

Parametric Representation of Scored Multinomial Distribution

Shin-ichi Mayekawa

Department of Human System Science,
Graduate School of Decision Science and Technology,
Tokyo Institute of Technology.

990704,10, 030524,29,30,31,0602,03,13,15,18,19,21,22,23,25,27,28,050713

1 Introduction

When the outcomes of a multinomial distribution are numbered, it is called a scored multinomial distribution. (Rosenbaum and Thayer,1987) The purpose of this note is to develop a general method to estimate the probabilities of a scored multinomial distribution as a function of its score.

2 Multinomial Distribution

Let Y be the random variable whose value is one of the elements in the following list: $\{A_1, A_2, \dots, A_m\}$. Using the dummy variable notation, the above random variable can be denoted by a $m \times 1$ vector \mathbf{Z} whose elements are

$$Z_j = \begin{cases} 1 & \text{if } Y = A_j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

This type of random variable is known to have the m-categorie multinomial distribution. Let us denote the parameter of this this distribution, namely, the probability of the event $Z_j = 1$ by $p_j, j = 1, 2, \dots, m$, satisfying the constraint

$$\sum_{j=1}^m p_j = 1, \quad j = 1, 2, \dots, m. \quad (2)$$

Denoting the result of n independent trials by Y_i or by $m \times 1$ vector $\mathbf{Z}_i, i = 1, 2, \dots, n$ whose elements consist of Z_{ij} , the frequency of each categorie can be calculated by tabulating the Y_i s or, using the dummy variable notation, as the sum

$$F_j = \sum_{i=1}^n Z_{ij}, \quad j = 1, 2, \dots, m. \quad (3)$$

Note that

$$\sum_{j=1}^m F_j = n. \quad (4)$$

This random variable, the sum of n independent multinomial trials, has the following joint probability function

$$Pr(F_1 = f_1, F_2 = f_2, \dots, F_m = f_m) = \frac{n!}{f_1! f_2! \dots f_m!} p_1^{f_1} p_2^{f_2} \dots p_m^{f_m} \quad (5)$$

and denoted as

$$\mathbf{F} \sim MN(n, \mathbf{p}) \quad (6)$$

where \mathbf{F} is the $m \times 1$ vector of the cell frequencies, and \mathbf{p} , the corresponding probabilities. The mean and the dispersion matrix of \mathbf{F} is known to be

$$\mathcal{E}(\mathbf{F}) = n\mathbf{p}, \quad \text{and} \quad \mathcal{D}(\mathbf{F}) = n(\text{Diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}'), \quad (7)$$

where $\text{Diag}(\mathbf{p})$ is the $m \times m$ diagonal matrix whose i -th element is p_i .

Note that the same data may be given as the frequency distribution of variable Y_i , $i = 1, 2, \dots, n$,

$$(A_j, f_j), \quad j = 1, 2, \dots, m, \quad (8)$$

where f_j is the frequency of the event $Y_i = A_j$ satisfying

$$\sum_{j=1}^m f_j = n. \quad (9)$$

Given the frequency of each category, the likelihood can be written as

$$L = \prod_{j=1}^m p_j^{f_j} \quad (10)$$

and the log likelihood,

$$\ln L = \sum_{j=1}^m f_j \ln p_j \quad (11)$$

3 Scored Multinomial Distribution

Suppose that the categories are scored, that is, each category A_j has an associated score, x_j , $j = 1, 2, \dots, m$, say, which can be vector valued. In this case, we may state that the discrete random variable Y has a scored multinomial distribution denoted as

$$Y \sim SMN(\mathbf{x}, \mathbf{p}) \quad (12)$$

where \mathbf{x} is the $m \times 1$ vector of scores, meaning that the frequency distribution of Y_i , $i = 1, 2, \dots, n$,

$$(x_j, f_j), \quad j = 1, 2, \dots, m, \quad \text{and} \quad \sum_{j=1}^m f_j = n \quad (13)$$

is distributed as $MN(n, \mathbf{p})$.

The mean and the variance of this random variable are, respectively,

$$\mathcal{E}(Y) = \sum_{j=1}^m x_j p_j, \quad \text{and} \quad \text{var}(Y) = \sum_{j=1}^m x_j^2 p_j - \left(\sum_{j=1}^m x_j p_j \right)^2. \quad (14)$$

For the scored multinomial distribution, it is natural to assume that the score and the probability are related. That is, we may assume that each probability is proportional to some positive function of the score,

$$p_j \propto g(\mathbf{x}_j|\boldsymbol{\xi}) \quad (15)$$

where $\boldsymbol{\xi}$ is the parameters of this function of the size $a \times 1$. Considering the constraint Eq.(2), this is equivalent to

$$p_j = \frac{g(\mathbf{x}_j|\boldsymbol{\xi})}{\sum_{k=1}^m g(\mathbf{x}_k|\boldsymbol{\xi})}. \quad (16)$$

We call this the parametric representation of scored multinomial probabilities.

Some of the examples are:

- linear model:

$$g(\mathbf{x}_j|\boldsymbol{\xi}) = \mathbf{x}'_j \boldsymbol{\xi} \quad (17)$$

See Mayekawa(1996) for the details of this model. If \mathbf{x} is univariate, a polynomial expansion of x such as

$$\mathbf{x}_j = (1, x_j, x_j^2, \dots, x_j^k), \quad j = 1, 2, \dots, m, \quad (18)$$

can be used. For the multivariate \mathbf{x} , the cross product terms may be included in the expansion.

$$\mathbf{x}_j = (1, x_{j1}, x_{j2}, \dots, x_{j1}^2, x_{j2}^2, \dots, x_{j1}x_{j2}, \dots), \quad j = 1, 2, \dots, m, \quad (19)$$

- log linear model:

$$g(\mathbf{x}_j|\boldsymbol{\xi}) = \exp(\mathbf{x}'_j \boldsymbol{\xi}) \quad (20)$$

See Rosenbaum & Thayer(1987) for the details of this model. As with the linear model, it is often the case that a polynomial expansion of a scalar score is used as \mathbf{x} .

- PDF model:

$$g(\mathbf{x}_j|\boldsymbol{\xi}) = f(\mathbf{x}_j|\boldsymbol{\xi}) \quad (21)$$

where $f(\mathbf{x}|\boldsymbol{\xi})$ is the probability density function (PDF) of some distribution with the parameters $\boldsymbol{\xi}$.

- CDF model(univariate case):

$$\begin{aligned} g(x_1|\boldsymbol{\xi}) &= F(b_1|\boldsymbol{\xi}) \\ g(x_j|\boldsymbol{\xi}) &= F(b_j|\boldsymbol{\xi}) - F(b_{j-1}|\boldsymbol{\xi}), \quad j = 2, 3, \dots, m-1, \\ g(x_m|\boldsymbol{\xi}) &= 1 - F(b_{m-1}|\boldsymbol{\xi}) \end{aligned} \quad (22)$$

where $F(x|\boldsymbol{\xi})$ is the cumulative distribution function (CDF) of some distribution with the parameters $\boldsymbol{\xi}$, and $b_j, j = 2, 3, \dots, m-1$ are the known (upper) boundaries satisfying

$$b_{j-1} < x_j < b_j. \quad (23)$$

- CDF model(biivariate case):

$$\begin{aligned}
g(\mathbf{x}_j|\boldsymbol{\xi}) &= \Pr(b_{j-1,1} < x_{j1} < b_{j1} \ \& \ b_{j-1,2} < x_{j2} < b_{j2}), \ j = 1, 2, 3, \dots, m, \\
&= F(b_{j1}, b_{j2}|\boldsymbol{\xi}) - F(b_{j-1,1}, b_{j2}|\boldsymbol{\xi}) \\
&\quad - F(b_{j1}, b_{j-1,2}|\boldsymbol{\xi}) + F(b_{j-1,1}, b_{j-1,2}|\boldsymbol{\xi})
\end{aligned} \tag{24}$$

where $F(x_1, x_2|\boldsymbol{\xi})$ is the cumulative distribution function (CDF) of some bivariate distribution with the parameters $\boldsymbol{\xi}$, and $b_j, j = 0, 1, 2, \dots, m$ are the known boundaries satisfying

$$b_{j-1} < x_j < b_j, \quad b_0 = -\infty, \quad b_m = +\infty. \tag{25}$$

Note that the above PDF or CDF can be of a mixture form

$$f(\mathbf{x}_j|\boldsymbol{\xi}) = \sum_{q=1}^Q \rho_q f(\mathbf{x}_j|\boldsymbol{\xi}_q) \tag{26}$$

$$F(\mathbf{x}_j|\boldsymbol{\xi}) = \sum_{q=1}^Q \rho_q F(\mathbf{x}_j|\boldsymbol{\xi}_q) \tag{27}$$

where $f(x|\boldsymbol{\xi}_q), q = 1, 2, \dots, Q$, are the PDFs whose parameters are denoted as $\boldsymbol{\xi}_q$, $F(\mathbf{x}|\boldsymbol{\xi}_q), q = 1, 2, \dots, Q$, are the corresponding CDFs, and $\rho_q, q = 1, 2, \dots, Q$, are the positive mixing proportions satisfying $\sum_{q=1}^Q \rho_q = 1$.

Note also that when there are zero-frequency cells, that is, for some $j, f_j = 0$, the value of the likelihood changes depending on if we keep those cells in the model or not. This is because of the restrictions Eq.(2) or Eq.(16). When the values of $g(\mathbf{x}_j|\boldsymbol{\xi})$ associated with the zero-frequency cells are not zero, the denominator changes and, accordingly, the probabilities. The decision to include or exclude the zero-frequency cells depends upon the definition of the probabilities to be estimated.

In the following sections, we derive the maximum likelihood solution of those parameters ($\boldsymbol{\xi}$) in general, and show some detailed analysis of Normal PDF and CDF models.

