1.指数族分布的推导：

考虑离散概率模型。我们希望概率模型不仅要无偏，即对于数据及其充分统计量满足：



而且要有最大的信息熵，我们认为更大的信息熵涵盖更多的可能性：



用拉格朗日乘子法改写：



对取每一个特定值时的求导，有：



移项得到：



式子两边关于求和并利用约束，可得



可见是关于的函数。这样我们就得到了指数族分布的一般形式：



我们还未求出。按理来说我们应该利用拉格朗日条件求出来，但这样直接求很困难。事实上我们也不需要直接求出的显式解，实际应用中只要能验证满足约束条件就是符合要求的。我们已经利用了约束得到了，只需找一个利用并且含为变量的式子即可。我们将目光转向，对其求导：



这样，只要满足，就满足了这个约束。

进一步有：



2.指数族分布和共轭先验：

贝叶斯公式两边取对数：



若数据集*D*有*M*个独立数据点，假设似然为指数族分布：



不妨认为先验也为指数族分布，维度为*K*，且不同维度间独立：



这样，后验分布为：



可见，后验仍然是指数族分布，这说明指数族分布的共轭先验为指数族分布。