## 随机事件和概率

### 事件的关系与运算

**1．事件的包含**

若事件**发生必然导致事件发生，则称事件包含事件**（或称事件**包含于事件），记为．

**2．事件相等**

若且，则称事件**与事件相等，记为**.

**3．事件的和（并）**

当*，*至少有一个发生时，称为事件**与事件的和（并）事件，记为.

称为个事件的和（并）事件，它表示中至少有一个发生.

称为可列个事件**的和（并）事件.

**4．事件的积（交）**

当，同时发生时，称为事件与事件的积（交）事件，记为.

称为个事件的积事件，它表示都发生.

称为可列个事件**的积事件.

**5．事件的差**

当发生，不发生时，称为事件与事件的差事件，记为.

**6．互斥事件和对立事件**

，即与不能同时发生，则称，为互斥事件（或互不相容事件）；若且，其中为不可能事件，为必然事件，则称事件与事件互为对立事件（或互为逆事件），记，即的对立事件为，注意，对立事件一定是互斥事件，但互斥事件不一定是对立事件.

**7．事件的运算律**

①交换律，.

②结合律，.

③分配律，.

④德·摩根律，.

，.

⑤.

注意：事件的运算律与集合的运算律完全一样.

**8．完全事件组**

事件满足



则称为一完全（完备）事件组.

### 事件的概率及其性质

**1．概率的定义**

设是随机试验，是它的样本空间，对于的每一个事件**赋于一个实数**，且**满足

①**（非负性）；

②**（规范性）；

③设**是两两互不相容的事件，即（），，则

（可列可加性），

则称**为事件的概率.

**2．概率的性质**

①若**两两互斥，则

.

②.

③．特别地，时，.

④对任意事件**恒有

．

对任意事件**恒有

．

一般地，对个任意事件有

.

### 条件概率与事件的独立性

**1．关于条件概率的重要公式有**

①，则．

②．

③．

注：也是一种概率，从而满足概率的定义及概率的所有基本性质．

**2．关于独立性的结论有**

①对于任意事件独立的充要条件是.

②对于**，独立的充要条件是（或）；

对于*，*独立的充要条件是（或）.

③对于，独立的充要条件是；

对于，独立的充要条件是.

④对于，独立的充要条件是；

对于，独立的充要条件是.

⑤相互独立分相互独立相互独立相互独立，其中符号“”表示等价.

⑥不可能事件与任何事件独立；必然事件与任何事件独立.

注意：“两两独立”与“相互独立”的区别．两两独立（，）；相互独立（），即任何两个独立且（互不相等），即任何三个独立且任何四个、五个，，个均独立，共满足个等式．如当且当，，时，两两独立；当且仅当，，，时，**相互独立.

### 三大公式

**1．乘法公式**

．推广：为个任意事件，则

.

**2．全概率公式**

设是一个完备事件组，即满足（，）且，则有全概率公式

.

**3．贝叶斯公式（又称逆概公式）**

设为一完备事件组，则有叶斯公式

.

## 随机变量及其分布

### 随机变量的分布函数

**1．随机变量的分布函数的概念**

（）.

**2．随机变量分布函数*F*（*x*）的性质**

①（）；

②，；

③**在是右连续函数，即.

**3．一个重要公式**

**为任一随机变量的分布函数，则**.

### 离散型随机变量

**1．离散型随机变量的概率分布（分布列或分布律）**

，

或用列表法表示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

其中，.

**2．离散型随机变量的分布函数*F*（*x*）**

，．

**为右连续的阶梯形函数．

**3．一些重要的离散型分布**

①分布（两点分布）

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

（0＜＜1）

②二项分布，记为**，**，则在次独立重复试验中出现次的概率为

，.

③泊松分布

，．

### 连续型随机变量

**1．****为连续型随机变量密度函数的充要条件

且．

**2．**连续型随机变量的分布函数**与密度函数的关系

，．

**3．计算概率的方法**

．

注意：由于连续型随机变量在任何一点外的概率为0，因此，若为连续型，则，这一点与离散随机变量不同．

**4．一些重要的连续型随机变量**

①服从正态分布，记为，其密度函数为

，．

当**，时，称为标准正态分布，记为**．

②**服从的均匀分布，记为，其密度函数为

.

