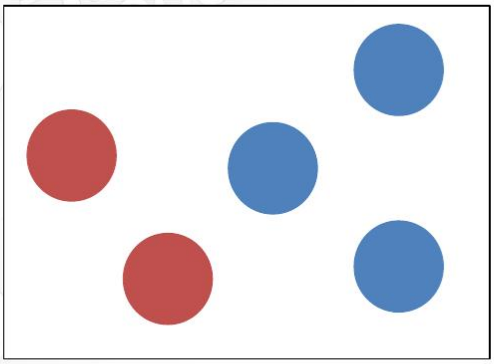
概率论

概率论

- 基本概念
- 随机变量及其分布
- 期望、方差、协方差
- 常见概率分布
- 大数定律

基本概念

◆ 试验、事件与样本空间



试验：本场游戏

样本空间：整个方框

事件：游戏中可能发生的事情

基本概念

◆ 条件概率与独立事件

条件概率是指在某个特定条件或情况已经发生的前提下，另一个事件发生的概率。

如打雷下雨时，外出带伞的几率会增加

独立事件是指两个（或更多）事件发生时，一个事件的结果不会影响另一个事件的结果。

如两次掷硬币的结果之间没有相互影响

案例实践

◆ 使用Python模拟随机实验

随机变量及其分布

◆ 随机变量的定义

随机变量是一个数值结果，它的值取决于随机现象的结果。

随机变量可以是离散的，既有一个可数的值集合（如掷骰子的结果
也可是连续的，既是一个实数区间的任何值（如测量一个人的身高）

每个随机变量都有一个与之关联的概率分布，描述了随机变量取各种可能值的概率。

随机变量及其分布

◆ 离散随机变量与连续随机变量

离散随机变量是指其可能的结果值是分开的，可以逐个列举的。例如，掷骰子的结果

连续随机变量，顾名思义，它的可能结果是连续的，不能逐个列举。例如，一个人的身高就

随机变量及其分布

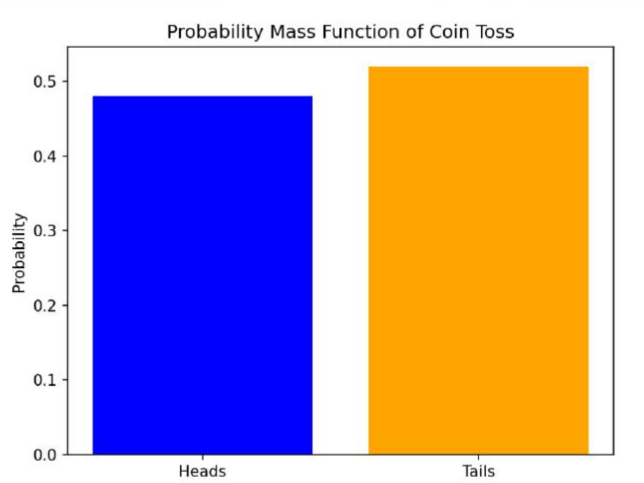
◆ 概率质量函数与概率密度函数

概率质量函数（PMF）是离散随机变量的概率表示。它给出了这个随机变量取特定值的概率。例如，掷一个公平的六面骰子，每个面上的数字出现的概率都是1/6，这就可以通过一个概率质量函数来描述。

概率密度函数（PDF）是连续随机变量的概率表示。由于连续随机变量取任何一个具体数值的概率都是0，所以概率密度函数不直接给出概率，而是给出了随机变量落在某一区间内的概率密度。PDF下某区间的面积代表随机变量落在这个区间的概率。例如，描述人的身高分布的随机变量就会用到概率密度函数。

案例实践

◆ 制作概率质量函数示意图



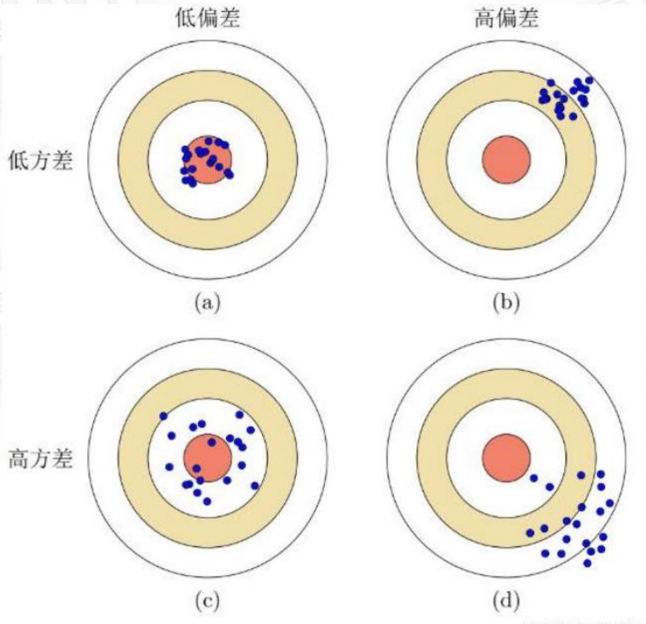
期望、方差与协方差

◆ 期望的定义与性质

期望，又称为数学期望、均值，是指随机事件在多次重复实验中的平均值。在射击比赛中，如果一个射击运动员多次射击同一目标，他的子弹击中目标的平均位置就可以看作是期望值。期望值代表的是射手的准确度，一个射手的期望命中点越靠近靶心，说明他越准确。

