 是取自总体的简单随机样本，$\overline{X}$ 为样本均值. 问样本容量 n 取多大时有：

（1）$E(|\overline{X} - \mu|^2) \le 0.1$；（2）$P\{|\overline{X} - \mu| \le 0.1\} \ge 0.95$.

解：（1）要使 $E(|\overline{X} - \mu|^2) = D(\overline{X}) = D(X)/n = 4/n \le 0.1$，即有 $n \ge 40$，故取 $n = 40$.

（2）由中心极限定理，要使

$$P\{|\overline{X} - \mu| \le 0.1\} = P\{|\overline{X} - \mu|/\sqrt{D(X)} \le 0.1\sqrt{n/4}\} \approx \Phi(0.05\sqrt{n})$$

$$-\Phi(-0.05\sqrt{n}) = 2\Phi(0.05\sqrt{n}) - 1 \ge 0.95$$，即有

$$\Phi(0.05\sqrt{n}) \ge 0.975, 0.05\sqrt{n} \ge 1.96, n \ge 1536.64$$，故取 $n = 1537$.

第七章　参数估计

内 容 提 要

1、参数的点估计及其求法

根据总体 X 的一个样本来估计参数的真值称为参数的点估计.

（1）估计量

根据总体 X 的一个样本 $X_1, X_2, \cdots, X_n$ 构造的用其观察值来估计参数 θ 真值的统计量

$\hat{\theta}(X_1, X_2, \cdots, X_n)$ 称为估计量，$\hat{\theta}(x_1, x_2, \cdots, x_n)$ 称为估计值.

(2)矩估计法

用样本矩作为相应的总体矩估计来求出估计量的方法. 其思想是：如果总体中有 k 个未知参数，可以用前 k 阶样本矩估计相应的前 k 阶总体矩，然后利用未知参数与总体矩的函数关系，求出参数的估计量.

（3）极大似然估计法

设总体 X 的密度函数为 $p(x,\theta)$，其中 θ 为未知参数，$X_1, X_2, \cdots, X_n$ 是取自总体 X 的样本，$x_1, x_2, \cdots, x_n$ 为一组样本观测值，则总体 X 的联合密度函数称为似然函数，记作

$L = \prod_{i=1}^{n} p(x_i, \theta)$，取对数 $\ln L = \sum_{i=1}^{n} \ln p(x_i, \theta)$，由 $\frac{d \ln L}{d\theta} = 0$，求得似然函数 L 的极大值 $\hat{\theta}$，即为未知参数 θ 的极大似然估计. 其思想是：在已知总体 X 概率分布时，对总体进行 n 次观测，得到一个样本，选取概率最大的 θ 值 $\hat{\theta}$ 作为未知参数 θ 的真值的估计是最合理的.

（4）估计量的优劣标准

1）无偏性. 设 $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n), E(\hat{\theta})$ 存在，且 $E(\hat{\theta}) = \theta$，则称值 $\hat{\theta}$ 是 θ 的无偏估计量. 否则称为有偏估计量.

2）有效性. 设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 均为参数 θ 的无偏估计量，如果 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$，则称估计量 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

3）一致性（相合性）. 设 $\hat{\theta}$ 为 θ 的估计量，$\hat{\theta}$ 与样本容量 n 有关，记为 $\hat{\theta} = \hat{\theta}_n$，对于任意给定的 $\varepsilon > 0$，都有 $\lim_{n\to\infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1$，则称 $\hat{\theta}$ 为参数 θ 的一致估计量.

2、参数的区间估计

设总体 X 的分布 $F(x;\theta)$ 中含有未知参数 θ，若存在样本的两个函数 $\underline{\theta}(X_1,X_2,\cdots,X_n)$ 和 $\overline{\theta}(X_1,X_2,\cdots,X_n)$，使对于给定的 $\alpha(0<\alpha<1)$，有 $P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha$，则随机区间（$\underline{\theta},\overline{\theta}$）称为参数 θ 的置信度为 $1-\alpha$ 的双侧置信区间.

若有 $P\{\underline{\theta}<\theta\}=1-\alpha$ 或 $P\{\theta<\overline{\theta}\}=1-\alpha$，则定义 $(\underline{\theta},\infty)$ 或 $(-\infty,\overline{\theta})$ 为 θ 的置信度为 $1-\alpha$ 的单侧置信区间.

