Journal of Statistical and Econometric Methods, vol.5, no.1, 2016, 59-86
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2016

Using Halton Sequences in Random Parameters Logit Models

Tong Zeng[1]

Abstract

Quasi-random numbers that are evenly spread over the integration domain have become used as alternatives to pseudo-random numbers in maximum simulated likelihood problems to reduce computational time. In this paper, we carry out Monte Carlo experiments to explore the properties of quasi-random numbers, which are generated by the Halton sequence, in estimating the random parameters logit model. We vary the number of Halton draws, the sample size and the number of random coefficients. We show that increases in the number of Halton draws influence the efficiency of the random parameters logit model estimators only slightly. The maximum simulated likelihood estimator is consistent. We find that it is not necessary to increase the number of Halton draws when the sample size increases for this result to be evident.

JEL Classification: C02; C13; C15; C25
Mathematics Subject Classification: 11B83; 11K45

[1] Department of Economics and Finance, Black Hills State University, Spearfish, SD 57799. Email: tong.zeng@bhsu.edu, Tel: +1 605 642 6286, Fax: +1 605 642 6273

Article Info: *Received* : November 11, 2015. *Revised* : December 11, 2015.
Published online : March 3, 2016

Keywords: Maximum simulated likelihood; Random parameters logit model; Quasi-random numbers; Halton sequence; Monte Carlo experiments

1 Introduction

In this paper, we construct Monte Carlo experiments to explore the properties of quasi-random numbers, which are generated by the Halton sequence, in estimating the random parameters logit (RPL) model. The random parameters logit model is a generalization of the conditional logit model for multinomial choices. It has become more frequently used in many fields, such as agricultural economics, marketing, labor economics, health economics and transportation study, because of its high flexibility. Unlike the multinomial logit (MNL) model, this model is not limited by the *Independence from Irrelevant Alternatives* (IIA) assumption. It can capture the random preference variation among individuals and allows unobserved factors of utility to be correlated over time. However, the choice probability in the RPL model cannot be calculated exactly because it involves a multi-dimensional integral which does not have closed form. The use of pseudo-random numbers to approximate the integral during the simulation requires a large number of draws and leads to long computational times.

To reduce the computational cost, it is possible to replace the pseudo-random numbers by a set of fewer, evenly spaced points and still achieve the same, or even higher, estimation accuracy. Quasi-random numbers are evenly spread over the integration domain. They have become popular alternatives to pseudo-random numbers in maximum simulated likelihood problems. Bhat (2001) compared the performance of quasi-random numbers (Halton draws) and pseudo-random numbers in the context of the maximum simulated likelihood estimation of the RPL model. The root mean squared error (RMSE) and the mean absolute error ratio across parameters were used to evaluate the proximity of estimated and true parameters. He found that using 100 Halton draws the RMSE of the RPL model estimates was smaller than using 1000 pseudo-random numbers. However, Bhat (2001) also mentioned that the error measures of the estimated parameters do not always become smaller as the number of Halton draws increases. Train (2009, p.231) summarizes some numerical experiments comparing the use of 100 Halton draws with 125 Halton draws. He says,"...the standard deviations are greater with 125 Halton draws than with 100 Halton draws. The reason for this anomaly has not been determined. Its occurrence indicates the need for further investigation of the properties of Halton sequences in simulation-based estimation." It is our purpose to further the understanding of these properties through extensive simulation experiments. How does the number of quasi-random numbers, which are generated by the Halton draws, influence the efficiency of the estimated parameters? How many number of Halton draws should be chosen in the application of Halton sequences with the maximum simulated likelihood estimation? To make the maximum simulated likelihood estimator asymptotically equivalent to the maximum likelihood estimator, should we increase the number of points generated by the Halton sequence with increases in the sample size as using the pseudo-random numbers? In our experiments, we vary the number of Halton draws, the sample size and the number of random coefficients to explore the properties of the Halton sequences in estimating the RPL model. Unlike Bhat (2001), we analyze the RMSE and the ratio of the average nominal standard error to the Monte Carlo standard deviation of each estimated parameter. The results of our experiments confirm the efficiency of the quasi-random numbers in the context of the RPL model. We show that increases in the number of Halton draws influence the efficiency of the random parameters logit model estimators by a small amount. The maximum simulated likelihood estimator is consistent. We find that it is not necessary to increase the number of Halton draws when the sample size increases for this result to be evident.

The plan of the paper is as follows. In the following section, we discuss the random parameters logit specification. Section 3 introduces Halton sequences. Section 4 describes our Monte Carlo experiments. Section 5 presents the experimental results. Some conclusions are given in Section 6.

2 The Random Parameters Logit Model

The RPL model is described in Train (2009, p.134-150). Consider individual n facing M alternatives. The random utility associated with alternative

i is $U_{ni} = \beta_n' x_{ni} + \varepsilon_{ni}$, where x_{ni} are K observed explanatory variables for alternative i, ε_{ni} is an iid type I extreme value error which is independent of β_n and x_{ni}. The random coefficients β_n vary over individuals in the population with density function $f(\beta)$ and can be regarded as being composed of mean b and deviations $\widetilde{\beta}_n$. The RPL model decomposes the unobserved part of the utility into the extreme value term and the random part $\widetilde{\beta}_n x_{ni}$. Conditional on β_n the probability that individual n chooses alternative i is of the usual logistic form, $L_{ni}(\beta_n) = e^{\beta_n' x_{ni}} / \sum_i e^{\beta_n' x_{ni}}$. The probability that individual n chooses alternative i is