4 Maximum Likelihood Estimation by the Fisher Scoring Method

The derivative of the log likelihood with respect to the parameter ξ can be written as

$$\frac{\partial \ln L}{\partial \xi} = \sum_{j=1}^m (f_j - np_j) \frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi} \quad (28)$$

$$= \sum_{j=1}^m \frac{(f_j - np_j)}{g(\mathbf{x}_j | \xi)} \frac{\partial g(\mathbf{x}_j | \xi)}{\partial \xi}, \quad (29)$$

where $\frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi}$ and $\frac{\partial g(\mathbf{x}_j | \xi)}{\partial \xi}$ are the $a \times 1$ vectors consisting of the derivative of, respectively, $\ln g(\mathbf{x}_j | \xi)$ and $g(\mathbf{x}_j | \xi)$ with respect to ξ .

Proof: By substituting Eq.(16) into Eq.(11) and using Eq.(9), we have

$$\ln L = \sum_{j=1}^m f_j \ln p_j \quad (30)$$

$$= \sum_{j=1}^m f_j \ln \frac{g(\mathbf{x}_j | \xi)}{\sum_{k=1}^m g(\mathbf{x}_k | \xi)} \quad (31)$$

$$= \sum_{j=1}^m f_j \left(\ln g(\mathbf{x}_j | \xi) - \ln \sum_{k=1}^m g(\mathbf{x}_k | \xi) \right) \quad (32)$$

$$= \sum_{j=1}^m f_j \ln g(\mathbf{x}_j | \xi) - n \ln \sum_{k=1}^m g(\mathbf{x}_k | \xi). \quad (33)$$

Therefore,

$$\frac{\partial \ln L}{\partial \xi} = \frac{\partial \sum_{j=1}^m f_j \ln g(\mathbf{x}_j | \xi) - n \ln \sum_{k=1}^m g(\mathbf{x}_k | \xi)}{\partial \xi} \quad (34)$$

$$= \sum_{j=1}^m f_j \frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi} - \frac{n}{\sum_{k=1}^m g(\mathbf{x}_k | \xi)} \sum_{k=1}^m \frac{\partial g(\mathbf{x}_k | \xi)}{\partial \xi} \quad (35)$$

$$= \sum_{j=1}^m f_j \frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi} - \frac{n}{\sum_{k=1}^m g(\mathbf{x}_k | \xi)} \sum_{k=1}^m g(\mathbf{x}_k | \xi) \frac{\partial \ln g(\mathbf{x}_k | \xi)}{\partial \xi}$$

$$= \sum_{j=1}^m f_j \frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi} - n \sum_{k=1}^m \frac{g(\mathbf{x}_k | \xi)}{\sum_{k=1}^m g(\mathbf{x}_k | \xi)} \frac{\partial \ln g(\mathbf{x}_k | \xi)}{\partial \xi} \quad (36)$$

$$= \sum_{j=1}^m f_j \frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi} - n \sum_{k=1}^m p_k \frac{\partial \ln g(\mathbf{x}_k | \xi)}{\partial \xi} \quad (37)$$

$$= \sum_{j=1}^m (f_j - np_j) \frac{\partial \ln g(\mathbf{x}_j | \xi)}{\partial \xi} \quad (38)$$

End of Proof

Note that this implies, for log linear model, that the estimated probabilities satisfy,

$$\sum_{j=1}^m f_j \mathbf{x}_j = \sum_{j=1}^m np_j \mathbf{x}_j \quad \text{or} \quad \mathbf{f}' \mathbf{X} = n \mathbf{p}' \mathbf{X} \quad (39)$$

where \mathbf{f} and \mathbf{p} are, respectively, the vectors of the observed frequencies and the model probabilities, and \mathbf{X} is the matrix of the scores. Therefore, if \mathbf{X} consists of the k -th degree polynomial expansions of the univariate raw score x_j ,

$$\mathbf{x}_j = (1, x_j, x_j^2, \dots, x_j^k), \quad j = 1, 2, \dots, m, \quad (40)$$

the log linear model preserves up to the k -th moment of the original distribution of the score.

In order to derive the information matrix, let us define the following two quantities: The discrepancy vector

$$\boldsymbol{\delta}(\boldsymbol{\xi}) = (\mathbf{f} - n\mathbf{p}) \quad (41)$$

where \mathbf{f} and \mathbf{p} are the $m \times 1$ vectors consisting of, respectively, the observed frequencies and the model probabilities defined as Eq.(16), and the $m \times a$ Jacobian matrix,

$$\mathbf{S}(\boldsymbol{\xi}) = \frac{\partial \ln g(\mathbf{X}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \frac{1}{g(\mathbf{X}|\boldsymbol{\xi})} \frac{\partial g(\mathbf{X}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}, \quad (42)$$

whose (j, k) element, which is $1 \times a$, is

$$\left[\frac{\partial \ln g(\mathbf{X}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \right]_{(jk)} = \frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \xi_k}. \quad (43)$$

With these, the derivative matrix can be written as

$$\frac{\partial \ln L}{\partial \boldsymbol{\xi}} = \left(\frac{\partial \ln g(\mathbf{X}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \right)' (\mathbf{f} - n\mathbf{p}) = \mathbf{S}(\boldsymbol{\xi})' \boldsymbol{\delta}(\boldsymbol{\xi}), \quad (44)$$

and the associated information matrix, by Eq.(7),

$$\mathbf{H}(\boldsymbol{\xi}) = \boldsymbol{\mathcal{E}} \left(\left(\frac{\partial \ln L}{\partial \boldsymbol{\xi}} \right) \left(\frac{\partial \ln L}{\partial \boldsymbol{\xi}} \right)' \right) \quad (45)$$

$$= \boldsymbol{\mathcal{E}}(\mathbf{S}(\boldsymbol{\xi})' \boldsymbol{\delta}(\boldsymbol{\xi}) \boldsymbol{\delta}(\boldsymbol{\xi})' \mathbf{S}(\boldsymbol{\xi})) = \mathbf{S}(\boldsymbol{\xi})' \boldsymbol{\mathcal{E}}(\boldsymbol{\delta}(\boldsymbol{\xi}) \boldsymbol{\delta}(\boldsymbol{\xi})') \mathbf{S}(\boldsymbol{\xi}) \quad (46)$$

$$= n \mathbf{S}(\boldsymbol{\xi})' (\text{Diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}') \mathbf{S}(\boldsymbol{\xi}). \quad (47)$$

Let $\boldsymbol{\xi}^\ell$ be the current estimate of $\boldsymbol{\xi}$. The updating equation for the Fisher scoring method is

$$\boldsymbol{\xi}^{\ell+1} = \boldsymbol{\xi}^\ell + \alpha \mathbf{H}(\boldsymbol{\xi}^\ell)^{-1} \mathbf{S}(\boldsymbol{\xi}^\ell)' (\mathbf{f} - n\mathbf{p}) \quad (48)$$

where α is the step size and $\mathbf{H}(\boldsymbol{\xi})$, $\mathbf{S}(\boldsymbol{\xi})$ and \mathbf{p} are evaluated at $\boldsymbol{\xi}^\ell$. By appropriately choosing the step size, each iteration should increase the log likelihood until the local maximum is obtained.

At convergence, the sampling distribution of the estimated parameter vector, $\hat{\boldsymbol{\xi}}$ can be approximated by

$$\hat{\boldsymbol{\xi}} \sim N(\boldsymbol{\xi}, \mathbf{H}^{-1}(\hat{\boldsymbol{\xi}})), \quad (49)$$

where $\mathbf{H}(\hat{\boldsymbol{\xi}})$ is the information matrix evaluated at $\hat{\boldsymbol{\xi}}$.

5 Normal Model

5.1 Univariate Normal PDF Model

Let (x_j, f_j) $j = 1, 2, \dots, m$, be the frequency distribution of a univariate random variable and $\boldsymbol{\xi} = (\mu, \sigma)'$. The number of parameters to be estimated is $a = 2$. By substituting

$$f(x_j|\boldsymbol{\xi}) \propto \exp\left(-\frac{1}{2}(x_j - \mu)^2/\sigma^2\right) \quad (50)$$

for $f(x_j|\boldsymbol{\xi})$ in Eq.(21), the jk , $j = 1, 2, \dots, m, k = 1, 2$, element of the Jacobian matrix shown in Eq.(42) becomes

$$\frac{\partial \ln g(x_j|\boldsymbol{\xi})}{\partial \xi_1} = \frac{x_j - \mu}{\sigma^2} \quad (51)$$

$$\frac{\partial \ln g(x_j|\boldsymbol{\xi})}{\partial \xi_2} = \frac{(x_j - \mu)^2}{\sigma^3} \quad (52)$$

Note that this is a special case of the log linear model where the score vector consists of the quadratic expansion of x_j

$$\mathbf{x}_j = (1, x_j, x_j^2) \quad (53)$$

and the parameters have the following constraint

$$\xi_1 = \frac{1}{4}\xi_2^2. \quad (54)$$

As noted previously, the mean and the variance are preserved with this model. That is

$$\frac{1}{n} \sum_{j=1}^m x_j f_j = \sum_{j=1}^m x_j p_j \quad \text{and} \quad \frac{1}{n} \sum_{j=1}^m x_j^2 f_j = \sum_{j=1}^m x_j^2 p_j \quad (55)$$

In the above development, σ is treated as the parameter to be estimated. If σ^2 is preferred instead, divide Eq.(93) by $2\sigma_q$ to get the derivative with respect to σ^2 .

