Bayesian Inference for Health Inequality and Welfare Using Qualitative Data

David Gunawan

University of New South Wales

William E. Griffiths

University of Melbourne

Duangkamon Chotikapanich

Monash University

June 26, 2017

Abstract

We show how to use Bayesian inference to compare two ordinal categorical distributions commonly occurring with data on self-reported health status. Procedures for computing probabilities for first and second order stochastic dominance and equality or S-dominance are described, along with methodology for obtaining posterior densities for health inequality indices. The techniques are applied to four years of data on Australian self-reported health status.

JEL classification numbers: C11, I14, I31.

Keywords: Dominance probabilities; ordinal data; inequality indices.

Corresponding Author
William Griffiths
Department of Economics
University of Melbourne 3010
Australia
Email: wegrif@unimelb.edu.au

1. Introduction

We are concerned with Bayesian inference for measuring welfare improvements based on ordinal qualitative data. We focus particularly on self-reported health although our proposed methods can be used for other dimensions of welfare that employ similar data such as perceptions of happiness or well-being. Welfare measures that we consider are:

1. First and second order stochastic dominance.
2. Mean health based on a specified cardinal measure.
3. A spread dominance concept introduced by Allison and Foster (2004).
4. A measure of inequality proposed by Abul Naga and Yalcin (2008).
5. A measure of inequality proposed by Cowell and Flachaire (2017).

In Section 2 we describe each of these measures. Bayesian inference procedures for estimating the measures are outlined in Section 3. An example using data on Australian health distributions is provided in Section 4 and some conclusions are drawn in Section 5.

2. Welfare and Inequality Measures

Consider an ordinal categorical distribution represented by a random variable X that takes discrete values $X = \{1, 2, \ldots, k\}$. The population proportion of individuals in category j is given by $p_X(j)$, with $0 \le p_X(j) \le 1$, and $\sum_{j=1}^{k} p_X(j) = 1$. The distribution function will be denoted by $F_X(j) = \sum_{i=1}^{j} p_X(i)$. Apart from being ordered, the values of X are arbitrary with no cardinal values. An example is self-assessed health on a scale where $X = 1$ for poor, $X = 2$ for fair, $X = 3$ for good, $X = 4$ for very good and $X = 5$ for excellent. Our focus is on univariate distributions such as health or well-being. There is a large literature on bivariate and multivariate distributions which relate variables such as health to income and other socio-economic measures. See, for example, many of the papers in Rasa Dias and O' Donald (2013).

First and second order stochastic dominance are concerned with whether one distribution represents a higher level of welfare or an improvement in welfare relative to another. Given any two well-being distributions, X and Y, with the same number of categories k, we say X first order stochastic dominates Y, written as, $X >_{FSD} Y$, if and only if

$$F_X(j) \le F_Y(j) \text{ for all } j \qquad \text{and } F_X(j) < F_Y(j) \text{ for some } j \tag{1}$$

where $j = 1, 2, \ldots, k-1$. Here, distribution X has a lower percentage of its population in the lowest j categories, than distribution Y, and hence a higher average level of well-being for any increasing scale. Only $k-1$ categories need to be compared since the cumulative proportions of populations at the k-th category $F_X(k)$ and $F_Y(k)$ are always equal to one. This condition for first order stochastic dominance was described by Allison and Foster (2004) and Yalonetzky (2013). Within the context of welfare functions that are additively separable and symmetric with respect to individuals, Yalonetzky shows that first order stochastic dominance implies a welfare improvement for the class of individual welfare functions that are monotonically increasing. Also, if $X >_{FSD} Y$, then the mean of X is greater than the mean of Y for all cardinal scalings of the categories. Given sample observations from distributions X and Y, we describe how to use Bayesian inference to find posterior probabilities $\Pr(X >_{FSD} Y)$, $\Pr(Y >_{FSD} X)$, and Pr(neither X nor Y is dominant).

