

Exponential Family Distribution 03

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:

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

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

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:

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

下一步则是在 $\exp\{A(\eta)\}$ 中对 η 进行求导。

$$\begin{aligned} \frac{\partial \exp\{A(\eta)\}}{\partial \eta} &= \exp\{A(\eta)\} \cdot A'(\eta) \\ &= \frac{\partial}{\partial \eta} \int h(x) \exp\{\eta^T \varphi(x)\} dx \\ &= \int \frac{\partial}{\partial \eta} h(x) \exp\{\eta^T \varphi(x)\} dx \\ &= \int h(x) \exp\{\eta^T \varphi(x)\} \varphi(x) dx \end{aligned} \quad (4)$$

将等式的左边的 $\exp\{A(\eta)\}$ 移到等式的右边可得,

$$A'(\eta) = \int h(x) \exp\{\eta^T \varphi(x) - A(\eta)\} \varphi(x) dx \quad (5)$$

$$A'(\eta) = \int p(x|\eta) \varphi(x) dx \quad (6)$$

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

其实通过同样的方法可以证明出，

$$A''(\eta) = \text{Var}_{x \sim p(x|\eta)}[\varphi(x)] \quad (8)$$

又因为，方差总是恒大于等于零的，于是有 $A''(\eta) \geq 0$ 。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 $\text{Var}_{x \sim p(x|\eta)}[\varphi(x)]$ 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，

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

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： $D = \{x_1, x_2, x_3, \dots, x_N\}$ ，那么我们的求解目标为：

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

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

$$\sum_{i=1}^N \varphi(x_i) = N A'(\eta) \quad (12)$$

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

或者说，我们可以认为是： $A'(\eta_{MLE}) = \frac{1}{N} \sum_{i=1}^N \varphi(x_i)$ 。并且， $A'(\eta_{MLE})$ 是一个关于 η_{MLE} 的函数。那么反解，我们就可以得到 $\eta_{MLE} = (A'^{-1}(\eta))'$ 。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。