Homework2

1

解

假定对于二项分布参数p我们采用均匀先验分布，并且在10次试验中观察到了4次正性结果，

a) 给出先验贝塔分布的参数值(2分)

先验贝塔分布的参数值为 $\alpha = 1$ 和 $\beta = 1$.

b) 给出后验贝塔分布的参数值(2分)

后验贝塔分布的参数值为 $\alpha' = 5$ 和 $\beta' = 7$.

c) 给出在先验分布下二项分布参数p的期望值（2分）

先验分布下二项分布参数 p 的期望值为 $\frac{\alpha}{\alpha+\beta} = \frac{1}{2} = 0.5$.

d) 给出样本中正性结果的比例(2分)

样本中正性结果的比例为 $\frac{4}{10} = \frac{2}{5} = 0.4$.

e) 给出二项分布参数p的极大似然估计值(2分)

二项分布参数 p 的极大似然估计值为 $\frac{4}{10} = \frac{2}{5} = 0.4$.

f) 给出在后验分布下二项参数p的期望值，并以先验分布下该参数的期望值和该参数的极大似然估计值的加权平均形式表达(4分)

f) 后验分布下二项分布参数 p 的期望值为 $\frac{\alpha'}{\alpha'+\beta'} = \frac{5}{5+7} = \frac{5}{12}$，可以表示为先验分布下该参数的期望值和该参数的极大似然估计值的加权平均形式，即

$\frac{1}{2} \cdot \frac{2}{5} + \frac{2}{5} \cdot \frac{3}{5} = 0.44$.

2

解

假定对于二项分布参数p，我们有三个假设，H0: p = 0.5, H1: p = 0.4, H2: p ~ unif(0,1)。另外，我们进行了50次伯努利试验，得到了20次正性结果，

a) 求对于每个假设的P(D|H)(6分)
对于每个假设H0: p = 0.5，H1: p = 0.4和H2: p ~ unif(0,1)，我们可以计算给定假设下观察到数据的概率P(D|H)。

对于H0: p = 0.5，我们可以使用二项分布来计算概率。在50次试验中，观察到20次正性结果的概率为：

$$P(D|H_0) = \binom{50}{20} \times (0.5^{20}) \times (0.5^{30}) = 0.0419$$

对于H1: p = 0.4，同样使用二项分布来计算概率。观察到20次正性结果的概率为：

$$P(D|H_1) = \binom{50}{20} \times (0.4^{20}) \times (0.6^{30}) = 0.1146$$

对于H2: p ~ unif(0,1)，我们采用均匀先验分布，因此概率为：

$$P(D|H_2) = \int_0^1 P(D|p) \times P(p)dp$$

观察到20次正性结果的概率为：

$$P(D|H_2) = \int_0^1 \binom{50}{20} \times (p^{20}) \times ((1-p)^{30})dp = 0.0196$$

b) 求各对假设间的贝叶斯因子(3分)
贝叶斯因子可以通过计算每对假设的P(D|H)之间的比值得到。

贝叶斯因子

$$BF_{01} = P(D|H_0)/P(D|H_1) = 0.365$$

这意味着数据更倾向于支持 H_1 相对于 H_0。
贝叶斯因子

$$BF_{02} = P(D|H_0)/P(D|H_2) = 2.135$$

这意味着数据更倾向于支持 H_0 相对于 H_1。
贝叶斯因子

$$BF_{12} = P(D|H1)/P(D|H2) = 5.842$$

这意味着数据更倾向于支持 H_1 相对于 H_2。
c) 根据以上结果，我们能够做出怎样的统计推断(2分)
基于以上结果，我们可以推断 $H_1 : p = 0.4$ 相对于其他两个假设获得了更多的支持。尽管 H_0 与 H_2 的贝叶斯因子也表明 H_0 相对于 H_2 有更多的支持，但在所有比较中，H_1 显示出了最强的相对支持度。因此，我们可以说数据更倾向于支持 $p = 0.4$ 的假设。

d) 为了能够做出接受上述三个假设中某一假设的统计推断，在假定样本正性结果比例不变的情况下，还需进行多少次额外的伯努利试验(5分)

为了估计为了能够做出接受上述三个假设中某一假设的统计推断，在假定样本正性结果比例不变的情况下，设置贝叶斯因子的阈值为10
我们可以通过迭代增加试验次数，并计算贝叶斯因子，并且我们只关注H1相对于H0和H2的贝叶斯因子，因为这是当前数据倾向支持的假设。
为了使得H1相对于H0和H2的贝叶斯因子至少达到10，我们需要进行大约100次额外的伯努利试验（基于原始的50次试验）。在这种情况下：

$H_1 vs. H_0$ 的贝叶斯因子约为20.5
$H_1 vs. H_2$ 的贝叶斯因子约为10.0

这意味着，在进行了150次总试验 (原始的50次加上额外的100次)之后，我们有强证据支持假设H1 ($p = 0.4$) 相对于假设HO ($p = 0.5$) 和假设H2 (p遵循$unif(0,1)$)
。