Working Paper 92-41
Statistics and Econometrics Series 29
December 1992

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

REJECTION SAMPLING IN DEMAND SYSTEMS

Eduardo Ley and Mark F.J. Steel*

Abstract

We illustrate the method of rejection sampling in a Bayesian application of a new approach to estimating Demand Systems. This approach, suggested by Varian (1990), is based on a generalization of Afriat's (1967) efficiency index. Rejection sampling is applied to the prior-to-posterior mapping enabling us to obtain posterior results in a nonstandard model.

Key Words

Rejection Sampling, Bayesian Methods, Demand Systems, Efficiency.

*Ley, Departamento de Economía, Universidad Carlos III de Madrid; Steel, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid.

Rejection Sampling in Demand Systems

Eduardo Ley
Mark F.J. Steel
Universidad Carlos III de Madrid

September, 1992
Version: November 18, 1992

Abstract. We illustrate the method of rejection sampling in a Bayesian application of a new approach to estimating Demand Systems. This approach, suggested by Varian (1990), is based on a generalization of Afriat's (1967) efficiency index. Rejection sampling is applied to the prior-to-posterior mapping enabling us to obtain posterior results in a nonstandard model.

Keywords. Rejection Sampling, Bayesian Methods, Demand Systems, Efficiency

Address. Eduardo Ley, Departamento de Economía, Universidad Carlos III, 28903–Getafe, Madrid. Email: `edley@eco.uc3m.es`. Mark Steel, Departamento de Estadística y Econometría, Universidad Carlos III, 28903–Getafe, Madrid. Email: `steel@eco.uc3m.es`

Rejection Sampling in Demand Systems

Eduardo Ley
Mark F.J. Steel

1. Introduction

Varian (1990) proposes the use of the money-metric utility function[1] to estimate the parameters of a demand system. He arrives to that formulation from a parametric generalization of Afriat's (1967) efficiency index. The main motivation for such a novel approach is the utilization of a sensible norm of goodness of fit.

Let the utility function be characterized by the functional form $u(\cdot)$ and the parameter $\alpha \in R^p$. Suppose that, for $t = 1, \ldots, T$, we observe the pairs (x_t, p_t), where $x_t = (x_{1t}, x_{2t}, \ldots, x_{nt})'$ is a vector of the n-good bundle chosen when prices are $p_t = (p_{1t}, p_{2t}, \ldots, p_{nt})$; $m_t = p_t x_t$ will be the consumer's expenditure at t.[2] We postulate that if x_t is not the optimal choice at prices p_t, it must provide a level of utility *close* to the optimal. The ordinal character of the utility function makes it difficult to operationalize this closeness concept. The most widespread practice in applied demand studies is to use a norm—usually a quadratic norm—on the goods' or the expenditure shares' space. This norm lacks any economic content as pointed out by Varian. Bundles providing similar levels of utility might be far away in the goods' space and vice-versa. Instead, we focus on a measure of efficiency based on money-metric utility.

Bayesian methods of inference will be used to formally treat parameter uncertainty and actually derive posterior densities for the out-of-sample or "average" efficiencies under different sampling models and prior assumptions. [Ley and Steel (1992) studied the behavior of within-sample efficiencies for the exponential model presented here.]

[1] Denote by x the bundle of goods, by p the vector of prices, and by m the individual's income. Assuming the utility function $u(x)$, we define the indirect utility function as $v(p, m) \equiv \max_x\{u(x) : px \leq m\}$; the expenditure function as $c(p, u) \equiv \min_x\{px : u(x) \geq u\}$; and the money-metric utility function as $e(p, x) \equiv c(p, u(x))$. See Varian (1992) for the derivation and properties of these functions.

[2] In some subsequent discussion, the subscript t is suppressed for notational convenience.

2. The Sampling Model

If $e(p, x)$ is the minimum expenditure required to achieve, at p, the utility level given by the chosen bundle, x, the *wasted* expenditure will be $m - e(p, x)$. Suppose that this wasted expenditure is *randomly proportional* to the *minimum* expenditure

$$m - e(p, x) = e(p, x)\eta, \tag{1}$$

where $\eta \geq 0$ is a random variable. Manipulating (1) and making explicit that the minimum expenditure is parametrized by α, we can write the inverse efficiency measure

$$\frac{m}{e(p, x; \alpha)} = 1 + \eta \equiv e^{\varepsilon} \geq 1, \tag{2}$$

where $\varepsilon = \log(1 + \eta) \geq 0$. Taking natural logs, we can rewrite (2) more conveniently as $\log m - \log e(p, x; \alpha) = \varepsilon \geq 0$. Basing the estimation of the relevant parameters on this equation offers several advantages over the methods traditionally employed in applied demand analysis. In addition to the motivation just presented in favor of using a norm with economic content, one only needs to impose homogeneity and concavity restrictions on $e(p, x; \alpha)$ to obtain estimates consistent with economic theory as opposed to the cross-equation restrictions imposed on a demand system to obtain a symmetric and negative-semidefinite Slutsky matrix which are far more difficult to implement in practice.