Conditions for second order stochastic dominance of distribution X over Y, written as $X >_{SSD} Y$, are

$$\sum_{i=1}^{j} F_X(i) \le \sum_{i=1}^{j} F_Y(i) \quad \text{for all } j \qquad \text{and } \sum_{i=1}^{j} F_X(i) < \sum_{i=1}^{j} F_Y(i) \text{ for some } j \tag{2}$$

for $j = 1, 2, \ldots, k-1$. These conditions are presented by Yalonetzky (2013) who shows that $X >_{SSD} Y$ implies greater welfare for distribution X relative to distribution Y for the class of individual welfare functions which are both monotonically increasing and concave – welfare increases between successively higher ordinal categories are smaller. Following similar procedures to those for first order stochastic dominance, we explore how to find posterior probabilities for $\Pr\left(X >_{SSD} Y\right)$, $\Pr\left(Y >_{SSD} X\right)$, and Pr(neither X nor Y is SSD).

Computing mean health from an ordinal categorical distribution requires the assignment of a cardinal scale to the categories, and, as Allison and Foster (2004) point out, comparisons can be very sensitive to that scale. Nevertheless, mean health for a specified cardinal scale is of interest to a number of researchers. Some examples appear in van Doorslaer and Jones (2013) who compare a number of alternative procedures for imposing cardinality on ordered responses. Given a set of scaling factors $c_1, c_2, c_3, \ldots, c_k$, we describe how to compute the posterior mean and standard deviation for the mean $\mu = \sum_{j=1}^{k} p_X(j) c_j x_j$.

Recognising that the mean is no longer a suitable reference point for comparing the extent of health inequality in two ordinal categorical distributions, Allison and Foster (2004) introduce a spread dominance criterion based on the median as a reference point. Because of a potential ambiguity about whether SD refers to stochastic dominance or spread dominance, we will refer to this concept as equality dominance. It can be viewed as an alternative to Lorenz dominance for data that are not continuous; it is limited to comparing distributions that have the same median category. We say X equality dominates Y (distribution Y has a greater spread than distribution X), written $X >_{ED} Y$ if

- Distribution X and Y have the same median category m
- $F_X\left(j\right) \leq F_Y\left(j\right)$ for all categories of $j < m$, (3)
- $F_X\left(j\right) \geq F_Y\left(j\right)$ for all categories of $j \geq m$.

As with the stochastic dominance criteria, we examine how to compute posterior probabilities for equality dominance in either direction, and the posterior probability that there is no equality dominance.

Equality dominance cannot provide a complete ordering of distributions based on their level of health inequality; it is only useful for distributions with the same median and even then neither one of two distributions may dominate. To overcome this deficiency aggregate indices of inequality have been developed paralleling the use of a Gini index as an alternative to Lorenz dominance for continuous distributions. We consider the posterior distributions of two indices, those suggested by Abul Naga and Yalcin (2008) and Cowell and Flachaire (2017). Abul Naga and Yalcin develop an inequality index based on axioms of continuity, scale invariance, normalisation, and an aversion to greater inequality in the sense introduced by Allison and Foster (2004). Their measure is

$$I_X(\alpha,\beta)=\frac{\sum_{j<m}\left(F_X(j)\right)^{\alpha}-\sum_{j\geq m}\left(F_X(j)\right)^{\beta}+(k+1-m)}{(m-1)(0.5)^{\alpha}-\left(1+(k-m)(0.5)^{\beta}\right)+(k+1-m)} \tag{4}$$

where α and β reflect value judgements of society, chosen by the analyst. When $\alpha=\beta=1$, the cumulative distributions for the lower half and upper half are given equal weight in the overall inequality. For a given value of β, as $\alpha\to\infty$, less weight is given to the inequality below the median, while for a given value of α and as $\beta\to\infty$, less weight is given to inequality above the median. When everyone is in the median category $I_X(1,1)=0$ and when half of the population is in the lowest category and the other half in the highest category, then $I_X(1,1)=1$.

