Exponential Family Distribution 03 Property

Chen Gong

24 October 2019

本小节主要介绍 Exponential Distribution 中对数配分函数和充分统计量，还有极大似然估计和充分统计量的关系。

指数族分布的基本形式可以表示为：

$$p(x|\eta) = h(x)exp\left\{\eta^T\varphi(x) - A(\eta)\right\} \tag{1}$$

$$p(x|\eta) = \frac{1}{exp\{A(\eta)\}}h(x)exp\left\{\eta^T\varphi(x)\right\} \tag{2}$$

1 对数配分函数和充分统计量

现在有一个问题，那就是我们如何求得对数配分函数 $exp\{A(\eta)\}$，或者说我们可不可以简单的求得对数配分函数。于是，就可以很自然的想到，前面所提到的充分统计量 $\varphi(x)$ 的概念。对数配分函数的目的是为了归一化，那么我们很自然的求出对数配分函数的解析表达式：

$$\begin{aligned}
\int p(x|\eta)dx &= \int \frac{1}{exp\{A(\eta)\}}h(x)exp\left\{\eta^T\varphi(x)\right\}dx \\
\int p(x|\eta)dx &= \frac{\int h(x)exp\left\{\eta^T\varphi(x)\right\}dx}{exp\{A(\eta)\}} = 1 \\
exp\{A(\eta)\} &= \int h(x)exp\left\{\eta^T\varphi(x)\right\}dx
\end{aligned} \tag{3}$$

下一步则是在 $exp\{A(\eta)\}$ 中对 η 进行求导。

$$\begin{aligned}
\frac{\partial exp\{A(\eta)\}}{\partial\eta} &= \nabla_\eta A(\eta)exp\{A(\eta)\} \\
&= \frac{\partial}{\partial\eta}\int h(x)exp\left\{\eta^T\varphi(x)\right\}dx \\
&= \int \frac{\partial}{\partial\eta}h(x)exp\left\{\eta^T\varphi(x)\right\}dx \\
&= \int h(x)exp\left\{\eta^T\varphi(x)\right\}\varphi(x)dx
\end{aligned} \tag{4}$$

将等式的左边的 $exp\{A(\eta)\}$ 移到等式的右边可得，

$$\nabla_\eta A(\eta) = \int h(x)exp\left\{\eta^T\varphi(x) - A(\eta)\right\}\varphi(x)dx \tag{5}$$

$$\nabla_\eta A(\eta) = \int p(x|\eta)\varphi(x)dx \tag{6}$$

$$\nabla_\eta A(\eta) = \mathbb{E}_{x\sim p(x|\eta)}[\varphi(x)] \tag{7}$$

其实通过同样的方法可以证明出，

$$\nabla\eta^2 A(\eta) = Var_{x\sim p(x|\eta)}[\varphi(x)] \tag{8}$$

又因为，协方差矩阵总是正定的矩阵，于是有 $\nabla_\eta^2 A(\eta) \succeq 0$。所以，由此得出 $A(\eta)$ 是一个凸函数。并且，由 $\mathbb{E}_{x\sim p(x|\eta)}[\varphi(x)]$ 和 $Var_{x\sim p(x|\eta)}[\varphi(x)]$ 就可以成功的求解得到 $A(\eta)$ 函数。那么我们做进一步思考，知道了 $\mathbb{E}[x]$ 和 $\mathbb{E}[x^2]$，我们就可以得到所有想要的信息。那么：

$$\mathbb{E}[\varphi(x)] = \begin{pmatrix} \mathbb{E}[x] \\ \mathbb{E}[x^2] \end{pmatrix} \tag{9}$$

2 极大似然估计和充分统计量

假设有一组观察到的数据集：$D = \{x_1, x_2, x_3, \cdots, x_N\}$，那么我们的求解目标为：

$$\begin{aligned}
\eta_{MLE} =& argmax \log \prod_{i=1}^N p(x_i|\eta) \\
=& argmax \sum_{i=1}^N \log p(x_i|\eta) \\
=& argmax \sum_{i=1}^N \log h(x_i) exp\left\{\eta^T\varphi(x_i) - A(\eta)\right\} \\
=& argmax \sum_{i=1}^N \log h(x_i) + \sum_{i}^N \left(\eta^T\varphi(x_i) - A(\eta)\right)
\end{aligned} \tag{10}$$

$$\frac{\partial}{\partial\eta}\left\{\sum_{i=1}^N \log h(x_i) + \sum_{i=1}^N \left(\eta^T\varphi(x_i) - A(\eta)\right)\right\} = 0 \tag{11}$$

$$\sum_{i=1}^N \varphi(x_i) = N \cdot \nabla_\eta A(\eta) \tag{12}$$

$$\nabla_\eta A(\eta) = \frac{1}{N}\sum_{i=1}^N \varphi(x_i) \tag{13}$$

或者说，我们可以认为是：$\nabla_\eta A(\eta_{MLE}) = \frac{1}{N}\sum_{i=1}^N \varphi(x_i)$。并且，$\nabla_\eta A(\eta_{MLE})$ 是一个关于 η_{MLE} 的函数。那么反解，我们就可以得到 η_{MLE}。所以我们要求 η_{MLE}，我们只需要得到 $\frac{1}{N}\sum_{i=1}^N \varphi(x_i)$ 即可。所以，$\varphi(x)$ 为一个充分统计量。

3 总结

在本小节中，我们使用了极大似然估计和对数配分函数来推导了，充分统计量，这将帮助我们理解 Exponential Distribution 的性质。