The sampling model then becomes

$$\log m_t - \log e(p_t, x_t; \alpha) = \varepsilon_t, \tag{3}$$

where $\varepsilon_t \geq 0$ has some distribution defined over the positive real line. With exponentially distributed error terms[3] ε_t (i.i.d. for $t = 1, 2, \ldots, T$), $p(\varepsilon_t|\mu) = f_\gamma(\varepsilon_t|1, \mu)$, we obtain for all T observations:

$$p(\log m|e(p, x; \alpha), \mu) = \exp\left\{-T\log\mu - \mu^{-1}\sum_{t=1}^{T}[\log m_t - \log e(p_t, x_t; \alpha)]\right\}, \tag{4}$$

where m, $e(p, x; \alpha)$, p and x are all straightforwardly extended to the case of T observations. Under the asssumption that the error terms follow a half-Normal distribution $p(\varepsilon_t|\sigma^2) = \sigma^{-1}\sqrt{2/\pi}\exp\{-\varepsilon^2/(2\sigma^2)\}$, we have the sampling density

$$p(\log m|e(p, x; \alpha), \sigma^2) = \left(\frac{2}{\pi}\right)^{T/2} \exp\left\{-T\log\sigma - \frac{1}{2\sigma^2}\sum_{t=1}^{T}[\log m_t - \log e(p_t, x_t; \alpha)]^2\right\}. \tag{5}$$

[3] The random variable $z > 0$ has a gamma distribution with parameters α and β if its density function is given by $f_\gamma(z|\alpha, \beta) = z^{\alpha-1}e^{-z/\beta}\beta^{-\alpha}/\Gamma[\alpha]$, with $E[x] = \alpha\beta$, and $Var[x] = \alpha\beta^2$. A gamma distribution with shape parameter $\alpha = 1$ is also known as an exponential distribution.

In the simple Cobb-Douglas case with only three goods ($n = 3$), the normalized utility function is $u(x;\alpha) = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$, with $\alpha_3 = 1 - \alpha_1 - \alpha_2$, and the minimum expenditure—*i.e.*, the money-metric utility function—becomes $e(p_t, x_t; \alpha) = \prod_{i=1}^{3} (p_{it} x_{it}/\alpha_i)^{\alpha_i}$ which can be substituted into (4) and (5).

3. Prior Densities

To complete the Bayesian model we need to specify a prior density on the parameters in (4) or (5). The complicated form of the likelihood functions precludes an analytical analysis and we are obliged to follow a numerical approach. As the method we plan to use involves drawing from the prior, we shall only consider proper priors here. However, a sensitivity analysis will be conducted.

The prior on α will be of the Dirichlet form[4]

$$p(\alpha) = f_D^3(\alpha | k), \tag{6}$$

and on the parameter of the exponential model we specify

$$p(\mu^{-1}|\alpha) = p(\mu^{-1}) = f_\gamma(\mu^{-1}|1, (-\log r^*)^{-1}). \tag{7}$$

The reason for choosing (7) is its analytical tractability and its associated ease in elicitation. In particular, it leads to such a prior density for ε_t that the prior median of the efficiency $r_t = e^{-\varepsilon_t}$ is exactly equal to r^* [see van den Broeck *et al.* (1992)]. A similar reasoning for the half-Normal model leads to

$$p(\sigma^{-2}|\alpha) = p(\sigma^{-2}) = f_\gamma(\sigma^{-2}|5, 0.1(-\log r^*)^{-2}), \tag{8}$$

again based on the prior median efficiency r^*.

[4] The vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)'$ with $\alpha_i > 0$, $\forall i$, and $\sum_{i=1}^{n} \alpha_i = 1$ has a Dirichlet distribution with parameters $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)'$ if its density function is given by $f_D^n(\alpha|\gamma) = \Gamma(g) \prod_{i=1}^{n} [\alpha_i^{\gamma_i - 1}/\Gamma(\gamma_i)]$, where $\gamma_i > 0$, $\forall i$, and $g = \sum_{i=1}^{n} \gamma_i$. The first two moments are given by $E[\alpha_i] = \gamma_i/g$, and $\text{Var}[\alpha_i] = \gamma_i(g - \gamma_i)/(g^2(g+1))$.