对应的分布函数为



③参数为的指数分布，其密度函数为



其中，对应的分布函数为



### 随机变量函数的分布

**1．**设为离散随机变量，其概率分布为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

考虑的概率分布．当，，，，互不相等时，的概率分布为，；当有相等的值时，应合并，如

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

则**有两个不同值*，*，于是，．

**2．**设为连续型随机变量，其密度函数为**，若也为连续型随机变量，求的密度函数，一般有两种方法

**方法一：**基本方法，先用分布函数的定义求出的分布函数，再求的密度函数．

**方法二：**利用下列定理直接求**的密度函数：设函数处处可导且有（或恒有）则的密度函数为



其中，，为的反函数．

## 多维随机变量及其分布

### 二维随机变量的分布函数及其性质

**1．定义**

的分布函数为．

**2．**对任意函数**为分布函数的充要条件是同时满足下列四个性质

①**是关于或的不减函数；

②**，，，；

③**是关于**或的右连续函数；

④对任意，，恒有．

**3．利用分布函数计算概率的公式**

**.**

**4．边缘分布函数**

**的分布函数为，的分布函数为．

### 二维离散型随机变量

**1．**设为二维离散型随机变量，则其概率分布为

，.

或写成表格形式：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**2．的边缘分布**

关于**的边缘分布律为，.即上述表格中第行的和．

关于的边缘分布律为，.即上述表格中第列的和．

**3．**与独立的充要条件是对一切基本事件中的和恒有．

### 二维连续型随机变量

**1．**设的分布函数为**，若存在二元非负函数，使得

，

则称**为的概率密度，从而也有.

**2．**对任意函数**为密度函数的充分必要条件是且．

**3．**设为一平面区域，则

．

**4．**关于或的边缘密度函数分别为

，．

**5．**独立的充分必要条件是**或，这里的密度函数分别为，分布函数分别为**，的密度函数为，分布函数为.

**定理**：设是二维连续型随机变量，则与独立的充要条件是其联合密度函数，且，，其中为常数或．

**6．**在条件下的条件密度为

.

对应的条件分布函数为

．

在**的条件下**的条件密度为

．

对应的条件分布函数为

．

因此，有下列公式

．

独立的充要条件是条件密度（分布）等于无条件密度（分布），即,（或，）.

### 重要的二维分布

**1．二维均匀分布**

服从平面区域上的均匀分布，则其密度函数为



其中**为的面积.

**2．二维正态分布**

即，其密度函数为

.

**性质：**，则，；独立的充要条件是**与**不相关，即；一定服从正态分布；关于或关于的条件分布也是正态分布．

### 两个随机变量的函数的分布

**1．**求**的密度函数最基本的方法

先求的分布函数**，然后再求密度函数．

**2．**有些简单函数的密度函数也可利用下列公式求解．

设为连续型随机变量，其概率密度为．

①**（，），则的密度函数为

.

特别地，当**时，.

当**时，.

②**，则**的密度函数为

.

③，则的密度函数为

.

④设**相互独立，其分布函数分别为**，则

，

的分布函数分别为

，.

**3．**对于为离散型随机变量的情形，采用分解法

如已知的概率分布为,，求**的分布列，.

**4．常用结论**

①，且**独立，则.

**推广：**，，则.

②服从参数为的泊松分布，**服从参数为的泊松分布，且**独立，则**服从参数为的泊松分布.

**推广：**服从参数为的泊松分布，，则**服从参数为的泊松分布．

③，，且**独立，则．

**推广：**，且**相互独立，则.

④服从参数为的指数分布，，且**相互独立，则**服从参数为的指数分布.

## 随机变量的数字特征

### 一维随机变量的数字特征

**1．数学期望**

①为离散型，其分布列为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

若，则．

②为连续型，设其密度函数为，若，则．

**2．数学期望的性质**

①设为常数，则．

②．

③对任意常数，．

**推广：****为任意常数，则

.

④若独立，则．

**推广：**相互独立，则．

注意：一般说来*，*相互独立能推出，但反之，不能推出独立，事实上的充要条件是不相关．

⑤若**，则；若，则．

⑥．

⑦柯西–瓦茨（Cauchy-Schwarz）不等式，若**均存在，则**存在，且

．

**3．方差**

①设为离散型，其分布列为，，若，则均存在，，的方差．

②设为连续型，其密度函数为，若，则，均存在，，的方差为．

**4．方差的性质**

①设为常数，则．

②设为常数，则，特别地，时，．

③若独立，则．

**推广：****为常数，相互独立，则

．

注意：一般说来**，若相互独立，则能推出**，但反之，不能推出独立，事实上**的充要条件是不相关．

④设**相互独立，则

．

⑤对一切实数，

．

**5．**一维随机变量函数的期望和方差

①若离散，设其分布列为，，

**方法一：**先求出的分布列，再求的期望和方差．

**方法二：**利用下列公式

，，.

②若*X*是连续型，设其密度为**，若也是连续型，求，．

**方法一：**先求出**的密度函数**，则

，，

．

**方法二：**直接利用下列公式

，，

.