$$P_{ni} = \int L_{ni}(\beta) f(\beta|\theta) d\beta \tag{1}$$

The density function $f(\beta)$ provides the weights, and the choice probability is a weighted average of $L_{ni}(\beta)$ over all possible values of β_n. Even though the integral in (1) does not have a closed form, the choice probability in the RPL model can be estimated through simulation. The unknown parameters (θ), such as the mean and variance of the random coefficient distribution, can be estimated by maximizing the simulated log-likelihood function. With simulation, a value of β labelled as β^r representing the rth draw, is selected randomly from a previously specified distribution. The standard logit $L_{ni}(\beta)$ in equation (1) can be calculated with β^r. Repeating this process R times, the simulated probability of individual n choosing alternative i is obtained by averaging $L_{ni}(\beta^r)$:

$$\check{P}_{ni} = \frac{1}{R} \sum_{r=1}^{R} L_{ni}(\beta^r) \tag{2}$$

The simulated log-likelihood function is:

$$SLL(\theta) = \sum_{n=1}^{N} \sum_{i=1}^{M} d_{ni} ln \check{P}_{ni} \tag{3}$$

where the indicator variable d_{ni}=1 if individual n chooses alternative i. The simulated log-likelihood is then maximized numerically with respect to θ.

The method used to estimate the probability P_{ni} in (2) is called the classical Monte Carlo method. It reduces the integration problem to the problem of estimating the expected value on the basis of the strong law of large numbers. In general terms, the classical Monte Carlo method is described as a numerical method based on random sampling. The random sampling here is pseudo-random numbers. In terms of the number of pseudo-random numbers N, it gives us a probabilistic error bound, also called the convergence rate, $O(N^{-1/2})$ for numerical integration, since there is never any guarantee that the expected accuracy is achieved in a concrete calculation (Niederreiter, 1992, p.7). It represents the stochastic character of the classical-Monte Carlo method. The useful feature of the classical Monte Carlo method is that the convergence rate of the numerical integration does not depend on the dimension of the integration. With the classical Monte Carlo method, it is not difficult to get an unbiased simulated probability $\check{P}_{ni}$ for P_{ni}. The problem is the simulated log-likelihood function in (2) is a logarithmic transformation, which causes a simulation bias in the SLL which translates into bias in the MSL estimator. To decrease the bias in the MSL estimator and get a consistent and efficient MSL estimator, Train (2009, p.255) shows that, with an increase in the sample size N, the number of pseudo-random numbers should rise faster than $\sqrt{N}$. The disadvantage of the classical Monte Carlo method in the RPL model estimation is the requirement of a large number of pseudo-random numbers, which leads to long computational times.

3 The Halton Sequences

To reduce the computational cost, quasi-random numbers are being used to replace the pseudo-random numbers in MSL, leading to the same or even higher accuracy estimation with much fewer points. The essence of the number theoretic method (NTM) is to find a set of uniformly scattered points over an s-dimensional unit cube. Such set of points obtained by NTM is usually called a set of quasi-random numbers, or a number theoretic net. Sometimes it can be used in the classical Monte Carlo method to achieve a significantly higher accuracy. The Monte Carlo method with using quasi-random numbers is called a quasi-Monte Carlo method. In fact, there are several classical methods to construct the quasi-random numbers. Here we use the Halton sequences proposed by Halton (1960).

The Halton sequences are based on the base-p number system which implies that any integer n can be written as:

$$n \equiv n_M n_{M-1} \cdots n_2 n_1 n_0 = n_0 + n_1 p + n_2 p^2 + \cdots + n_M p^M \tag{4}$$

where $M = [\log_p n] = [\ln n / \ln p]$ and $M + 1$ is the number of digits of n, square brackets denoting the integral part, p is base and can be any integer except 1, n_i is the digit at position i, $0 \leq i \leq M$, $0 \leq n_i \leq p - 1$ and p^i is the weight of the digit at position i. For example, with the base $p = 10$, the integer $n = 468$ has $n_0 = 8$, $n_1 = 6$, $n_2 = 4$. The weights for n_0, n_1 and n_2 are 10^0, 10^1 and 10^2 respectively.

Using the base-p number system, we can construct one and only one fraction φ which is smaller than 1 by writing n with a different base number system and reversing the order of the digits in n. It is also called the radical inverse function defined as the follows:

$$\varphi = \varphi_p(n) = 0.n_0 n_1 n_2 \cdots n_M = n_0 p^{-1} + n_1 p^{-2} + \cdots + n_M p^{-M-1} \tag{5}$$

Based on the base-p number system, the integer $n = 468$ can be converted into the binary number system by successively dividing by the new base $p = 2$:

$$\begin{aligned} 468_{10} &= 1 \times 2^8 + 1 \times 2^7 + 1 \times 2^6 + 0 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 \\ &= 111010100_2 \end{aligned}$$

Applying the radical inverse function (5), we can get an unique fraction for the integer $n = 468$ with base $p = 2$:

$$\begin{aligned} \varphi_2(111010100) &= 0.001010111_2 = 1 \times 2^{-3} + 1 \times 2^{-5} + 1 \times 2^{-7} + 1 \times 2^{-8} + 1 \times 2^{-9} \\ &= 0.169921875_{10} \end{aligned}$$

The value 0.169921875_{10} is the corresponding fraction of 111010100_2 in the decimal number system.

