1. 学科地位

概率论与数理统计是研究1随机现象的2统计规律的3数学学科。

这个定义中有三个要素：一是随机现象，大体上我们可以将日常中的事件分为确定性事件和不确定性事件两类，而概率论研究的对象是后者；第二，但虽然随机现象的结果是不确定的，但在大量重复实验中却又体现出一定的**统计规律**，这也就是我们研究的价值之所在；对以上两点，我们可以举一个简单的例子来解释：抛硬币是一个随机事件，因为每一次是哪一面朝上是不确定的，但是如果不断抛硬币，正面朝上的频率会接近0.5（以后我们会看到这是由于大数定律），这就是统计规律。最后，概统是数学的一个分支，和数学的其他方面有着紧密的联系。

2. 随机试验

为了研究随机现象，我们需要对事件进行观察，这个观察的过程被称为**随机试验**，简称**试验**，通常记作。随机试验需要满足以下三个特点：

①在相同的条件下，试验可以重复进行；

②每一次试验的可能结果不止一个（否则就是确定性事件了），并且能够**事先明确**试验的所有可能结果。值得注意的是，事先明确所有可能的结果并不意味着试验的结果只能是有限个，无限个（包括可数无穷[[1]](#footnote-1)和不可数无穷）也是可以的；

③在每一次试验之前不能确定哪一个结果会出现。

满足以上三个条件的试验才能够被称为是随机试验。

3. 样本空间和随机事件

我们知道，随机试验的所有可能结果是事先明确的，我们将随机试验的所有可能结果所构成的集合称为的**样本空间**，通常记作或. 样本空间的每一个元素（即随机试验的每一个可能的结果）称为**样本点**，通常记作, 则我们有

举个例子，对于抛一枚硬币，记正面为反面为, 则样本空间可以写作

再看一个不可数无穷的例子（这种情况无法用列举法表示）：在一批已知最长寿命为的灯泡中任意抽取一只，测试它的寿命, 则样本空间可以写作

随机试验的样本空间的子集称为的**随机事件**(random event), 简称**事件**；对于事件我们定义发生的概念：当中的某一样本点出现时，我们称事件**发生**(happen).  
 仅由一个样本点构成的单点集称为**基本事件**(elementary event)[[2]](#footnote-2). 例如，对于上面的抛一次硬币，其基本事件有. 我们注意到，所有事件（除不可能事件外）都可以被表示为基本事件的并，这也就是基本事件“基本”的原因，由若干个基本事件组合成的事件称为**复合事件**。

4. 必然事件和不可能事件

我们注意到，样本空间本身可以作为一个事件（集合是自身的子集），其中包含了事件的所有样本点，在每一次随机试验中必然发生，我们称样本空间为必然事件(certain event). 而空集（空集是任何集合的子集）中不包含任何样本点，在每一次随机试验中都不会发生，因此我们称空集为不可能事件(impossible event).

5. 事件个数的计算

由于n元集合的子集个数为2^n, 因此一个具备n个样本点的试验就具有2^n个不同的事件。例如，抛硬币事件有2个样本点，则该试验有2^2=4个基本事件，其中包括一个不可能事件、一个“出现正面”事件、一个“出现反面”事件，和一个必然事件。

6. 事件的关系和运算

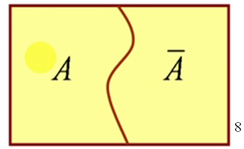
由于事件被定义为样本空间的自己，因此事件的关系和运算实际上就是集合的关系和运算，完全是一样的，这里不再赘述。

值得注意的一点是，在事件的运算中有一些特殊的表述方法：事件的并又被称为事件的**和**，对于有限n个事件的和可以用大型运算符表达为

这两种写法是等效的。

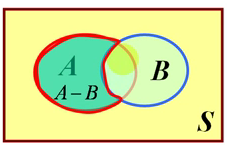
另外，事件的交又被称为事件的**积**，对于有限n个事件的积可以表达为

7. 对立事件

 如果在每一次试验中，A和B中必有一个发生，但又不同时发生，即满足：

则称事件A与事件B互为**对立事件**（**逆事件**）。事件的对立事件通常记作, 实际上也可以理解为集合中的补集概念：. 因而可以用右边的Veen图表示对立事件。

和补集的性质一样，对事件求两次对立事件就会像这样回到原事件: .

8. 事件的差

事件的差就是集合的差。它服从以下两条性质：

9. 互斥事件

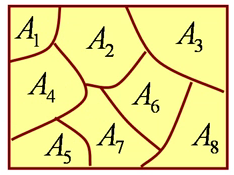
如果事件A和事件B不能同时发生，即, 则称A与B互不相容（互斥）.

事件A与事件B互斥等价于.

