

ST3233: Tutorial 7

October 2016

1 Identification

Write down the expression for the following time series model:

1. SARIMA(1,1,1)(2,0,2)[4]
2. SARIMA(1,0,1)(1,0,2)[12]

2 KL divergence is non-negative

A convex function $F : (0, \infty) \rightarrow \mathbb{R}$ is a function such that its second derivative is positive everywhere; in other words, $F''(x) \geq 0$ for all $x \in (0, \infty)$. The Jensen's inequality states that if $F : (0, \infty) \rightarrow \mathbb{R}$ is a convex function and X is a positive random variable then $\mathbb{E}[F(X)] \geq F(\mathbb{E}[X])$. Prove that if $p(x)$ and $q(x)$ are two positive densities on $(0, \infty)$ then $\mathbf{KL}(p, q) \geq 0$.

3 KL divergence: an example

Compute the KL divergence between the two distributions $\mathcal{N}(\mu = 0, \sigma^2 = 1)$ and $\mathcal{N}(\mu = \alpha, \sigma^2 = 1)$.

4 Durians, again..

Use SARIMA modeling to predict the number of google queries containing the word Durian during the whole year after the data collection? Explain your approach. Use cross validation to choose between this approach and a triple-exponential smoothing method.