### 二维随机变量的数字特征

**1．数学期望和方差**

①若为离散型随机变量，设其概率分布为

,，

求．

**方法一：**先求和的分布列（边缘分布），再求.

如的分布列为，于是

，，.

**方法二：**直接利用下列公式求

，，

，，

，.

②若是连续型，设其联合密度为**.

**方法一：**先求出的边缘密度，，再求.

**方法二：**直接利用下列公式

，，

，，

，.

③二维随机变量函数的数学期望和方差．

**方法一：**先求出的分布列或密度函数，再求**和.

**方法二：**直接利用下列公式.

若为离散型，其分布列为，则

.

.

.

若为连续型，设其密度函数为，则

，

，

.

**2．协方差：**

①的协方差为

．

②协方差的性质

，

，

，

，

.

对任意随机变量*X*，*Y*恒有：

，

，

．

**3．相关系数**

①的相关系数：．

②相关系数的性质：

，

的充要条件是；的充要条件是.

③与不相关的相关系数，其中“”表示相互等价．

注：**独立则一定不相关，但不相关不能推出独立．

### 一些常用分布的数学期望和方差

**1．**服从二项分布，即，则，，，．

**2．**服从参数为的泊松分布，则，，，.

**3．****服从几何分布，则，，，.

**4．****服从超几何分布，则，，，，.

**5．****服从正态分布，则其密度函数为

，，

，.

**6．****服从上均匀分布，则其密度函数为



，.

**7．****服从参数为的指数分布，则其密度函数为



，.

**8．****服从自由度为的分布，，．

## 大数定律和中心极限定理

### 大数定理

**1．切比雪夫不等式**

设随机变量的方差存在，则对任意**，有

，

即

.

**2．几个重要的大数定理**

①切比雪夫大数定理

设为两两不相关的随机变量序列，且每个随机变量的方差存在而且有公共上界，即，，则对有

.

②伯努利大数定理

设为重独立伯努利试验中事件出现的次数，设每次试验事件出现的概率为，即，则对任意，有

.

③辛钦大数定律

设是独立同分布的随机变量序列，且数学期望存在，，则对任意，有

.

### 中心极限定理

**1．列维-林德伯格（Levy-Lindberg）定理**

设**为独立同分布的随机变量序列，，，，则对任意实数，有

.

**2．棣莫弗–拉普拉斯定理**

设为重独立伯努利试验中事件出现的次数，设每次试验事件出现的概率为（），即，则对任意实数，有

.

## 数理统计的基本概念

### 基本概念

**1．总体、个体和样本**

在数理统计中，我们把研究对象的全体所组成的集合称为总体（或母体），通常用随机变量**表示，组成总体的每一个元素称为个体；从总体中随机抽取若干个个体，这些个体称为样本，个体的个数称为样本容量，容量为的样本用表示．

**2．**样本的性质及其联合分布

容量为的样本相互独立且与总体同分布．

若总体的分布函数为，则的联合分布函数为**．

若总体是连续型，其概率密度为**，则的联合概率密度为.

若总体是离散型，其分布列为，则的联合概率分布为

.

**3．统计量**

①不含总体分布中任何未知参数的样本函数称为统计量

②一些重要的统计量

设是总体*X*的样本，则

，称为样本均值．

，称为样本方差，称为样本标准差.

，称为样本阶原点矩.

，称为样本阶中心矩.

③定理

设为任一总体，为来自的样本，,，则

，，.

### 三个抽样分布

**1．–分布**

①设独立同分布，均服从．则服从自由度为**的分布，记为．

②****分布具有下列性质：

**，，且独立，则；

，，．

③分位数：

若，且，则称数**为自由度为**的分布的**分位数．

**2．**–分布**

①设，，且独立，则服从自由度为的分布，记为．

②性质：

分布的密度函数关于**轴对称，且其极限分布为标准正态分布；，则，（）.

③分位数：

若，且，则称实数为自由度为的分布的分位数.

**3．*F*分布**

①设，，且独立，则称为自由度为的分布，并记为.

②性质：

，则.

③分位数：

若，且，则称实数为自由度为的**分布的分位数．

.

### 正态总体的样本均值与样本方差的分布

**定理1：**设是来自正态总体**的样本，则

，，，

，且与独立.

**定理2：**设和分别是取自两个相互独立的正态总体和的两个样本，其样本均值分别为，样本方差分别为，则

，，.

当时，，且相互独立，

其中.

## 参数估计

### 矩估计与最大似然估计

**1．点估计的概念，估计量与估计值**

用一个数值来估计某个参数，这类问题称为参数的点估计；设为总体**的待估计的参数，用样本的一个统计量来估计，则称为的估计量；对应于样本的一次观察值，估计量的值（*x*1，*x*2，…，*xn*）称为的估计值.