（1）单个正态总体均值与方差的置信区间（见表 7-1）

表 7-1

估计的参数	参数的情况	统计量	置信度为 $1-\alpha$ 的置信区间
μ	σ^2 已知	$U=\dfrac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\sim N(0,1)$	$(\overline{X}-Z_{\frac{\alpha}{2}}\cdot\dfrac{\sigma}{\sqrt{n}},\overline{X}+Z_{\frac{\alpha}{2}}\cdot\dfrac{\sigma}{\sqrt{n}})$
	σ^2 未知	$t=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$	$(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\cdot\dfrac{S}{\sqrt{n}},\overline{X}+t_{\frac{\alpha}{2}}(n-1)\cdot\dfrac{S}{\sqrt{n}})$
σ^2	μ 未知	$\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$	$\left(\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$
	μ 已知	$\chi^2=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$	$\left(\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)},\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right)$

（2）两个正态总体均值差与方差比的置信区间（见表 7-2）

表 7-2

估计的参数	参数的情况	置信度为 $1-\alpha$ 的置信区间
$\mu_1-\mu_2$	σ_1^2,σ_2^2 已知	$(\overline{X}-\overline{Y}-Z_{\frac{\alpha}{2}}\cdot\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}},\overline{X}-\overline{Y}+Z_{\frac{\alpha}{2}}\cdot\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}})$
	$\sigma_1^2=\sigma_2^2$ 未知	$(\overline{X}-\overline{Y}-t_{\frac{\alpha}{2}}(n_1+n_2-1)\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}},\overline{X}-\overline{Y}+t_{\frac{\alpha}{2}}(n_1+n_2-1)\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}})$
$\dfrac{\sigma_1^2}{\sigma_2^2}$	μ_1,μ_2 未知	$\left(\dfrac{S_1^2}{S_2^2\cdot F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\dfrac{S_1^2}{S_2^2\cdot F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$
	μ_1,μ_2 已知	$\left(\dfrac{n_2\sum_{i=1}^{n}(X_i-\mu_1)^2}{n_1F_{\frac{\alpha}{2}}(n_1,n_2)\sum_{i=1}^{n}(Y_i-\mu_2)^2},\dfrac{n_2\sum_{i=1}^{n}(X_i-\mu_1)^2}{n_1F_{1-\frac{\alpha}{2}}(n_1,n_2)\sum_{i=1}^{n}(Y_i-\mu_2)^2}\right)$

疑 难 分 析

1、有了点估计为什么还要引入区间估计？

点估计是利用样本值求得参数θ的一个近似值，对了解参数θ的大小有一定的参考价值，但没有给出近似值的精确程度和可信程度，因此在使用中意义不大. 而区间估计是通过两个（或一个）统计量$\underline{\theta},\overline{\theta}$ $(\underline{\theta}\le\overline{\theta})$，构成随机区间（$\underline{\theta},\overline{\theta}$），使此区间包含未知参数$\theta$的概率不小于事先设定的常数$\alpha(0<\alpha<1)$. $1-\alpha$的值越大，则（$\underline{\theta},\overline{\theta}$）包含$\theta$真值的概率越大，即由样本值得到的区间（$\underline{\theta},\overline{\theta}$）覆盖未知参数$\theta$的可信程度越大，而（$\underline{\theta},\overline{\theta}$）的长度越小，又反映估计$\theta$的精确程度越高. 所以区间估计不仅是提供了$\theta$的一个估计范围，还给出了估计范围的精确与可信程度，弥补了点估计的不足，有广泛的应用意义.

2、怎样理解置信度$1-\alpha$的意义？

置信度$1-\alpha$有两种方式的理解.

对于一个置信区间（$\underline{\theta},\overline{\theta}$）而言， $1-\alpha$表示随机区间（$\underline{\theta},\overline{\theta}$）中包含未知参数的概率不小于事先设定的数值$1-\alpha$.

对于区间估计而言，$1-\alpha$表示在样本容量不变的情况下反复抽样得到的全部区间中，包含θ真值的区间不少于$100(1-\alpha)\%$.

3、怎样处理区间估计中精度与可靠性之间的矛盾？

区间估计量（$\underline{\theta},\overline{\theta}$）的长度称为精度，$1-\alpha$称为（$\underline{\theta},\overline{\theta}$）的可靠程度. 长度越短，精确程度越高；$1-\alpha$越大，可靠程度越大. 但在样本容量固定时，两者不能兼顾. 因此，奈曼指出的原则是，先照顾可靠程度，在满足可靠性$P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha$时，再提高精度. 否则，只有增加样本容量，才能解决.

