

Mètodes Bayesians

octubre 2016

Hi ha dos autors candidats d'haver escrit un text, d'autor desconegut, de 10 paraules. Es té la certesa que el text ha estat escrit per un d'aquests dos autors, però no se sap per quin d'ells. El nombre d'adjectius utilitzats en un text de 10 paraules escrites per l'autor candidat i -èssim es pot modelar mitjançant una $Poisson(\lambda_i)$,

$$p(y|\lambda_i) = e^{-\lambda_i} \frac{\lambda_i^y}{y!},$$

on el λ_i , per $i = 1, 2$, se suposa completament conegut, $\lambda_1 = 1$ i $\lambda_2 = 3$. El text d'autor desconegut té $Y = 2$ adjectius. L'objectiu és atribuir el text a un dels dos autors candidats utilitzant aquesta informació.

1. Especifica el model bayesià per abordar aquesta qüestió.
2. Calcula i dibuixa la funció de versemblança estandarditzada.
3. Calcula i dibuixa la distribució a posterior per λ .
4. Calcula la probabilitat de que el nombre d'adjectius en un altre text de 10 paraules escrit pel mateix autor (desconegut) sigui $\tilde{y} = 3$.
5. Proposa un altre model estadístic per modelar el nombre d'adjectius en un text de $n = 10$ paraules per l'autor candidat i -èssim.

Bayesian Methods

october 2016

There are two candidate authors for a text of unknown authorship that has 10 words. One has the certainty that the text is by one of these two authors, but one does not know by which one of them. The number of adjectives used in a text of 10 words written by the i -th candidate author can be modeled through a $Poisson(\lambda_i)$,

$$p(y|\lambda_i) = e^{-\lambda_i} \frac{\lambda_i^y}{y!},$$

where the λ_i , for $i = 1, 2$ are assumed to be completely known, $\lambda_1 = 1$ and $\lambda_2 = 3$. The text of unknown authorship has $Y = 2$ adjectives, and one would like to use that information to attribute the text to one of the two candidate authors.

1. Specify the Bayesian model to address that question.
2. Compute and draw the standardized likelihood function.
3. Compute and draw the posterior distribution for λ .
4. Compute the probability that the number of adjectives in another text of 10 words by the same (unknown) author was $\tilde{y} = 3$.
5. Propose another statistical model to model the number of adjectives in a text of $n = 10$ words by the i -th candidate author.