5.2 Univariate Normal CDF Model

Let $\boldsymbol{\xi} = (\mu, \sigma)'$. By substituting $\Phi\left(\frac{b_j - \mu}{\sigma}\right)$ for $F(b_j|\boldsymbol{\xi})$ in Eq.(22), where $\Phi(x)$ is the CDF of the standard normal distribution, we have

$$\begin{aligned} g(x_1|\mu, \sigma^2) &= \Phi\left(\frac{b_1 - \mu}{\sigma}\right) \\ g(x_j|\mu, \sigma^2) &= \Phi\left(\frac{b_j - \mu}{\sigma}\right) - \Phi\left(\frac{b_{j-1} - \mu}{\sigma}\right), \quad j = 2, 3, \dots, m-1, \\ g(x_m|\mu, \sigma^2) &= 1 - \Phi\left(\frac{b_{m-1} - \mu}{\sigma}\right) \end{aligned} \quad (56)$$

Therefore, the jk , $j = 1, 2, \dots, m, k = 1, 2$, element of the Jacobian matrix shown in Eq.(42) becomes $\frac{1}{g(x_j|\boldsymbol{\xi})}$ times:

$$\frac{\partial g(x_1|\boldsymbol{\xi})}{\partial \xi_1} = -\phi\left(\frac{b_1 - \mu}{\sigma}\right) \frac{1}{\sigma} \quad (57)$$

$$\frac{\partial g(x_1|\boldsymbol{\xi})}{\partial \xi_2} = -\phi\left(\frac{b_1 - \mu}{\sigma}\right) \frac{b_1 - \mu}{\sigma^2} \quad (58)$$

for $j = 2, 3, \dots, m - 1$

$$\frac{\partial g(x_j|\boldsymbol{\xi})}{\partial \xi_1} = -\phi\left(\frac{b_j - \mu}{\sigma}\right) \frac{1}{\sigma} + \phi\left(\frac{b_{j-1} - \mu}{\sigma}\right) \frac{1}{\sigma} \quad (59)$$

$$\frac{\partial g(x_j|\boldsymbol{\xi})}{\partial \xi_2} = -\phi\left(\frac{b_j - \mu}{\sigma}\right) \frac{b_j - \mu}{\sigma^2} + \phi\left(\frac{b_{j-1} - \mu}{\sigma}\right) \frac{b_{j-1} - \mu}{\sigma^2} \quad (60)$$

$$\frac{\partial g(x_m|\boldsymbol{\xi})}{\partial \xi_1} = \phi\left(\frac{b_{m-1} - \mu}{\sigma}\right) \frac{1}{\sigma} \quad (61)$$

$$\frac{\partial g(x_m|\boldsymbol{\xi})}{\partial \xi_2} = \phi\left(\frac{b_{m-1} - \mu}{\sigma}\right) \frac{b_{m-1} - \mu}{\sigma^2} \quad (62)$$

where $\phi(x)$ is the pdf of the standard normal distribution.

5.3 Multivariate Normal PDF Model

Let (\mathbf{x}_j, f_j) $j = 1, 2, \dots, m$, be the frequency distribution of a R -variate random variable and $\boldsymbol{\xi} = (\boldsymbol{\mu}', \text{vec}(\boldsymbol{\Sigma}^{-1}))'$, where $\text{vec}(\boldsymbol{\Sigma}^{-1})$ is the $R(R+1)/2 \times 1$ vector consisting of the lower half of the inverse of the $R \times R$ symmetric matrix $\boldsymbol{\Sigma}$. Total number of parameters to be estimated is $a = R + R(R+1)/2 = R(R+3)/2$. By substituting

$$f(\mathbf{x}_j|\boldsymbol{\xi}) \propto \exp\left(-\frac{1}{2}(\mathbf{x}_j - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_j - \boldsymbol{\mu})\right) \quad (63)$$

for $f(\mathbf{x}_j|\boldsymbol{\xi})$ in Eq.(21), and taking the derivative with respect to $\boldsymbol{\Sigma}^{-1}$, the j -th row of the $m \times a$ Jacobian matrix shown in Eq.(42) becomes

$$\frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \left[\frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\mu}}, \frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \text{vec}(\boldsymbol{\Sigma}^{-1})} \right] \quad (64)$$

where

$$\frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\mu}} = (\mathbf{x}_j - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} \quad (65)$$

$$\frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \text{vec}(\boldsymbol{\Sigma}^{-1})} = -\frac{1}{2} \text{vec} (2(\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})' - \text{diag}((\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})')) \quad (66)$$

The multivariate normal method presented here can be used for a univariate case. The difference is that, in the univariate case, standard deviation σ is treated as a parameter, while in the multivariate case, the inverse of the dispersion matrix $\boldsymbol{\Sigma}$ is treated as the parameter.

5.4 Bivariate Normal CDF Model

Let \mathbf{X} be the bivariate normal random variable and $\Phi(z_1, z_2|\rho)$ be the distribution function of the standardized bivariate normal distribution with correlation coefficient ρ . With this, we have, if

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (67)$$

then

$$\Pr(X_1 \leq x_1 \ \& \ X_2 \leq x_2) = \Pr(Z_1 \leq z_1 \ \& \ Z_2 \leq z_2) \quad (68)$$

$$= \Phi(z_1, z_2 | \rho) \quad (69)$$

where \mathbf{Z} is the standardized bivariate normal variable defined as

$$\mathbf{Z} = \text{diag}(\boldsymbol{\Sigma})^{-1}(\mathbf{X} - \boldsymbol{\mu}) \quad (70)$$

or

$$z_{jk} = \frac{x_{jk} - \mu_k}{\sigma_k}, \quad j = 1, 2, \dots, m, \quad k = 1, 2 \quad (71)$$

Let $\boldsymbol{\xi} = (\boldsymbol{\mu}', \sigma_1, \sigma_2, \rho)'$. By substituting $\Phi(\frac{b_{j1} - \mu_1}{\sigma_1}, \frac{b_{j2} - \mu_2}{\sigma_2} | \rho)$ for $F(b_{j1}, b_{j2} | \boldsymbol{\xi})$ in Eq.(24) we have

$$\begin{aligned} g(\mathbf{x}_j | \boldsymbol{\xi}) &= \Phi((b_{j1} - \mu_1)/\sigma_1, (b_{j2} - \mu_2)/\sigma_2 | \rho) - \Phi((b_{j-1,1} - \mu_1)/\sigma_1, (b_{j2} - \mu_2)/\sigma_2 | \rho) \\ &\quad - \Phi((b_{j1} - \mu_1)/\sigma_1, (b_{j-1,2} - \mu_2)/\sigma_2 | \rho) + \Phi((b_{j-1,1} - \mu_1)/\sigma_1, (b_{j-1,2} - \mu_2)/\sigma_2 | \rho) \end{aligned} \quad (72)$$

Therefore, we need to evaluate 1×5 vector

$$\frac{\partial h(\mathbf{b}_j | \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \frac{\partial \Phi((b_{j1} - \mu_1)/\sigma_1, (b_{j2} - \mu_2)/\sigma_2 | \rho)}{\partial \boldsymbol{\xi}} \quad (73)$$

where

$$\boldsymbol{\xi} = \{\mu_1, \mu_2, \sigma_1, \sigma_2, \rho\}' \quad (74)$$

The required derivatives are as follows: (See the appendix for details.)

$$\begin{aligned} \frac{\partial h(\mathbf{b}_j | \boldsymbol{\xi})}{\partial \mu_1} &= \frac{-1}{4\pi} \int_0^\rho \frac{1}{(1-t^2)^{3/2}} \exp\left(\frac{-1}{2(1-t^2)} Q_{j12}\right) \times (-2(z_{j1} - tz_{j2})/\sigma_1) dt \\ &\quad - (1 - \Phi(-z_{j2})) \phi(z_{j1}) \frac{1}{\sigma_1} \end{aligned} \quad (75)$$

$$\begin{aligned} \frac{\partial h(\mathbf{b}_j | \boldsymbol{\xi})}{\partial \mu_2} &= \frac{-1}{4\pi} \int_0^\rho \frac{1}{(1-t^2)^{3/2}} \exp\left(\frac{-1}{2(1-t^2)} Q_{j12}\right) \times (-2(z_{j2} - tz_{j1})/\sigma_2) dt \\ &\quad - (1 - \Phi(-z_{j1})) \phi(z_{j2}) \frac{1}{\sigma_2} \end{aligned} \quad (76)$$

$$\begin{aligned} \frac{\partial h(\mathbf{b}_j | \boldsymbol{\xi})}{\partial \sigma_1} &= \frac{-1}{4\pi} \int_0^\rho \frac{1}{(1-t^2)^{3/2}} \exp\left(\frac{-1}{2(1-t^2)} Q_{j12}\right) \times (-2(z_{j1} - tz_{j2})z_{j1}/\sigma_1) dt \\ &\quad - (1 - \Phi(-z_{j2})) \phi(z_{j1}) \frac{z_{j1}}{\sigma_1} \end{aligned} \quad (77)$$

$$\begin{aligned} \frac{\partial h(\mathbf{b}_j | \boldsymbol{\xi})}{\partial \sigma_2} &= \frac{-1}{4\pi} \int_0^\rho \frac{1}{(1-t^2)^{3/2}} \exp\left(\frac{-1}{2(1-t^2)} Q_{j12}\right) \times (-2(z_{j2} - tz_{j1})z_{j2}/\sigma_2) dt \\ &\quad - (1 - \Phi(-z_{j1})) \phi(z_{j2}) \frac{z_{j2}}{\sigma_2} \end{aligned} \quad (78)$$

$$\frac{\partial h(\mathbf{b}_j | \boldsymbol{\xi})}{\partial \rho} = \frac{1}{2\pi} \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)} (z_{j1}^2 - 2\rho z_{j1} z_{j2} + z_{j2}^2)\right) \quad (79)$$

where

$$Q_{j12} = z_{j1}^2 - 2tz_{j1}z_{j2} + z_{j2}^2 \quad (80)$$

and

$$z_{jk} = \frac{b_{jk} - \mu_k}{\sigma_k}, \quad j = 1, 2, \dots, m, \quad k = 1, 2 \quad (81)$$

5.5 Mixture of Univariate Normal PDF Model

Let us consider the Q component mixture of the univariate normal PDF:

$$g(x_j|\boldsymbol{\xi}) = \sum_{q=1}^Q \rho_q g_q(x_j|\boldsymbol{\xi}) \quad (82)$$

$$= \sum_{q=1}^Q r_q^2 \frac{1}{\sigma_q} \exp\left(-\frac{1}{2}(x_j - \mu_q)^2/\sigma_q^2\right), \quad \sum_{q=1}^Q r_q^2 = 1 \quad (83)$$

An alternative formulation is

$$g(x_j|\boldsymbol{\xi}) = \sum_{q=1}^Q \rho_q g_q(x_j|\boldsymbol{\xi}) \quad (84)$$

$$= \sum_{q=1}^Q r_q^2 \exp\left(-\frac{1}{2}(x_j - \mu_q)^2/\sigma_q^2\right), \quad r_1 = 1 \quad (85)$$

With this formulation, the parameter \mathbf{r} and the mixing proportions are related as

$$\rho_q = \frac{r_q^2 \sigma_q}{\sum_{q'=1}^Q r_{q'}^2 \sigma_{q'}} \quad (86)$$

However, we prefer to use the former formulation.

