

# Using a Discrete Population Distribution

Ray Adams. 9 July 2010

ConQuest 3 is able to implement either joint maximum likelihood (JML) or marginal maximum likelihood (MML) estimation. In the case of MML estimation either a Gaussian or a discrete distribution can be requested. This note is concerned with the ConQuest 3 implementation of MML when a discrete distribution is requested.

The ConQuest model normally requires the specification of two components, the item response model and the population model; the population model describes the assumed distribution of the underlying latent variable(s). If left unspecified a (multivariate) Gaussian distribution is assumed and the mean and variance of that distribution are estimated from the data. In some circumstances however the assumption of a Gaussian distribution for the latent variable may not be desirable. In this case either JML<sup>1</sup> or MML with a discrete distribution can be utilised (see Adams, Wilson and Wang (1997))

In the case of a discrete distribution a  $D$ -dimensional uniform grid of nodes is chosen and then the density at each of the nodes is estimated. For example in the case of a one-dimensional model with 15 nodes running from  $-2.5$  to  $+2.5$  logits the estimated shape of an example latent distribution is given in Figure 1. The height of each bar is the estimated density at each of the nodes. In this case note that the distribution is almost uniform in shape.

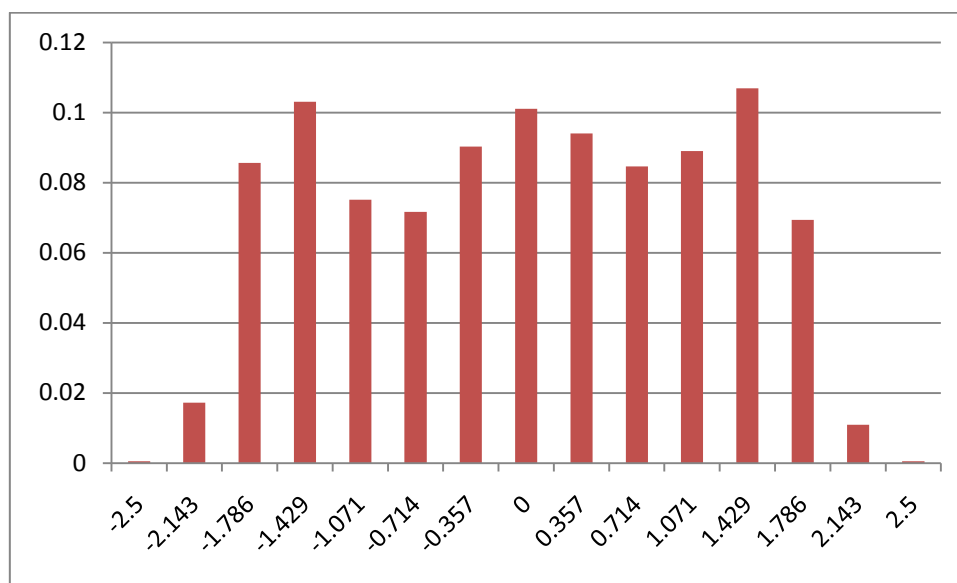
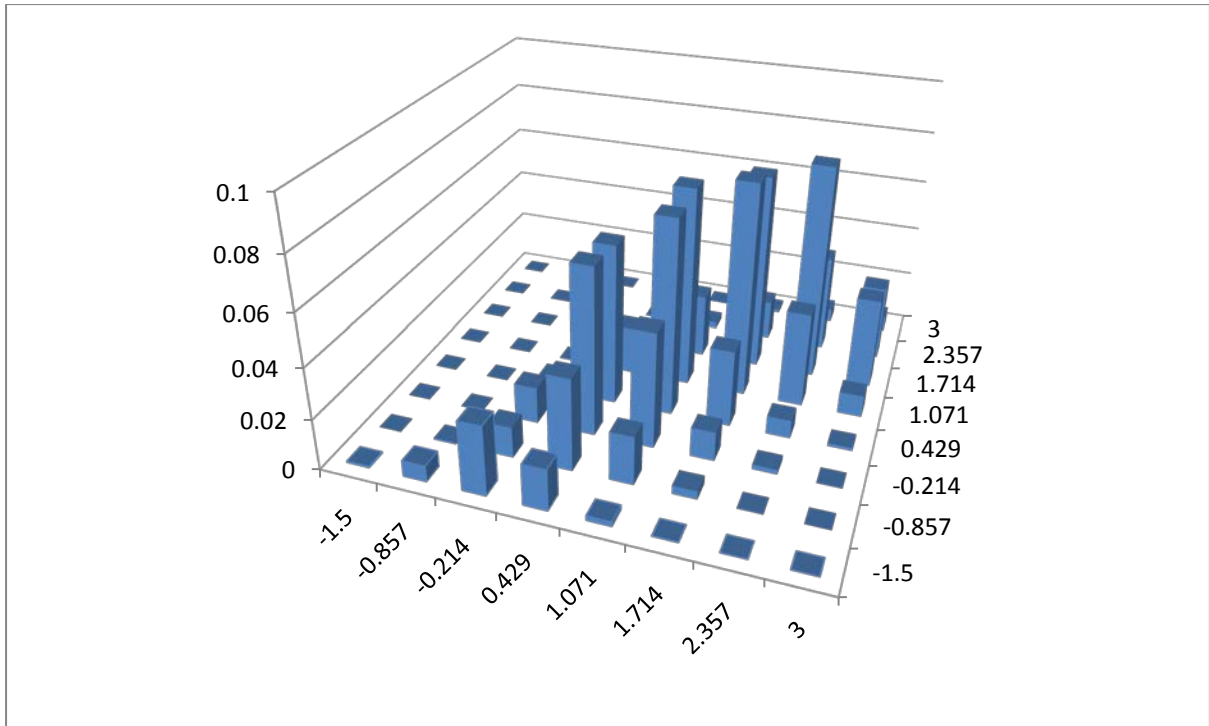


Figure 1: Example estimated discrete unidimensional latent distribution

An example in the case of a two-dimensional model with a grid of  $8 \times 8 = 64$  nodes running from  $-1.5$  to  $3.0$  on each dimension is shown in Figure 2.

<sup>1</sup> see ConQuest note 3



**Figure 2: Example estimated discrete two-dimensional latent distribution**

A discrete distribution assumption is requested through the use of the option `population=discrete` with the `estimate` command. Bock/Aitken quadrature is the estimation method that must be used with a discrete distribution.

### Controlling the Estimation

The `nodes=n1`, `minnode=n2` and `maxnode=n3` options of the `estimate` command can be used to control the number and range of the nodes. Note that the number of nodes is the number per dimension and as a consequence the number of actual nodes increases exponentially with the number of dimensions. For instance for the model shown in Figure 2, the options `nodes=8`, `minnode=-1.5` and `maxnode=3` were used so that 64 weights were estimated.

Users should explore the impact of choosing different numbers of nodes and node ranges on their results.

### Model Limitations for MML with a discrete distribution

A discrete distribution cannot be used in conjunction with conditioning variables – that is, the `regression` statement cannot be used. Further, all latent ability estimation is restricted to the use of WLEs and MLEs and as a consequence all plots and fit tests will rely upon the use of WLEs or MLEs

To obtain estimates of population characteristics (eg means, variances and percentiles) the `descriptive` command can be used. Note, however, that it is widely recognised that the estimates of population parameters derived from WLEs or MLEs can be flawed.

### **Output Implications**

When a discrete distribution is selected the output for the population model will consist of a list of the nodes and estimates for the weights at each node.

### **References**

Adams, R. J., Wilson, M. R., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*, 1 – 23.