Based on the search and Nash bargaining model developed in class, (which was used as a basis for the estimation in Flinn (2006)), you are to generate a simulated data set of observations that contains information similar to that found in the Current Population Survey (CPS). In this analysis there will be no binding minimum wage in order to simplify things. The labor market environment is described in terms of the parameters

\[
\begin{align*}
\rho &= 0.005 \\
b &= ? \\
\mu &= 2.5 \\
\sigma &= 0.8 \\
\lambda &= 0.3 \\
\eta &= 0.03 \\
\alpha &= 0.4 \\
\end{align*}
\]

where the match productivity distribution is assumed to be lognormal with parameters \(\mu\) and \(\sigma\) (i.e., \(\ln(\theta)\) is normal with mean \(\mu\) and standard deviation \(\sigma\)). The critical match value is equal to the critical wage, or \(\theta^* = w^*\), where

\[
\theta^* = b + \frac{\alpha \lambda}{\rho + \eta} \int_{\theta^*}^{\theta} (\theta - \theta^*) dG(\theta).
\]

In this labor market environment, \(\theta^* = 6\). Any \(\theta\) match greater than or equal to \(\theta^*\) is accepted by the worker and firm, and the Nash-bargained wage rate is

\[
w(\theta) = \alpha \theta + (1 - \alpha) \theta^*.
\]

Then the cumulative distribution function of accepted wage offers is given by

\[
F(w; \alpha, \theta^*, \mu, \sigma) = \frac{G\left(\frac{w - (1 - \alpha) \theta^*}{\alpha}; \mu, \sigma\right) - G(\theta^*; \mu, \sigma)}{\tilde{G}(\theta^*; \mu, \sigma)}, \ w \geq \theta^*,
\]

and where \(\tilde{G}(x) \equiv 1 - G(x)\) is the survivor function. Then the density of observed wage offers is given by

\[
f(w; \alpha, \theta^*, \mu, \sigma) = \frac{\alpha^{-1} g\left(\frac{w - (1 - \alpha) \theta^*}{\alpha}; \mu, \sigma\right)}{\tilde{G}(\theta^*; \mu, \sigma)}, \ w \geq \theta^*.
\]
The steady state unemployment rate is, as usual, given by
\[ p_{SS}(U) = \frac{h_e}{h_e + h_u}, \]
where \( h_i \) is the hazard rate out of state \( i, i = U, E \). We know that \( h_e = \eta \) and \( h_u = \lambda G(\theta^*; \mu, \sigma) \).

Given this model structure, you are to investigate the performance of maximum likelihood estimators using Monte Carlo simulation techniques.

1. Given the labor market environment and \( \theta^* \), solve for the unknown parameter \( b \) (this is not an estimate, correct?).

2. Find the steady state probability of unemployment given the parameters and the decision rule \( \theta^* \).

3. Generate CPS-type data using this model structure for a sample of size \( N = 100,000 \). Use the following algorithm to accomplish this:

(a) For each “observation,” draw a uniform random number, \( z_1 \). In Gauss, this involves the use of the operator `rndu(r,c)`, where \( r \) and \( c \) denote the row and column dimension of the matrix of pseudo-random uniform variates. If \( z_1 < p_{SS}(U) \), then the individual is considered to be unemployed. If not, the individual is employed.

(b) If “individual \( i \)” is unemployed at the sampling time, draw another uniform random number for them, \( z_2 \). The right-censored, length-biased spell of unemployment in this case is just a draw from the population distribution of unemployment spells, as we showed. Since
\[ P(T \leq t_u) = 1 - \exp(-h_u t_u), \]
we consider the uniform draw to be the probability on the left-hand side, and this is used to generate the random duration time. Then
\[ z_2 = 1 - \exp(-h_u t_u) \]
\[ \Rightarrow t_u = -\frac{\ln(1 - z_2)}{h_u}. \]

Create the duration of on-going, point-sampled spells in this way.

If “individual \( i \)” is employed, we create an accepted wage draw. Take a pseudo-random number draw from the \( U[0,1] \) once again, \( z_3 \). Then define
\[ \bar{\theta} = \exp\{\mu + \sigma \Phi^{-1}(z_3 \Phi\left(\frac{\ln \theta^* - \mu}{\sigma}\right) + \Phi\left(\frac{\ln \theta^* - \mu}{\sigma}\right))\}, \]
where $\Phi^{-1}$ is the inverse of the standard normal c.d.f. [In GAUSS, this operator is cdfni]. The wage is then

$$w = \alpha \tilde{\theta} + (1 - \alpha) \theta^*.$$  

4. Assume that you know that the true value of $\alpha$ is 0.4 and that the value of the discount rate, $\rho$, is equal to 0.005. Implement the maximum likelihood estimator for all of the identified parameters of the model following the estimation strategy that you used in Assignment 1. *Use only the first 2000 observations you have created to represent the sample at your disposal.*

5. Now assume that you don’t know $\alpha$. You are to attempt to find a maximum likelihood estimator of it. Try this under 3 separate scenarios. First, use only the first 300 observations from your data. Next, use the first 2000 observations from your data, and, finally, use all 100,000 observations. Describe your experiences in attempting to estimate the model.

6. *(Method of Simulated Moments).* Assume that you know $\alpha$ again (and use its true value) for purposes of implementing this estimator. From the sample you have generated *(the first 2000 observations, as in Part 4)*, compute the following data characteristics:

(a) The **percentage** of individuals unemployed ($m_1$)

(b) The mean wage of the employed ($m_2$)

(c) The standard deviation of wages of the employed ($m_3$)

(d) The **percentage** of wages between 10-12 dollars ($m_4$)

(e) The **percentage** of wages between 15-18 dollars ($m_5$)

7. Draw a matrix of uniform random numbers, $Z$, that is $50,000 \times 3$. These are 50,000 simulated observations used in the estimation part of the problem. Column 1 corresponds to $z_1$, the random number used to determine employment status. Column 2 corresponds to $z_2$, the random number used to determine duration of the on-going unemployment spell if the individual is unemployed. Column 3 corresponds to $z_3$, which is used to generate an accepted wage if the individual is determined to be employed. Given a vector of guesses of the parameters of the model, denoted by $\varphi$, generate a new “CPS-type” data set from these 50,000 observations in the same way you generated the “real” data set in Part 3. From these new simulated data, compute the sample characteristics (a)-(e) in Part 6. Call this vector of 5 data characteristics $\tilde{m}(\varphi)$.

8. Define

$$\hat{\varphi} = \arg \min_{\varphi} (m - \tilde{m}(\varphi))'(m - \tilde{m}(\varphi)).$$
The problem with finding the solution in this case is that $\tilde{m}(\varphi)$ is not necessarily a continuously differentiable function of $\varphi$ (this issue will be discussed in the lecture and lab). In this case one should employ an optimization technique that is not derivative-based, the most popular one being the simplex (or Nelder-Mead) algorithm. Attempt to compute the estimator $\hat{\varphi}$, and compare it with the m.l.e. from Part 4.