The derivative of $\ln(g(x_j|\boldsymbol{\xi}))$ can be expressed as

$$\frac{\partial \ln g(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \frac{1}{g(x_j|\boldsymbol{\xi})} \frac{\partial g(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \frac{1}{g(x_j|\boldsymbol{\xi})} \frac{\partial \left(\sum_{q=1}^Q g_q(x_j|\boldsymbol{\xi}) \right)}{\partial \boldsymbol{\xi}} \quad (87)$$

$$= \frac{1}{g(x_j|\boldsymbol{\xi})} \sum_{q=1}^Q \frac{\partial g_q(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \quad (88)$$

$$= \frac{1}{g(x_j|\boldsymbol{\xi})} \sum_{q=1}^Q g_q(x_j|\boldsymbol{\xi}) \frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \quad (89)$$

where

$$\ln g_q(x_j|\boldsymbol{\xi}) = 2 \ln r_q - \ln \sigma_q - \frac{1}{2}(x_j - \mu_q)^2/\sigma_q^2 \quad (90)$$

Now let us define $(Q \times 2 + Q - 1) \times 1$ vector $\boldsymbol{\xi}$ as

$$\boldsymbol{\xi} = \begin{pmatrix} (\mu_1, \sigma_1)' \\ (\mu_2, \sigma_2)' \\ \vdots \\ (\mu_Q, \sigma_Q)' \\ (r_1, r_2, \dots, r_{Q-1})' \end{pmatrix} \quad (91)$$

The required derivatives are as follows:

$$\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \mu_q} = \frac{(x_j - \mu_q)}{\sigma_q^2}, q = 1, 2, \dots, Q \quad (92)$$

$$\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \sigma_q} = \frac{(x_j - \mu_q)^2}{\sigma_q^3} - \frac{1}{\sigma_q}, q = 1, 2, \dots, Q \quad (93)$$

$$\frac{\partial g_q(x_j|\boldsymbol{\xi})}{\partial r_q} = 2r_q (g_q(x_j|\boldsymbol{\xi}) - g_Q(x_j|\boldsymbol{\xi})) \quad , q = 1, 2, \dots, Q - 1 \quad (94)$$

It was found that the above method sometimes diverges depending on the initial values of the parameters. To avoid this problem, it is better to divide the set of parameters into $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ and \mathbf{r} , and maximize the log likelihood conditionally upon each other. That is, alternate the maximization of the log likelihood with respect to $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ conditional upon the current value of \mathbf{r} , and the maximization of the log likelihood with respect to \mathbf{r} conditional upon the current value of $(\boldsymbol{\mu}, \boldsymbol{\sigma})$.

5.6 Mixture of Multivariate Normal PDF Model

Let us consider the Q component mixture of the multivariate normal PDF:

$$g(\mathbf{x}_j|\boldsymbol{\xi}) = \sum_{q=1}^Q \rho_q g_q(\mathbf{x}_j|\boldsymbol{\xi}) \quad (95)$$

$$= \sum_{q=1}^Q r_q^2 \frac{1}{|\boldsymbol{\Sigma}_q|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x}_j - \boldsymbol{\mu}_q)' \boldsymbol{\Sigma}_q^{-1} (\mathbf{x}_j - \boldsymbol{\mu}_q)\right), \quad \sum_{q=1}^Q r_q^2 = 1 \quad (96)$$

The derivative of $\ln(g(\mathbf{x}_j|\boldsymbol{\xi}))$ can be expressed as

$$\frac{\partial \ln g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \frac{1}{g(\mathbf{x}_j|\boldsymbol{\xi})} \frac{\partial g(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \frac{1}{g(\mathbf{x}_j|\boldsymbol{\xi})} \frac{\partial \left(\sum_{q=1}^Q g_q(\mathbf{x}_j|\boldsymbol{\xi}) \right)}{\partial \boldsymbol{\xi}} \quad (97)$$

$$= \frac{1}{g(\mathbf{x}_j|\boldsymbol{\xi})} \sum_{q=1}^Q \frac{\partial g_q(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \quad (98)$$

$$= \frac{1}{g(\mathbf{x}_j|\boldsymbol{\xi})} \sum_{q=1}^Q g_q(\mathbf{x}_j|\boldsymbol{\xi}) \frac{\partial \ln g_q(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \quad (99)$$

where

$$\ln g_q(\mathbf{x}_j|\boldsymbol{\xi}) = 2 \ln r_q - \frac{1}{2} \ln |\boldsymbol{\Sigma}_q| - \frac{1}{2} (\mathbf{x}_j - \boldsymbol{\mu}_q)' \boldsymbol{\Sigma}_q^{-1} (\mathbf{x}_j - \boldsymbol{\mu}_q) \quad (100)$$

$$= 2 \ln r_q + \frac{1}{2} \ln |\boldsymbol{\Sigma}_q^{-1}| - \frac{1}{2} (\mathbf{x}_j - \boldsymbol{\mu}_q)' \boldsymbol{\Sigma}_q^{-1} (\mathbf{x}_j - \boldsymbol{\mu}_q) \quad (101)$$

Now let us define $(Q \times R(R+3)/2 + Q - 1) \times 1$ vector $\boldsymbol{\xi}$ as

$$\boldsymbol{\xi} = \begin{pmatrix} (\boldsymbol{\mu}'_1, \text{vec}(\boldsymbol{\Sigma}_1^{-1})')' \\ (\boldsymbol{\mu}'_2, \text{vec}(\boldsymbol{\Sigma}_2^{-1})')' \\ \vdots \\ (\boldsymbol{\mu}'_Q, \text{vec}(\boldsymbol{\Sigma}_Q^{-1})')' \\ (r_1, r_2, \dots, r_{Q-1})' \end{pmatrix} \quad (102)$$

The required derivatives are as follows:

$$\frac{\partial \ln g_q(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \boldsymbol{\mu}_q} = (\mathbf{x}_j - \boldsymbol{\mu}_q)' \boldsymbol{\Sigma}_q^{-1}, \quad q = 1, 2, \dots, Q \quad (103)$$

$$\begin{aligned} \frac{\partial \ln g_q(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \text{vec}(\boldsymbol{\Sigma}_q^{-1})} &= -\frac{1}{2} \text{vec} \left(2(\mathbf{x}_j - \boldsymbol{\mu}_q)(\mathbf{x}_j - \boldsymbol{\mu}_q)' - \text{diag} \left((\mathbf{x}_j - \boldsymbol{\mu}_q)(\mathbf{x}_j - \boldsymbol{\mu}_q)' \right) \right)' \\ \frac{\partial \ln g_q(\mathbf{x}_j|\boldsymbol{\xi})}{\partial \text{vec}(\boldsymbol{\Sigma}_q)} &= \frac{1}{2} \text{vec} \left(2\boldsymbol{\Sigma}_q^{-1}(\mathbf{x}_j - \boldsymbol{\mu}_q)(\mathbf{x}_j - \boldsymbol{\mu}_q)' \boldsymbol{\Sigma}_q^{-1} - \text{diag} \left(\boldsymbol{\Sigma}_q^{-1}(\mathbf{x}_j - \boldsymbol{\mu}_q)(\mathbf{x}_j - \boldsymbol{\mu}_q)' \boldsymbol{\Sigma}_q^{-1} \right) \right)' \\ &\quad + -\frac{1}{2} \text{vec} \left(2\boldsymbol{\Sigma}_q^{-1} - \text{diag} \left(\boldsymbol{\Sigma}_q^{-1} \right) \right) \quad , q = 1, 2, \dots, Q \end{aligned} \quad (104)$$

$$\frac{\partial g_q(\mathbf{x}_j|\boldsymbol{\xi})}{\partial r_q} = 2r_q (g_q(\mathbf{x}_j|\boldsymbol{\xi}) - g_Q(\mathbf{x}_j|\boldsymbol{\xi})) \quad , q = 1, 2, \dots, Q - 1 \quad (105)$$

As with the univariate model, we could omit the determinant factor from the definition of $g_q(\mathbf{x}_j|\boldsymbol{\xi})$.

Also, the alternating the conditional maximization seems to work better than the simultaneous maximization

5.7 Linear Constraints

In this section we consider enforcing some linear constraints on the means and the variances of the normal mixture distribution.

For example, we may assume that some of the μ_q 's, σ_q 's, and ρ_q 's are equal. This type of constraints can easily be expressed in the following form B linear constraints (Mayekawa, 1996.)

$$\mathbf{B}\boldsymbol{\mu} = \mathbf{c}_1 \quad \text{or} \quad \mathbf{B}\boldsymbol{\sigma}^2 = \mathbf{c}_2 \quad \text{or} \quad \mathbf{B}\boldsymbol{\rho} = \mathbf{c}_3 \quad (106)$$

where the \mathbf{B} matrix is defined as the contrast matrix such as

$$\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \end{bmatrix} \quad (107)$$

and $\mathbf{c} = \mathbf{0}$.

