一、我提出的问题：

问题一：是否两个分布的分布形式一致，就称为共轭？

讨论后的理解：不是，二手一个分布加上另一个分布，结果跟该分布形式相同，才叫共轭。

直观上讲，一个分布加上该分布的样本，当然还是该分布。

问题二：LDA中，多项式分布和狄利克雷分布的 形式一致，所以称为共轭，但是 这两个分布的区别是什么呢？

讨论后的理解：多项分布中，底数是分布参数，指数是随机变量；  
Dirichlet分布中，底数是随机变量，指数是分布参数。

在贝叶斯估计中，多项分布的参数也被看成随机变量，它们的分布就由Dirichlet分布来描述了。

二、别人提出的问题：

问题一：Dirichlet分布和Beta分布的关系是水母？

Dirichlet分布是Beta分布的多元推广。Beta分布是二项式分布的共轭分布，Dirichlet分布是多项式分布的共轭分布。通常情况下，我们说的分布都是关于某个参数的函数，把对应的参数换成一个函数（函数也可以理解成某分布的概率密度）就变成了关于函数的函数。

问题二：先验分布、后验分布它们之间的关系是什么？

先验概率，就是常识、经验所透露出的“因”的概率，即瓜熟的概率。应该很清楚。

后验概率，就是在知道“果”之后，去推测“因”的概率，也就是说，如果已经知道瓜蒂脱落，那么瓜熟的概率是多少。后验和先验的关系可以通过贝叶斯公式来求。

三、读书计划：

本周二十章

下周二十一章

四、读书小结：

20.1 潜在狄利克雷分配(LDA)，作为基于贝叶斯学习的话题模型，是潜在语义分析、概率潜在语义分析的扩展，于2002年由Blei等提出。LDA在文本数据挖掘、图像处理、生物信息处理等领域被广泛使用。

20.2 LDA模型是文本集合的生成概率模型。假设每个文本由话题的一个多项式分布表示，每个话题由单词的一个多项式分布表示，特别假设文本的话题分布的先验分布是狄利克雷分布，话题的单词分布的先验分布也是狄利克雷分布。先验分布的导入使LDA能够更好地应对话题模型学习的过拟合现象。

20.3 LDA的文本集合的生成过程如下：首先随机生成一个文本话题分布，之后再该文本的每个位置，依据该文本的话题分布随机生成一个话题，然后在该位置依据该话题的单词分布随机生成一个单词，直至文本的最后一个位置，生成整个文本。重复以上的过程生成所有文本。

20.4 LDA模型是含隐变量的概率图模型。模型中，每个话题的单词分布，每个文本的话题分布，文本的每个位置的话题是隐变量；文本的每个文职的单词是观测变量。LDA模型的学习与推理无法直接求解，通常使用吉布斯抽样和变分EM算法。前者是蒙特卡洛法，后者是近似计算。