例 题 解 析

【例 1】设总体X服从几何分布，分布律为：$P\{X=x\}=(1-p)^{x-1}p, x=1,2,\cdots$，其中$p$为未知参数，且$0\le p\le 1$. 设$X_1,X_2,\cdots,X_n$为$X$的一个样本，求$p$的矩估计与极大似然估计.

分析：根据矩估计与极大似然估计方法直接进行估计.

解：（1）因为$E(X)=1/p$ ，所以p的矩估计为$\hat{p}=1/\overline{X}$；

（2）似然函数为：$L(x_1,x_2,\cdots x_n;p)=\prod_{i=1}^{n}[p(1-p)^{x_i-1}]=(1-p)^{\sum_{i=1}^{n}x_i-n}p^n$，

取对数：$\ln L=(\sum_{i=1}^{n}x_i-n)\ln(1-p)+n\ln p$，

求导，令$\frac{d\ln L}{dp}=\frac{-(\sum_{i=1}^{n}x_i-n)}{1-p}+\frac{n}{p}=0$，

解得，p的极大似然估计为$\hat{p}=1/\overline{X}$.

【例 2】设$\hat{\theta}$是参数θ的无偏估计，且有$D(\hat{\theta})>0$，试证明$\hat{\theta}^2$不是θ^2的无偏估计.

分析：证明无偏性，可直接按定义：$E(\hat{\theta})=\theta$进行证明.

证明：由$D(\hat{\theta})=E(\hat{\theta}^2)-(E\hat{\theta})^2$，及$E(\hat{\theta})=\theta$（由题意），

而$D(\hat{\theta})>0$，可以得出 $E(\hat{\theta}^2)=D(\hat{\theta})+(E\hat{\theta})^2=\theta^2+D(\hat{\theta})\neq\theta^2$，

因此，$\hat{\theta}^2$不是θ^2的无偏估计.

【例 3】某厂生产的钢丝. 其抗拉强度$X\sim N(\mu,\sigma^2)$，其中μ,σ^2均未知，从中任取 9 根钢丝，测得其强度（单位：kg）为：

578，582，574．568，596，572，570，584，578

求总体方差σ^2、均方差σ的置信度为 0.99 的置信区间.

分析：由于参数μ,σ^2均未知，故取统计量$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，从而得$\sigma^2$、$\sigma$置信度为$1-\alpha$的置信区间分别为

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)、\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}},\sqrt{\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}\right).$$

解：$\bar{x}=\frac{1}{9}\sum_{i=1}^{9}x_i=578$，$S^2=\frac{1}{8}\sum_{i=1}^{9}(x_i-\bar{x})^2=\frac{1}{8}\times592=74$，

$\alpha=0.01,\chi^2_{\frac{\alpha}{2}}(n-1)=\chi^2_{0.005}(8)=21.955,\chi^2_{1-\frac{\alpha}{2}}(n-1)=\chi^2_{0.995}(8)=1.344$，

所以方差σ^2的置信度为 0.99 的置信区间为：

$\left(\frac{592}{21.955},\frac{592}{1.344}\right)$，即（26.96，440.48）；

均方差σ的置信度为 0.99 的置信区间为：

$\left(\sqrt{\frac{592}{21.955}},\sqrt{\frac{592}{1.344}}\right)$，即（5.19，20.99）.

【例 4】设有两个正态总体，$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$. 分别从 X 和 Y 抽取容量为 $n_1=25$ 和 $n_2=8$ 的两个样本，并求得 $S_1=8,S_2=7$. 试求两正态总体方差比 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信度为 0.98 的置信区间.

分析：由于 μ_1,μ_2 均未知，故取统计量 $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$，$\frac{\sigma_1^2}{\sigma_2^2}$ 的置信度为 $1-\alpha$ 的置信区间为：$\left(\frac{S_1^2}{S_2^2\cdot F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2\cdot F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$.

解：由 $\alpha=0.02$，查表得：$F_{0.01}(24,7)=6.07,F_{0.99}(24,7)=\frac{1}{F_{0.01}(7,24)}=0.2857$，所以，$\frac{\sigma_1^2}{\sigma_2^2}$ 的置信度为 0.98 的置信区间为：（0.2152，4.5714）.

第八章　假设检验

内 容 提 要

1、假设检验的基本概念

（1）假设检验

对总体的分布提出某种假设，然后利用样本所提供的信息，根据概率论的原理对假设作出“接受”还是“拒绝”的判断，这一类统计推断问题统称为假设检验.

假设检验所依据的原则是：小概率事件在一次试验中是不该发生的.