Cowell and Flachaire (2017) consider inequality from the standpoint of an individual's perception of their status in the health distribution. Indices based on "downward looking" and "upward looking" assessment of health status and which satisfy several desirable axioms are developed. Costa Font and Cowell (2013) use these measures to compute health inequality for several countries, and tentatively recommend the upward looking index

$$U_X(\alpha) = \begin{cases} \dfrac{1}{\alpha(\alpha-1)}\left[\sum_{j=1}^{k} p_X(j)\left[s_X(j)\right]^{\alpha} - 1\right] & \text{for } \alpha < 1 \text{ and } \alpha \neq 0 \\ -\sum_{j=1}^{k} p_X(j) \log s_X(j) & \text{for } \alpha = 0 \end{cases} \tag{5}$$

where $s_X(j) = \sum_{i=j}^{K} p_X(i) = 1 - F_X(j) + p_X(j)$. With perfect inequality where $s_X(j) = 1$ for everybody, $U_X(\alpha) = 0$. Otherwise, $U_X(\alpha) > 0$. The parameter α governs the sensitivity of the index to particular parts of the distribution. We are interested in finding posterior densities for $I_X(\alpha,\beta)$ and $U_X(\alpha)$, given a sample of observations.

3. Bayesian Estimation

Let $p_j = p_X(j),\ j = 1,2,\ldots,k$ denote the population proportions in each category for a distribution of interest. The objective is to use a random sample of size n to estimate $\boldsymbol{p} = (p_1, p_2, \ldots, p_k)'$. Let $\boldsymbol{n} = (n_1, n_2, \ldots, n_k)'$ denote the number of sample observations in each category. The distribution for $\boldsymbol{n}$ is multinomial with density

$$f(\boldsymbol{n} \mid \boldsymbol{p}) \propto p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k} \tag{6}$$

with the restriction $\sum_{j=1}^{k} n_j = n.$ A non-informative prior for $\boldsymbol{p}$ is (Gelman et al 1995)

$$f(\boldsymbol{p}) \propto p_1^{-1} p_2^{-1} \ldots p_k^{-1}$$

The combination of this prior pdf with the multinomial likelihood function yields the Dirichlet posterior density

$$f(\boldsymbol{p} \mid \boldsymbol{n}) \propto f(\boldsymbol{n} \mid \boldsymbol{p})\, f(\boldsymbol{p}) \propto p_1^{n_1-1} p_2^{n_2-1} \cdots p_k^{n_k-1} \tag{7}$$

This density can be used to make inferences about the various measures described in the preceeding section. For FSD, SSD and ED we take a large number of draws of $\boldsymbol{p}$ from each of the two posterior distributions, that conditional on X and that conditional on Y, and count the number of draws of $\boldsymbol{p}_X$ and $\boldsymbol{p}_Y$ for which each of the dominance criteria is satisfied. The proportion of draws for which a criterion is satisfied is an estimate of the posterior probability of dominance. The proportion of draws for which dominance in either direction is not satisfied is an estimate of the posterior probability that neither distribution dominates.[1]

For situations where interest centres on the mean $\mu = \sum_{j=1}^{k} c_j x_j p_j = \sum_{j=1}^{k} d_j p_j$ for a cardinal scaling c_j, and $d_j = c_j x_j$, we note that $E\left(p_i \mid \boldsymbol{n}\right) = n_j / n$ and the covariance matrix for $\boldsymbol{p}$ is given by

$$\operatorname{cov}\left(\boldsymbol{p} \mid \boldsymbol{n}\right) = \frac{1}{n^2\left(n+1\right)} \begin{bmatrix} n_1\left(n-n_1\right) & -n_1 n_2 & \cdots & -n_1 n_k \\ -n_2 n_1 & n_2\left(n-n_2\right) & \cdots & -n_2 n_k \\ \vdots & \vdots & \ddots & \vdots \\ -n_k n_1 & -n_k n_2 & \cdots & n_k\left(n-n_k\right) \end{bmatrix}$$

Hence the posterior mean and variance for $\mu = \boldsymbol{d}'\boldsymbol{p}$, where $\boldsymbol{d}' = \left(d_1, d_2, \ldots, d_k\right)$, are given by

$$E\left(\mu \mid \boldsymbol{n}\right) = \sum_{j=1}^{k} d_j \frac{n_j}{n}$$

$$\operatorname{var}\left(\mu \mid \boldsymbol{n}\right) = \boldsymbol{d}' \operatorname{cov}\left(\boldsymbol{p}\right) \boldsymbol{d}$$

Closed form expression for posterior means and variances of $I_X\left(\alpha,\beta\right)$ and $U_X\left(\alpha\right)$ are not available, but, using the sample of draws of $\boldsymbol{p}$ from its Dirichlet posterior density, they can be estimated, along with their marginal posterior densities.

