Averaging Income Distributions

William E. Griffiths
University of Melbourne

Duangkamon Chotikapanich
Curtin University of Technology

D. S. Prasada Rao
University of New England

Various inequality and social welfare measures often depend heavily on the choice of a distribution of income. Picking a distribution that best fits the data in some sense involves throwing away information and does not allow for the fact that, by chance, a wrong choice can be made. It also does not allow for the statistical inference implications of making the wrong choice. Instead, Bayesian model averaging utilises a weighted average of the results from a number of income distributions, with each weight given by the probability that a distribution is 'correct'. In this study prior densities are placed on mean income, the mode of income and the Gini coefficient for Australian income units with one parent (1997-98). Then, using grouped sample data on incomes, posterior densities for the mean and mode of income, and the Gini coefficient are derived for a variety of income distributions. The model-averaged results from these income distributions are obtained.

Key words: Bayesian model averaging; Gini coefficient; grouped data.

JEL classification: C11, D31

1. Introduction

Considerable research has been devoted to specifying and estimating alternative models for income distributions, and to making consequent inferences about the level of inequality in those distributions. See, for example, McDonald (1984), Dagum (1996), Bordley et al (1996), Creedy et al (1997), and references given in those papers. A typical approach for assessing inequality is to fit a number of alternative models for an income distribution, pick the best-fitting model, and, from that estimated model, compute values of inequality measures of interest, such as the Gini coefficient or one of Atkinson's inequality measures. The best-fitting model could be one that maximizes the likelihood function, or minimizes an information criterion, or, it could be chosen via a sequence of formal hypothesis tests. In any event, one problem with this practice is that, once a particular model has been chosen, the fact that a number of other models have been discarded is usually ignored. No allowance is made for the possibility of sample statistics yielding an incorrect choice. Assessment of the precision of estimation via standard errors makes no provision for the preliminary-test implications for inference. Although the preliminary-test problem has received considerable attention in the sampling-theory literature (see, for example, Judge and Bock 1978, and Danilov and Magnus 2001), most of the solutions are for particular special cases and do not carry over to the model selection problem considered here.

An alternative way to proceed, and one that does yield results and measures of precision that reflect model uncertainty, is to adopt Bayesian model averaging along the lines described by Geweke (1999). In this approach the results from each model are combined, as a weighted average, with the weight attached to each model being the posterior probability that that model is “correct”. If one model is vastly superior to the others, then its posterior probability will be close to one, and the averaged results will not be distinguishable from those of the best-fitting superior model. On the other hand, if the choice between models is a less definite one, then each of the models will contribute information to the averaged-results, and measures of precision, such as posterior standard deviations, will reflect the model uncertainty. In this paper we illustrate how Bayesian model averaging can be applied to alternative income distribution models. The quantities of interest, for which averaged results are presented, are taken as mean income, modal income and the Gini coefficient. Seven income distribution models are considered; they are the lognormal, gamma, Weibull, Fisk, Singh-Maddala, beta type 2 (beta2) and Burr type 3 (Burr3) distributions. The methods are applied to a 1997-98 sample of gross weekly income for one-parent income units in Australia (Australian Bureau of Statistics 1999).

In Section 2 an overview of the methodology is given; some aspects of the methodology are spelt out in more detail in subsequent sections. Specification of the prior probability density functions (pdf’s) is described in Section 3. Derivation of the posterior pdf’s and estimation of quantities from them is considered in Section 4. The technique for computing posterior model probabilities appears in Section 5. The results are presented and discussed in Section 6 and concluding remarks are made in Section 7.

2. Overview of Methodology

Characteristics of the seven distributions that were considered are given in an Appendix. There are four 2-parameter distributions (lognormal, gamma, Weibull and Fisk) and three 3-parameter distributions (Singh-Maddala, beta2 and Burr3). Of course, many other distributions could be added to this list. These seven were chosen as being reasonably representative of popular choices, and sufficient to illustrate the procedure without becoming excessively complex. In what follows they are denoted by M_i, $i = 1, 2, ..., 7$.

Let θ_i be a 2 or 3-dimensional vector containing the unknown parameters of the *i*-th income distribution, $i = 1, 2, ..., 7$. Also, let $\delta_{(2)}$ and $\delta_{(3)}$ be 2 and 3-dimensional vectors, respectively, containing what we term "economic quantities of interest". The vector $\delta_{(2)}$ contains mean income and the Gini coefficient; $\delta_{(3)}$ contains mean income, modal income and the Gini coefficient. The first step is to assign prior pdf's to $\delta_{(2)}$ and $\delta_{(3)}$. Details of the specification that we employ are given in Section 3. The prior pdf for $\delta_{(2)}$ is the same for all 2-parameter distributions and the prior pdf for $\delta_{(3)}$ is the same for all 3-parameter distributions. Specifying priors in this way means that the same prior information is used for all distributions that have the same number of parameters. The mean, mode and Gini coefficient were chosen as quantities of interest for which some prior information is likely to be available. One of these quantities becomes redundant (we chose the mode) when 2-parameter distributions are considered. That is, knowledge of mean income and the Gini coefficient implies knowledge of the 2 distributional parameters which in turn implies knowledge of the mode.

Using $f(.)$ as generic notation for any pdf, the prior pdf for the parameters of the *i*-th income distribution, $f(\theta_i \mid M_i)$, is obtained by transformation of the prior pdf on the economic quantities of interest, $f(\delta_{(k)})$. That is,

$$f(\theta_i \mid M_i) = f(\delta_{(k)}) \left| \frac{\partial \delta_{(k)}}{\partial \theta_i} \right| \qquad i = 1,2,3,4 \text{ for } k = 2 \qquad (1)$$

$$i = 5,6,7 \text{ for } k = 3$$

Derivatives useful for computing the Jacobian term $\left|\partial\delta_{(k)}/\partial\theta_i\right|$ appear in the Appendix. Sometimes it was more convenient to use the derivative of the log of an element of $\delta_{(k)}$ (or the log of a function of an element) and then adjust it appropriately to get the correct Jacobian.