（2）两类错误

在根据样本作推断时，由于样本的随机性，难免会作出错误的决定. 当原假设 H_0 为真时，而作出拒绝 H_0 的判断，称为犯第一类错误；当原假设 H_0 不真时，而作出接受 H_0 的判断，称为犯第二类错误.

控制犯第一类错误的概率不大于一个较小的数 $\alpha(0<\alpha<1)$ 称为检验的显著性水平.

（3）假设检验的基本步骤

1）建立原假设 H_0；

2）根据检验对象，构造合适的统计量；

3）求出在假设 H_0 成立的条件下，该统计量服从的概率分布；

4）选择显著性水平 α，确定临界值；

5）根据样本值计算统计量的观察值，由此作出接受或拒绝 H_0 的结论.

2、单个正态总体的假设检验

设总体 $X\sim N(\mu,\sigma^2)$.

（1）关于均值 μ 的检验（见表 8-1）

（2）关于方差 σ^2 的检验（见表 8-2）

表 8-1

	H_0	H_1	统计量	拒绝域

μ 检验法（σ^2 已知）	$\mu=\mu_0$ $\mu\le\mu_0$ $\mu\ge\mu_0$	$\mu\ne\mu_0$ $\mu>\mu_0$ $\mu<\mu_0$	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$	$\lvert U\rvert>z_{\alpha/2}$ $U>z_\alpha$ $U<-z_\alpha$
t 检验法（σ^2 未知）	$\mu=\mu_0$ $\mu\le\mu_0$ $\mu\ge\mu_0$	$\mu\ne\mu_0$ $\mu>\mu_0$ $\mu<\mu_0$	$T=\dfrac{\overline{X}-\mu_0}{S_n/\sqrt{n}}\sim t(n-1)$	$\lvert T\rvert>t_{\alpha/2}(n-1)$ $T>t_\alpha(n-1)$ $T<-t_\alpha(n-1)$

表 8-1

	H_0	H_1	统计量	拒绝域
χ^2 检验法（μ 已知）	$\sigma^2=\sigma_0^2$ $\sigma^2\le\sigma_0^2$ $\sigma^2\ge\sigma_0^2$	$\sigma^2\ne\sigma_0^2$ $\sigma^2>\sigma_0^2$ $\sigma^2<\sigma_0^2$	$k^2=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n)$	$k^2>x^2_{\alpha/2}(n)$或 $k^2<x^2_{1-\alpha/2}(n)$ $k^2>x^2_\alpha(n)$ $k^2<x^2_{1-\alpha}(n)$
χ^2 检验法（μ 未知）	$\sigma^2=\sigma_0^2$ $\sigma^2\le\sigma_0^2$ $\sigma^2\ge\sigma_0^2$	$\sigma^2\ne\sigma_0^2$ $\sigma^2>\sigma_0^2$ $\sigma^2<\sigma_0^2$	$k^2=\dfrac{(n-1)S_n^2}{\sigma^2}\sim\chi^2(n-1)$	$k^2>x^2_{\alpha/2}(n-1)$或 $k^2<x^2_{1-\alpha/2}(n-1)$ $k^2>x^2_\alpha(n-1)$ $k^2<x^2_{1-\alpha}(n-1)$

3、两个正态总体的假设检验

设总体 $X\sim N(\mu_1,{\sigma_1}^2)$，样本容量为 n_1；$Y\sim N(\mu_2,{\sigma_2}^2)$，样本容量为 n_2.

（1）两个正态总体均值的检验（见表 8-3）

（2）两个正态总体方差的检验（见表 8-4）

表 8-3

	H_0	H_1	统计量	拒绝域
μ 检验法（σ_1^2,σ_2^2 已知）	$\mu_1=\mu_2$ $\mu_1\le\mu_2$ $\mu_1\ge\mu_2$	$\mu_1\ne\mu_2$ $\mu_1>\mu_2$ $\mu_1<\mu_2$	$U=\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$\lvert U\rvert>z_{\alpha/2}$ $U>z_\alpha$ $U<-z_\alpha$

t 检验法 $(\sigma_1^2=\sigma_2^2=\sigma^2$ 未知)	$\mu_1=\mu_2$ $\mu_1\le\mu_2$ $\mu_1\ge\mu_2$	$\mu_1\ne\mu_2$ $\mu_1>\mu_2$ $\mu_1<\mu_2$	$T=\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$	$\lvert T\rvert>t_{\alpha/2}(n_1+n_2-2)$ $T>t_\alpha(n_1+n_2-2)$ $T<-t_\alpha(n_1+n_2-2)$