The Halton sequence of length N is developed from the radical inverse function and the points of the Halton sequence are $\varphi_p(n)$ for $n = 1, 2, \cdots, N$, where p is a prime number. The k-dimensional sequence is defined as:

$$\phi_n = (\varphi_{p_1}(n), \varphi_{p_2}(n), \cdots, \varphi_{p_k}(n)) \tag{6}$$

Where $p_1, p_2, \cdots, p_k$ are prime to each other and are chosen from the first k primes. By setting $p_1, p_2, \cdots, p_k$ to be prime to each other we avoid the correlation among the points generated by any two Halton sequences with different base-p.

In applications, Halton sequences are used to replace random number generators to produce points in the interval [0, 1]. The points of the Halton sequence are generated iteratively. As far as a one-dimensional Halton sequence is concerned, the Halton sequence based on prime p divides the 0-1 space into p segments and systematically fills in the empty space by dividing each segment into smaller p segments iteratively. This is illustrated below. The numbers above the line represents the order of points filling in the space.

The position of the points is determined by the base which is used to construct the iteration. A large base implies more points in each iteration or longer cycle. Due to the high correlation among the initial points of the Halton sequence, the first ten points of the sequences are usually discarded in applications. Compared to the pseudo-random numbers, the coverage of the points of the Halton sequence are more uniform, since the pseudo-random numbers may cluster in some areas and leave some areas uncovered. This can be seen from Figure 1, which is similar to the figure from Bhat (2001). Figure 1(a) is a plot of 200 points taken from uniform distribution of two dimensions using pseudo-random numbers. Figure 1(b) is a plot of 200 points obtained by the Halton sequence. The latter scatters more uniformly on the unit square than the former. Since the points generated from the Halton sequences are deterministic points, unlike the classical-Monte Carlo method, quasi-Monte Carlo provides a deterministic error bound instead of probabilistic error bound. It is also called the discrepancy in the literature of number theoretic methods. The smaller the discrepancy, the more evenly the quasi-random numbers are spread over the domain. The deterministic error bound of quasi-Monte Carlo method with the k-dimensional Halton sequence, which is represented in terms of the number of points used, was shown [Halton, 1960] smaller than the probabilistic error bound of classical-Monte Carlo method as $O(N^{-1}(\ln N)^k)$. It means that with much fewer points generated by the Halton sequence we can achieve the same or even higher accuracy estimation than that with using pseudo-random numbers. However, some researchers pointed out the correlation problem among the points generated by the Halton sequence with two adjacent large prime number in high dimensional integral.

With high dimensional Halton sequences, usually $k \geq 10$, a large number of points is needed to complete the long cycle with large prime numbers. In addition to increasing the computational time, it will also cause a correlation between two adjacent large prime-based sequences, such as the thirteenth and fourteenth dimension generated by prime number 41 and 43 respectively. The correlation coefficient between two close large prime-based sequences is almost equal to one. This is shown in Figure 2, which is based on a graph from Bhat (2003). To solve this problem, number theorists such as Wang and Hickernell (2000) scramble the digits of each number of the sequences, which is called a scrambled Halton sequences. In this paper, we only focus on the normal Halton sequences with relatively low dimentional integral.

4 The Quasi-Monte Carlo Experiments with Halton Sequences

Our experiments begin from the simple RPL model which has no intercept term and only one random coefficient. Then, we expand the number of random coefficient to four by adding the random coefficient one by one. In our experiments, each individual faces four mutually exclusive alternatives with only one choice occasion. The associated utility for individual n choosing alternative i is:

$$U_{ni} = \beta_n' x_{ni} + \varepsilon_{ni} \tag{7}$$

The explanatory variables for each individual and each alternative x_{ni} are generated from independent standard normal distributions. The coefficients for each individual β_n are generated from normal distribution $N(\overline{\beta}, \overline{\sigma}_\beta^2)$. These values of x_{ni} and β_n are held fixed over each experiment design. The choice probability for each individual is generated with the logit-smoothed accept-reject simulator suggested by McFadden (1989). We set λ as 0.125.

$$\check{P}_{ni} = \frac{1}{R}\sum_{r=1}^{R}\frac{e^{U_{ni}^r/\lambda}}{\sum_j e^{U_{nj}^r/\lambda}} \tag{8}$$

The dependent variables y_{ni} are determined by these values of simulated choice probabilities. Our generated data is composed of the explanatory and dependent variables x_{ni} and y_{ni} which are used to estimate the RPL model parameters. In our experiments, we generate 1000 Monte Carlo samples ($NSAM$) with specific true values that we set for the RPL model parameters. During the estimation process, the random coefficients β_n in (7) are generated by the Halton sequences instead of pseudo-random numbers. First, we generate the k-dimensional Halton sequences of length $N \times R + 10$, where N is sample size, R is the number of the Halton draws assigned to each individual and 10 is the number of Halton draws that we discard due to the high correlation [Morokoff and Caflisch (1995), Bratley, et al. (1992)]. Then we transform these Halton draws into a set of numbers β_n with normal distribution using discrepancy-preserving transformation. Based on the discrepancy-preserving transformation, the independent multivariate normal distribution β_n which is transformed from the k-dimensional Halton sequences, has the same discrepancy as the Halton sequences generated from the k-dimensional unit cube. So the smaller discrepancy of the Halton sequences leads to the smaller discrepancy of β_n. To calculate the corresponding simulated probability $\check{P}_{ni}$ in (2), the first R points are assigned to the first individual, the second R points are assigned to the second individual, and so on. They are used to calculate the simulated probability $\check{P}_{ni}$ of each individual respectively.

