1. （必填）自己提出的问题的理解（罗列全部）：
2. 提出的问题1：参数对于狄利克雷分布意味着什么？

讨论后的理解：是狄利克雷分布的分布参数，在贝叶斯模型中，狄利克雷后验分布的参数等于狄利克雷先验分布参数加上多项分布的观测计数。

1. 提出的问题2：如何理解LDA的文本生成过程？

讨论后的理解：首先基于单词分布的狄利克雷先验分布生成多个单词分布，决定多个话题内容；之后基于话题分布的狄利克雷先验分布生成多个话题分布，决定多个文本内容；然后再基于每一个话题分布生成话题序列，针对每一个话题，基于话题的单词分布生成单词，整体构成一个单词序列，即生成文本，重复这个过程生成所有文本。

1. （必填）别人提出的问题的理解（选择几个问题罗列，并给出理解）：
2. 问题3：LDA与潜在语义分析和概率潜在语义分析之间有什么区别和联系？

自己的理解：两者相同点为都假设话题是单词的多项分布，文本是话题的多项分布。不同点在于先验分布的假设分布不同：LDA为狄利克雷分布，PLSA可以理解成均匀分布；学习过程不同：LDA基于贝叶斯学习，而PLSA基于极大似然估计。

1. 问题4：吉布斯抽样和变分EM推理分别适合在什么情况使用？

自己的理解：吉布斯抽样实现较简单但迭代次数可能会很多，变分EM算法具有推理和学习效率高的优点。

5、问题5：LDA中，狄利克雷分布的重要性是什么？能否使用其他分布来充当先验分布？

自己的理解：狄利克雷分布是多项分布的共轭分布，这一点在贝叶斯学习中比较重要，所以说应该是不可以讲狄利克雷分布改为其他分布。当然了，感觉先验分布改为贝塔分布，后验分布改为二项分布也是可以的，是简化情况（二分类）。

1. （必填）读书计划

1、本周完成的内容章节：第二十章

2、下周计划：第二十一章