**6.1 介绍** 2020年3月23日09点25分

**例题6.1.1** 硬币正面的比例 计算一定范围内的概率

**例题6.1.2** 平均等待时间 计算一定范围内的概率

**6.2 大数定理** 2020年3月23日09点48分

**定理6.2.1 马尔可夫不等式** 假设是一个随机变量且满足.则对于每一个实数,

**定理6.2.2 切比雪夫不等式** 设是一个随机变量且存在. 则对于每一个,

**定理6.2.3 样本均值的期望和方差** 设是来源于期望为,方差为的总分布的一组样本.设是样本均值.则,.

对施加切比雪夫不等式，根据公式(6.2.3)得到

**例题6.2.1** 确定观察结果的所需数量 对公式的应用

**例题6.2.2** 模拟 **定理6.2.2**和**习题14**的应用

**例题6.2.3** 抛硬币 比较切比雪夫不等式与二项式对范围概率计算结果的精度差异,**需要注意二项分布的计算方式是通过查表得到的**.

**定义6.2.1 概率收敛** 一组随机变量的在概率上收敛于仅当对于每一个值,满足

这种属性被定义为,有时也被简单的称为在概率上收敛于.

**定理6.2.4 大数定理** 假设构成了某个期望为,方差有限的分布的样本空间.设是样本均值.则

**定理6.2.5 随机变量的连续函数** 如果,假设函数在处连续,则.(**需要手动证明**)

**定理6.2.6 直方图** 设是一组i.i.d.随机变量.设是两个常数.当,,否则.则与位于区间的个数成正比,且.

**例题6.2.4** 服务频率 对直方图的说明 指数分布

**例题6.2.5** 服务频率 对直方图进一步说明 **习题17有更进一步的说明**

**例题6.2.6** 二项随机变量 计算二项分布的范围概率,引出问题

**定理6.2.7 切尔诺夫界** 设是一个随机变量且它的距量母函数为.则对于每一个值,

当是n个i.i.d的和时,定理6.2.7最有用.每个随机变量的有限m.g.f.

**例题6.2.7** 几何随机样本的均值 **定理6.2.7**的应用,是**一道非常复杂的综合题.**

**6.3 中心极限定理** 2020年3月30日09点50分

**例题6.3.1** 大量样本 引出现象——二项分布为何与正态分布相似

**定理6.3.1 中心极限定理(Lindeberg和Levy)** 如果随机变量构成了任意给定分布大小为随机样本,该分布的均值为,方差为,则对于每一个

其中表示标准正态分布的c.d.f.

公式的解释如下:如果从均值和方差的任何分布中获取大的随机样本,不论该分布是离散的还是连续的,则随机变量将近似为标准正态分布.因此,的分布将近似为具有均值和方差的正态分布,或者等效地,总和的分布将近似为具有均值和方差的正态分布.

**例题6.3.2** 抛硬币 **定理6.3.1**对伯努利分布的应用

**例题6.3.3** 均匀分布 **定理6.3.1**对均匀分布的应用

**例题6.3.4** 泊松分布 **定理6.3.1**对泊松分布的应用

**定义6.3.1 分布/渐近分布的收敛性** 设是一组随机变量,其中,设为的c.d.f..同样,设是一个c.d.f..则称序列分布收敛于如果

在所有的x上是连续的.有时,会简称分布收敛于,被称为的渐进分布.

因此,根据定理6.3.1,如等式(6.3.1)所示,随机变量的分布收敛到标准正态分布,或等效地,的渐近分布是标准正态分布.

中心极限定理为物理实验中研究的许多随机变量的分布近似正态这一事实提供了合理的解释.例如,一个人的身高受许多随机因素的影响.如果通过将这些各个因素的值相加来确定每个人的身高,那么许多人的身高分布将近似为正态.通常,中心极限定理表明,即使每个随机变量的总和的分布与正态不同,许多随机变量的总和的分布也可以近似正态.

**例题6.3.5** 确定模拟样本的大小 **例题6.2.2**的扩展 **定理6.3.1**的应用

**例题6.3.6** 服务频率 引出问题

假设从具有有限均值和有限方差的分布中形成随机样本.中心极限定理说具有近似标准正态分布.现在假设我们对的某些函数的分布感兴趣.我们将假定是一个微分函数,其导数在处为非零.我们将通过统计中称为delta方法的方法来近似估计的分布.