表 8-4

	H_0	H_1	统计量	拒绝域
F 检验法 $(\mu_1,\mu_2$ 已知)	$\sigma_1^2=\sigma_2^2$ $\sigma_1^2\le\sigma_2^2$ $\sigma_1^2\ge\sigma_2^2$	$\sigma_1^2\ne\sigma_2^2$ $\sigma_1^2>\sigma_2^2$ $\sigma_1^2<\sigma_2^2$	$F=\dfrac{n_1\sum_{i=1}^{n_1}(x_i-\mu_1)^2}{n_2\sum_{j=1}^{n_2}(y_j-\mu_2)^2}$	$F>F_{\frac{\alpha}{2}}(n_1,n_2)$或 $F<F_{1-\frac{\alpha}{2}}(n_1,n_2)$ $F>F_{1-\alpha}(n_1,n_2)$ $F<F_\alpha(n_1,n_2)$
F 检验法 $(\mu_1,\mu_2$ 未知)	$\sigma_1^2=\sigma_2^2$ $\sigma_1^2\le\sigma_2^2$ $\sigma_1^2\ge\sigma_2^2$	$\sigma_1^2\ne\sigma_2^2$ $\sigma_1^2>\sigma_2^2$ $\sigma_1^2<\sigma_2^2$	$F=\dfrac{S_1^2}{S_2^2}$	$F>F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$或 $F<F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$ $F>F_{1-\alpha}(n_1-1,n_2-1)$ $F<F_\alpha(n_1-1,n_2-1)$

疑 难 分 析

1、什么是显著性检验？其基本思想是什么？有什么缺陷？

显著性检验是指只考虑一个假设是否成立的检验. 其原则是，只要求犯第一类错误的概率不大于设定的$\alpha(0<\alpha<1)$.

基本思想是：根据小概率事件在一次试验中一般不应该发生的实际推断原理来检验假设是否成立.

其缺陷是：由于只有一个假设，不能评判显著性检验方法本身的好坏，因而对同一假设的众多显著性检验法难以评定优劣.

2、对于实际问题的择一检验中，原假设与备择假设地位是否相等？应如何选择原假设与备择假设？

假设检验是控制犯第一类错误的概率，所以检验发本身对原假设起保护的作用，决不轻易拒绝原假设，因此原假设与备择假设的地位是不相等的，正因为如此，常常把那些保守的、历史的、经验的取为原假设，而把那些猜测的、可能的、预期的取为备择假设.

3、参数的假设检验与区间估计之间有什么关系？

常见的区间估计与相应的参数的假设检验有着密切联系，一般某个参数的置信区间可以确定关于此参数的假设检验的接受域. 如 $X \sim N(\mu,\sigma^2),\sigma^2$ 已知，

$X_1,X_2,\cdots,X_n$ 为一个样本. 对于给定置信度 $1-\alpha$ ， μ 的置信区间为：

$(\overline{X}-Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}},\overline{X}+Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}})$. 而 μ 的显著性水平为 α 的拒绝域为（假设 $H_0:\mu=\mu_0$ ）为 $(\overline{X}-\mu_0)\sqrt{n}/\sigma<Z_{\frac{\alpha}{2}}$. 从以上结果可以看出，置信度 $1-\alpha$ 的 μ 的置信区间与关于 μ 的假设的显著性水平为 α 的接受域是相呼应的，由它们中的一个可以确定另一个.

例 题 解 析

【例 1】根据长期资料分析，钢筋强度服从正态分布. 今测得六炉钢生产出钢的强度分别为：48.5, 49.0，53.5，49.5，56.0，52.5；能否认为其强度的均值为 52.0（ $\alpha=0.05$ ）？

分析：问题为在 σ^2 未知的条件下，检验 $\mu=52.0$.

解：检验假设 $H_0:\mu=52.0$

取统计量 $T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(6-1)$

当 $\alpha=0.05$ ，自由度 $n-1=5$ ，查 t 分布表得临界值 $t_{0.025}=2.57$

由题意得统计量 T 的观察值 $t=-0.41$

由于 $|t|=0.41<2.57=t_{0.025}$ ，所以接受假设 H_0 ，即认为钢筋的强度的均值为 52.0.

【例 2】两台机床加工同一种零件，分别取 6 个和 9 个零件测量其长度，计算得 $S_1^2=0.345,S_2^2=0.357$ ，假设零件长度服从正态分布，问：是否认为两台机床加工的零件长度的方差无显著差异（ $\alpha=0.05$ ）？

分析：问题为在 μ_1,μ_2 未知的条件下，检验 $\sigma_1^2=\sigma_2^2$.