4. Rejection Sampling

Here we adopt the rejection sampling technique, which can generally be used if we need to draw from a density $f(y)$ which is too complicated to draw from directly. If we can find a density $g(y)$ such that

$$M = \max_y \frac{f(y)}{g(y)} < \infty$$

and we can generate drawings from $g(y)$, then we can use the following algorithm [see Ripley (1986), Johnson (1987)]:

[S1] Generate y_d from $g(y)$ and compute $R_d = f(y_d)/Mg(y_d)$.

[S2] Generate u_d from a uniform on $(0,1)$.

[S3] If $u_d \leq R_d$ accept y_d; otherwise reject y_d.

The accepted drawings y_d will be distributed according to $f(y)$.

This general principle is applied to the prior-to-posterior mapping in Smith and Gelfand (1992). If we take the prior $p(\alpha, \theta)$ as $g(\cdot)$ where θ is either μ or σ^2, we can obtain drawings from the posterior with kernel $f(\cdot)$ provided that $M = \max_{\alpha,\theta} p(\log m|\epsilon(p,x;\alpha),\theta)$ is finite. The ratio R_d then becomes the ratio of the likelihood value at the drawn parameter vector and the maximum value of the likelihood, M. In both cases treated here M will be finite.

5. Posterior Results

We use U.S. aggregate consumption data of three groups of goods: durables, nondurables and services from 1947 to 1987.[5] This implies $n = 3$ and $T = 41$. Varian (1990) lists the data.

As the prior on α we assume the Dirichlet in (6) with parameter vector $k = (30, 90, 80)$ which is elicited postulating mean expenditure shares of 15%, 45%, and 40% with standard deviations of 2.5%, 3.5% and 3.5% for x_1, x_2 and x_3 respectively. Alternatively, we take parameters $(3, 9, 8)$ leading to standard deviations of 7.8%, 10.9% and 10.7%. We also perform a sensitivity analysis over θ using the values of 0.90, 0.95 and 0.98 for the prior median efficiency r^* in (7) and (8). All the results are based on 2,000 accepted drawings and the acceptance rates are reported in table 1.

[5] We are well aware that the i.i.d. assumption for ε is a bit suspicious for these time-series data. However, we believe that sophisticating the sampling model might not add much credibility to this simple aggregate demand system and might distract the attention from what we want to focus on—i.e., the rejection sampling method.

Table 1. Rejection Sampling Acceptance Rates

Dirichlet Hyperparameters	(30, 90, 80)			(3, 9, 8)		
Prior Median Efficiency	0.90	0.95	0.98	0.90	0.95	0.98
Exponential	0.1034%	0.6162%	1.1148%	0.0121%	0.0675%	0.1337%
Half-Normal	0.0000%	0.0013%	3.7225%	0.0000%	0.0002%	0.4822%

It seems that acceptance rates increase as the prior information is more in line with the data. We draw from the prior in order to sample from the posterior. Clearly, if the two are close, acceptance rates will be high. Therefore, if the prior densities are not too flat (as the tighter Dirichlet on a) and have their mass concentrated in areas with high likelihood values ($r^* = 0.98$) we can expect relatively high rates of acceptance. Note that Varian (1990) finds an average "wasted expenditure" of 2% using classical nonlinear least squares methods.[6] As found for the truncated Normal in van den Broeck *et al.* (1992), the prior for the half-Normal in (7) is much less conservative than (8), so if (7) corresponds to the data information it leads to higher acceptance rates than in the exponential model, whereas these rates are much lower for cases with conflicting prior information.

5.1. *Efficiency*

We are particularly interested in the posterior distribution of the efficiency measure given by

$$r_t = \frac{e(p_t, x_t; a)}{m_t} = e^{-\varepsilon_t} \in (0, 1] \tag{9}$$

which is a parametric generalization of Afriat's efficiency index [Varian (1990)]. The posterior density of this efficiency measure can be examined by drawing from the posterior density of ε_f, where f is an unobserved year, $p(\varepsilon_f|p,x) = \int p(\varepsilon_f|\theta)p(\theta|p,x)d\theta$. Thus, at each drawing, θ_d, of θ from $p(\theta|p,x)$ we can draw ε_f from $p(\varepsilon_f|\theta_d)$ and evaluate $r_f = e^{-\varepsilon_f}$ to study the distribution of this ***average*** efficiency measure. The corresponding prior efficiency can be found by integrating out θ from the joint density of (ε_t, θ) which leads, in both models, to an analytical expression.[7] Table 2 groups some characteristics of the posterior densities of r_f, whereas figure 1 shows the prior and posterior densities for both

[6] Varian's NLLS point estimates correspond to the maximum likelihood estimates of the half-Normal model.

