Production Frontiers with Counts

Anastasios D. Varias
Department of Management Science and Technology, School of Business, Athens University of Economics and Business, 76 Patission street, 10434, Athens, Greece
{ vareias@aueb.gr }

Dimitris Karlis
Department of Statistics, School of Information Sciences and Technology, Athens University of Economics and Business, 76 Patission street, 10434, Athens, Greece
{ karlis@aueb.gr }

Abstract

The aim of this work is to present the concept of production frontiers and to examine the relationship between the number of patents awarded to a firm and the expenditure in Research and Development (R&D). Taking into consideration that the number of patents is a count, we apply count data models, and more specifically mixed Poisson regression models, Karlis and Xekalaki (2005), in order to evaluate the production function and the inefficiency that takes place in the production process. Furthermore, the Expectation – Maximization (EM) algorithm is employed in order to estimate the parameters of the proposed models.

According to the efficiency analysis, a set of inputs is transformed into a set of outputs by a production technology which is modeled by a production function, implying the production frontier, minus a positively defined inefficiency term, Aigner et al. (1977). More specifically, the number of patents is modeled via a Poisson regression model where the log-link function is the linear predictor of the production function and the inefficiency term serves to shape the mixing distribution. In the current work, according to the distribution function that is assumed for the inefficiency term, we present the Poisson – half Normal model, the Poisson – Exponential model, the Poisson – Inverse Gaussian model and the Poisson – Inverse Gamma model.

To that end, effort is given to estimate the coefficients of the production function and the parameters of the distribution regarding the inefficiency term. The EM algorithm is employed which constitutes a powerful approach for maximum likelihood estimates and takes into account data that are missing or are treated like missing. The implementation of the EM algorithm is iterative and consists of two steps, the Expectation step (E-step) and the Maximization step (M-step). According to the E-step the mean values of the missing data are obtained, based on the observed data and the values of the parameters so far. The M-step replaces the missing data with their mean values and maximizes the likelihood function so as to obtain the updated values of the parameters. The algorithm is repeated until a stopping criterion is satisfied, i.e. the degree that the likelihood function has converged, Karlis (2001).

Finally, for the proposed models, we adopt the production functions that are presented in the work of Fé and Hofler (2013). These are evaluated and tested on a data set of 70 pharmaceutical and biomedical firms, illustrated in the National Bureau of Economic Research R&D Masterfile, Hall et al. (1988).

Keywords
Production Frontiers, Counts, EM, Poisson regression

Acknowledgements
The authors would like to thank The George and Victoria Karelias Foundation for the financial support [Scholarship for Graduate Studies].

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. Journal of Econometrics, 6(1), 21-37.

Fé, E., & Hofler, R. (2013). Count data stochastic frontier models, with an application to the patents–R&D relationship. Journal of Productivity Analysis, 39(3), 271-284.

Hall, B. H., Cummins, C., Laderman, E. S., & Mundy, J. (1988). The R&D master file documentation.

Karlis, D. (2001). A general EM approach for maximum likelihood estimation in mixed Poisson regression models. Statistical Modelling, 1(4), 305-318.

Karlis, D., & Xekalaki, E. (2005). Mixed poisson distributions. International Statistical Review, 73(1), 35-58.