Obligatory assignment 3 MSA101/MVE187, autumn 2020

Petter Mostad

October 1, 2020

This third assignment builds on the second part of assignment 2, and the data and models presented there. To review, in a study of the number of death notices in the London Times on each day in the interval 1910-1912 for women aged 80 and over, it was noticed that the number of deaths did not follow a Poisson distribution as was expected. The counts Y_i of days with i death notices are given in the table below:

Daily death notices (i)	0	1	2	3	4	5	6	7	8	9
Day count (Y_i)	162	267	271	185	111	61	27	8	3	1

We assume that, each day, there is a choice between two Poisson distributions, one with intensity λ_1 and the other with intensity λ_2. The probability for choosing λ_1 is p and the probability for choosing λ_2 is $1-p$. Thus we have a model with the parameter vector $\theta = (p, \lambda_1, \lambda_2)$. We use the prior $\pi(\theta) \propto \exp(-\lambda_1 - \lambda_2)$.

In addition to the observed data $Y = (Y_0, \dots, Y_9)$, we also consider the unobserved counts $Z = (Z_0, \dots, Z_9)$, where Z_i represents the number of days where λ_1 was used and where i deaths were observed. We would like to use the EM algorithm to compute a maximum posterior estimate for θ in this model:

(a) Review your solution to the second assignment: Write down the log posterior for the total data Y and Z given θ, and describe the distribution $Z \mid Y, \theta^{old}$, for a given value θ^{old}.

(b) Do the E-step of the algorithm, i.e., compute the expectation of the log posterior above under the conditional distribution for Z given above.

(c) Do the M-step of the algoritm, i.e., the maximization step. Remember to include the prior for θ.

(d) Implement the EM-algorithm in this case. Compare your results to those of assignment 2. Make a plot showing for each i the actual counts Y_i, the predicted counts using the model with one λ, fitted in assignment 2, and the predicted counts using the model fitted above, with two λ's.