4. Example

The data used to illustrate the methodology are the Self Reported Health Status (SRHS) obtained from the Household, Income, and Labour Dynamics in Australia (HILDA) survey for the years 2002, 2005, 2008, and 2010. The HILDA survey is a national representative

[1] Chotikapanich and Griffiths (2006) use a similar procedure to compute dominance probabilities for continuous distributions.

longitudinal survey, which began in Australia in 2001; it is designed, managed, and maintained by the Melbourne Institute of Applied Economic and Social Research, University of Melbourne. Individuals aged 15 years or above answer a question of the form: in general, would you say your health is poor, fair, good, very good or excellent? Sample sizes are between 11000 to 12000 for each year being considered and the median category is good for all years. Summary statistics for the data are presented in Table 1. They reveal negatively skewed distributions for all years.; there are more people in the better health categories and less people in the inferior health categories.

Table 1: Sample Sizes and Proportions for Categorical Self-Reported Health Status Distributions

	2002		2005		2008		2010	
Category	*n*	prop	*n*	prop	*n*	prop	*n*	prop
Poor	381	0.0323	362	0.0320	318	0.0286	359	0.0300
Fair	1626	0.1377	1644	0.1451	1576	0.1419	1722	0.1441
Good	4154	0.3519	4116	0.3633	4063	0.3658	4337	0.3628
Very good	4218	0.3573	4037	0.3563	3931	0.3539	4265	0.3568
excellence	1426	0.1208	1169	0.1032	1220	0.1098	1271	0.1063
Sample size	11805	1.0000	11328	1.0000	11108	1.0000	11954	1.0000

Table 2: Posterior means and standard deviations for the population proportions of the SRHS

Category		2002	2005	2008	2010
Poor	p_1	0.0323 (0.0016)	0.0320 (0.0017)	0.0286 (0.0016)	0.0300 (0.0016)
Fair	p_2	0.1377 (0.0032)	0.1451 (0.0033)	0.1419 (0.0033)	0.1441 (0.0032)
Good	p_3	0.3519 (0.0044)	0.3633 (0.0045)	0.3658 (0.0046)	0.3628 (0.0044)
Very good	p_4	0.3573 (0.0044)	0.3564 (0.0045)	0.3539 (0.0045)	0.3568 (0.0044)
excellence	p_5	0.1208 (0.0030)	0.1032 (0.0029)	0.1098 (0.0030)	0.1063 (0.0028)

Posterior means and standard deviations for the population proportions for all the years are given in Table 2. Since $E\left(p_i \mid \boldsymbol{n}\right) = n_i/n$, the posterior means are the same as the sample proportions. The posterior standard deviations, given by $\left[n_i\left(n-n_i\right)\Big/\left(n^2\left(n+1\right)\right)\right]^{0.5}$, are relatively small, suggesting that the population proportions are well estimated. For estimating posterior probabilities of dominance and posterior densities of the inequality measures, we draw $M = 10000$ independent draws from the Dirichlet posterior distribution.

The posterior probabilities for FSD are reported in Table 3 for the 6 possible pairwise comparisons of the 4 years. The numbers in parentheses are numerical standard errors calculated as $\left[prob\left(1-prob\right)/10000\right]^{1/2}$. With the exception of the 2005/2008 comparison, the probabilities for no dominance are always greater than 0.5. There are, however, several probabilities of dominance that are moderately large in the sense that they are greater than 0.33, namely, $\left(2002 >_{FSD} 2005\right)$, $\left(2008 >_{FSD} 2005\right)$, $\left(2010 >_{FSD} 2005\right)$ and $\left(2008 >_{FSD} 2010\right)$. These results suggest that the self-assessed health distribution in 2005 was a poor one. There is also little evidence to suggest that the distribution has improved over time.