**定理6.3.2 Delta方法** 设是一组随机变量,设是一个连续c.d.f..设是一个实数,设是一组递增到无穷大的正数.假设分布收敛至.设是具有连续导数的函数且.则分布收敛至.(**证明过程需要理解,书上证明过程中公式(6.3.5)书写错误**)

**推论6.3.1** **随机样本平均值的Delta方法** 设是一组独立同分布随机变量,其分布均值为,方差为.设是导数连续的函数且.则

渐进分布是标准正态分布.

在**推论6.3.1**结果的一种常用方法是说,的分布近似为均值为,方差为正态分布.

**例题6.3.7** 服务频率 应用**推论6.3.1**求解**例题6.3.6**

**方差稳定变换** 如果像**例题6.3.4**那样观察泊松随机变量的随机样本,我们将假定是未知的.在这种情况下,由于的近似方差取决于,因此我们无法计算公式(6.3.2)中的概率.因此,有时希望通过函数对进行变换,以使的近似分布具有已知值的方差.这种函数称为*方差稳定变换*.我们通常可以通过反向执行delta方法来找到方差稳定化变换.通常,我们注意到,的近似分布具有方差.为了使该方差恒定,我们需要将设为乘以某个常数.如果是函数,那么我们通过使

来实现目标,其中是使得该积分有限的任意常数.

**例题6.3.8** 泊松分布 应用**方差稳定变换**求解**例题6.3.4**

我们现在将中心极限定理应用到一组随机变量,假设这些随机变量是独立但不一定是相同分布.假设，.设

则.接下来说明的定理给出了使该随机变量的分布近似为标准正态分布的充分条件.

**定理6.3.3** 假设随机变量是独立的并且.同样假设

最后,我们设随机变量如公式(6.3.8)中定义.则对于每一个固定值,

该定理的解释如下:如果满足公式,则对于的每个大值,的分布将近似为均值,并且方差 正态分布.应该注意的是,当随机变量分布相同且变量的第三阶矩存在时,将自动满足公式,然后公式简化为公式.

**定理6.3.4** 假设随机变量是独立的并且是参数为的伯努利分布.假设无穷级数是发散的，设(**该定理的证明过程存在两处书写错误**)

则对于每一个固定值,

**定理6.3.4**意味着如果无限级数是发散的,那么大量独立伯努利随机变量之和的分布将近似于均值为,方差为的正态分布.然而,应牢记,典型的实际问题将仅涉及有限数量的随机变量,而不是无限随机变量序列.在这样的问题中,考虑无穷级数是否发散是没有意义的,因为仅有限数量的值将在问题中指定.因此,从某种意义上说,总和的分布总是可以通过正态分布来近似.关键是该正态分布是否提供了对实际分布的良好近似.答案取决于值.

由于当时将越来越接近正态分布,因此当的值较大时,正态分布可提供良好的近似值.此外,由于当时项的值最大,因此当比较大且接近1/2时能得到最好的近似.

**例题6.3.9** 成绩测验 **定理6.3.4**的应用

**定理6.3.5** 设是一组随机变量序列.当,令是的c.d.f.,令是的m.g.f..同样,令表示其它随机变量,其c.d.f.为,m.g.f.为.假设m.g.f.和存在.如果对于所有值,在点周围的某个区间上,则序列分布收敛到.

换句话说,如果收敛到,则必须收敛到.

**6.4 连续性校正** 2020年4月7日09点43分

**例题6.4.1** 大量样本 引出现象 二项分布和正态分布更进一步的观察

**例题6.4.2** 大量样本 根据连续性校正来优化离散随机变量取值区间

**例题6.4.3** 成绩测验 根据连续性校正来优化**例题6.3.9**

**例题6.4.4** 抛硬币 连续性校正的应用

**总结** 令为仅包含整数值的随机变量.假设近似具有均值和方差的正态分布.令和为整数,并假设我们希望近似.对连续性的正态分布近似的校正是使用而不是作为近似值.