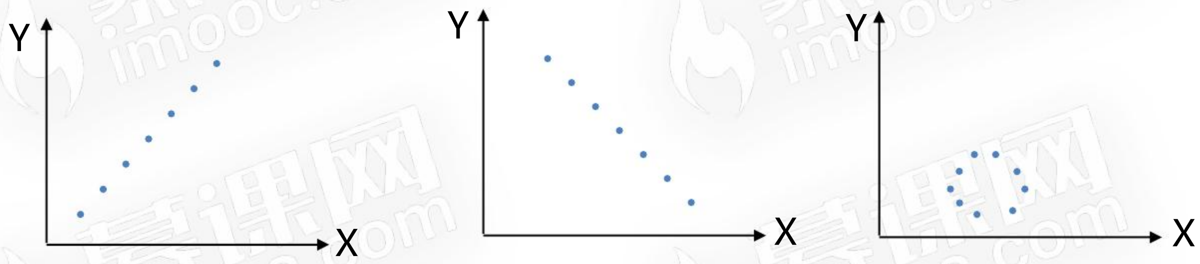
期望、方差与协方差

◆ 方差与偏差



期望、方差与协方差

◆ 协方差与相关系数



正相关 Conv(X,Y)>0　　负相关 Conv(X,Y)<0　　不相关 Conv(X,Y)=0

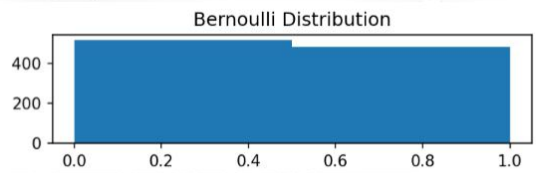
案例实践

◆ 计算期望、方差与协方差

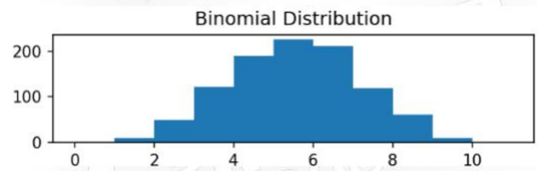
常见的概率分布

伯努利分布：描述只有两种可能结果的单次随机试验

二项分布：描述固定n次的独立伯努利试验中成功的次数。



伯努利分布

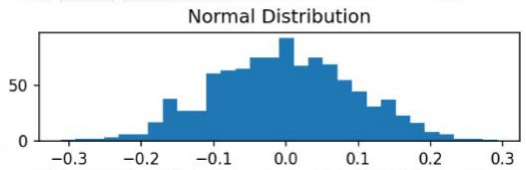


二项分布

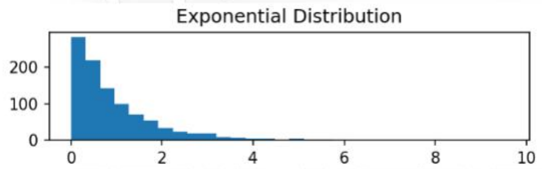
常见的概率分布

正态分布：描述自然和社会现象中出现的许多随机变量的分布情况

指数分布：描述独立随机事件发生的时间间隔



正态分布



指数分布

案例实践

◆ 模拟常见的概率分布

大数定律与中心极限定理

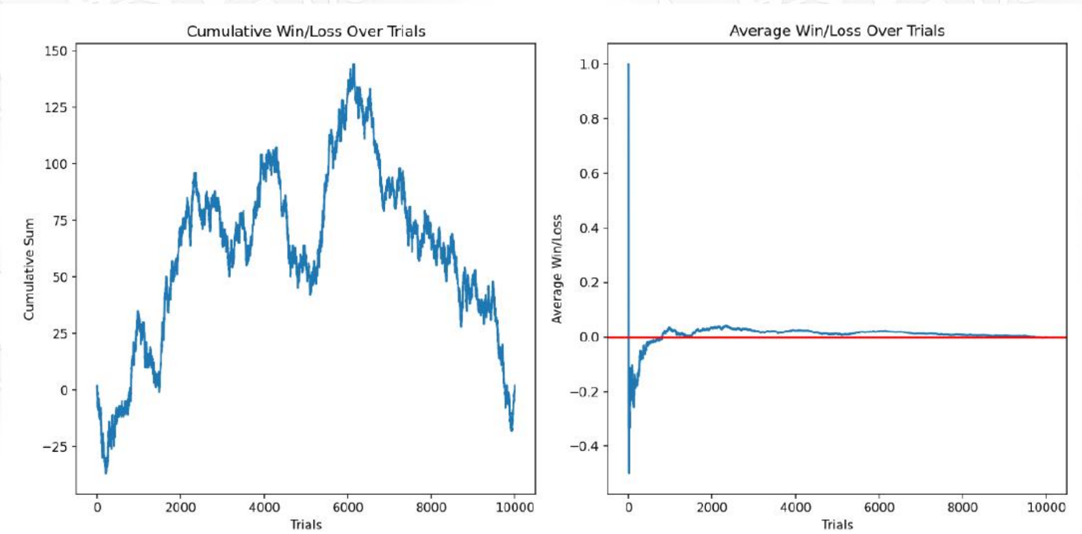
◆ 大数定率

随着尝试次数的增加，实验的平均结果将趋近于期望值。

例如抛硬币，随着你扔的次数越来越多，你会注意到，正面和反面出现的次数开始趋近于各占一半。如果你扔了成千上万次，你几乎可以肯定，正反面各自出现的次数会非常接近于总扔硬币次数的50%

案例实践

◆ 模拟大数定律



大数定律与中心极限定理

◆ 中心极限定理

当独立随机变量的数量足够大时，它们的总和将趋近于正态分布，无论原始随机变量的分布如何。

即使单个骰子的结果是完全随机的，但是当许多随机结果加在一起时，总和的分布将趋向于形成一个特定的、可预测的形状，也就是正态分布。

案例实践

◆ 模拟中心极限定理