显然，基本事件两两互斥。

互斥事件和对立事件概念值得加以比较，我们知道互斥要求, 而对立只要求, 这就说明对立事件一定是互斥的，但互斥事件不一定对立。

我们可以在互斥概念的基础上定义完备事件组：

 如果事件两两互斥，且满足, 则称事件是一个**完备事件组**。一个完备事件组也可以看作是对样本空间的一个划分。

容易看出，在每一次试验中完备事件组中有且只有一个事件发生。事实上完备事件组是对对立事件的一个推广。（从2个事件对立推广到了多个事件“对立”）

**随机变量及其分布**

1. 随机变量

通常，为了便于研究随机事件，我们可以将随机试验的结果映射到一个数字上，像这样随着试验结果变化而变化的变量称为**随机变量**，通常用字母。因此随机变量事实上相当于函数的因变量，其取值范围相当于函数的值域，随机试验的结果则相当于函数的定义域。

所有取值可以一一列出的变量称为**离散型随机变量**。通过适当地定义从试验结果到随机变量的映射，我们可以将一些连续的试验结果转化为离散型随机变量以便于研究。例如：电灯泡的寿命的数值并不是离散量，但如果我们只关心电灯泡的寿命是否不少于1000小时，就可以换个方式来定义映射：

从而将连续的值转化为离散量来进行处理。

2. 随机变量的分布列

若随机变量可能的不同取值为, 则取得每一个值的概率, 称为随机变量的**概率分布列**，简称**分布列**。根据概率的基本性质可得：

分布列**通常**都使用表格形式来表现：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | ... |  | ... |  |
|  |  |  | ... |  | ... |  |

分布列也可以用图像形式表现，但用的较少。

若随机变量的分布列具有下表所示形式，则称服从**两点分布**/**0—1分布**/**伯努利分布**，并将其中的概率称为**成功概率**.

|  |  |  |
| --- | --- | --- |
|  | 0 |  |
|  |  |  |

考虑从含有件次品的件产品中任取件，其中恰有件次品的概率，根据古典概型

写成表格形式则是

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 |  | ... |  |
|  |  |  | ... |  |

若随机变量的分布列具有以上形式，则称随机变量服从**超几何分布**。

两点分布很简单，就是有且只有两种情况，其分布列如tab1所示，值得注意的是这里的成功概率被定义为X=1时的概率P(X)=p而非相反，这一点在方差计算的过程M1中非常重要.

超几何分布常常在一堆有两类物品的混合物中随机取物产生，若A个物品A中混有少量的B个物品B, 从中取B<n<A个，含有B的件数记为X, 那么所产生的分布列为Tab2. 由于复杂度较高，我们不给出超几何分布的期望和方差的计算通式.

二项分布是重复的两点分布事件，在成功概率为p的两点分布中重复n次，记成功的次数为X, 其分布列为Tab3, 这样的二项分布记作, 因此对于题中的, 即重复3次且成功概率为0.5的二项分布，则

二项分布的期望

这里右边的形式很类似二项式定理。请注意这里不是一个等差等比复合数列求和，只能试图利用组合数的性质。我们知道，数列求和的关键（以及众多函数处理的本质）在于把变元化归到单一项中，这里变元同时出现在系数k、二项式系数和两个指数中，我们首先试图把系数为了试图实现这一点显然后面两项是没有希望的了，只能对展开：

如果把k拆出来，那么就可以成功把系数消去，但这样就破坏了它作为一个“组合数”的特点，因此我们同时考虑把n拆出来

这样一来

这样我们先减少了一个k的存在，而n是常数项可以提到求和符号外. 接下来只需要集中注意力处理这个类似“二项式定理的展开”即可，如果可以将它化归成二项式定理的形式，那么就可以把事情变成一个二项式！由于组合数形式是不能动的（否则都会造成变元k变成常数项），因此我们必须处理指数，让两个指数之和为, 并且让其中一个指数的形式和C的上面一项一致（显然处理p的指数方便，直接挪一个下来就可以了），于是我们得到：

注意到, 于是

从而

事实上从**期望的可加性**也可以直接得到这个结论. 另外还可以得到方差

最后是正态分布，其定义为

记作. 只需要注意到图像对称性以及即可，图像关于直线对称, 值只影响到正态分布图像的左右平移，而方差只影响函数的形状, 越小，函数图像越瘦高，表明数据分布越集中.

**概率论**

1. 条件概率

设,是两个事件且, 则在事件发生的条件下，事件发生的概率称为**条件概率**，记作. 根据古典概型概率公式，设概率空间为, 则

1. 能够和自然数集建立一一对应关系的集合称为可数无穷集合。 [↑](#footnote-ref-1)
2. 徐小湛视频中认为也可以将样本点本身看成基本事件，我们认为这样不妥，因为事件被定义为子集，不应当用元素来理解。 [↑](#footnote-ref-2)