[7] In the exponential model, the marginal prior density for $r_t = e^{-\varepsilon_t}$ is $p(r_t) = \frac{-1}{r_t \log r^*}\left(1 + \frac{\log r_t}{\log r^*}\right)^{-2}$ for all years. In the half-Normal model, we have $p(r_t) = \frac{63\sqrt{5}}{256}\frac{1}{r_t|\log r^*|}\left(1 + \frac{(\log r_t)^2}{20(\log r^*)^2}\right)^{-11/2}$ for all years.

Table 2. Posterior Characteristics of Efficiency

	Exponential						Half-Normal			
Dirichlet Hyperparameters	(30, 90, 80)			(3, 9, 8)			(30, 90, 80)		(3, 9, 8)	
Prior Median Efficiency	0.90	0.95	0.98	0.90	0.95	0.98	0.95	0.98	0.95	0.98
Mean	0.9785	0.9797	0.9805	0.9779	0.9791	0.9798	0.9664	0.9768	0.9656	0.9763
Median	0.9853	0.9861	0.9865	0.9848	0.9854	0.9859	0.9714	0.9802	0.9708	0.9796
Mode	0.9995	0.9989	0.9996	0.9970	0.9986	0.9958	0.9835	0.9913	0.9770	0.9970
S.D.	0.0219	0.0205	0.0197	0.0233	0.0211	0.0204	0.0255	0.0179	0.0256	0.0180

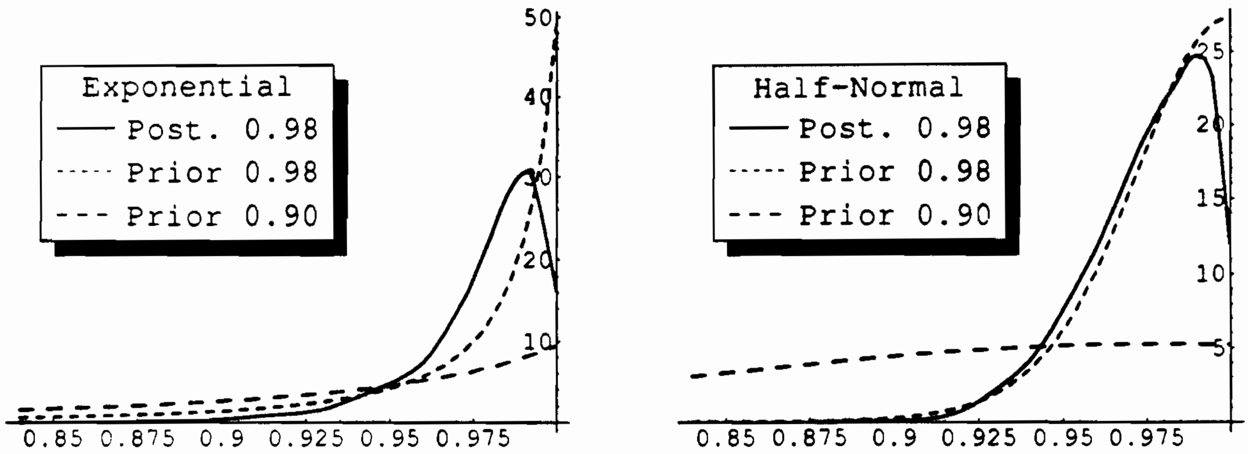

Figure 1. Prior and Posterior Densities of Efficiency: $k = (30, 90, 80)$, various r^*.

models for $r^* = 0.98$ and $k = (30, 90, 80)$ (for the half-Normal we do not present results for $r^* = 0.90$ as acceptance rates are virtually zero).