To examine the efficiency of the estimated parameters using Halton sequences, we use the error measures: the ratio of the average nominal standard error to the Monte Carlo standard deviation of the estimated parameters and the root mean squared error (RMSE) of the RPL model estimates. They are calculated as follows using one estimated parameter $\hat{\beta}$ as an example:

$$\text{Monte Carlo average } \bar{\hat{\beta}}_i = \sum \hat{\beta}_i / NSAM \tag{9}$$

$$\text{Monte Carlo standard deviation (s.d.) of } \hat{\beta}_i = \sqrt{\sum (\hat{\beta}_i - \bar{\hat{\beta}})^2 / (NSAM - 1)} \tag{10}$$

$$\text{Average nominal standard error (s.e.) of } \hat{\beta}_i = \sum \sqrt{v\hat{a}r(\hat{\beta}_i)} / NSAM \tag{11}$$

$$\text{Root mean square error (RMSE) of } \hat{\beta}_i = \sqrt{\sum (\hat{\beta}_i - \bar{\beta})^2 / NSAM} \tag{12}$$

where $\bar{\beta}$ and $\hat{\beta}_i$ are the true parameter and estimates of parameter, respectively. To explore the properties of the Halton sequences in estimating the RPL model, we vary the number of Halton draws, the sample size and the number of random coefficients. we also do the same experiments using the pseudo-random numbers to compare the performance of the Halton sequence and pseudo-random numbers in estimating the RPL model. To avoid different simulation errors from the different process of probability integral transformation, we use the same probability integral transformation process with Halon draws and pseudo-random numbers.

5 The Experimental Results

In our experiments, we increase the number of random coefficients one by one. For each case, the RPL model is estimated by 25, 100, 250 and 500 Halton draws. We use 2000 pseudo-random numbers to get the benchmark results of the error measures which are based on the RPL model estimators. The mean and standard deviation of the random coefficient are set as 1.5 and 0.8 respectively. Table 1 and Table 2 show the results of the one random coefficient parameter logit model using Halton draws. Tables 3 and 4 present the results using 1000 and 2000 pseudo-random numbers. From Table 1 and Table 2, with the given number of observations, increasing the number of Halton draws from 25 to 500 only changes the RMSE of the estimated mean of the random coefficient distribution by less than 4%, and influences the RMSE of the estimated standard deviation of the random coefficient distribution by no more than 10%. When the number of observations increases to 500 and 800, increasing the number of Halton draws from 100 to 500 only influences of the RMSE and the ratio the average nominal standard deviations to the Monte Carlo standard deviations of each estimated parameter very slightly. The RMSE of the estimated parameter mean is lower using 25 Halton draws than that using more Halton draws and pseudo-random numbers. With 100 Halton draws, we can reach almost the same efficiency of the RPL model estimators as using 2000 pseudo-random numbers. The results are consistent with Bhat (2001). The ratios of the average nominal standard deviations to the Monte Carlo standard deviations of the estimated parameters are stable with increases in the number of Halton draws.

Tables 5-12 present the results of two independent random coefficients logit model using Halton draws and pseudo-random numbers. We set the mean and the standard deviation of the new random coefficient as 1.0 and 0.5 respectively. The same error measures are used to analyze the efficiency of each estimator for each case. After including another random coefficient, the mean of each random coefficient distribution is overestimated. The RMSE of the RPL estimator is stable in the number of Halton draws. Again, with increases in the number of observations, increasing the number of Haton draws doesn't influence the efficiency of the estimated parameters significantly. The $\hat{\beta}$ has the lowest RMSE with 25 Halton draws. In the two random coefficients case, the RMSEs with 500 Halton draws are the closest ones to the according benchmark results. The results with 100 Halton draws are also very close to the benchmark and there is no significantly difference between them.

As the number of random coefficients increases, the computational time increases greatly using pseudo-random numbers rather than using quasi-random numbers. However, we can get almost the same efficiency of the estimated parameters using 100 Halton draws as using 1000 pseudo-random numbers. Tables showing the results of three and four independent random coefficients logit model are available upon request. With three and four independent random coefficients, using 25 Halton draws doesn't always provide the lowest RMSE of the estimated parameter mean. When the number of random coefficients is increased, the effect of rising Halton draws on the efficiency of the RPL model estimators is still slightly, especially with 500 and 800 observations. The results are similar to the one and two random coefficients cases. Train (2009, p.225) discusses that the negative correlation between the average of two adjacent observation's draws can reduce errors in the simulated log-likelihood function, like the method of antithetic variates. However, this negative covariance across observations declines with increases in the number of observations N, since the length of Halton sequences in estimating the RPL model is determined by the number of observations and the number of Halton draws assigned to each observation. The increases in the number of observations will decrease the gap between two adjacent observation's coverage. Train (2009, p.225) suggests increasing the number of Halton draws for each individual when the number of observations increases. But, based on our experimental results, we find that increasing the number of Halton draws for each individual does not significantly affect the RMSE of the RPL model estimators as the number of observations increases.

6 Conclusions

In this paper, we study the properties of the Halton sequences in estimating the RPL model, which is a very flexible and a generalization of the conditional logit model. With one or two independent random coefficients, using only 25 Halton draws can provides smaller RMSE of the estimated parameters than that using 2000 pseudo-random numbers. When the number of random coefficients is increased to three and four, with 100 Halton draws can achieve almost the same efficiency of the estimated parameters as using 1000 pseudo-random numbers. However, the computational time is reduced greatly. The most important thing is, as the number of observations increases, we find it is not necessary to increase the number of Halton draws to get the efficient and consistent maximum simulated likelihood estimators. Our experimental results can also provide the guidance of using quasi-random numbers generated by the Halton sequence in estimating other discrete choice model, like the probit model.

