Asha Bharadwaj, Caitlin Dutta, John Higgins, Alexis Smith
Econ 899
4 November, 2022

We consider a dynamic model of loan repayment. $T_i \in \{1, 2, 3, 4\}$ denotes the observed loan duration, and Y_{it} denotes an indicator variable equal to one if the loan is pre-paid at the end of period t. T_i is related to Y_{it} in the following manner:

$$T_i = \begin{cases} 1 & \text{if } Y_{i0} = 0 \\ 2 & \text{if } Y_{i0} = 0 \text{ and } Y_{i1} = 1 \\ 3 & \text{if } Y_{i0} = Y_{i1} = 0 \text{ and } Y_{i2} = 1 \\ 4 & \text{if } Y_{i0} = Y_{i1} = Y_{i2} = 0 \end{cases}$$

We assume that, at each period t, a loan is repaid only if

$$\alpha_t + X_i\beta + Z_{it}\gamma + \varepsilon_{it} < 0$$

Rearranging this, we find that the loan will be repaid only If

$$\varepsilon_{it} < -\alpha_t - X_i\beta - Z_{it}\gamma$$

This means that a loan will *not* be repaid in period t if $\varepsilon_{it} \geq -\alpha_t - X_i\beta - Z_{it}\gamma$.

We assume that X_i is a time-invariant vector of borrow characteristics, Z_{it} is a vector of time varying characteristics (such as the borrower's FICO score), and $\varepsilon_{it} = \rho\,\varepsilon_{it-1} + \eta_{it}$ if $t > 1$ and $\eta_{it} \sim N(0, 1)$. We assume that $\varepsilon_{it} \sim N(0, \sigma_0^2)$, with $\sigma_0^2 = \frac{1}{(1-\rho)^2}$.

Given this setup, we must derive the likelihood associated with loan duration T_i. I believe that the likelihood given in the problem set is not quite correct, so I have attempted to derive it based on the above assumptions.

We first wish to find the likelihood of observing $T_i = 1$ given X_i, Z_i, and θ. $T_i = 1$

corresponds to the loan being paid off at $t = 0$. This means that

$$\begin{aligned} P(T_i = 1 \mid X_i, Z_i, \theta) &= P(\alpha_0 + X_i\beta + Z_{i0}\gamma + \varepsilon_{i0} < 0) \\ &= P(\varepsilon_{i0} < -\alpha_0 + -X_i\beta - Z_{i0}\gamma) \\ &= \Phi\left(\frac{-\alpha_0 + -X_i\beta - Z_{i0}\gamma}{\sigma_0}\right) \end{aligned}$$

since $\varepsilon_{i0} \sim N(0, \sigma_0^2)$ implies that $\frac{\varepsilon_{i0}}{\sigma_0} \sim N(0,1)$. Hence, we have that

$$P(T_i = 1 X_i, Z_i, \theta) = \Phi\left(\frac{-\alpha_0 + -X_i\beta - Z_{i0}\gamma}{\sigma_0}\right)$$

We now must find $P(T_i = 2 \mid X_i, Z_i, \theta)$. $T_i = 2$ corresponds to the loan not being repaid in $t = 0$ and being repaid in $t = 1$. Thus, we have that

$$\begin{aligned} P(T_i = 2 \mid X_i, Z_i, \theta) &= P(\alpha_0 + X_i\beta + Z_{i0}\gamma + \varepsilon_{i0} \geq 0, \alpha_1 + X_i\beta + Z_{i1}\gamma + \varepsilon_{i1} < 0) \\ &= P(\varepsilon_{i0} \geq -\alpha_0 - X_i\beta - Z_{i0}\gamma, \varepsilon_{i1} < -\alpha_1 - X_i\beta - Z_{i1}\gamma) \\ &= P(\varepsilon_{i0} \geq -\alpha_0 - X_i\beta - Z_{i0}\gamma, \rho\,\varepsilon_{i0} + \eta_{i1} < -\alpha_1 - X_i\beta - Z_{i1}\gamma) \\ &= P(\varepsilon_{i0} \geq -\alpha_0 - X_i\beta - Z_{i0}\gamma, \eta_{i1} < -\alpha_1 - X_i\beta - Z_{i1}\gamma - \rho\,\varepsilon_{i0}) \\ &= \int_{-\alpha_0 - X_i\beta - Z_{i0}\gamma}^{\infty} \Phi(-\alpha_1 - X_i\beta - Z_{i1}\gamma - \rho\,\varepsilon_{i0}) \frac{\phi(\varepsilon_{i0}/\sigma_0)}{\sigma_0}\, d\,\varepsilon_{i0} \end{aligned}$$

We now consider $P(T_i = 3 \mid X_i, Z_i, \theta)$. This corresponds to the loan not being repaid in $t = 0$ or $t = 1$ and being repaid in $t = 2$. For compactness, we define $b_0 = -\alpha_0 - X_i\beta - Z_{i0}\gamma$,