The posterior pdf's of the parameters for each of the income distribution models are obtained by combining the prior pdf in (1) with the likelihood functions as prescribed by Bayes' theorem. To specify the likelihood functions, we begin by recognizing that the data are grouped into 14 classes. They take the form of the number of sampled income units in each of 14 income classes, as depicted in Table 1. The income classes refer to weekly gross income, measured in dollars, of one-parent income units. For estimation purposes the largest possible income in the last group is taken as 100,000. For all income distributions, the likelihood function is the pdf for a potential sample of numbers of income units in each of the groups, $x = (x_1, x_2, \ldots, x_{14})'$. It is given by the multinomial distribution

Table 1. Income Classes and Number of Income Units

Income class (z_{j-1}, z_j)	Number of income units (x_j)
1 - 119	7
200 - 159	5
160 - 199	14
200 - 299	154
300 – 399	120
400 – 499	76
500 – 599	54
600 – 699	49
700 - 799	22
800 - 999	43
1000 - 1199	12
1200 - 1499	10
1500 - 1999	1
≥ 2000	3

$$f\left(x \mid \theta_i, M_i\right) \propto \prod_{j=1}^{14}\left[F_i\left(z_j;\theta_i\right) - F_i\left(z_{j-1};\theta_i\right)\right]^{x_j} \qquad i = 1, 2, \ldots, 7 \tag{2}$$

where $F_i(.)$ is the cumulative distribution function (cdf) for the *i*-th income distribution. Expressions for the cdf's for each of the income distributions are given in the Appendix. The proportionality constant that is omitted from equation (2) is the same for all models.

The posterior pdf's for the parameters of each income distribution are given by

$$\begin{aligned} f\left(\theta_i \mid M_i, x\right) &\propto f\left(x \mid \theta_i, M_i\right) f\left(\theta_i \mid M_i\right) \\ &= f\left(x \mid \theta_i, M_i\right) f\left(\delta_{(k)}\right)\left|\frac{\partial \delta_{(k)}}{\partial \theta_i}\right| \qquad i = 1, 2, ..., 7 \end{aligned} \tag{3}$$

These pdf's are not sufficiently tractable for derivation of the moments of the elements in each θ_i or for deriving marginal posterior pdf's. Nor is it possible to derive marginal posterior pdf's and posterior moments for the economic quantities of interest that are elements in the $\delta_{(k)}$ vectors. However, it is straightforward to use a Metropolis-Hastings algorithm to draw observations $\theta_i^{(1)}, \theta_i^{(2)}, \ldots, \theta_i^{(N)}$ from each of the posterior pdfs $f(\theta_i \mid M_i, x)$. From these draws we can compute observations on mean income, modal income and the Gini coefficient. These observations can then be used to estimate marginal posterior pdfs and moments for δ. Expressions for the quantities of interest δ, in terms of the θ_i, are given in the Appendix. At this point it is convenient to drop the subscript (*k*) from δ and to assume that δ is a 3-dimensional vector containing the mean, mode and Gini coefficient. Even although one of these quantities can always be found from the other two in the case of 2-parameter distributions, we can generate observations on all three quantities, and all three are of interest for all models.

To draw observations from the posterior pdf's, a random-walk Metropolis-Hastings algorithm, similar to that employed by Griffiths and Chotikapanich (1997) was used. In each case a total of 45,000 observations were generated, with 5,000 being discarded as a burn-in . Plots of the observations were taken to confirm the convergence of the Markov Chain. For each θ_i, corresponding values for δ were obtained. The observations were used to estimate posterior means and standard deviations for (θ_i, δ) from each model, and to obtain histograms used to estimate posterior pdfs for δ, conditional on each income distribution model.

Following the above procedure yields seven posterior pdfs for mean income, modal income and the Gini coefficient. Examining these pdfs and their moments reveals the sensitivity of estimation to choice of a specific income distribution model. To proceed with averaging the results from the different models, we use the result

$$E\left[g(\delta) \mid x\right] = \sum_{i=1}^{7} E\left[g(\delta) \mid M_i, x\right] P\left(M_i \mid x\right) \quad (4)$$

where g(.) is a function of interest. Several things in equation (4) need explanation.

1. We consider three functions for $g(\delta)$. Setting $g(\delta) = \delta_m$ and $g(\delta) = \delta_m^2$ $(m = 1, 2, 3)$ gives us the opportunity to compute posterior means and variances (and standard deviations). For the third function we divide the range of each δ_m into a number of histogram classes, and make $g(\delta)$ a series of indicator functions, equal to unity when an observation falls into a histogram class, and zero otherwise. In this case equation (4) can be viewed as an averaging of the numbers in each histogram class over the seven income distributions. With suitable scaling, the average histogram is an estimate of the average posterior pdf.

2. The expectations conditional on each model, $E\left[g(\delta) \mid M_i, x\right]$ are estimated from the Metropolis-Hastings generated observations. Thus, equation (4) is applied to these estimated quantities.

3. The posterior probabilities attached to each model, $P\left(M_i \mid x\right)$, are obtained from the discrete version of Bayes theorem

$$P\left(M_i \mid x\right) = \frac{f\left(x \mid M_i\right) P\left(M_i\right)}{\sum_{j=1}^{7} f\left(x \mid M_j\right) P\left(M_j\right)} \quad (5)$$

where the $P\left(M_j\right)$ are prior probabilities assigned to each of the models, and the $f\left(x \mid M_j\right)$ are the marginal likelihoods

$$f\left(x \mid M_j\right) = \int f\left(x \mid \theta_j, M_j\right) f\left(\theta_j \mid M_j\right) d\theta_j \tag{6}$$

The marginal likelihoods are also estimated from the posterior draws, as described in Section 4.

3. The Prior Specification

The prior specification has two components. Prior pdf's are required for $\delta_{(2)}$ and $\delta_{(3)}$ and prior probabilities $P\left(M_i\right)$ need to be assigned to each of the seven models. We first consider $\delta_{(2)} = (\mu, \gamma)'$ where μ denotes mean income and γ denotes the Gini coefficient. Independent gamma and beta distributions were chosen for μ and γ, respectively. The parameters of these distributions were chosen so that the priors were consistent with the views of applied researchers in the income distribution area, as well as relatively noninformative, so they would not conflict with a wide range of prior opinions. Specifically,

$$f\left(\delta_{(2)}\right) = f(\mu, \gamma) = f(\mu)\, f(\gamma) \tag{7}$$

where

$$f(\mu) = \frac{1}{b^c \Gamma(c)} \mu^{c-1} e^{-\mu/b} \qquad c = 2,\ b = 300 \tag{8}$$

and

$$f(\gamma) = \frac{1}{B(v, w)} \gamma^{v-1} (1-\gamma)^{w-1} \qquad v = 1.1,\ w = 2 \tag{9}$$

Some characteristics of these distributions are summarised in Table 2, along with characteristics of the prior pdf for the mode to which we now turn.