Table 1: The mixed logit model with one random coefficient

$\bar{\beta} = 1.5$, $\bar{\sigma}_\beta = 0.8$

Quasi-Monte Carlo Estimation

	Number of Halton Draws			
Estimator $\hat{\beta}$	25	100	250	500
Observations=200				
Monte Carlo average	1.468	1.477	1.477	1.477
Monte Carlo s.d.	0.226	0.233	0.232	0.233
Average nominal s.e.	0.236	0.237	0.237	0.237
Average nominal s.e./MC s.d.	1.044	1.017	1.022	1.017
RMSE	0.228	0.234	0.233	0.234
Observations=500				
Monte Carlo average	1.578	1.582	1.585	1.585
Monte Carlo s.d.	0.163	0.163	0.163	0.163
Average nominal s.e.	0.165	0.166	0.165	0.165
Average nominal s.e./MC s.d.	1.012	1.018	1.012	1.012
RMSE	0.181	0.183	0.184	0.183
Observations=800				
Monte Carlo average	1.521	1.533	1.535	1.534
Monte Carlo s.d.	0.125	0.125	0.125	0.125
Average nominal s.e.	0.128	0.129	0.129	0.129
Average nominal s.e./MC s.d.	1.024	1.032	1.032	1.032
RMSE	0.127	0.129	0.129	0.129

Table 2: The mixed logit model with one random coefficient

$\bar{\beta} = 1.5$, $\bar{\sigma}_\beta = 0.8$

Quasi-Monte Carlo Estimation

	Number of Halton Draws			
Estimator $\hat{\sigma}_\beta$	25	100	250	500
Observations=200				
Monte Carlo average	0.594	0.606	0.602	0.601
Monte Carlo s.d.	0.337	0.372	0.375	0.377
Average nominal s.e.	0.417	0.447	0.465	0.473
Average nominal s.e./MC s.d.	1.237	1.202	1.240	1.255
RMSE	0.395	0.419	0.424	0.426
Observations=500				
Monte Carlo average	0.728	0.740	0.743	0.743
Monte Carlo s.d.	0.236	0.243	0.242	0.243
Average nominal s.e.	0.245	0.249	0.248	0.249
Average nominal s.e./MC s.d.	1.038	1.025	1.025	1.025
RMSE	0.246	0.250	0.249	0.250
Observations=800				
Monte Carlo average	0.741	0.763	0.766	0.766
Monte Carlo s.d.	0.177	0.173	0.172	0.172
Average nominal s.e.	0.183	0.182	0.181	0.182
Average nominal s.e./MC s.d.	1.034	1.052	1.052	1.058
RMSE	0.187	0.177	0.176	0.176

Table 3: The mixed logit model with one random coefficient

$\bar{\beta} = 1.5$, $\bar{\sigma}_\beta = 0.8$

Classical-Monte Carlo Estimation

	Number of Random Draws	
Estimator $\hat{\beta}$	1000	2000
Observations=200		
Monte Carlo average	1.479	1.483
Monte Carlo s.d.	0.229	0.233
Average nominal s.e.	0.236	0.239
Average nominal s.e./MC s.d.	1.031	1.026
RMSE	0.230	0.234
Observations=500		
Monte Carlo average	1.584	1.590
Monte Carlo s.d.	0.162	0.163
Average nominal s.e.	0.165	0.166
Average nominal s.e./MC s.d.	1.019	1.018
RMSE	0.182	0.187
Observations=800		
Monte Carlo average	1.531	1.536
Monte Carlo s.d.	0.124	0.125
Average nominal s.e.	0.129	0.129
Average nominal s.e./MC s.d.	1.040	1.032
RMSE	0.128	0.130

Table 4: The mixed logit model with one random coefficient

$\bar{\beta} = 1.5$, $\bar{\sigma}_\beta = 0.8$

Classical-Monte Carlo Estimation

	Number of Random Draws	
Estimator $\hat{\sigma}_\beta$	1000	2000
Observations=200		
Monte Carlo average	1.479	1.483
Monte Carlo s.d.	0.229	0.233
Average nominal s.e.	0.236	0.239
Average nominal s.e./MC s.d.	1.031	1.026
RMSE	0.230	0.234
Observations=500		
Monte Carlo average	0.614	0.618
Monte Carlo s.d.	0.354	0.368
Average nominal s.e.	0.424	0.435
Average nominal s.e./MC s.d.	1.198	1.182
RMSE	0.400	0.410
Observations=800		
Monte Carlo average	0.758	0.768
Monte Carlo s.d.	0.172	0.173
Average nominal s.e.	0.182	0.181
Average nominal s.e./MC s.d.	1.058	1.046
RMSE	0.177	0.175

Table 5: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0$, $\bar{\sigma}_{\beta_1} = 0.5$; $\bar{\beta}_2 = 1.5$, $\bar{\sigma}_{\beta_2} = 0.8$