Another type of linear constraints is the one on the mean and the variance of the mixture distribution. The mean and the variance of the mixture distribution are:

$$\mathcal{E}(X) = \sum_{q=1}^Q \rho_q \mu_q \quad (108)$$

$$\mathcal{E}(X^2) = \sum_{q=1}^Q \rho_q (\sigma_q^2 + \mu_q^2) \quad (109)$$

$$\text{var}(X) = \sum_{q=1}^Q \rho_q (\sigma_q^2 + \mu_q^2) - (\mathcal{E}(X))^2 \quad (110)$$

where μ_q, σ_q^2 are, respectively, the mean and the variance of each component. Therefore, the constraints that the mean and the variance of the mixture distribution are some particular values, say, μ_0, σ_0^2 , can be written as

$$\sum_{q=1}^Q \rho_q \mu_q = \mu_0 \quad (111)$$

$$\sum_{q=1}^Q \rho_q (\sigma_q^2 + \mu_q^2) - \mu_0^2 = \sigma_0^2 \quad (112)$$

or

$$\sum_{q=1}^Q \rho_q \sigma_q^2 = \sigma_0^2 + \mu_0^2 - \sum_{q=1}^Q \rho_q \mu_q^2 \quad (113)$$

and these can be expressed as the form B constraints as

$$\mathbf{B}\boldsymbol{\mu} = \mathbf{c}_1 \quad \text{and} \quad \mathbf{B}\boldsymbol{\sigma}^2 = \mathbf{c}_2 \quad (114)$$

where

$$\mathbf{B} = \boldsymbol{\rho}' \quad \text{and} \quad \mathbf{c}_1 = \mu_0 \quad \text{and} \quad \mathbf{c}_2 = \sigma_0^2 + \mu_0^2 - \sum_{q=1}^Q \rho_q \mu_q^2 \quad (115)$$

Note that the above constraints are not strictly linear in the parameters because the \mathbf{B} and \mathbf{c} matrices contain the parameters themselves. However, experience shows that if we update those constraint matrices during the course of iteration, it will result in the desired constrained solution when the iteration converges. Especially, when we enforce those constraints, it is desirable to alternate the conditional maximization with respect to $\boldsymbol{\mu}$, $\boldsymbol{\sigma}^2$, and $\boldsymbol{\rho}$. Also, it is much easier to use the equivalent form A constraints as shown below. The equivalent form A constraints of the form

$$\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\lambda}_1 + \mathbf{d}_1 \quad \text{and} \quad \boldsymbol{\sigma}^2 = \mathbf{A}\boldsymbol{\lambda}_2 + \mathbf{d}_2 \quad (116)$$

can be found as

$$\mathbf{A} = \mathbf{B}_L, \quad \mathbf{d}_i = (\mathbf{B}'\mathbf{B})^{-1} \mathbf{B}'\mathbf{c}_i, \quad (117)$$

where the \mathbf{B}_L is a $(Q \times Q - \text{rank}(\mathbf{B}))$ matrix of full column rank satisfying

$$\mathbf{Q}_{\mathbf{B}'} = \mathbf{I}_p - \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1} \mathbf{B} = \mathbf{B}_L \mathbf{B}_R. \quad (118)$$

With the form A constraints, the constrained maximization now can be translated in the unconstrained maximization with respect to a new parameter $\boldsymbol{\lambda}$.

From the previous section, we have

$$\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \mu_q} = \frac{(x_j - \mu_q)}{\sigma_q^2}, \quad q = 1, 2, \dots, Q \quad (119)$$

$$\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \sigma_q^2} = \frac{(x_j - \mu_q)^2}{2(\sigma_q^2)^2} - \frac{1}{2\sigma_q^2}, \quad q = 1, 2, \dots, Q \quad (120)$$

$$\frac{\partial g_q(x_j|\boldsymbol{\xi})}{\partial r_q} = 2r_q (g_q(x_j|\boldsymbol{\xi}) - g_Q(x_j|\boldsymbol{\xi})) \quad , q = 1, 2, \dots, Q - 1 \quad (121)$$

Let $\boldsymbol{\xi}_1$ be either $\boldsymbol{\mu}$, $\boldsymbol{\sigma}^2$, or $\boldsymbol{\rho}$. Also let $\left[\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_1} \right]$ be the $Q \times 1$ vector consisting of $\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \xi_q}$. Then

$$\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\lambda}} = \frac{\partial \text{tr} \left(\left[\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_1} \right] \right)' \mathbf{A}\boldsymbol{\lambda}}{\partial \boldsymbol{\lambda}} \quad (122)$$

$$= \mathbf{A}' \left[\frac{\partial \ln g_q(x_j|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_1} \right] \quad (123)$$

5.8 Analysis of Simulated Data

The following three tables shows the result of the univariate Normal model analysis. $n = 100$ observations were generated from $N(0, 1)$, and categorized to $m = 21$ categories $-3.0, -2.7, \dots, 2.7, 3.0$. The sample mean and standard deviation are, respectively, $\bar{x} = -0.12300$ and $s = 0.92767$. Table 1 shows the frequency distribution, in which each x point is augmented by the boundaries.

Starting from the method of moment estimators (see below), the iteration converged, for both models, after 3 cycles with the convergence criterion for the successive difference between ps set to $1e-8$.

The estimated parameters are shown in Table 2 . For PDF model, we have $\mu = -0.12409, \sigma = 0.93200$ and for CDF model, we have $\mu = -0.12300, \sigma = 0.92362$.

For the PDF model, the loglikelihood (LLH1) is larger than the initial value (LLHM1) which is calculated by substituting sample mean and std for μ and σ in Eq.(50), Eq.(21), and Eq.(16). For CDF model, the loglikelihood (LLH2) is larger than the initial value (LLHM2) which is calculated by substituting sample meand and std for μ and σ in Eq.(56), Eq.(22), and Eq.(16). Note that the number of parameters estimated for the unrestricted model (NPARAM0) is 15 because there are 5 zero-frequency cells. Since we included those zero-frequency cells in the analysis, the estimated 21 probabilities sum to unity.

The estimated probabiltieis are shown in Table 3 . The column labeled p_{M1}, p_{M2}, p_1, p_2 are, respectively, the moment estimates calculated by substituting sample mean and std for μ and σ in Eq.(50), Eq.(21), and Eq.(16), the moment estimates calculated by substituting sample meand and std for μ and σ in Eq.(56), Eq.(22), and Eq.(16), and the maximum likelihood estimates under PDF and CDF models.

For the multivariate Normal model, $n = 200$ observations were generated from the standard bi-variate normal distribution with $\rho = .6$, and categorized to $m = 9$ categories $-3.0, -2.25, \dots, 2.25, 3.0$. The upper and the lower boundaries for the CDF model are shown in the first six columns of Table 5 . Out of possible 9×9 cells, there are 37 nonzero cells were observed as shown in the seven-th columns of Table 5 . The sample mean and the dispersion matrix are

$$\bar{\mathbf{x}} = (-0.09000, -0.09375)' \quad \text{and} \quad \mathbf{D} = \begin{bmatrix} 1.12253 & 0.69750 \\ 0.69750 & 1.06277 \end{bmatrix} \quad (124)$$

and the method of moment estimate log-likelihood was LLHM1=-647.51371 and LLHM2=-648.21821.

The iteration converged after four cycles and reslted in LLH1=-647.49334 for PDF model and LLH2= -647.62132 for CDF model. The estimated parameters are shown in Table 4 . For PDF model,

$$\boldsymbol{\mu} = (-0.09192, -0.09552)' \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1.14377 & 0.71493 \\ 0.71493 & 1.08039 \end{bmatrix} \quad (125)$$

and for CDF model,

$$\boldsymbol{\mu} = (-0.08919, -0.09375)' \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1.08163 & 0.69978 \\ 0.69978 & 1.01731 \end{bmatrix} \quad (126)$$

As observed with the univariate data, the PDF method seems to produce slightly larger variance estimates.

The estimated cell probabilities are shown in Table 5 .

The last example is the Normal Mixture PDF model. 1000 observations were generated from the mixture of the following two bivariate normal distributions

$$N\left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{bmatrix} 0.8 & 0.3 \\ 0.3 & 0.8 \end{bmatrix}\right) \quad \text{and} \quad N\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{bmatrix} 0.8 & -0.3 \\ -0.3 & 0.8 \end{bmatrix}\right) \quad (127)$$

with the mixing proportions being $\rho_1 = 0.2$ and $\rho_2 = 0.8$.

The results of the analysis are shown in Table 6 and Table 7 . The 18 iterations were required for the method to converge where each iteration contains two sub-iteration for $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and five sub-iterations for $\boldsymbol{\rho}$ resulting in the log likelihood value of -3966.24818. The true value of the parameters were correctly recovered as shown in Table 7 .

Finally, the mean/variance restricted analysis was performed using the same dataset as shown in Table 1 . The dataset was analysed using the two component normal mixture model with the constraints that mean=0 and variance=1. The resulting estimates are $\mu_1 = 3.35526, \mu_2 = -0.08807, \sigma_1^2 = 0.41358, \sigma_2^2 =$ with the value of the log likelihood being -254.86957.