For 3-parameter distributions we specified the prior pdf

$$f\left(\delta_{(3)}\right) = f\left(\mu, \gamma, m_o\right) = 1.4979\, f\left(\mu\right) f\left(\gamma\right) f\left(m_o\right) \qquad m_o < \mu \tag{10}$$

where m_o denotes modal income, $f\left(\mu\right)$ and $f\left(\gamma\right)$ are given in equations (8) and (9), respectively, and

$$f\left(m_o\right) = \frac{1}{b^c \Gamma(c)} m_o^{c-1} e^{-m_o/b} \qquad c = 1.3,\ b = 300 \tag{11}$$

Characteristics of the pdf in equation (11) appear in Table 2. Note that these and the other characteristics in Table 2 do not hold for the joint prior pdf in equation (10). Imposing the *a priori* restriction that the mode is less than the mean restricts the parameter supports and changes at least the means and the probability intervals. Also, the value 1.4979 that appears in equation (10) is the additional constant necessary to make the joint prior pdf integrate to unity, after imposing the truncation $m_o < \mu$.

Table 2. Characteristics of Prior pdfs

	Mean income μ	Gini coefficient γ	Modal income m_o
mean	600	0.355	390
mode	300	0.020	90
95% probability interval	(107, 1423)	(0.030, 0.815)	(35, 1066)
80% probability interval	(247, 898)	(0.143, 0.616)	(115,613)

For prior probabilities for each of the models, we took all 2-parameter distributions as equally likely, and all 3-parameter distributions as equally likely. To penalize the additional complexity of 3-parameter distributions, the prior probabilities for the 2-parameter distributions were set equal to 1.5 times the prior probabilities for the 3-parameter distributions. We therefore have

$$P\left(M_i\right)=\begin{cases}0.16667 & \text{for } i=1,2,3,4\\ 0.11111 & \text{for } i=5,6,7\end{cases} \tag{12}$$

4. Posterior Model Probabilities

Estimation of the marginal likelihood for each model (defined in equation (6)) is the major task to be completed before posterior model probabilities can be calculated. Once the marginal likelihood values have been found, it is straightforward to apply equation (5) for computation of the $P\left(M_i \mid x\right)$. The modified harmonic mean approach suggested by Gelfand and Dey (1994), and described in Geweke (1999) was used. Specifically, an estimate of the inverse of the marginal likelihood is given by

$$\left[\hat{f}\left(x \mid M_i\right)\right]^{-1}=\frac{1}{N}\sum_{n=1}^{N}\frac{h\left(\theta_i^{(n)}\right)}{f\left(x \mid \theta_i^{(n)}, M_i\right) f\left(\theta_i^{(n)} \mid M_i\right)} \tag{13}$$

where

1. $h\left(\theta_i\right)$ is a truncated normal distribution

$$h\left(\theta_i\right)=P^{-1}\left(2\pi\right)^{-k/2}\left|\hat{\Sigma}_i\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\theta_i-\overline{\theta}_i\right)'\hat{\Sigma}_i^{-1}\left(\theta_i-\overline{\theta}_i\right)\right\}$$

   truncated such that $\left(\theta_i-\overline{\theta}_i\right)'\hat{\Sigma}_i^{-1}\left(\theta_i-\overline{\theta}_i\right)\le\chi_P^2(k)$.

2. $\chi_P^2(k)$ is a value from a chi-square distribution for cdf probability P and k degrees of freedom.

3. $\overline{\theta}_i$ and $\hat{\Sigma}_i$ are the sample mean and covariance matrix from the retained N = 40,000 draws $\theta_i^{(n)}, n=1,2,\ldots,N$.

For each model (13) was evaluated for $P = 0.1, 0.2, \ldots, 0.9$. Similar values were obtained for all *P*.

5. The Application

The techniques described in the preceding sections were applied to the data in Table 1. Maximum likelihood estimates and standard errors and posterior means and standard deviations of the parameters for each of the seven income distributions appear in Table 3. This table also contains the maximum log-likelihood function value for each of the models along with the log of the marginal likelihood, estimated using equation (13).

From Table 3 we observe that:

1. There is little difference between the maximum likelihood estimates and the posterior means. This outcome suggests that the posterior pdf has been dominated by the sample data, as was our original intention.
2. The standard errors from maximum likelihood estimation and the posterior standard deviations are sufficiently small to suggest relatively precise estimation. A possible exception is the results for the parameters *b* and *p* for the beta2 distribution; in this case the posterior standard deviations are larger relative to the posterior means.
3. Based on the log-likelihood value the best fitting distributions are Burr3, beta2, Singh-Maddala and lognormal. After penalizing the 3-parameter distributions for their additional parameter (through the assignment of prior model probabilities), the Singh-Maddala distribution turns out to be favoured less.

In Table 4 we present the posterior means and standard deviations for mean income, modal income and the Gini coefficient, conditional on each of the income distribution models. The unconditional posterior means and standard deviations, obtained via model averaging all the income distributions, are also reported in this Table, along with the posterior probabilities used in the averaging process. The complete posterior pdf's, conditional and unconditional, are graphed in Figures 1 through 6. From Table 4 and the figures, we observe that:

1. Averaging is dominated by three distributions, the lognormal, the Burr3 and the beta2. The gamma and Weibull distributions are relatively poor fits and do not contribute at all to the averaging process. The Fisk and Singh-Maddala distributions have a marginal impact.
2. Our posterior information about the economic quantities of interest, particularly modal income and the Gini coefficient, depends heavily on the income distribution assumption. For example, using a beta2 distribution, we find that modal income lies somewhere between \$275 and \$320 with a probability close to one. With a Fisk distribution, it lies between \$305 and \$360 with a probability close to one. A lognormal distribution suggests the Gini coefficient will lie between 0.27 and 0.32; its possible range implied by a Burr3 distribution is 0.29 to 0.36. The implied knowledge about mean income is fairly insensitive to the distribution, except when the Burr3 and Singh-Maddala distributions are used. The Burr3 exception is an important one, however, because it is one of the best fitting distributions.