Quasi-Monte Carlo Estimation

	Number of Halton Draws			
Estimator $\hat{\beta}_1$	25	100	250	500
Observations=200				
Monte Carlo average	1.002	1.011	1.007	1.009
Monte Carlo s.d.	0.168	0.176	0.174	0.175
Average nominal s.e.	0.188	0.190	0.188	0.188
Average nominal s.e./MC s.d.	1.119	1.080	1.080	1.074
RMSE	0.168	0.176	0.174	0.175
Observations=500				
Monte Carlo average	1.018	1.029	1.029	1.031
Monte Carlo s.d.	0.107	0.111	0.111	0.111
Average nominal s.e.	0.122	0.125	0.125	0.125
Average nominal s.e./MC s.d.	1.140	1.126	1.126	1.126
RMSE	0.108	0.115	0.115	0.115
Observations=800				
Monte Carlo average	1.007	1.020	1.018	1.019
Monte Carlo s.d.	0.083	0.086	0.086	0.086
Average nominal s.e.	0.095	0.097	0.097	0.097
Average nominal s.e./MC s.d.	1.145	1.128	1.128	1.128
RMSE	0.083	0.089	0.088	0.089

Table 6: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0,\ \bar{\sigma}_{\beta_1} = 0.5;\ \bar{\beta}_2 = 1.5,\ \bar{\sigma}_{\beta_2} = 0.8$

Quasi-Monte Carlo Estimation

	Number of Halton Draws			
Estimator $\hat{\sigma}_{\beta_1}$	25	100	250	500
Observations=200				
Monte Carlo average	0.433	0.431	0.409	0.414
Monte Carlo s.d.	0.315	0.350	0.358	0.358
Average nominal s.e.	0.460	0.515	0.544	0.542
Average nominal s.e./MC s.d.	1.460	1.471	1.520	1.514
RMSE	0.322	0.357	0.369	0.368
Observations=500				
Monte Carlo average	0.487	0.503	0.504	0.506
Monte Carlo s.d.	0.221	0.229	0.230	0.230
Average nominal s.e.	0.282	0.290	0.290	0.292
Average nominal s.e./MC s.d.	1.276	1.266	1.261	1.270
RMSE	0.222	0.229	0.230	0.230
Observations=800				
Monte Carlo average	0.460	0.478	0.474	0.473
Monte Carlo s.d.	0.184	0.191	0.194	0.196
Average nominal s.e.	0.222	0.222	0.228	0.234
Average nominal s.e./MC s.d.	1.207	1.162	1.175	1.194
RMSE	0.189	0.192	0.196	0.197

Table 7: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0,\ \bar{\sigma}_{\beta_1} = 0.5;\ \bar{\beta}_2 = 1.5,\ \bar{\sigma}_{\beta_2} = 0.8$

Quasi-Monte Carlo Estimation

	Number of Halton Draws			
Estimator $\hat{\beta}_2$	25	100	250	500
Observations=200				
Monte Carlo average	1.557	1.566	1.561	1.562
Monte Carlo s.d.	0.260	0.264	0.260	0.261
Average nominal s.e.	0.279	0.280	0.278	0.277
Average nominal s.e./MC s.d.	1.073	1.061	1.069	1.061
RMSE	0.266	0.272	0.267	0.268
Observations=500				
Monte Carlo average	1.518	1.533	1.531	1.532
Monte Carlo s.d.	0.167	0.167	0.166	0.167
Average nominal s.e.	0.176	0.179	0.178	0.178
Average nominal s.e./MC s.d.	1.054	1.072	1.072	1.066
RMSE	0.168	0.170	0.169	0.170
Observations=800				
Monte Carlo average	1.511	1.534	1.531	1.533
Monte Carlo s.d.	0.124	0.127	0.127	0.128
Average nominal s.e.	0.137	0.141	0.140	0.141
Average nominal s.e./MC s.d.	1.105	1.110	1.102	1.102
RMSE	0.124	0.132	0.131	0.132

Table 8: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0,\ \bar{\sigma}_{\beta_1} = 0.5;\ \bar{\beta}_2 = 1.5,\ \bar{\sigma}_{\beta_2} = 0.8$

Quasi-Monte Carlo Estimation

	Number of Halton Draws			
Estimator $\hat{\sigma}_{\beta_2}$	25	100	250	500
Observations=200				
Monte Carlo average	0.874	0.894	0.882	0.883
Monte Carlo s.d.	0.338	0.330	0.326	0.328
Average nominal s.e.	0.369	0.367	0.367	0.369
Average nominal s.e./MC s.d.	1.092	1.112	1.126	1.125
RMSE	0.345	0.343	0.336	0.338
Observations=500				
Monte Carlo average	0.816	0.843	0.834	0.838
Monte Carlo s.d.	0.221	0.212	0.213	0.213
Average nominal s.e.	0.237	0.232	0.233	0.233
Average nominal s.e./MC s.d.	1.072	1.094	1.094	1.094
RMSE	0.222	0.216	0.215	0.216
Observations=800				
Monte Carlo average	0.771	0.811	0.804	0.807
Monte Carlo s.d.	0.163	0.161	0.161	0.161
Average nominal s.e.	0.185	0.185	0.185	0.185
Average nominal s.e./MC s.d.	1.135	1.149	1.149	1.149
RMSE	0.165	0.161	0.161	0.161

Table 9: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0$, $\bar{\sigma}_{\beta_1} = 0.5$; $\bar{\beta}_2 = 1.5$, $\bar{\sigma}_{\beta_2} = 0.8$