References

- Mayekawa, Shin-ichi. (1996) Maximum likelihood estimation of the cell probabilities under linear constraints. *Behaviormetrika*, Vol.23, No.1, 111-128
- Rosenbaum, P.R. and Thayer, D. (1987) Smoothing the joint and marginal distributions of scored two-way contingency tables in test equating. *British Journal of Mathematical and Statistical Psychology* vol. 40, p43-39.
- Yamauchi, J (Eds.) (1972) Statistical Tables JSA-1972, Section E, Page 81, (1.12)

Appendix Bivariate Normal CDF

Let us denote the standard bivariate normal distribution as

$$Z \sim N(\mathbf{0}, \Sigma^*) \quad (128)$$

where

$$\text{diag}(\Sigma^*) = \mathbf{I} \quad \text{and} \quad \text{off}(\Sigma^*) = \rho \quad (129)$$

Using Yamauchi(1972) we have, if $\rho > 0$,

$$\Phi(z_1, z_2|\rho) = FF(z_1, z_2|\rho) + (1 - \Phi(-z_1))(1 - \Phi(-z_2)) \quad (130)$$

and if $\rho < 0$,

$$\Phi(z_1, z_2|\rho) = (1 - \Phi(-z_2)) - \Phi(z_1, -z_2|-\rho) \quad (131)$$

where

$$FF(z_1, z_2, \rho) = \frac{1}{2\pi} \int_0^\rho \frac{1}{\sqrt{1-t^2}} \exp\left(\frac{-1}{2(1-t^2)}(z_1^2 - 2tz_1z_2 + z_2^2)\right) dt, \quad \rho > 0 \quad (132)$$

By assuming that the integration and the differentiation can be interchangeable, we have, for $\xi = \mu_j, \sigma_j, \quad j = 1, 2$,

$$\frac{\partial FF(z_1, z_2, \rho)}{\partial \xi} = \frac{-1}{4\pi} \int_0^\rho \frac{1}{(1-t^2)^{3/2}} \exp\left(\frac{-1}{2(1-t^2)}Q_{12}\right) \times \frac{\partial Q_{12}}{\partial \xi} dt \quad (133)$$

where

$$Q_{12} = z_1^2 - 2tz_1z_2 + z_2^2 \quad (134)$$

Since

$$\frac{\partial Q_{12}}{\partial z_1} = 2(z_1 - tz_2) \quad \text{and} \quad \frac{\partial Q_{12}}{\partial z_2} = 2(z_2 - tz_1) \quad (135)$$

and

$$\frac{\partial z_j}{\partial \mu_j} = -\frac{1}{\sigma_j} \quad \text{and} \quad \frac{\partial z_j}{\partial \sigma_j} = -\frac{x_j - \mu_j}{\sigma_j^2} = -\frac{z_j}{\sigma_j} \quad (136)$$

we have

$$\frac{\partial Q_{12}}{\partial \mu_1} = -2(z_1 - tz_2)/\sigma_1 \quad (137)$$

$$\frac{\partial Q_{12}}{\partial \mu_2} = -2(z_2 - tz_1)/\sigma_2 \quad (138)$$

$$\frac{\partial Q_{12}}{\partial \sigma_1} = -2(z_1 - tz_2)z_1/\sigma_1 \quad (139)$$

$$\frac{\partial Q_{12}}{\partial \sigma_2} = -2(z_2 - tz_1)z_2/\sigma_2 \quad (140)$$

Also

$$\frac{\partial (1 - \Phi(-z_1))(1 - \Phi(-z_2))}{\partial \mu_1} = -(1 - \Phi(-z_2))\phi(z_1)\frac{1}{\sigma_1} \quad (141)$$

$$\frac{\partial (1 - \Phi(-z_1))(1 - \Phi(-z_2))}{\partial \mu_2} = -(1 - \Phi(-z_1))\phi(z_2)\frac{1}{\sigma_2} \quad (142)$$

$$\frac{\partial (1 - \Phi(-z_1))(1 - \Phi(-z_2))}{\partial \sigma_1} = -(1 - \Phi(-z_2))\phi(z_1)\frac{z_1}{\sigma_1} \quad (143)$$

$$\frac{\partial (1 - \Phi(-z_1))(1 - \Phi(-z_2))}{\partial \sigma_2} = -(1 - \Phi(-z_1))\phi(z_2)\frac{z_2}{\sigma_2} \quad (144)$$

For ρ

$$\frac{\partial FF(z_1, z_2, \rho)}{\partial \rho} = \frac{1}{2\pi} \frac{1}{\sqrt{1 - \rho^2}} \exp\left(\frac{-1}{2(1 - \rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right) \quad (145)$$

$$\frac{\partial (1 - \Phi(-z_1))(1 - \Phi(-z_2))}{\partial \rho} = 0 \quad (146)$$

Table 1: Univariate Normal Model Data

	b_{j-1}	x	b_j	f
1	.	-3.000	-2.850	0
2	-2.850	-2.700	-2.550	0
3	-2.550	-2.400	-2.250	1
4	-2.250	-2.100	-1.950	1
5	-1.950	-1.800	-1.650	3
6	-1.650	-1.500	-1.350	6
7	-1.350	-1.200	-1.050	5
8	-1.050	-0.900	-0.750	4
09	-0.750	-0.600	-0.450	13
10	-0.450	-0.300	-0.150	22
11	-0.150	0.000	0.150	10
12	0.150	0.300	0.450	8
13	0.450	0.600	0.750	10
14	0.750	0.900	1.050	6
15	1.050	1.200	1.350	5
16	1.350	1.500	1.650	4
17	1.650	1.800	1.950	0
18	1.950	2.100	2.250	1
19	2.250	2.400	2.550	0
20	2.550	2.700	2.850	1
21	2.850	3.000	.	0

Table 2: Univariate Normal Model Statistics

LLH0	M2LLH0	NPARAMO	AIC0		
-244.50082	489.00165	15	519.00165		
LLH1	M2LLH1	NPARAM	AIC1	MU1	SIGMA1
-254.70888	509.41776	2	513.41776	-0.12409	0.93200
LLH2	M2LLH2	NPARAM	AIC2	MU2	SIGMA2
-254.78325	509.56650	2	513.56650	-0.12300	0.92362
LLHMOM1	M2LLHMOM1	NPARAM	AICMOM1	MEANO	STD0
-254.71099	509.42199	2	513.42199	-0.12300	0.92767
LLHMOM2	M2LLHMOM2	NPARAM	AICMOM2	MEANO	STD0
-254.78513	509.57025	2	513.57025	-0.12300	0.92767

Table 3: Univariate Normal Model Result

x	f	p_{UR}	p_{M1}	p_{M2}	p_1	p_2
-3.00	0	0.0000	0.0011	0.0016	0.0011	0.0016
-2.70	0	0.0000	0.0027	0.0028	0.0028	0.0027
-2.40	1	0.0100	0.0063	0.0065	0.0065	0.0063
-2.10	1	0.0100	0.0133	0.0135	0.0136	0.0133
-1.80	3	0.0300	0.0252	0.0254	0.0255	0.0252
-1.50	6	0.0600	0.0429	0.0431	0.0432	0.0429
-1.20	5	0.0500	0.0658	0.0659	0.0660	0.0658
-0.90	4	0.0400	0.0909	0.0907	0.0909	0.0908
-0.60	13	0.1300	0.1131	0.1127	0.1128	0.1130
-0.30	22	0.2200	0.1268	0.1262	0.1262	0.1267
0.00	10	0.1000	0.1280	0.1273	0.1274	0.1279
0.30	8	0.0800	0.1164	0.1159	0.1159	0.1163
0.60	10	0.1000	0.0953	0.0951	0.0950	0.0952
0.90	6	0.0600	0.0703	0.0703	0.0703	0.0702
1.20	5	0.0500	0.0467	0.0469	0.0468	0.0467
1.50	4	0.0400	0.0279	0.0282	0.0282	0.0279
1.80	0	0.0000	0.0151	0.0153	0.0153	0.0151
2.10	1	0.0100	0.0073	0.0075	0.0075	0.0073
2.40	0	0.0000	0.0032	0.0033	0.0033	0.0032
2.70	1	0.0100	0.0013	0.0013	0.0013	0.0013
3.00	0	0.0000	0.0004	0.0007	0.0005	0.0006

Table 4: Biivariate Normal Model Statistics

LLHO	M2LLHO	NPARAMO	AICO			
-627.69992	1255.39985	34	1323.39985			
LLH1	M2LLH	NPARAM	AIC	MU	SIGMA	
-647.49335	1294.98670	5	1304.98670	-0.09192	1.14377	0.71493
				-0.09552	0.71493	1.08039
LLH2	M2LLH	NPARAM	AIC	MU	SIGMA	
-647.62132	1295.24263	5	1305.24263	-0.08919	1.08163	0.69978
				-0.09375	0.69978	1.01731
LLHMOM1	M2LLHMO1	NPARAM	AICMOM1	MEANO	DISPO	
-647.51371	1295.02742	5	1305.02742	-0.09000	1.12253	0.69750
				-0.09375	0.69750	1.06277
LLHMOM2	M2LLHMO2	NPARAM	AICMOM2	MEANO	DISPO	
-648.21821	1296.43642	5	1306.43642	-0.09000	1.12253	0.69750
				-0.09375	0.69750	1.06277