3. A comparison of an averaged posterior pdf (for μ, m_o or γ) with the conditional posterior pdf of the best fitting distribution indicates the extent by which estimation precision is overstated by following a strategy of picking the best distribution and ignoring the preliminary test implications of discarding the remaining distributions. The best fitting 3-parameter distributions is the Burr3; the best fitting 2-parameter distribution is the lognormal. Both the mean and the spread of the average pdf's, particularly the spread, are quite different from those for the Burr3 and lognormal distributions. For example, conditional on the lognormal distribution, one would conclude that mean income lies between $445 and $500 with probability close to one. The average posterior pdf for mean income suggests it lies between $440 and $520. If the Burr3 distribution is used, one would conclude the Gini coefficient lies between 0.29 and 0.36. After averaging the distributions, the inferred range is 0.27 to 0.36.

6. Concluding Remarks

When a particular income distribution model is chosen to make inferences about quantities such as mean income or the level of inequality; as measured by the Gini coefficient, the inferences drawn are conditional on the model that is selected. Different models can lean to quite different conclusions. Choosing the best-fitting model from a number of alternative models helps reduce the chance of making mistaken inferences, but, because this strategy typically ignores discarded models, it overstates the precision with which economic quantities of interest are estimated. In this paper we have described and illustrated a Bayesian model averaging procedure that solves these problems.

References

Australian Bureau of Statistics (1999), *Income Distribution 1997-98*, Document No. 6523.0, Canberra.

Bordley, R.F., J.B. McDonald and A. Mantrala (1996), "Something New, Something Old: Parametric Models for the Size Distribution of Income", *Journal of Income Distribution* 6, 91-103.

Creedy, J., J.N. Lye and V.L. Martin (1997), " A Model of Income Distribution", in J. Creedy and V.L. Martin, editors, *Nonlinear Economic Models: Cross-sectional, Time Series and Neural Network Applications*, Cheltenham: Edward Elgar, 29-45.

Dagum, C. (1996), " A Systematic Approach to the Genreration of Income Distribution Models", *Journal of Income Distribution* 6, 105-126.

Danilov, D.L. and J.R. Magnus (2001), "On the Harm that Pretesting Does", CentER Discussion Paper No. 2001-37, Tilberg University.

Geweke, J. (1999), "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication, *Econometric Reviews* 18, 1-74.

Gelfand, A.E. and D.K. Dey (1994), "Bayesian Model Choice: Asymptotics and Exact Calculations", *Journal of the Royal Statistical Society (Series B*) 56, 501-514.

Griffiths, W.E. and D. Chotikapanich (1997), "Bayesian Methodology for Imposing Inequality Constraints on a Linear Expenditure Function with Demographic Factors", *Australian Economic Papers* 36, 321-341.

Judge, G.G. and M.E. Bock (1978), *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*, Amsterdam: North-Holland.

McDonald, J.D. (1984), "Some Generalized Functions for the Size Distribution of Income", *Econometrica* 52, 647-664.

Appendix: Characteristics of Income Distributions

1. Weibull Distribution

cdf: $F(y;\, a,\beta) = 1 - e^{-(y/\beta)^a}$ $\quad a > 1,\ \beta > 0$

pdf: $f(y;\, a,\beta) = \dfrac{a\, y^{a-1} e^{-(y/\beta)^a}}{\beta^a}$

The mean: $\mu = \beta\ \Gamma\left(1 + \dfrac{1}{a}\right).$ $\quad$ The mode: $m_o = \left(\dfrac{(a-1)}{a}\ \beta^a\right)^{1/a}$

Gini coefficient: $\gamma = 1 - \dfrac{1}{2^{1/a}}$

The derivatives:

$$\frac{\partial\mu}{\partial a} = -\frac{\beta}{a^2}\ \Gamma'\left(1 + \frac{1}{a}\right) \qquad \frac{\partial\gamma}{\partial a} = \frac{-\log(2)}{2^{1/a} a^2}$$

$$\frac{\partial\mu}{\partial\beta} = \Gamma\left(1 + \frac{1}{a}\right) \qquad \frac{\partial\gamma}{\partial\beta} = 0$$

2. Fisk Distribution

cdf: $F(y;\, a,b) = 1 - \dfrac{1}{(1 + (y/b)^a)}$ $\quad a > 0,\ \beta > 0$

pdf: $f(y;\, a,b) = \dfrac{a\, y^{a-1}}{b^a (1 + (y/b)^a)^2}$

The mean: $\mu = b\ \Gamma(1 + 1/a)\ \Gamma(1 - 1/a)\,.$ $\quad$ The mode: $m_o = \left(\dfrac{a-1}{a+1}\right)^{1/a} b\,.$

Gini coefficient: $\gamma = \dfrac{1}{a}$

The derivatives:

$$\frac{\partial\mu}{\partial a} = \frac{b}{a^2}\big(\Gamma(1+1/a)\,\Gamma'(1-1/a) - \Gamma(1-1/a)\,\Gamma'(1+1/a)\big)$$

$$\frac{\partial\mu}{\partial b} = \Gamma(1+1/a)\,\Gamma(1-1/a)$$

$$\frac{\partial\gamma}{\partial a} = -\frac{1}{a^2} \qquad \frac{\partial\gamma}{\partial b} = 0$$

3. Gamma Distribution

cdf: $F(y;\beta,p) = \dfrac{1}{\beta^p\Gamma(p)}\int_0^y t^{p-1}e^{-(t/\beta)}dt = G(y;\beta,p)$

pdf: $f(y;\beta,p) = \dfrac{y^{p-1}e^{-(y/\beta)}}{\beta^p\,\Gamma(p)}$ $\qquad \beta > 0,\ p > 1$

The mean $\mu = \beta p$. The mode $m_o = (p-1)\beta$.

