**极大似然估计 (maximum likelihood estimation)**

考虑一个投掷硬币的实验：现在我们有两枚硬币A和B，这两枚硬币和普通的硬币不一样，他们投掷出正面的概率和投掷出反面的概率不一定相同。我们将A和B投掷出正面的概率分别记为和。我们现在独立地做5次试验：随机的从这两枚硬币中抽取1枚，投掷10 次，统计出现正面的次数。那么我们就得到了如表格1的实验结果。

|  |  |  |
| --- | --- | --- |
| 试验代号 | 投掷的硬币 | 出现正面的次数 |
| 1 | B | 5 |
| 2 | A | 9 |
| 3 | A | 8 |
| 4 | B | 4 |
| 5 | A | 7 |

Table 1: 硬币投掷实验的结果

在这个实验中，我们记录两组随机变量 X = (X1, X2, X3, X4, X5), Z = (Z1, Z2, Z3, Z4, Z5)，其中 ∈ {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} 代表试验i中出现正面的次数， ∈ {A, B} 代表这次试验投掷的是硬币 A 还是硬币 B。 我们的目标是通过这个实验来估计 θ= (, ) 的数值。这个实验中的参数估计就是有完整数据的参数估计，这是因为我们不仅仅知道每次试验中投掷出正面的次数，我们还知道每次试验中投掷的是硬币A还是B。

一个很简单也很直接的估计θ的方法如下(1)所示。

实际上这样的估计就是统计上的极大似然估计 (maximum likelihood estimation) 的结果。

用 P(X, Z|θ) 来表示 X,Z 的联合概率分布（其中带有参数θ），那么对于上面的实验，我们可以计算出他们出现我们观察到的结果即= (5, 9, 8, 4, 7), = (B, A, A, B, A)的概率。函数 P(X =,Z=|θ) 就叫做θ的似然函数。我们将它对θ求偏导并令偏导数为 0，就可以 得到上面(1)的结果。

P(X =,Z=|θ) = P× (1-)

×

×

×

×