Table 5: Bivariate Normal Model Analysis

b_{L1}	x_1	b_{U1}	b_{L2}	x_2	b_{U2}	f	p_{UR}	p_{M1}	p_{M2}	p_1	p_2
-999999	-3.000	-2.625	-999999	-3.000	-2.625	0	0.0000	0.0009	0.0015	0.0010	0.0014
-999999	-3.000	-2.625	-2.625	-2.250	-1.875	0	0.0000	0.0022	0.0023	0.0024	0.0025
-999999	-3.000	-2.625	-1.875	-1.500	-1.125	0	0.0000	0.0022	0.0021	0.0023	0.0023
-999999	-3.000	-2.625	-1.125	-0.750	-0.375	0	0.0000	0.0009	0.0009	0.0009	0.0009
-999999	-3.000	-2.625	-0.375	0.000	0.375	0	0.0000	0.0001	0.0002	0.0001	0.0001
-999999	-3.000	-2.625	0.375	0.750	1.125	0	0.0000	0.0000	0.0000	0.0000	0.0000
-999999	-3.000	-2.625	1.125	1.500	1.875	0	0.0000	0.0000	0.0000	0.0000	0.0000
-999999	-3.000	-2.625	1.875	2.250	2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
-999999	-3.000	-2.625	2.625	3.000	999999.0	0	0.0000	0.0000	0.0000	0.0000	0.0000
-2.625	-2.250	-1.875	-999999	-3.000	-2.625	0	0.0000	0.0019	0.0027	0.0020	0.0021
-2.625	-2.250	-1.875	-2.625	-2.250	-1.875	2	0.0100	0.0079	0.0079	0.0082	0.0079
-2.625	-2.250	-1.875	-1.875	-1.500	-1.125	4	0.0200	0.0133	0.0122	0.0137	0.0134
-2.625	-2.250	-1.875	-1.125	-0.750	-0.375	0	0.0000	0.0092	0.0088	0.0094	0.0093
-2.625	-2.250	-1.875	-0.375	0.000	0.375	2	0.0100	0.0026	0.0029	0.0027	0.0026
-2.625	-2.250	-1.875	0.375	0.750	1.125	0	0.0000	0.0003	0.0004	0.0003	0.0003
-2.625	-2.250	-1.875	1.125	1.500	1.875	0	0.0000	0.0000	0.0000	0.0000	0.0000
-2.625	-2.250	-1.875	1.875	2.250	2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
-2.625	-2.250	-1.875	2.625	3.000	999999.0	0	0.0000	0.0000	0.0000	0.0000	0.0000
-1.875	-1.500	-1.125	-999999	-3.000	-2.625	1	0.0050	0.0017	0.0027	0.0017	0.0018
-1.875	-1.500	-1.125	-2.625	-2.250	-1.875	3	0.0150	0.0120	0.0134	0.0123	0.0120
-1.875	-1.500	-1.125	-1.875	-1.500	-1.125	6	0.0300	0.0352	0.0337	0.0357	0.0351
-1.875	-1.500	-1.125	-1.125	-0.750	-0.375	7	0.0350	0.0424	0.0394	0.0426	0.0423
-1.875	-1.500	-1.125	-0.375	0.000	0.375	5	0.0250	0.0208	0.0214	0.0209	0.0209
-1.875	-1.500	-1.125	0.375	0.750	1.125	2	0.0100	0.0042	0.0053	0.0042	0.0042
-1.875	-1.500	-1.125	1.125	1.500	1.875	0	0.0000	0.0003	0.0006	0.0004	0.0003
-1.875	-1.500	-1.125	1.875	2.250	2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
-1.875	-1.500	-1.125	2.625	3.000	999999.0	0	0.0000	0.0000	0.0000	0.0000	0.0000
-1.125	-0.750	-0.375	-999999	-3.000	-2.625	0	0.0000	0.0006	0.0012	0.0006	0.0006
-1.125	-0.750	-0.375	-2.625	-2.250	-1.875	2	0.0100	0.0078	0.0101	0.0079	0.0078
-1.125	-0.750	-0.375	-1.875	-1.500	-1.125	8	0.0400	0.0400	0.0415	0.0400	0.0398
-1.125	-0.750	-0.375	-1.125	-0.750	-0.375	24	0.1200	0.0838	0.0784	0.0832	0.0832
-1.125	-0.750	-0.375	-0.375	0.000	0.375	11	0.0550	0.0718	0.0689	0.0712	0.0715
-1.125	-0.750	-0.375	0.375	0.750	1.125	4	0.0200	0.0252	0.0281	0.0251	0.0252
-1.125	-0.750	-0.375	1.125	1.500	1.875	1	0.0050	0.0036	0.0053	0.0036	0.0036
-1.125	-0.750	-0.375	1.875	2.250	2.625	0	0.0000	0.0002	0.0005	0.0002	0.0002
-1.125	-0.750	-0.375	2.625	3.000	999999.0	0	0.0000	0.0000	0.0000	0.0000	0.0000
-0.375	0.000	0.375	-999999	-3.000	-2.625	0	0.0000	0.0001	0.0002	0.0001	0.0001
-0.375	0.000	0.375	-2.625	-2.250	-1.875	0	0.0000	0.0022	0.0034	0.0022	0.0022
-0.375	0.000	0.375	-1.875	-1.500	-1.125	3	0.0150	0.0195	0.0227	0.0194	0.0195
-0.375	0.000	0.375	-1.125	-0.750	-0.375	11	0.0550	0.0711	0.0695	0.0704	0.0708
-0.375	0.000	0.375	-0.375	0.000	0.375	26	0.1300	0.1062	0.0988	0.1049	0.1055
-0.375	0.000	0.375	0.375	0.750	1.125	11	0.0550	0.0649	0.0652	0.0643	0.0646
-0.375	0.000	0.375	1.125	1.500	1.875	1	0.0050	0.0162	0.0200	0.0162	0.0162
-0.375	0.000	0.375	1.875	2.250	2.625	1	0.0050	0.0017	0.0028	0.0017	0.0017
-0.375	0.000	0.375	2.625	3.000	999999.0	0	0.0000	0.0001	0.0002	0.0001	0.0001

Table 5: Bivariate Normal Model Analysis (cont)

b_{L1}	x_{-1}	b_{U1}	b_{L2}	x_{-2}	b_{U2}	f	p_{UR}	p_{M1}	p_{M2}	p_1	p_2
0.375	0.750	1.125	-999999	-3.000	-2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
0.375	0.750	1.125	-2.625	-2.250	-1.875	0	0.0000	0.0003	0.0005	0.0003	0.0003
0.375	0.750	1.125	-1.875	-1.500	-1.125	0	0.0000	0.0041	0.0055	0.0041	0.0041
0.375	0.750	1.125	-1.125	-0.750	-0.375	6	0.0300	0.0259	0.0274	0.0257	0.0260
0.375	0.750	1.125	-0.375	0.000	0.375	12	0.0600	0.0674	0.0632	0.0668	0.0673
0.375	0.750	1.125	0.375	0.750	1.125	11	0.0550	0.0718	0.0675	0.0713	0.0714
0.375	0.750	1.125	1.125	1.500	1.875	8	0.0400	0.0313	0.0335	0.0314	0.0312
0.375	0.750	1.125	1.875	2.250	2.625	3	0.0150	0.0056	0.0077	0.0057	0.0056
0.375	0.750	1.125	2.625	3.000	999999.0	0	0.0000	0.0004	0.0009	0.0004	0.0004
1.125	1.500	1.875	-999999	-3.000	-2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
1.125	1.500	1.875	-2.625	-2.250	-1.875	0	0.0000	0.0000	0.0000	0.0000	0.0000
1.125	1.500	1.875	-1.875	-1.500	-1.125	0	0.0000	0.0004	0.0006	0.0004	0.0004
1.125	1.500	1.875	-1.125	-0.750	-0.375	1	0.0050	0.0040	0.0048	0.0041	0.0041
1.125	1.500	1.875	-0.375	0.000	0.375	6	0.0300	0.0183	0.0180	0.0184	0.0185
1.125	1.500	1.875	0.375	0.750	1.125	7	0.0350	0.0340	0.0311	0.0342	0.0341
1.125	1.500	1.875	1.125	1.500	1.875	4	0.0200	0.0259	0.0250	0.0262	0.0258
1.125	1.500	1.875	1.875	2.250	2.625	1	0.0050	0.0080	0.0093	0.0083	0.0080
1.125	1.500	1.875	2.625	3.000	999999.0	0	0.0000	0.0010	0.0017	0.0011	0.0011
1.875	2.250	2.625	-999999	-3.000	-2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
1.875	2.250	2.625	-2.625	-2.250	-1.875	0	0.0000	0.0000	0.0000	0.0000	0.0000
1.875	2.250	2.625	-1.875	-1.500	-1.125	0	0.0000	0.0000	0.0000	0.0000	0.0000
1.875	2.250	2.625	-1.125	-0.750	-0.375	0	0.0000	0.0003	0.0004	0.0003	0.0003
1.875	2.250	2.625	-0.375	0.000	0.375	0	0.0000	0.0021	0.0023	0.0022	0.0022
1.875	2.250	2.625	0.375	0.750	1.125	2	0.0100	0.0069	0.0064	0.0071	0.0070
1.875	2.250	2.625	1.125	1.500	1.875	1	0.0050	0.0092	0.0083	0.0095	0.0093
1.875	2.250	2.625	1.875	2.250	2.625	1	0.0050	0.0050	0.0050	0.0052	0.0050
1.875	2.250	2.625	2.625	3.000	999999.0	0	0.0000	0.0011	0.0016	0.0012	0.0012
2.625	3.000	999999.0	-999999	-3.000	-2.625	0	0.0000	0.0000	0.0000	0.0000	0.0000
2.625	3.000	999999.0	-2.625	-2.250	-1.875	0	0.0000	0.0000	0.0000	0.0000	0.0000
2.625	3.000	999999.0	-1.875	-1.500	-1.125	0	0.0000	0.0000	0.0000	0.0000	0.0000
2.625	3.000	999999.0	-1.125	-0.750	-0.375	0	0.0000	0.0000	0.0000	0.0000	0.0000
2.625	3.000	999999.0	-0.375	0.000	0.375	0	0.0000	0.0001	0.0001	0.0001	0.0001
2.625	3.000	999999.0	0.375	0.750	1.125	0	0.0000	0.0006	0.0006	0.0006	0.0006
2.625	3.000	999999.0	1.125	1.500	1.875	0	0.0000	0.0014	0.0013	0.0015	0.0015
2.625	3.000	999999.0	1.875	2.250	2.625	2	0.0100	0.0013	0.0013	0.0014	0.0015
2.625	3.000	999999.0	2.625	3.000	999999.0	0	0.0000	0.0005	0.0008	0.0006	0.0008

Table 6: Biivariate Normal Mixture Model Statistics

LLHO	M2LLHO	NPARAMO	AICO					
-3922.86102	7845.72205	91	8027.72205					
LLH	M2LLH	NPARAM	AIC	MUALL	SIGMAALL	RALL		
-3966.24818	7932.49635	11	7954.49635	1.04830	0.87208	-0.35190	0.88323	
				1.00422	-0.35190	0.87526	0.46895	
				-0.94329	0.80071	0.31987	RHO	
				-1.05991	0.31987	0.75303	0.7801	
				0.2199				
LLHMOM1	M2LLHMOM1	NPARAM	AICMOM1	MEANO	DISPO			
-4157.72022	8315.44044	11	8337.44044	0.64200	1.43104	0.44916		
				0.58560	0.44916	1.47435		
LLHMOM2	M2LLHMOM2	NPARAM	AICMOM2	MEANO	DISPO			
-4164.23985	8328.47971	11	8350.47971	0.64200	1.43104	0.44916		
				0.58560	0.44916	1.47435		