解：检验假设 $H_0:\sigma_1^2=\sigma_2^2$

选择统计量 $F=\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$ ，因为 $F_0=\dfrac{0.345}{0.357}=0.9664$ ，

而 $F_{0.975}(5,8)=1/F_{0.025}(8,5)=0.1479, F_{0.05}(5,8)=4.82$，所以有

$F_{0.975}(5,8)<F_0<F_{0.05}(5,8)$，故接受 H_0，即认为两台机床加工的零件长度的方差无显著差异.

第九章　方差分析和回归分析

内 容 提 要

1、方差分析

（1）基本概念

方差分析：通过随机抽样及数据处理，检验试验结果是否受试验条件这一类可控制因素显著影响，从而确认对质量指标影响主要来自哪一类因素，即用来鉴别所谓因素效应的有效统计分析方法.

因素（因子）：人为可以控制的实验条件称为因素或因子.

水平：因素或因子的不同等级或因素所处的不同状态称为因素的不同水平.

单因素试验：试验中如果只有一个因素或因子在变化，其它可控条件保持不变，这样的方差试验称为单因素试验.

多因素试验：试验中不止一个因素或因子在变化，称为多因素试验.若只有二个因素在变化就叫双因素试验.

（2）单因素试验的方差分析

设因素 A 有 j 个不同水平（ $j=1,2,\cdots,r$ ），在总的 r 个水平下均重复试验 i 次 $(i=1,2,\cdots,m)$.每一个水平视为一个独立总体 $X_j \sim N(\mu_j,\sigma_j^2)$ ，每个水平下总的 m 次试验结果视为取自 X_j 的容量为 m 的样本 $(X_{1j},X_{2j},\cdots X_{kj},\cdots,X_{mj})$.单因素方差分析的一般方法步骤如下：

1）提出待检假设 H_0： $\mu_1=\mu_2=\cdots=\mu_r=\mu$ ；

2）列方差计算表 9-1，计算 S_A^2 、 S_E^2 ；

3）选取建立 F 统计量

$$F=\frac{mr-r}{r-1}\cdot\frac{S_A^2}{S_E^2}\sim F(r-1,mr-r)$$ ，并计算 F 统计量的值；

4）对给定的检验水平 α ，查 F 分布表，找到 F 统计量的临界值（表值）；

5）比较得出结论：

① 若计算值 $F>F$ 临界值 $F_\alpha(r-1,mr-r)$ ，拒绝 H_0，即因素水平影响显著，或有显著影响；

② 若计算值 $F<F$ 临界值 $F_\alpha(r-1,mr-r)$ ，接受 H_0，即因素水平影响不显著或没有显著影响.

单因素方差分析见表 9-2.

表 9-1

水平 / 试验序号	A_1	A_2	…	A_j	…	A_r	
1	x_{11}	x_{12}	…	x_{1j}	…	x_{1r}	
2	x_{12}	x_{22}	…	x_{2j}	…	X_{2r}	
…	…	…	…	…	…	…	
i	x_{i1}	x_{i2}	…	x_{ij}	…	x_{ir}	
…	…	…	…	…	…	…	$(\sum)$
m	x_{m1}	x_{m2}	…	x_{mj}	…	x_{mr}	
$T_{\cdot j}=\sum_{i=1}^{m}x_{\cdot j}$	$T_{\cdot 1}$	$T_{\cdot 2}$	…	$T_{\cdot j}$	…	$T_{\cdot m}$	$x_{\cdot\cdot}=\sum_{j=1}^{r}T_{\cdot j}$
$T_{\cdot j}^2=(\sum_{i=1}^{m}x_{\cdot j})^2$	$T_{\cdot 1}^2$	$T_{\cdot 2}^2$	…	$T_{\cdot j}^2$	…	$T_{\cdot m}^2$	$T^*=\sum_{j=1}^{r}T_{\cdot j}^2$
$T_j^2=\sum_{i=1}^{m}x_{\cdot j}^2$	T_1^2	T_2^2	…	T_j^2	…	T_m^2	$T^2=\sum_{j=1}^{r}T_j^2$

2、回归分析

（1）基本概念

回归分析：利用样本数据建立起相关变量之间相关关系的数学模型，并应用统计推断的一般法则，对相关关系进行有效的统计分析方法.

一元线性回归模型为 $Y=a+bx+\varepsilon$，$\varepsilon\sim N(0,\sigma^2)$ 其中，a、b 称为回归系数.