**2．矩估计法**

求矩估计的基本方法是：假定总体**的密度函数或分布列含有**个未知参数，为一简单随机样本，则建立个方程组：



从中解出，则分别为**的矩估计量．

特别地，当总体**的分布中含有一个未知参数时，直接利用方程解出的矩估计量，当总体的分布中含有两个未知参数时，可直接利用下列两个方程



求出的矩估计量．

**3．最大似然估计法**

求最大似然估计的基本方法：假定总体为连续型随机变量，其概率密度为，其中为未知参数，则考虑似然函数

（只考虑非零部分）.

取对数，然后分别对未知参数求偏导数，并令其导数等于，得到

．

解出，则分别为的最大似然估计.

注意：当对未知参数求偏导无解时，可考虑下列两种方法.

**方法一：**利用观察法直接求出当时最大，则即为的最大似然估计.

**方法二：**利用（或），判断是的单调增还是单调减函数，从而求出时，最大，则即为的最大似然估计.

在利用这两种方法时，一定要注意似然函数中非零部分的范围.

注意，若总体是离散型，则似然函数为

.

一定要用表示，用上述类似的方法求出最大似然估计.

### 估计量的评选原则

**1．无偏性**

设为的估计量，为的取值范围，若对任意，有，则称是的无偏估计量（或称是无偏的）．

**2．有效性**

设，均为的无偏估计量，若，则称比有效．

**3．一致性**

设为的估计量，对任意，有

，

即依概率收敛于，则称为的一致估计量（或相合估计量）.

### 区间估计

**1．置信区间**

设总体的分布函数含有一个未知参数．对于给定值**（），若由样本确定的两个统计量，满足，则称为的置信度为的置信区间，分别称为置信下限和置信上限．

**2．**整个正态总体总均值和方差的区间估计

设为总体的样本，，分别是样本均值和样本方差，设已给定置信度为列表如下：

|  |  |  |  |
| --- | --- | --- | --- |
| 待估参数 | 其它参数 | 统计量及其分布 | 待估参数的置信区间 |
|  | 已知 |  |  |
| 未知 |  |  |
|  | 已知 |  |  |
| 未知 |  |  |

**3．**两个正态总体，的情况

设已给定置信度为，并设**来自总体，来自第二个总体的样本，分别为第一、二个总体的样本均值，分别是第一、二个总体的样本方差，列表如下：

|  |  |  |  |
| --- | --- | --- | --- |
| 待估参数 | 其他参数 | 统计量及其分布 | 置信区间 |
|  | 已知 |  |  |
| 未知 |  |  |
|  | 均为已知 |  |  |
| 均为已知 |  |  |

## 假设检验

### 显著性检验的基本思想

设有某假设需要检验，先假设是正确的，在此假设之下，构造某一个事件，它在为正确的条件下发生的概率很小．现在进行一次试验，如果事件发生了，就说一个小概率事件居然发生了，就完全有理由拒绝的正确性，否则没有充分理由拒绝的正确性，从而接受，这就是显著性检验的基本思想．

### 假设检验的基本步骤

1．根据给定问题提出原假设利备选假设**.

2．寻找检验统计量，并在为真下推导统计量的分布.

3．给定显著性水平，确定检验的拒绝域*W*.

4．作判断：由样本值计算统计量的值．若统计量的值落入拒绝域，则拒绝；否则，接受.

### 两类错误

|  |  |  |
| --- | --- | --- |
| 决策 | 假设 | |
| 为真 | 不真 |
| 否定 | 第一类错误 | 正确 |
| 不否定 | 正确 | 二类错误 |

### 单个正态总体的均值和方差的假设检验

设总体服从正态分布，来自总体的样本，为样本观察值，给定显著水平.

**单个正态总体均值和方差的假设检验（双边检验）**

|  |  |  |
| --- | --- | --- |
| 原假设 | 检验统计量及为真时的分布 | 的拒绝域 |
| （已知） |  |  |
| （未知） |  |  |
| （已知） |  | 或 |
| （未知） |  | 或 |

### 两个正态总体的均值和方差的假设检验

设是来自正态总体的样本观察值，是来自正态总体的样本观察值，设两样本独立，它们的样本均值分别为，样本方差分别为.

**两个正态总体均值和方差的假设检验（双边检验）**

|  |  |  |
| --- | --- | --- |
| 原假设 | 检验统计量及为真时的分布 | 的拒绝域 |
| （已知） |  |  |
| （未知） |  |  |
| （未知） |  | 或 |