Classical-Monte Carlo Estimation

	Number of Random Draws	
Estimator $\hat{\beta}_1$	1000	2000
Observations=200		
Monte Carlo average	1.010	1.012
Monte Carlo s.d.	0.173	0.175
Average nominal s.e.	0.190	0.189
Average nominal s.e./MC s.d.	1.098	1.080
RMSE	0.173	0.176
Observations=500		
Monte Carlo average	1.026	1.034
Monte Carlo s.d.	0.110	0.111
Average nominal s.e.	0.124	0.126
Average nominal s.e./MC s.d.	1.127	1.135
RMSE	0.113	0.116
Observations=800		
Monte Carlo average	1.015	1.022
Monte Carlo s.d.	0.085	0.086
Average nominal s.e.	0.096	0.097
Average nominal s.e./MC s.d.	1.129	1.128
RMSE	0.086	0.089

Table 10: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0,\ \bar{\sigma}_{\beta_1} = 0.5;\ \bar{\beta}_2 = 1.5,\ \bar{\sigma}_{\beta_2} = 0.8$

Classical-Monte Carlo Estimation

	Number of Random Draws	
Estimator $\hat{\sigma}_{\beta_1}$	1000	2000
Observations=200		
Monte Carlo average	0.429	0.426
Monte Carlo s.d.	0.333	0.342
Average nominal s.e.	0.507	0.502
Average nominal s.e./MC s.d.	1.523	1.468
RMSE	0.341	0.350
Observations=500		
Monte Carlo average	0.499	0.516
Monte Carlo s.d.	0.219	0.220
Average nominal s.e.	0.281	0.276
Average nominal s.e./MC s.d.	1.283	1.255
RMSE	0.219	0.221
Observations=800		
Monte Carlo average	0.465	0.481
Monte Carlo s.d.	0.186	0.187
Average nominal s.e.	0.221	0.216
Average nominal s.e./MC s.d.	1.188	1.155
RMSE	0.189	0.188

Table 11: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0,\ \bar{\sigma}_{\beta_1} = 0.5;\ \bar{\beta}_2 = 1.5,\ \bar{\sigma}_{\beta_2} = 0.8$		
Classical-Monte Carlo Estimation		
	Number of Random Draws	
Estimator $\hat{\beta}_2$	1000	2000
Observations=200		
Monte Carlo average	1.562	1.562
Monte Carlo s.d.	0.258	0.261
Average nominal s.e.	0.277	0.278
Average nominal s.e./MC s.d.	1.074	1.065
RMSE	0.266	0.268
Observations=200		
Monte Carlo average	1.531	1.531
Monte Carlo s.d.	0.165	0.166
Average nominal s.e.	0.177	0.178
Average nominal s.e./MC s.d.	1.073	1.072
RMSE	0.168	0.169
Observations=200		
Monte Carlo average	1.532	1.532
Monte Carlo s.d.	0.126	0.127
Average nominal s.e.	0.140	0.140
Average nominal s.e./MC s.d.	1.111	1.102
RMSE	0.130	0.131

Table 12: The mixed logit model with two random coefficients

$\bar{\beta}_1 = 1.0,\ \bar{\sigma}_{\beta_1} = 0.5;\ \bar{\beta}_2 = 1.5,\ \bar{\sigma}_{\beta_2} = 0.8$

Classical-Monte Carlo Estimation

	Number of Random Draws	
Estimator $\hat{\sigma}_{\beta_2}$	1000	2000
Observations=200		
Monte Carlo average	0.881	0.889
Monte Carlo s.d.	0.316	0.327
Average nominal s.e.	0.357	0.369
Average nominal s.e./MC s.d.	1.130	1.128
RMSE	0.326	0.338
Observations=200		
Monte Carlo average	0.834	0.841
Monte Carlo s.d.	0.208	0.214
Average nominal s.e.	0.228	0.233
Average nominal s.e./MC s.d.	1.096	1.089
RMSE	0.210	0.218
Observations=200		
Monte Carlo average	0.807	0.808
Monte Carlo s.d.	0.158	0.161
Average nominal s.e.	0.182	0.185
Average nominal s.e./MC s.d.	1.152	1.149
RMSE	0.158	0.162