（2）最小二乘法

表 9-2

方差来源	离差平方和	方差	自由度	F 统计量计算值	F 临界值（表值）
组间	S_A^2	$\frac{S_A^2}{r-1}$	$r-1$	$F=\frac{(m-r)S_A^2}{(r-1)S_E^2}$	$F_\alpha(r-1,mr-r)$
组内	S_E^2	$\frac{S_E^2}{mr-r}$	$mr-r$		
总变差	S_T^2	$\frac{S_T^2}{mr-1}$	$n-1$ 或 $mr-1$		

线性回归方程可表示为 $\hat{y}=\hat{a}+\hat{b}x$，可用最小二乘法求得回归系数的估计值：

$$\begin{cases}\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}\\ \hat{a}=\bar{y}-\hat{b}\bar{x}\end{cases}\quad 或\quad \begin{cases}\hat{b}=\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\\ \hat{a}=\bar{y}-\hat{b}\bar{x}\end{cases}$$

令 $\begin{cases}L_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-\frac{1}{n}(\sum_{i=1}^{n}x_i)^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2\\ L_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}\\ L_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-n\bar{y}^2\end{cases}$ 则 $\begin{cases}\hat{b}=\dfrac{L_{xy}}{L_{xx}}\\ \hat{a}=\bar{y}-\hat{b}\bar{x}\end{cases}$

（3）线性相关显著性检验

检验回归方程 $\hat{y}=\hat{a}+\hat{b}x$ 是否有效，或 x、y 之间线性相关关系是否显著，只要检验回归系数 $\hat{b}=0$ 是否成立.这个问题只存在下面两种可能

$$\hat{b}=0\begin{cases}是，即\Rightarrow \hat{y}=\hat{a}; & 1)\\ 否，即b\neq 0\Rightarrow\begin{cases}\hat{y}=\hat{a}+\hat{b}x & (\hat{a}\neq 0)\\ \hat{y}=\hat{b}x & (\hat{a}=0)\end{cases} & 2)\end{cases}$$

1）表示 $\hat{y}$ 与 x 无关 即 y 与 x 没有线性相关关系，反过来，若不能否定 1），就表示线性相关显著；

2）有二种可能，都表示 y 与 x 之间线性相关关系成立.

运用 R 检验法进行线性相关显著性检验. R 检验法检验线性相关显著性或回归方程有效性的一般步骤为：

1）提出检验假设：H_0：$\hat{b}=0$；

2）选用统计量 $R=\dfrac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}\sim R(n-2)$，并计算 $|R|$ 值；

3）在给定 α 下，查相关系数表得到临界值 $R_\alpha(n-2)$；

4）比较得出结论：

① 若 $|R|>R_\alpha(n-2)$，拒绝 H_0，所求回归方程有效或线性相关显著；

② 若 $|R|<R_\alpha(n-2)$，接受 H_0，所求回归方程无效或线性相关不显著.

在实际应用中，经常出现 $|R|\approx 1.0$ 的情况，这时不用查表即可判断线性相关显著，即回归方

程有效.

疑 难 分 析

1、怎样区分讨论的问题是方差分析还是回归分析？

实际问题所考察的指标 y 往往既受因素 x_i 的影响，又受随机误差的影响. 而因素又分为属性的和数量的. 属性的因素一般无数量大小可言，只是性质的不同，如：种子的品种、机器的型号、加工的工艺、材料的品质等等. 数量的因素，可以在一定范围内取值，如：人的身高、体重，试验的温度，产品的合格率等等. 当所考虑的因素是属性时，问题属于方差分析的范围；当所考虑的因素是数量时，问题属于回归分析的范围.

例 题 解 析

【例 1】设某地区酿酒公司下属有 A_1、A_2、A_3、A_4 共 4 个酒厂.公司总经理为提高酒的质量，开展质量评优活动，随机地从 4 个酒厂各抽取 3 瓶样酒，指定同一名品酒员按事先规定的色、香、味质量标准评分，评分结果的原始数据如表 9-3 所示.

表 9-3

试验序号 \ 厂别	A_1	A_2	A_3	A_4
1	5	8	7	11
2	6	9	8	10
3	6	8	6	12

试问：不同酒厂对酒的质量有无显著影响（$\alpha = 0.05$）？

解：（1）提出待检假设 H_0：$\mu_1 = \mu_2 = \cdots = \mu_r = \mu$；

（2）列方差计算表，如表 9-4 所示.