Table 3: Probabilities (numerical standard errors) for First Order Stochastic Dominance

	02 *A* 05 *B*	02 *A* 08 *B*	02 *A* 10 *B*	05 *A* 08 *B*	05 *A* 10 *B*	08 *A* 10 *B*
$\Pr\left(A >_D B\right)$	0.4274 (0.0016)	0.0449 (0.0007)	0.1503 (0.0011)	0.0014 (0.0001)	0.0171 (0.0004)	0.3306 (0.0015)
$\Pr\left(B >_D A\right)$	0.0000 (0.0000)	0.0004 (0.0000)	0.0000 (0.0000)	0.6345 (0.0015)	0.4136 (0.0016)	0.0237 (0.0005)
Pr(no dominance)	0.5726 (0.0016)	0.9547 (0.0007)	0.8497 (0.0011)	0.3641 (0.0015)	0.5693 (0.0016)	0.6457 (0.0015)

Because FSD implies SSD, the posterior probabilities for SSD will be at least as great as the corresponding ones for FSD. The posterior probabilities for SSD reported in Table 4 generally support the conclusions about the SRHS distributions that were drawn from the FSD results, but because the dominance probabilities are larger, there is more evidence of welfare improvements. Three of the pairwise comparisons whose FSD probabilities were greater than 0.33 all have SSD probabilities greater than 0.5.

Table 4: Probabilities (numerical standard errors) for Second Order Stochastic Dominance

	02 *A*	05 *B*	02 *A*	08 *B*	02 *A*	10 *B*	05 *A*	08 *B*	05 *A*	10 *B*	08 *A*	10 *B*
$\Pr(A >_D B)$	0.4330 (0.0016)		0.0473 (0.0007)		0.1534 (0.0011)		0.0108 (0.0003)		0.0607 (0.0008)		0.5506 (0.0016)	
$\Pr(B >_D A)$	0.0004 (0.0000)		0.0414 (0.0006)		0.0065 (0.0003)		0.8434 (0.0011)		0.6306 (0.0015)		0.0861 (0.0009)	
Pr(no dominance)	0.5666 (0.0016)		0.9113 (0.0009)		0.8401 (0.0012)		0.1458 (0.0011)		0.3087 (0.0015)		0.3633 (0.0015)	

The posterior means and standard deviations for mean health using a linear cardinal scale with $c_i = 1,\ i = 1, 2, \ldots, 5,$ are given in Table 5. The results are generally in line with those for stochastic dominance. Mean health was lowest in 2005, the year which had high probabilities of being dominated by the remaining years.

Table 5: Posterior means (standard deviations) for mean health

	Mean (linear scale)
2002	3.3966 (0.00897)
2005	3.3537 (0.0090)
2008	3.3744 (0.0090)
2010	3.3653 (0.0087)

Turning to the equality dominance probabilities in Table 6, we find little evidence of any dominance. $\Pr(2008 >_{ED} 2002) = 0.43$ and $\Pr(2010 >_{ED} 2002) = 0.19,$ but all other dominance probabilities are less than 0.07.

Table 6: Probabilities (numerical standard errors) for Equality Dominance

	02 05 A B	02 08 A B	02 10 A B	05 08 A B	05 10 A B	08 10 A B
$\Pr(A >_D B)$	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0083 (0.0003)	0.0371 (0.0006)	0.0616 (0.0008)
$\Pr(B >_D A)$	0.0635 (0.0008)	0.4349 (0.0016)	0.1864 (0.0012)	0.0208 (0.0005)	0.0462 (0.0007)	0.0266 (0.0005)
Pr(no dominance)	0.9365 (0.0008)	0.5651 (0.0016)	0.8136 (0.0012)	0.9709 (0.0005)	0.9167 (0.0009)	0.9118 (0.0009)

Posterior means and standard deviations for the Abul Naga/Yalcin and Cowell/Flachaire indices are presented in Table 7. For the Abul Naga/Yalcin index we used $(\alpha = 1, \beta = 1)$, $(\alpha = 1, \beta = 4)$ and $(\alpha = 4, \beta = 1)$; for the Cowell/Flachaire index we set $\alpha = 0.1$ and $\alpha = 0.9.$ Posterior densities for the Abul Naga/Yalcin index for $(\alpha = 1, \beta = 1)$ are graphed in Figure 1; for the Cowell/Flachaire index posterior densities were graphed for both settings of α and presented as Figure 2(a) and (b). The posterior means and the complete densities in the figures all lead to the same conclusion. Inequality is similar in 2005, 2008 and 2010, but greater in 2002. This conclusion is consistent with the equality dominance probabilities. There is some chance that 2002 is dominated by the other three years, but very little probability of dominance among the other three years.