Gini coefficient: $\gamma = \dfrac{\Gamma(p+1/2)}{\Gamma(p+1)\sqrt{\pi}}$

The derivatives:

$$\frac{\partial\mu}{\partial\beta} = p \qquad \frac{\partial\gamma}{\partial\beta} = 0$$

$$\frac{\partial\mu}{\partial p} = \beta \qquad \frac{\partial\log\gamma}{\partial p} = \psi(p+1/2) - \psi(p+1)$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the "digamma" function.

4. Lognormal Distribution

cdf: $\Lambda(y;\mu_n,\sigma_n) = \Phi\left(\dfrac{\log y - \mu_n}{\sigma_n}\right)$ $\qquad \sigma_n > 0$

pdf $f(y;\mu_n,\sigma_n) = \dfrac{1}{\sqrt{2\pi}\,\sigma_n y}\exp\left\{-\dfrac{1}{2\sigma_n^2}(\log y - \mu_n)^2\right\}$

The mean: $\mu = e^{\mu_n + \sigma_n^2/2}$. The mode $m_o = e^{\mu_n - \sigma_n^2}$.

Gini coefficient: $\gamma = 2\Phi\left(\dfrac{\sigma_n}{\sqrt{2}}\right) - 1$

The derivatives

$$\frac{\partial\mu}{\partial\mu_n} = e^{\mu_n+\sigma_n^2/2} \qquad \frac{\partial\gamma}{\partial\mu_n} = 0$$

$$\frac{\partial\mu}{\partial\sigma_n} = \sigma_n e^{\mu_n+\sigma_n^2/2} \qquad \frac{\partial\gamma}{\partial\sigma_n} = \sqrt{2}\,\phi\left(\frac{\sigma_n}{\sqrt{2}}\right)$$

5. Singh-Maddala Distribution

cdf: $F(y;a,b,q) = 1 - \dfrac{1}{\left(1+(y/b)^a\right)^q}$ $\quad a>0,\ b>0,\ q>\dfrac{1}{a}$

pdf: $f(y;a,b,q) = \dfrac{q\,a\,y^{a-1}}{b^a\left(1+(y/b)^a\right)^{1+q}}$

The mean: $\mu = \dfrac{b\,\Gamma(1/a)\,\Gamma(q-1/a)}{a\,\Gamma(q)}$. The mode: $m_o = \left[\dfrac{(a-1)}{aq+1}\right]^{1/a} b$.

Gini coefficient: $\gamma = 1 - \dfrac{\Gamma(q)\,\Gamma(2q-1/a)}{\Gamma(q-1/a)\,\Gamma(2q)}$

The derivatives

$$\frac{\partial \log \mu}{\partial a} = \frac{1}{a^2}\left[\psi(q-1/a) - \psi(1/a) - a\right] \qquad \frac{\partial \log \mu}{\partial b} = \frac{1}{b}$$

$$\frac{\partial \log \mu}{\partial q} = \psi(q-1/a) - \psi(q)$$

$$\frac{\partial \log m_o}{\partial a} = \frac{1}{a^2}\left[\log(aq+1) - \log(a-1)\right] + \frac{q+1}{a(a-1)(aq+1)}$$

$$\frac{\partial \log m_o}{\partial q} = \frac{-1}{aq+1} \qquad \frac{\partial \log m_o}{\partial b} = \frac{1}{b}$$

$$\frac{\partial \log(1-\gamma)}{\partial q} = \psi(q) + 2\psi(2q-1/a) - \psi(q-1/a) - 2\psi(2q)$$

$$\frac{\partial \log(1-\gamma)}{\partial a} = \frac{1}{a^2}\left[\psi(2q-1/a) - \psi(q-1/a)\right] \qquad \frac{\partial \log(1-\gamma)}{\partial b} = 0$$

6. Beta2 Distribution

pdf: $$f(y;b,p,q) = \frac{y^{p-1}}{b^p B(p,q)(1+y/b)^{p+q}} \qquad b>0,\ p>1, q>1$$

cdf: $$F(y;b,p,q) = \frac{1}{B(p,q)}\int_0^{(y/b)/[1+(y/b)]} t^{p-1}(1-t)^{q-1}dt = B\left[\left(\frac{y/b}{1+(y/b)}\right), p, q\right]$$

The mean: $\mu = \frac{bp}{q-1}$. The mode: $m_o = \frac{(p-1)b}{q+1}$.

Gini coefficient: $$\gamma = \frac{2B(2p,2q-1)}{pB^2(p,q)}$$

Derivatives:

$$\frac{\partial\mu}{\partial b} = \frac{p}{q-1} \qquad \frac{\partial m_o}{\partial b} = \frac{p-1}{q+1}$$

$$\frac{\partial\mu}{\partial p} = \frac{b}{q-1} \qquad \frac{\partial m_o}{\partial p} = \frac{b}{q+1}$$

$$\frac{\partial\mu}{\partial q} = \frac{-bp}{(q-1)^2} \qquad \frac{\partial m_o}{\partial q} = \frac{-(p-1)b}{(q+1)^2}$$

$$\frac{\partial \log\gamma}{\partial p} = 2\psi(2p)+2\psi(p+q)-\frac{1}{p}-2\psi(2p+2q-1)-2\psi(p)$$

$$\frac{\partial \log\gamma}{\partial q} = 2\left[\psi(2q-1)+\psi(p+q)-\psi(2p+2q-1)-\psi(q)\right]$$

$$\frac{\partial \log\gamma}{\partial b} = 0$$

7. Burr3 Distribution

pdf: $$f\left(y;a,b,p\right) = \frac{a\, y^{ap-1} p}{b^{ap}\left[1+\left(y/b\right)^{a}\right]^{p+1}} \qquad b>0,\ a>1,\ p>\frac{1}{a}$$

cdf: $$F\left(y;a,b,p\right) = \frac{1}{\left[\left(b/y\right)^{a}+1\right]^{p}}$$

The mean: $$\mu = b\frac{\Gamma\left(p+1/a\right)\Gamma\left(1-1/a\right)}{\Gamma\left(p\right)}$$

The mode $$m_o = \left[\frac{(ap-1)}{(a+1)}\right]^{1/a} b$$

Gini coefficient: $$\gamma = \frac{\Gamma(2p+1/a)\Gamma\left(p\right)}{\Gamma\left(p+1/a\right)\Gamma\left(2p\right)} - 1$$