Table 7: Bivariate Normal Mixture Model Analysis

x	f	PUR	$PM1$	$PM2$	$p1$	x	f	PUR	$PM1$	$PM2$	$p1$		
-3.00	-3.00	1	0.0010	0.0000	0.0001	0.0004	0.00	0.00	13	0.0130	0.0348	0.0333	0.0142
-3.00	-2.40	1	0.0010	0.0001	0.0002	0.0009	0.00	0.60	27	0.0270	0.0362	0.0346	0.0219
-3.00	-1.80	1	0.0010	0.0002	0.0003	0.0012	0.00	1.20	24	0.0240	0.0287	0.0277	0.0304
-3.00	-1.20	1	0.0010	0.0004	0.0005	0.0008	0.00	1.80	30	0.0300	0.0174	0.0170	0.0282
-3.00	-0.60	0	0.0000	0.0004	0.0005	0.0003	0.00	2.40	20	0.0200	0.0080	0.0080	0.0162
-3.00	0.00	0	0.0000	0.0004	0.0005	0.0001	0.00	3.00	4	0.0040	0.0028	0.0039	0.0057
-3.00	0.60	0	0.0000	0.0002	0.0003	0.0000	0.60	-3.00	0	0.0000	0.0004	0.0005	0.0000
-3.00	1.20	0	0.0000	0.0001	0.0002	0.0000	0.60	-2.40	0	0.0000	0.0015	0.0016	0.0002
-3.00	1.80	0	0.0000	0.0001	0.0001	0.0000	0.60	-1.80	1	0.0010	0.0051	0.0052	0.0010
-3.00	2.40	0	0.0000	0.0000	0.0000	0.0000	0.60	-1.20	2	0.0020	0.0131	0.0128	0.0034
-3.00	3.00	0	0.0000	0.0000	0.0000	0.0000	0.60	-0.60	7	0.0070	0.0254	0.0245	0.0095
-2.40	-3.00	1	0.0010	0.0001	0.0002	0.0010	0.60	0.00	21	0.0210	0.0376	0.0359	0.0230
-2.40	-2.40	2	0.0020	0.0004	0.0004	0.0027	0.60	0.60	44	0.0440	0.0425	0.0405	0.0425
-2.40	-1.80	3	0.0030	0.0008	0.0008	0.0043	0.60	1.20	55	0.0550	0.0367	0.0351	0.0522
-2.40	-1.20	4	0.0040	0.0013	0.0013	0.0037	0.60	1.80	46	0.0460	0.0242	0.0234	0.0400
-2.40	-0.60	2	0.0020	0.0016	0.0017	0.0018	0.60	2.40	14	0.0140	0.0121	0.0119	0.0189
-2.40	0.00	1	0.0010	0.0016	0.0016	0.0005	0.60	3.00	8	0.0080	0.0047	0.0065	0.0055
-2.40	0.60	0	0.0000	0.0012	0.0012	0.0001	1.20	-3.00	0	0.0000	0.0002	0.0002	0.0000
-2.40	1.20	0	0.0000	0.0007	0.0007	0.0000	1.20	-2.40	0	0.0000	0.0009	0.0009	0.0000
-2.40	1.80	0	0.0000	0.0003	0.0003	0.0001	1.20	-1.80	1	0.0010	0.0033	0.0033	0.0005
-2.40	2.40	1	0.0010	0.0001	0.0001	0.0001	1.20	-1.20	4	0.0040	0.0090	0.0090	0.0029
-2.40	3.00	0	0.0000	0.0000	0.0000	0.0000	1.20	-0.60	15	0.0150	0.0190	0.0186	0.0121
-1.80	-3.00	3	0.0030	0.0003	0.0004	0.0013	1.20	0.00	32	0.0320	0.0307	0.0296	0.0321
-1.80	-2.40	5	0.0050	0.0008	0.0008	0.0047	1.20	0.60	68	0.0680	0.0378	0.0361	0.0534
-1.80	-1.80	8	0.0080	0.0020	0.0020	0.0091	1.20	1.20	53	0.0530	0.0355	0.0339	0.0550
-1.80	-1.20	9	0.0090	0.0036	0.0036	0.0100	1.20	1.80	36	0.0360	0.0254	0.0245	0.0348
-1.80	-0.60	6	0.0060	0.0050	0.0049	0.0062	1.20	2.40	14	0.0140	0.0139	0.0136	0.0135
-1.80	0.00	3	0.0030	0.0052	0.0052	0.0022	1.20	3.00	3	0.0030	0.0058	0.0083	0.0032
-1.80	0.60	1	0.0010	0.0042	0.0043	0.0005	1.80	-3.00	0	0.0000	0.0001	0.0001	0.0000
-1.80	1.20	0	0.0000	0.0026	0.0027	0.0003	1.80	-2.40	0	0.0000	0.0004	0.0004	0.0001
-1.80	1.80	0	0.0000	0.0012	0.0013	0.0005	1.80	-1.80	0	0.0000	0.0016	0.0016	0.0006
-1.80	2.40	1	0.0010	0.0004	0.0005	0.0005	1.80	-1.20	4	0.0040	0.0047	0.0048	0.0036
-1.80	3.00	2	0.0020	0.0001	0.0002	0.0003	1.80	-0.60	13	0.0130	0.0108	0.0107	0.0134
-1.20	-3.00	0	0.0000	0.0004	0.0006	0.0011	1.80	0.00	27	0.0270	0.0190	0.0185	0.0302
-1.20	-2.40	4	0.0040	0.0015	0.0015	0.0046	1.80	0.60	35	0.0350	0.0254	0.0246	0.0418
-1.20	-1.80	10	0.0100	0.0038	0.0038	0.0113	1.80	1.20	22	0.0220	0.0260	0.0250	0.0355
-1.20	-1.20	22	0.0220	0.0075	0.0074	0.0157	1.80	1.80	17	0.0170	0.0203	0.0196	0.0184
-1.20	-0.60	13	0.0130	0.0113	0.0110	0.0122	1.80	2.40	3	0.0030	0.0121	0.0118	0.0059
-1.20	0.00	4	0.0040	0.0130	0.0127	0.0056	1.80	3.00	2	0.0020	0.0055	0.0080	0.0011
-1.20	0.60	2	0.0020	0.0114	0.0112	0.0023	2.40	-3.00	0	0.0000	0.0000	0.0000	0.0000
-1.20	1.20	4	0.0040	0.0076	0.0076	0.0024	2.40	-2.40	0	0.0000	0.0001	0.0001	0.0001
-1.20	1.80	1	0.0010	0.0039	0.0040	0.0032	2.40	-1.80	0	0.0000	0.0006	0.0006	0.0006
-1.20	2.40	1	0.0010	0.0015	0.0016	0.0027	2.40	-1.20	3	0.0030	0.0019	0.0019	0.0031
-1.20	3.00	2	0.0020	0.0005	0.0006	0.0014	2.40	-0.60	10	0.0100	0.0047	0.0047	0.0095
-0.60	-3.00	1	0.0010	0.0005	0.0007	0.0005	2.40	0.00	18	0.0180	0.0089	0.0089	0.0177
-0.60	-2.40	2	0.0020	0.0020	0.0020	0.0027	2.40	0.60	23	0.0230	0.0130	0.0127	0.0201
-0.60	-1.80	8	0.0080	0.0056	0.0055	0.0082	2.40	1.20	20	0.0200	0.0144	0.0141	0.0140
-0.60	-1.20	11	0.0110	0.0119	0.0116	0.0143	2.40	1.80	6	0.0060	0.0123	0.0119	0.0060
-0.60	-0.60	15	0.0150	0.0196	0.0189	0.0144	2.40	2.40	1	0.0010	0.0079	0.0078	0.0016
-0.60	0.00	7	0.0070	0.0245	0.0236	0.0095	2.40	3.00	1	0.0010	0.0039	0.0060	0.0002
-0.60	0.60	7	0.0070	0.0233	0.0226	0.0080	3.00	-3.00	0	0.0000	0.0000	0.0000	0.0000
-0.60	1.20	12	0.0120	0.0170	0.0166	0.0109	3.00	-2.40	0	0.0000	0.0000	0.0000	0.0001
-0.60	1.80	4	0.0040	0.0094	0.0094	0.0121	3.00	-1.80	0	0.0000	0.0002	0.0002	0.0004
-0.60	2.40	11	0.0110	0.0040	0.0041	0.0085	3.00	-1.20	1	0.0010	0.0006	0.0008	0.0017
-0.60	3.00	3	0.0030	0.0013	0.0018	0.0037	3.00	-0.60	1	0.0010	0.0015	0.0021	0.0041
0.00	-3.00	0	0.0000	0.0005	0.0006	0.0001	3.00	0.00	13	0.0130	0.0032	0.0043	0.0063
0.00	-2.40	1	0.0010	0.0020	0.0020	0.0009	3.00	0.60	6	0.0060	0.0050	0.0070	0.0059
0.00	-1.80	5	0.0050	0.0062	0.0061	0.0035	3.00	1.20	1	0.0010	0.0061	0.0086	0.0034
0.00	-1.20	8	0.0080	0.0144	0.0140	0.0079	3.00	1.80	1	0.0010	0.0056	0.0082	0.0012
0.00	-0.60	15	0.0150	0.0256	0.0246	0.0112	3.00	2.40	0	0.0000	0.0040	0.0060	0.0003
							3.00	3.00	1	0.0010	0.0021	0.0055	0.0000