Table 7: Inequality measures, 2002-2010

	Abul Naga and Yalcin Index			Cowell and Flachaire Index	
	$\alpha = 1, \beta = 1$	$\alpha = 1, \beta = 4$	$\alpha = 4, \beta = 1$	$\alpha = 0.1$	$\alpha = 0.9$
2002	0.4006 (0.00314)	0.5324 (0.00326)	0.5331 (0.00565)	0.6153 (0.00176)	3.7164 (0.01049)
2005	0.3859 (0.00313)	0.5137 (0.00337)	0.5011 (0.00561)	0.6122 (0.00195)	3.6866 (0.01099)
2008	0.3863 (0.00318)	0.5177 (0.00339)	0.5106 (0.00571)	0.6125 (0.00188)	3.6859 (0.01101)
2010	0.3868 (0.00306)	0.5158 (0.00325)	0.5070 (0.00547)	0.6121 (0.00187)	3.6832 (0.01066)

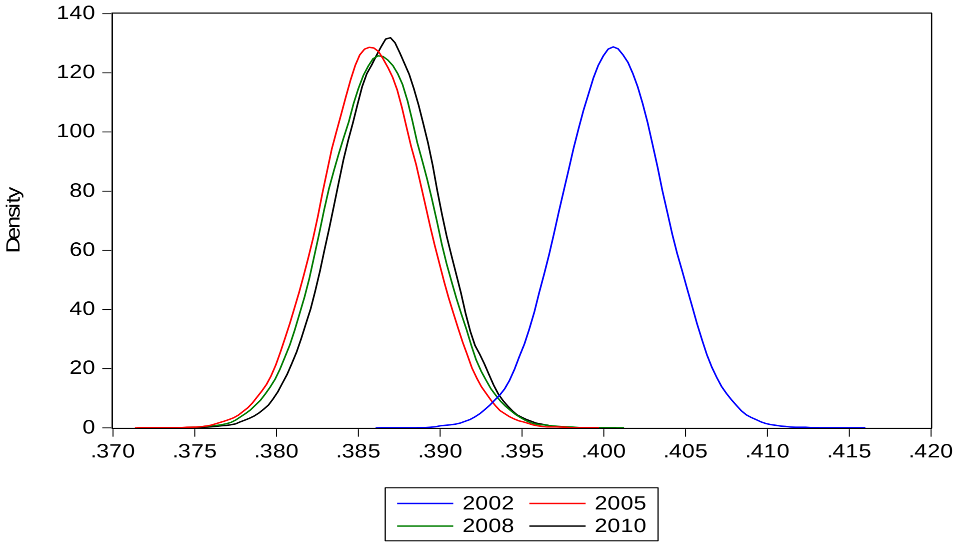

Figure 1 Posterior Densities for Abul Naga/Yalcin Index, $\alpha = 1, \beta = 1$

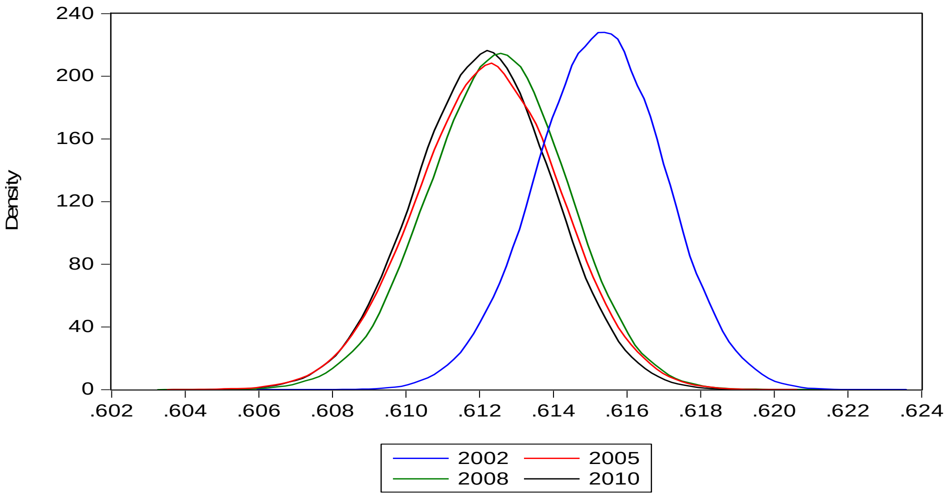

Figure 2(a) Posterior densities for Cowell/Flachaire Index $\alpha = 0.1$

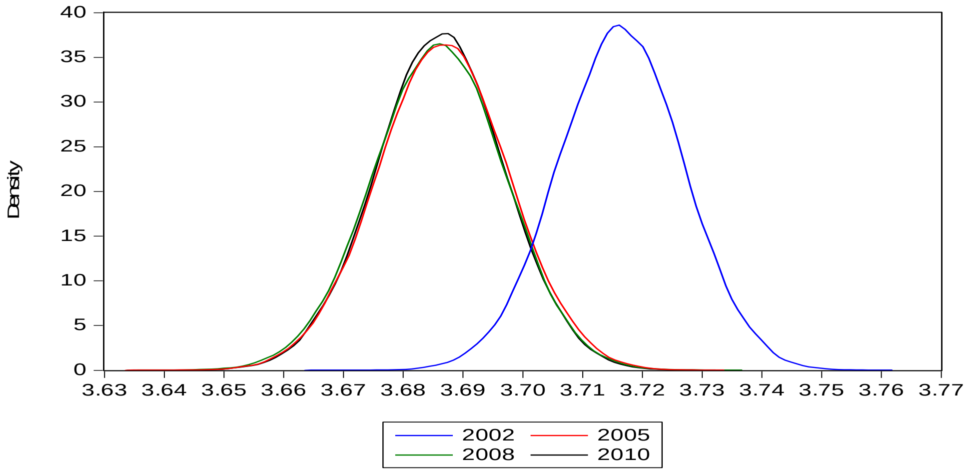

Figure 2(b) Posterior densities for Cowell/Flachaire Index $\alpha = 0.9$

5. Concluding Remarks

Various measures for comparing ordinal categorical distributions have been suggested in the literature for assessing improvements in welfare resulting from an increase in the level, and/or a decrease in inequality, of self-assessed health distributions. Bayesian inference is a convenient and straightforward method for providing probabilistic information on welfare improvements. We have demonstrated how to compute posterior probabilities of first and second order stochastic dominance, and equality dominance, and how to find posterior densities for indices designed to measure inequality.