Derivatives

$$\frac{\partial \log \mu}{\partial b} = \frac{1}{b} \qquad \frac{\partial \log \mu}{\partial p} = \psi\left(p+1/a\right)-\psi\left(p\right)$$

$$\frac{\partial \log \mu}{\partial a} = \frac{1}{a^2}\left[\psi\left(1-1/a\right)-\psi\left(p+1/a\right)\right]$$

$$\frac{\partial \log m_o}{\partial b} = \frac{1}{b} \qquad \frac{\partial \log m_o}{\partial p} = \frac{1}{ap-1}$$

$$\frac{\partial \log m_o}{\partial a} = \frac{1}{a^2}\left[\log(a+1)-\log(ap-1)\right]+\frac{(p+1)}{a(ap-1)(a+1)}$$

$$\frac{\partial \log(\gamma+1)}{\partial b} = 0 \qquad \frac{\partial \log(\gamma+1)}{\partial a} = \frac{1}{a^2}\left(\psi\left(p+1/a\right)-\psi\left(2p+1/a\right)\right)$$

$$\frac{\partial \log(\gamma+1)}{\partial p} = 2\psi\left(2p+1/a\right)+\psi\left(p\right)-\psi\left(p+1/a\right)-2\psi(2p)$$

Table 3: ML Estimates and Standard Errors and Posterior Means and Standard Deviations

	ML	Bayes
Gamma		
β	130.3082	131.03
	(8.0346)	(8.1618)
p	3.6155	3.6074
	(0.2046)	(0.2090)
log l'hood	-1244.24	-1251.349
Weibull		
a	1.8263	1.8235
	(0.0553)	(0.0560)
b	532.52	532.97
	(13.066)	(13.095)
log l'hood	-1281.457	-1288.536
Fisk		
a	3.3104	3.3033
	(0.1173)	(0.1189)
b	401.17	401.39
	(8.9892)	(9.0280)
log l'hood	-1223.464	-1230.272
Lognormal		
μ_n	6.0123	6.0121
	(0.0222)	(0.0221)
σ_n	0.5264	0.5276
	(0.0065)	(0.0162)
log l'hood	-1220.179	-1227.208

	ML	Bayes
Singh-M		
b	315.14	316.45
	(21.754)	(22.041)
a	4.3012	4.3420
	(0.4301)	(0.4248)
q	0.5604	0.5675
	(0.1024)	(0.1041)
log l'hood	-1219.092	-1229.634
Burr		
b	278.40	280.66
	(33.498)	(35.369)
a	2.7690	2.7771
	(0.1576)	(0.1589)
q	2.1137	2.1714
	(0.4840)	(0.5724)
log l'hood	-1216.487	-1226.880
Beta2		
b	94.2454	84.902
	(32.3546)	(34.396)
p	19.7376	24.104
	(7.0918)	(7.3725)
q	4.9260	4.8216
	(0.4296)	(0.4489)
log l'hood	-1216.019	-1226.993

Table 4: Posterior Means and Standard Deviations of the Means, Modes and Gini's

	Mean	Mode	Gini	Posterior Prob.
Gamma	471.08 (10.427)	340.05 (10.660)	0.2873 (0.0078)	0.0000
Weibull	473.76 (11.461)	344.23 (16.383)	0.3164 (0.0080)	0.0000
Fisk	469.30 (12.032)	331.96 (8.7104)	0.3031 (0.0109)	0.0163
Lognormal	469.53 (11.041)	309.17 (8.7257)	0.2909 (0.0085)	0.3499
Singh-Maddala	494,86 (19.677)	312.40 (9.8411)	0.3364 (0.0201)	0.0262
Burr	486.62 (15.391)	305.41 (10.084)	0.3218 (0.0136)	0.3238
Beta2	475.78 (12.755)	296.88 (8.2004)	0.3038 (0.0108)	0.2892
Average	477.30 (16.722)	304.77 (11.789)	0.3057 (0.0181)	