利用表中最后一列，即（$\sum$）列的数据计算

表 9-4

试验序号 \ 水平	A_1	A_2	A_3	A_4	
1	5	8	7	11	
2	6	9	8	10	
3	6	8	6	12	（$\sum$）
$T_{\cdot j} = \sum_{i=1}^{m} x_{\cdot j}$	17	25	21	33	$x_{\cdot\cdot} = \sum_{j=1}^{r} T_{\cdot j}$ =96

$T_{\cdot j}^2=(\sum_{i=1}^{m}x_{\cdot j})^2$	289	625	441	1089	$T^*=\sum_{j=1}^{r}T_{\cdot j}^2$ =2444
$T_j^2=\sum_{i=1}^{m}x_{\cdot j}^2$	97	209	149	365	$T^2=\sum_{j=1}^{r}T_j^2$ =820

$$S_E^2=T^2-\frac{1}{m}T^*=820-\frac{1}{3}\times 2444=5.33$$

$$S_A^2=\frac{1}{m}\cdot T^*-\frac{x_{\cdot\cdot}^2}{mr}=\frac{1}{3}\times 2444-\frac{96^2}{12}=46.67$$

(3)选 F 统计量并求 F 计算值和临界值

$$F=\frac{12-4}{4-1}\cdot\frac{S_A^2}{S_E^2}=\frac{8\cdot S_A^2}{3\cdot S_E^2}\sim F(3,8)$$

$$F=\frac{8\times 46.67}{3\times 5.33}=23.35$$

又查附表 5， $F_{0.05}(3.8)=4.07$

（4）比较得出结论

因为 $F=23.35>>F_{0.05}(3,8)$，拒绝 H_0，即表示不同酒厂对酒的质量有显著影响.这里 $F>>F_\alpha$，可认为因素水平影响特别显著，事实上由原始数据可见，A_4 评分特别高，直观上已可判断有显著差异，说明分析的结论是符合实际情况的，也证明了方差分析的科学性.

【例 2】设有某种创汇商品在国际市场上需求量 q（单位：万件），价格 p（单位：万美元/件）. 根据往年市场调查获悉 q 与 p 之间的一组调查数据如表 9-5 所示.

表 9-5

价格 p_i	2	4	4	4.5	3	4.2	3.5	2.5	3.3	3
需求量 q_i	6	2	2	1	4	1.5	2.8	5.1	3.4	4.2

如果今年该商品预定价为 p=4.6（万美元/件），要求根据往年资料建立的 q 对 p 的回归方程，进行线性相关性是否显著，并预测国际市场上今年的需求量大致为多大？（$\alpha=0.05$）

解：根据样本数据，用最小二乘法求 $\hat{a}$、$\hat{b}$ 的值.

$$\hat{b}=\frac{\sum_i p_iq_i-10\overline{pq}}{\sum_i p_i^2-10(\overline{p})^2}=\frac{97.17-10\times 3.4\times 3.2}{121.8-10\times 3.4^2}=-2.04$$

$$\hat{a}=\overline{q}-\hat{b}\overline{p}=3.2-(-2.04)\times 3.4=10.136$$

将 $\hat{a}$、$\hat{b}$ 的值代入得到所要求的引例中需求量 q 对价格 p 的回归方程为

$$\hat{q}=10.136-2.04p.$$

对所建立的 q 对 p 的回归方程进行线性相关性显著检验：

1）提出待检假设 H_0：$\hat{b}=0$；

2）选用统计量 $R=\dfrac{L_{qp}}{\sqrt{L_{pp}L_{qq}}}\sim R(n-2)$，并利用回归计算的结果计算 $|R|$：

因为 $L_{pp}=\sum_{i=1}^{10}p_i^2-10(\overline{p})^2=5.68$；$L_{pq}=\sum_{i=1}^{10}p_iq_i-10\overline{pq}=-11.63$；

$$L_{qq}=\sum_{i=1}^{10}q_i^2-10(\overline{q})^2=126.3-10\times(3.2)^2=23.9$$

所以 $|R|=\left|\dfrac{-11.63}{\sqrt{5.68\times23.9}}\right|=0.998$；

3）查附表 7 得到 $R_{0.05}(8)=0.632$；

4）结论

$\because |R|>R_{\alpha}(n-2)$，拒绝 H_0，即 q 对 p 的回归方程有效或线性相关性显著.

经检验说明：回归方程 $\hat{q}=10.136-2.04p$ 有效，可以用于预测.

当 p=4.6 时，国际市场上今年对该商品的需求量大致为：10.136-2.04×4.6=0.752（万件）.

概率论与数理统计