Departamento de Ingeniería Informática

Practical Lesson

Model Fitting

Data Science I

Cristóbal Donoso

12 Septiembre 2018

Exercise 3.6 MLE for the Poisson distribution

The Poisson pmf is defined as $\mathrm{Poi}(x|\lambda) = e^{-\lambda}\frac{\lambda^x}{x!}$, for $x \in \{0, 1, 2, \ldots\}$ where $\lambda > 0$ is the rate parameter. Derive the MLE.

Exercise 3.7 Bayesian analysis of the Poisson distribution

In Exercise 3.6, we defined the Poisson distribution with rate λ and derived its MLE. Here we perform a conjugate Bayesian analysis.

a. Derive the posterior $p(\lambda|D)$ assuming a conjugate prior $p(\lambda) = \mathrm{Ga}(\lambda|a, b) \propto \lambda^{a-1}e^{-\lambda b}$. Hint: the posterior is also a Gamma distribution.
b. What does the posterior mean tend to as $a \to 0$ and $b \to 0$? (Recall that the mean of a $\mathrm{Ga}(a, b)$ distribution is a/b.)

Exercise 3.14 Posterior predictive for Dirichlet-multinomial

(Source: Koller.).

a. Suppose we compute the empirical distribution over letters of the Roman alphabet plus the space character (a distribution over 27 values) from 2000 samples. Suppose we see the letter "e" 260 times. What is $p(x_{2001} = e|\mathcal{D})$, if we assume $\boldsymbol{\theta} \sim \text{Dir}(\alpha_1, \ldots, \alpha_{27})$, where $\alpha_k = 10$ for all k?

b. Suppose, in the 2000 samples, we saw "e" 260 times, "a" 100 times, and "p" 87 times. What is $p(x_{2001} = p, x_{2002} = a|\mathcal{D})$, if we assume $\boldsymbol{\theta} \sim \text{Dir}(\alpha_1, \ldots, \alpha_{27})$, where $\alpha_k = 10$ for all k? Show your work.