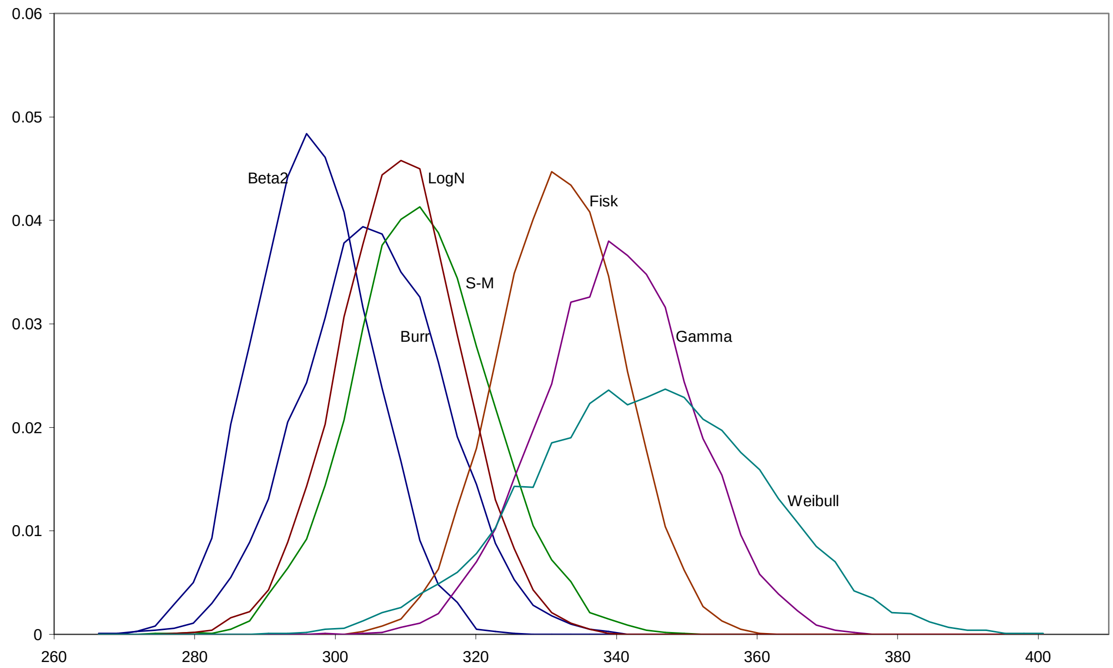

Figure 1: Posterior Densities for the Mode of Income

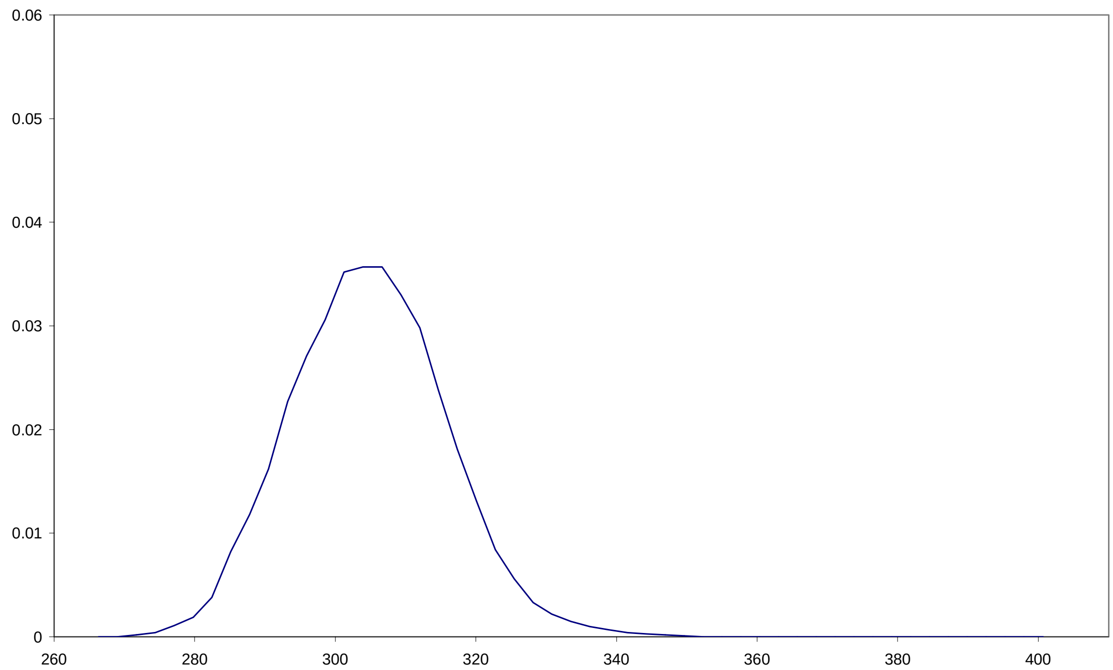

Figure 2: Average Posterior Density for the Mode

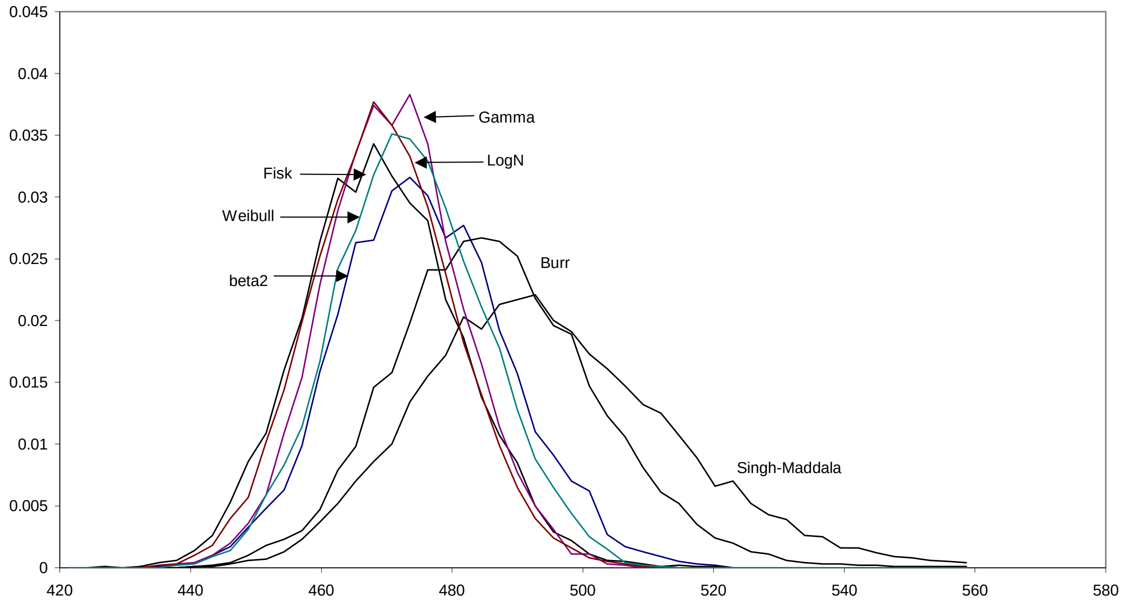

Figure 3: Posterior Densities for the Mean of Income

