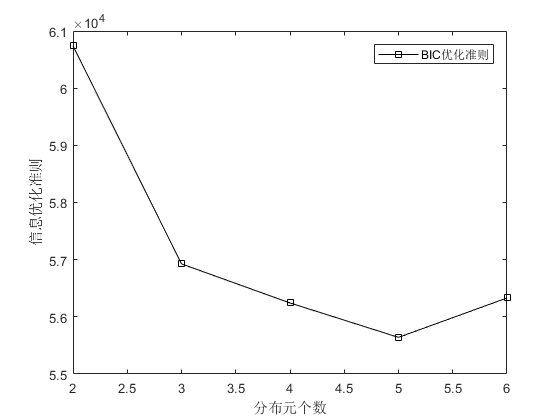
## 仿真及实验结果分析

### 点模式特征分布的分布元信息对性能的影响

某类点模式模型的基数分布服从泊松分布，特征分布服从混合高斯分布，参数为;;,,;,,。通过该模型生成140组点模式数据，将其中的100组数据作为训练集，剩余的40组作为测试集。

图表, 散点图

描述已自动生成

图（一） 图（二）

图（一）为该模型的真实分布，我们通过结合BIC信息准则的参数学习算法对100组训练数据进行参数学习.对分布元个数从2个递增至6个分别进行学习，并且以BIC信息准则作为判断依据，从图（二）的BIC信息准则曲线得，分布元为5时得到最小的BIC值。我们选取所有学习得的模型中分布元为5作为训练集所对应的模型。

我们用该模型的40组测试集数据对BIC信息准则的结果进行验证，分别通过Gibbs采样得的三分布元的参数、四分布元的参数、五分布元的参数和六分布元参数构建似然4个函数。将每个训练集分别代入4个似然函数，若基于BIC信息准则和Gibbs采样的有限混合模型学习算法所得结果合理，那么测试集数据在5分布元处取到最大值的比列应为最高。通过图（三）的实验结果，40组数据中有38组在5分布元处取到最值，说明5分布元的混合模型最符合样本数据，且基于BIC信息准则和Gibbs采样的有限混合模型学习算法的学习准确率高。

图表, 条形图

描述已自动生成 图表

描述已自动生成

图（三） 图（四）

图表, 散点图

描述已自动生成 图表, 条形图

描述已自动生成

图（五） 图（六）

### 点模式数据集分类实验

#### 实验一

第一类基数分布服从泊松分布，参数为，特征分布为高斯混合模型，参数为；，,;, , 。

第二类基数分布服从泊松分布，参数为，特征分布为高斯混合模型，参数为；，,;, , 。

第三类基数分布服从泊松分布，参数为，特征分布为高斯混合模型，参数为；，,;, , 。

每一类生成140个样本，其中100个用作模型学习，40个用作测试学习效果；从图（五）可以看出三类样本在特征空间中重叠严重，单纯使用点模式的特征信息进行分类，无法得到效果较为不错的分类结果；再从图（四）分析，三类点模式在基数分布上存在部分重叠现象，结合基数信息理论上可以提高分类效果。通过实验效果得，我们从图（六）可以看出，结合基数信息的Poisson模型相对于单纯使用特征分布信息的NB模型，在分类效果上存在显著的提升。

#### 实验二

第一类基数分布服从泊松分布，参数为，特征分布为高斯混合模型，参数为；，,;, , 。

第二类基数分布服从泊松分布，参数为，特征分布为高斯混合模型，参数为；，,;, , 。

第三类基数分布服从泊松分布，参数为，特征分布为高斯混合模型，参数为；，,;, , 。

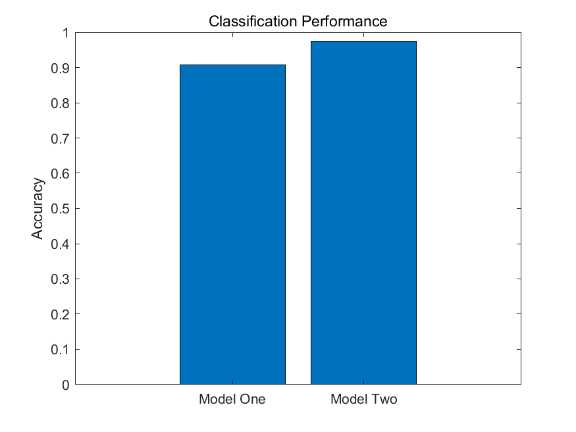
图表, 直方图

描述已自动生成 图表, 散点图

描述已自动生成

图（七） 图（八）

图表, 条形图

描述已自动生成 

图（九） 图（十）

每一类生成140个样本，其中100个用作模型学习，40个用作测试学习效果；从图（七）中可以看出类一与类二在基数信息上重叠，类三可以借助基数分布信息提高自身分类效率；

从图（八）中可以看出，类二和类三在特征分布信息上重叠，类一可以通过特征信息提高自身的分类效果，从图（九）的分类结果中显示结合基数信息的Poisson模型相对于单纯使用特征分布信息的NB模型在分类任务中效果更好。

图（十）中Model Two三类样本的特征分布均选取BIC信息最小混合模型，而Model One除了类二的特征分布，其余两类我们均选取BIC信息最小混合模型，进行对比实验得，Model Two的分类效果优于Model One，由此可以说明分布元信息的确影响分类效果。

## 3 泊松点过程模型的最大似然估计

学习点过程模型的计算高效算法很重要，因为机器学习与空间统计应用相比通常涉及大数据集，由于学习一般的点过程模型在计算上是非常困难的，而泊松点过程模型通过忽略点之间的相关性，在模型的通用性和模型参数的学习难度之间取得了一个良好的平衡。

泊松点过程模型由它的基数分布和特征分布共同决定，模型的参数表示形式如下



其中和分别