Easy and flexible mixture distributions

Mogens Fosgerau* Stefan L. Mabit†

March 11, 2013

1 Extended abstract

We propose a method to generate flexible mixture distributions that are useful for estimating models such as the mixed logit model using simulation. The method is easy to implement, yet it can approximate essentially any mixture distribution. We test it with good results in a simulation study and on real data.

The method is easy yet powerful for creating a mixture distribution for a random parameter in an econometric model that is estimated using simulation. The present the method using maximum simulated likelihood estimation of the mixed logit model as example, but the method can be applied in a wide range of circumstances. The advantages of the method are that essentially any distribution can be represented arbitrarily well, while implementation is very simple.

Consider a model that specifies the likelihood $P(y|x,\beta)$ of some outcome y conditional on variables x and an unobserved random parameter β having distribution F.[1] Assuming that x and β are independent, the likelihood $P(y|x)$ may be simulated given R independent draws β_r from F. This is the basis for estimation by simulation (Train, 2003; McFadden, 1989), which can be applied when the distribution F is considered as known.

Most applications of this method rely on the inversion method for generating draws from F: If u_r are draws from a standard uniform distribution, then $F^{-1}(u_r)$ are draws from F. In order to use this method, it is necessary to compute the

*corresponding author, Technical University of Denmark and Centre for Transport Studies, Sweden, mf@transport.dtu.dk, DTU Transport, Bygningstorvet, Building 116B, 2800 Kgs. Lyngby, Denmark, +45 45251510

†Technical University of Denmark, smab@transport.dtu.dk

[1]There will generally be other parameters to be estimated in the likelihood. They are suppressed in the notation here as the focus lies elsewhere.

inverse of F explicitly.[2]

There are many situations where it is not desirable to impose a specific functional form on F. Generally, this is the case whenever the choice of F has impact on the object of interest for the investigation but there is no a priori reason to choose a particular F. It is particularly undesirable to impose a specific form on F when F is the object of interest itself, e.g., when the purpose is to estimate a distribution of willingness-to-pay. Then it is preferable if the shape of F can be estimated. This can be accomplished by the method of sieves (see e.g. Chen, 2007; Gallant and Nychka, 1987), also known as series estimators. It is however necessary to guarantee that the approximation of F is actually a CDF and then it must be inverted in order to generate random draws from F using the inversion method.

Another idea is to approximate F^{-1} directly. Then inversion is unnecessary. It is however still necessary to ensure that F^{-1} is monotone, which might involve somewhat complicated restrictions on the deep parameters of F^{-1} in a series approximation.

The key insight of this paper is that approximating F or F^{-1} is actually an unnecessary complication for the present purpose. All that is required for simulating the likelihood is draws β_r from some distribution F that depends on some deep parameters to be estimated. The simulated likelihood is simply

$$\frac{1}{R}\sum_{r} P\left(y|x,\beta_r\right). \tag{1}$$

It is not necessary that the draws β_r are monotone functions of standard uniform draws. It is not even necessary to know explicitly the distribution of the draws β_r in order to compute (1); the ability to generate draws from the distribution is sufficient. Being able to obtain the draws, it is always possible to estimate their distribution.

In this paper we take draws u_r from some distribution and transform them using a power series

$$f\left(u|\alpha\right) = \sum_{k=0}^{K} \alpha_k u^k \tag{2}$$

to compute random draws $\beta_r = f\left(u_r|\alpha\right)$ that depend on deep parameters $\alpha = \left(\alpha_0, ..., \alpha_K\right)$ to be estimated. The random draws are inserted into (1) and the resulting expression is very easy to implement in software. For instance, if the model contains a term βx, then that is replaced by $\sum_{k=0}^{K} \alpha_k \left(x u_r^k\right)$. This is a convenient form, since it is linear in deep parameters α that are multiplied by

[2] Devroye (1986) provides a comprehensive treatment of techniques for random variable generation.

easily computed variables xu_r^k. In most cases the distribution of $f(u|\alpha)$ is not easily derived analytically. The distribution is by construction, however, very easy to simulate, which is all that is really needed.

A predecessor for our method is Fleishman (1978), who considers the problem of generating random variables with prespecified moments. He generates a random variable as a third-order polynomial in a standard normal random variable and provides formulae for the coefficients of the polynomial such that specific values of the first four moments are matched by such a variable. The present case is similar, except we are not concerned with matching given moments, but estimate coefficients in order to match a given dataset and may use polynomials of any degree. We present results using both uniform and normal draws.

In conclusion, the paper has developed and applied a simple method for creating flexible mixing distributions. It is easy to implement and mixing distributions can be arbitrarily flexible. The method has successfully been applied in a simulation study as well as to real data, both using the mixed logit model estimated with maximum simulated likelihood. The application to real data was carried out in a freely available and much used package for estimation of discrete choice models, demonstrating that the method is readily applicable and does not require specialised programming.

References

Chen, X., 2007. Large sample sieve estimation of semi-nonparametric models. In: Heckman, J. J., Leamer, E. E. (Eds.), Handbook of Econometrics. Vol. 6, Part B. pp. 5549–5632.

Devroye, L., 1986. Non-Uniform Random Variate Generation. Springer-Verlag, New York.

Fleishman, A. I., Dec. 1978. A method for simulating non-normal distributions. Psychometrika 43 (4), 521–532.

Gallant, A., Nychka, D., 1987. Semi-nonparametric maximum likelihood estimation. Econometrica 55 (2), 363–390.

McFadden, D., 1989. A method of simulated moments for estimation of the multinomial probit without numerical integration. Econometrica 57 (5), 995–1026.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, MA.