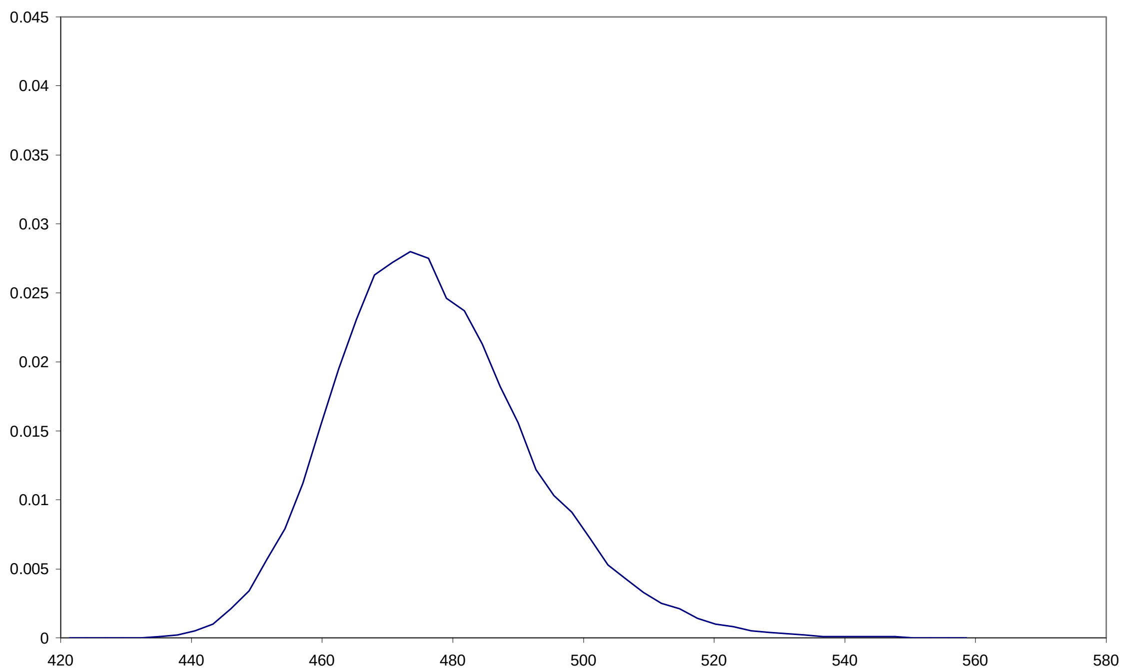

Figure 4: Average Posterior Density for the Mean

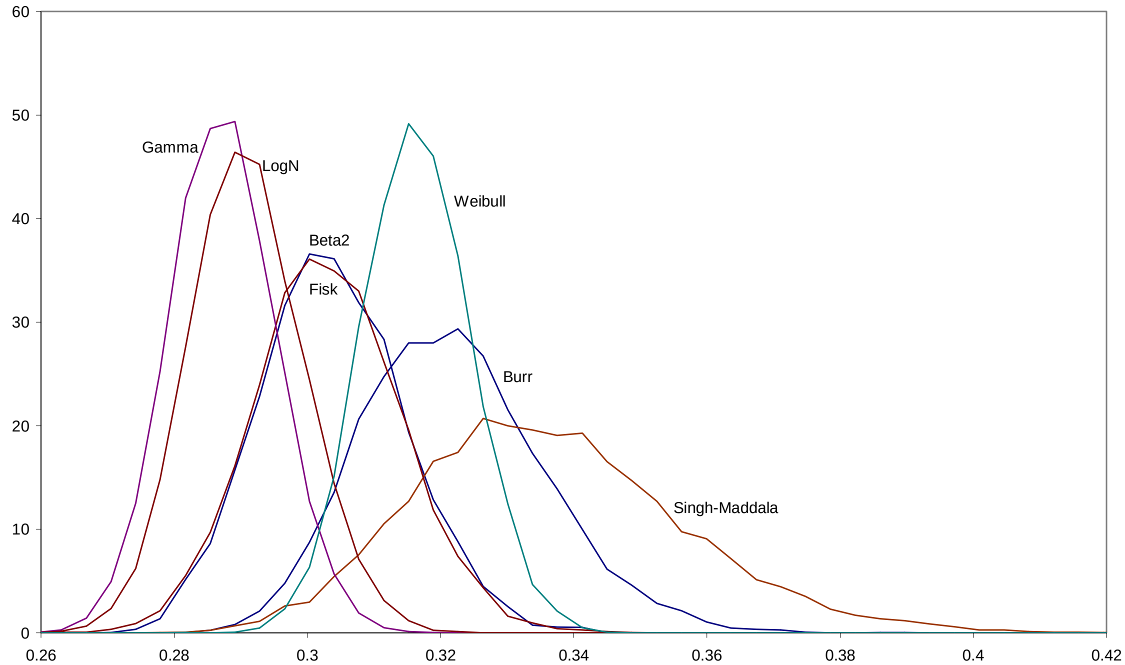

Figure 5: Posterior Densities for the Gini

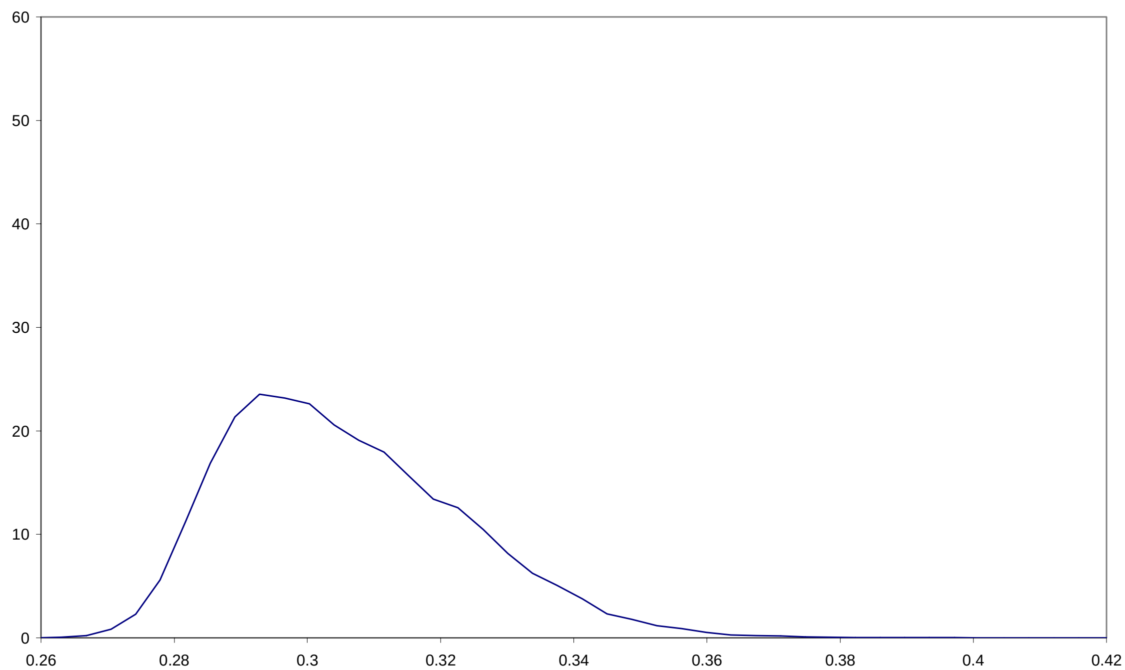

Figure 6: Average Posterior Density for the Gini