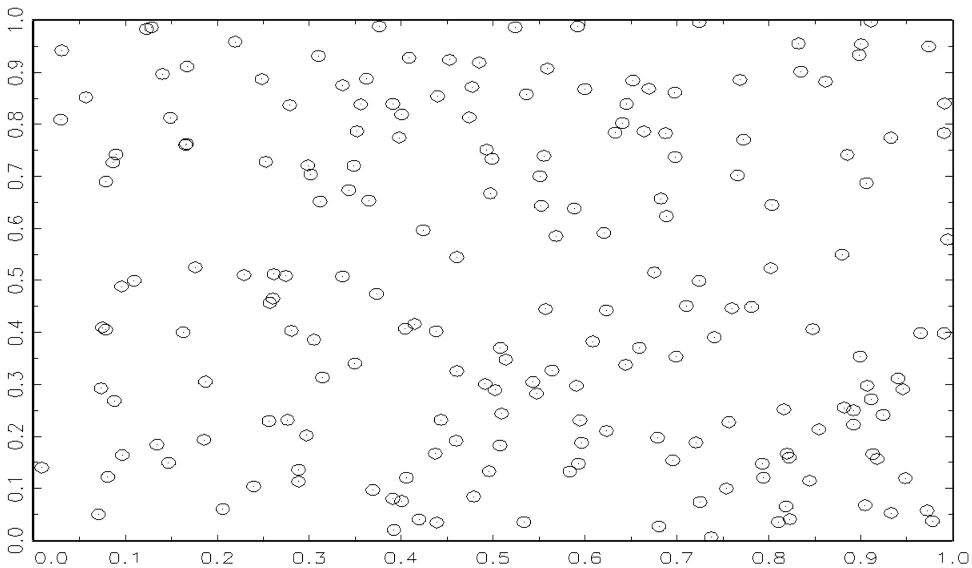

Figure 1(a): 200 points pseduo-random numbers in two-dimension

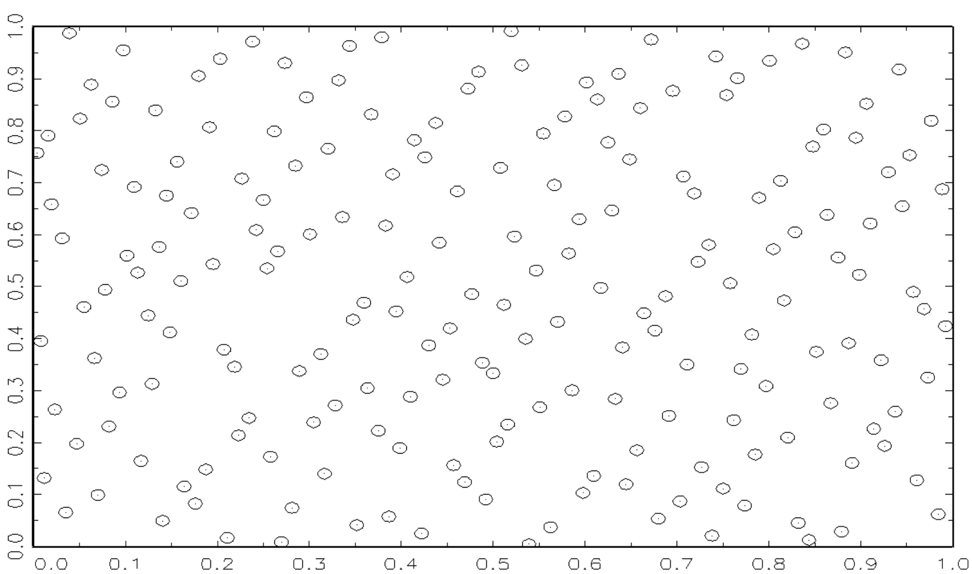

Figure 1(b): 200 points generated from two-dimension Halton sequence with prime 2 and 3

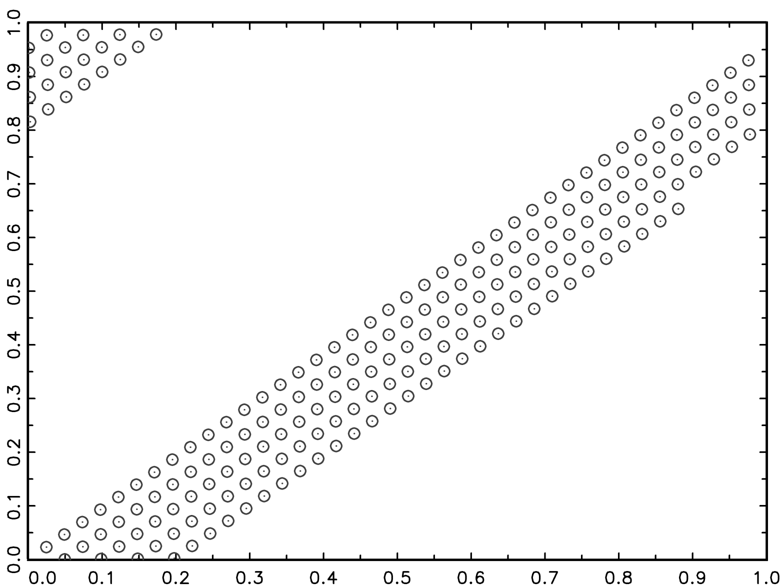

Figure 2: 200 points of two-dimension Halton sequence generated with prime 41 and 43

ACKNOWLEDGEMENTS. I thank R. Carter Hill for his help and insightful suggestion and thank John H. Halton for his kind encouragement. All mistakes are mine.

References

[1] Bhat, C.R., Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model, *Transportation Research Part B*, **35**(7), (2001), 677-693.

[2] Bratley, P., Fox, B.L., and Niederreiter,H., Implementation and Tests of Low-discrepancy Sequences, *ACM Transportations on Modeling and Computer Simulation*, **2**(3), (1992), 195-213.

[3] Fang, K.T. and Wang, Y., *Number-theoretic Methods in Statistics*, Chapman and Hall/CRC, 1994.

[4] Halton, J. H., On the Efficiency of Certain Quasi-random Sequences of Points in Evaluating Multi-dimensional Integrals, *Numerishe Mathematik*, **2**, (1960), 84-90.

[5] Keng, H.L. and Yan, W., *Application of Number Theory to Numerical Analysis*, Spring-Verlag Berlin Heidelberg New York, 1981.

[6] McFadden, D., A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration, *Econometrica*, **57**, (1989), 995-1026.

[7] Morokoff, W.J. and Caflisch, R.E.C., Quasi-Monte Carlo Integration, *Journal of Computational Physics*, **122**, (1995), 218-230.

[8] Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*, Society for Industrial and Applied Mathematics, 1992.

[9] Train, K. E., *Discrete Choice Methods with Simulation*, Cambridge University Press, 2009.

[10] Wang, X.Q. and Hickernell, Fred J., Randomized Halton Sequence, *Mathematical and Computer Modeling*, **32**, (2000), 887-899.

[11] Yan, S. Y., *Number Theory for Computing*, Springer-Verlag Berlin Heidelberg New York, 2002.