References

Abul Naga, R. H. and Yalcin, T (2008), "Inequality Measurement for Ordered Response Health Data", *Journal of Health Economics*, 27, 1614-1625.

Allison, R. A. and Foster, J. E. (2004), "Measuring Health Inequality using Qualitative Data", *Journal of Health Economics*, 23, 505-524.

Chotikapanich, D. and W.E. Griffiths (2006), "Bayesian Assessment of Lorenz and Stochastic Dominance in Income Distributions", in J. Creedy and G.Kalb, editors, *Research on Economic Inequality, Volume 13: Dynamics of Inequality and Poverty*, Amsterdam: Elsevier, 297-321.

Costa Font, J. and Cowell, F. (2013), "Measuring Health Inequality with Categorical Data: Some Regional Patterns", In P. Rosa Dias and O. O'Donnell, editors, *Research on Economic Inequality, Volume 21: Health and Inequality,* Amsterdam: Elsevier, 53-76

Cowell, F. and Flachaire, E. (2017), "Inequality with Ordinal Data", *Economica*, 84, 290-321.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995), *Bayesian Data Analysis*, Chapman & Hall

Rosa Dias, P. and O. O'Donnell (2013), *Research on Economic Inequality, Volume 21: Health and Inequality,* Amsterdam: Elsevier, 53-76

van Doorslaer, E. and Jones, A.M. (2003), "Inequalities in self-reported health: validation of a new approach to measurement", *Journal of Health Economics*, 22, 61-87.

Yalonetzky, G. (2013), "Stochastic Dominance with Ordinal Variables: Conditions and a Test", *Econometric Reviews*, 32:1, 126-163.