$b_1 = -\alpha_1 - X_i\beta - Z_{i1}\gamma$, and $b_2 = -\alpha_2 - X_i\beta - Z_{i2}\gamma$. We deduce the following:

$$\begin{aligned}
P(T_i = 3 \mid X_i, Z_i, \theta) &= P(\alpha_0 + X_i\beta + Z_{i0}\gamma + \varepsilon_{i0} \geq 0, \alpha_1 + X_i\beta + Z_{i1}\gamma + \varepsilon_{i1} \geq 0, \\
&\quad \& \ \alpha_2 + X_i\beta + Z_{i2}\gamma + \varepsilon_{i2} < 0) \\
&= P(\varepsilon_{i0} \geq b_0, \varepsilon_{i1} \geq b_1, \varepsilon_{i2} < b_2) \\
&= P(\varepsilon_{i0} \geq b_0, \varepsilon_{i1} \geq b_1, \eta_{i2} < b_2 - \rho\,\varepsilon_{i1}) \\
&= \int_{b_0}^{\infty} \int_{b_1}^{\infty} \Phi(b_2 - \rho\,\varepsilon_{i1})\phi(\varepsilon_{i1} - \rho\,\varepsilon_{i0}) \frac{\phi(\varepsilon_{i0}/\sigma_0)}{\sigma_0}\, d\,\varepsilon_{i1}\, d\,\varepsilon_{i0}
\end{aligned}$$

Finally, we find $P(T_i = 4 \mid X_i, Z_i, \theta)$. This corresponds to the loan not being paid off in $t = 0, 1$, or 2. The setup is very similar as the previous one, but with a flipped inequality for ε_{i2}:

$$\begin{aligned}
P(T_i = 4 \mid X_i, Z_i, \theta) &= P(\alpha_0 + X_i\beta + Z_{i0}\gamma + \varepsilon_{i0} \geq 0, \alpha_1 + X_i\beta + Z_{i1}\gamma + \varepsilon_{i1} \geq 0, \\
&\quad \& \ \alpha_2 + X_i\beta + Z_{i2}\gamma + \varepsilon_{i2} \geq 0) \\
&= P(\varepsilon_{i0} \geq b_0, \varepsilon_{i1} \geq b_1, \varepsilon_{i2} \geq b_2) \\
&= P(\varepsilon_{i0} \geq b_0, \varepsilon_{i1} \geq b_1, \eta_{i2} \geq b_2 - \rho\,\varepsilon_{i1}) \\
&= \int_{b_0}^{\infty} \int_{b_1}^{\infty} [1 - \Phi(b_2 - \rho\,\varepsilon_{i1})]\phi(\varepsilon_{i1} - \rho\,\varepsilon_{i0}) \frac{\phi(\varepsilon_{i0}/\sigma_0)}{\sigma_0}\, d\,\varepsilon_{i1}\, d\,\varepsilon_{i0}
\end{aligned}$$

Combining these and using the above definitions of b_0, b_1, and b_2, the likelihood is given by the following:

$$P(T_i \mid X_i, Z_i, \theta) = \begin{cases} \Phi(b_0/\sigma_0) & \text{if } T_i = 1 \\ \int_{b_0}^{\infty} \Phi(b_1 - \rho\,\varepsilon_{i0}) \frac{\phi(\varepsilon_{i0}\,/\sigma_0)}{\sigma_0}\, d\,\varepsilon_{i0} & \text{if } T_i = 2 \\ \int_{b_0}^{\infty} \int_{b_1}^{\infty} \Phi(b_2 - \rho\,\varepsilon_{i1}) \phi(\varepsilon_{i1} - \rho\,\varepsilon_{i0}) \frac{\phi(\varepsilon_{i0}\,/\sigma_0)}{\sigma_0}\, d\,\varepsilon_{i1}\, d\,\varepsilon_{i0} & \text{if } T_i = 3 \\ \int_{b_0}^{\infty} \int_{b_1}^{\infty} [1 - \Phi(b_2 - \rho\,\varepsilon_{i1})] \phi(\varepsilon_{i1} - \rho\,\varepsilon_{i0}) \frac{\phi(\varepsilon_{i0}\,/\sigma_0)}{\sigma_0}\, d\,\varepsilon_{i1}\, d\,\varepsilon_{i0} & \text{if } T_i = 4 \end{cases}$$

$$\text{with } b_0 = -\alpha_0 - X_i\beta - Z_{i0}\gamma$$

$$b_1 = -\alpha_1 - X_i\beta - Z_{i1}\gamma$$

$$b_2 = -\alpha_2 - X_i\beta - Z_{i2}\gamma$$

Problem 1.

Solution: We wrote a routine which evaluates the log-likelihood function using Gaussian quadrature. We used nodes and weights with precision of 20. We find that the log-likelihood of the initial parameter vector with $\alpha_0 = 0$, $\alpha_1 = -1$, $\alpha_2 = -1$, β_0, $\gamma = 0.3$, and $\rho = 0.5$ is equal to -13598.877. □