Various Notes

Bertolotti Francesco

February 26, 2021

Contents

Generalized Linear Models

In this chapter, we will discuss the the so called Generalized Linear Models (*GLM*)

Exponential Family

The Exponential Family ($\mathcal{E}.\mathcal{F}.$) is a parametric family of probability distribution. Many of the common distribution are part of the $\mathcal{E}.\mathcal{F}.$. Here are some examples:

- Bernulli Distribution.
- Normal Distribution.
- Categorical Distribution.
- Multinomial Distribution.
- Poisson Distribution.

$\mathcal{E}.\mathcal{F}.$ represents a convienient framework to represent many distributions at the same time. A distribution belongs to the $\mathcal{E}.\mathcal{F}.$ if its Probability Mass Function (*pmf*) $p(x|\theta)$ can be written in the following form:

$$p(x|\theta) = h(x)\exp(\eta(\theta)^T\phi(x) + A(\eta(\theta))) \tag{1}$$

Each of these terms has its own meaning. Firstly, $phi(x)$ is known as the *sufficient statistics* (why? we will discuss this later). Secondly, $\eta(\theta)$ is called the *canonical parameters* while θ is called *natural parameters* . Thirdly, $A(\eta(\theta))$ is known as *cumulant function* and it has the following form:

$$A(\eta(\theta)) = \log \int h(x) exp(\eta(\theta)^T\phi(x))dx \tag{2}$$

$A(\eta(\theta))$ is nothing more than the logarithm of the normalization factor of the distribution. That is, the term that makes sure that the integral sums up to 1.

Where θ represents the parameters. Let's see some examples starting with the Bernulli Distribution ($Ber(x|\theta)$ with $x \in 1, -1$). Usually, the Bernulli Distribution is written as follows:

$$Ber(x|\theta) = \theta^x(1-\theta)^{1-x} \tag{3}$$

In Figure 1, You have a plot for $Ber(x|\theta = 0.3)$. Now, we can proceeds by writing the the distribution in 3 in the $\mathcal{E}.\mathcal{F}.$ form (1).

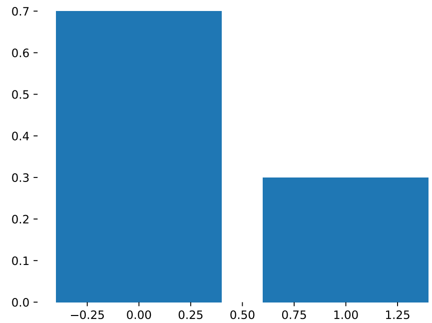

Figure 1: Bernulli distribution examples with $\theta = .3$

$$
\begin{aligned}
Ber(x|\theta) = \theta^x(1-\theta)^{1-x} = && (4)\\
\exp(\log(\theta^x(1-\theta)^{1-x})) = && (5)\\
\exp(x\log(\theta) + (1-x)log(1-\theta)) = && (6)\\
\exp(\eta(\theta)\phi(x) + A(\eta(\theta))) && (7)
\end{aligned}
$$

Where:

$$
\eta(\theta) = \begin{bmatrix} \log(\theta) \\ \log(1-\theta) \end{bmatrix}, \phi(x) = \begin{bmatrix} x \\ 1-x \end{bmatrix}, A(\eta(\theta)) = 0 \tag{8}
$$