From table 2 the sample information clearly dominates the prior information we assume in (7) and (8). Posterior densities of r_f are hardly affected by our choice of prior median efficiencies r^*. The choice of k in (6) is not found to matter either, so results are only graphed for $k = (30, 90, 80)$ and $r^* = 0.98$. The different models, however, do produce slightly different results, as the posterior varies more with r^* for the half-Normal. This finding is linked to the acceptance rate behavior in table 1 and illustrated in figure 1. Prior and posterior densities are closest for the half-Normal model for $r^* = 0.98$, whereas prior (8) for $r^* = 0.90$ moves away sharply from areas with considerable likelihood values.

5.2. Marginal Rate of Substitution

Finally, we can look at the posterior distribution of any transformation of the parameters. For instance, the marginal rate of substitution (MRS) between nondurables, x_2, and services, x_3, when equal amounts of both goods are consumed is given by α_2/α_3. We perform an informal test on the adequacy of the Cobb-Douglas utility function (which assumes constant expenditure shares) by splitting the sample in two groups of 20 observations—$t = 1, \dots, 20$ and $t = 22, \dots, 41$, omitting observation 21—and comparing the posterior distribution of this MRS. Figure 2 displays the prior and the two posterior densities of the MRS in the exponential model.

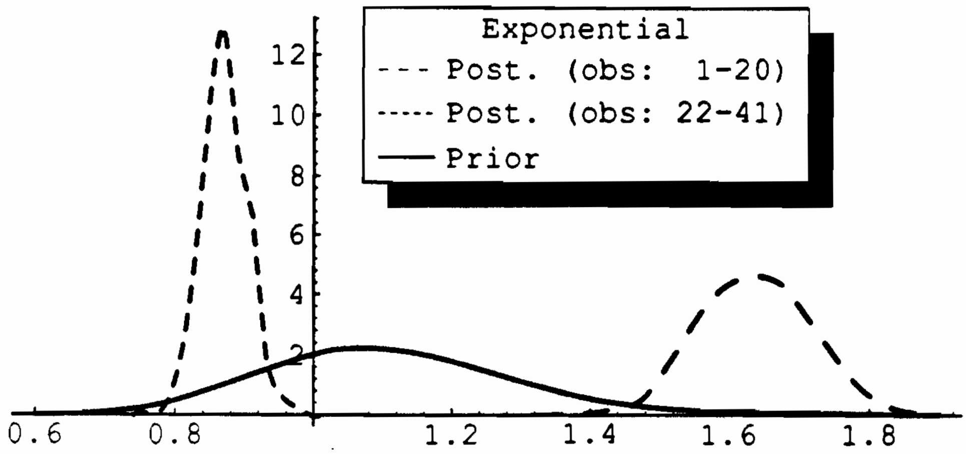

Figure 2. Prior and Posterior Densities of MRS (exponential model): $k = (30, 90, 80)$ and $r^* = 0.99$.

As expected from the evolving shares of nondurables and services on total expenditure along the years, the posterior densities lie very far apart. We take this as clear evidence that the constant-share constraint imposed by the Cobb-Douglas functional form is at odds with these data. As this analysis is merely meant to illustrate the feasibility and flexibility of Bayesian methods using rejection sampling we shall, however, not explore more flexible functional forms in this paper. We leave that task for subsequent, more challenging, applications.

6. Concluding Remarks

The method of rejection sampling seems an extremely flexible and easily understood mechanism for the application of Bayesian analysis in a very general context. Its main cost is in terms of computer time, but as more and more computing power becomes generally available, empirical applications of economic relevance can now be performed routinely. We hope this simple illustration will prompt applied researchers to use rejection sampling on a wide range of problems.

References

Afriat, S. (1967): "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, **8**, 67–77.

Johnson, M.E. (1987): *Multivariate Statistical Simulation*. New York: Wiley.

Ley, E. and M.F.J. Steel (1992): "Bayesian Econometrics: Conjugate Analysis and Rejection Sampling," forthcoming in *Economic and Financial Modelling using Mathematica*, ed. H.R. Varian. New York: SpringerVerlag.

Ripley, B. (1986): *Stochastic Simulation*. New York: Wiley.

Smith, A.F.M. and A.E. Gelfand (1992): "Bayesian Statistics Without Tears: A Sampling-Resampling Perspective," *American Statistician*, **46**, 84–88.

van den Broeck, J., G. Koop, J. Osiewalski and M.F.J. Steel (1992): "Stochastic Frontier Models: A Bayesian Perspective," Working Paper 92-12, Universidad Carlos III de Madrid.

Varian, H.R. (1990): "Goodness-of-Fit in Optimizing Models," *Journal of Econometrics*, **46**, 125–140.

Varian, H.R. (1992): *Microeconomic Analysis*. New York: W.W. Norton.