In the name of God

Sharif University of Technology

Department of Economics

Quantitative Economics

Problem Bank

1 Law of Large Numbers (LLN)

As you remember from your statistics classes, LLN tells us when sample averages will converge to their population means. The strongest version of the classical LLN, known as Kolmogorov's strong law, states that if $X_1, \ldots, X_n$ are independent and identically distributed scalar random variables, with common distribution F, with common mean of μ:

$$\mu = E[x] = \int xF(dx)$$

and sample average

$$\bar{X}_n = \sum_{i=1}^{n} X_i$$

moreover, if $E[|x|]$ is finite, then

$$P\{\bar{X}_n \to \mu \text{ as } n \to \infty\} = 1$$

A weaker version of the LLN is as follows: If $X_1, \ldots, X_n$ are i.i.d. with $E[X_i^2] < \infty$, then, for any $\epsilon > 0$, we have

$$P\{|\bar{X}_n - \mu| \geq \epsilon\} \to 0 \text{ as } n \to \infty \tag{1}$$

(This version is weaker because we claim only convergence in probability rather than almost sure convergence, and assume a finite second moment)

1.1 LLN in action

Draw $N = 100$ samples from $X \sim N(5, 3)$. Use *for*-loop to find average of the first i samples, for all i. Then draw sample average and population average.

Now do the same exercise without using *for*-loop and set $N = 1000$. (Hint: use lower triangular matrix of ones.) Then repeat this for the following population distributions: $X \sim Uniform(0, 1)$, $X \sim Gamma(5, 1/2)$, $X \sim LogNormal(1, 2)$. Plot the histogram and the kernel density function for each.

1.2 LLN fails?

Try to repeat the same exercise with a Cauchy distribution: $f(x) = \frac{1}{\pi(1+x^2)}$. Does the sample average converge to the population mean? Does the *population mean* even exist? Increase the number of observations to $N = 10^4$. Do you see any convergence? Interpret.

1.3 True story

Assume $X_A \sim N(30, 1)$ and $X_B \sim N(-70, 1)$. The economy is randomly populated with 70 percent of type-A people and the rest are type-B people. A researcher is drawing from this population. What would be the true average and variance of the population?

Draw $N = 1000$ observations from this population. Plot the histogram and the kernel density function. Then use your code from previous part to see whether the sample average converges to the population average. Does LLN fail? Interpret.

2 CLT in action

Now do yourself!

Setup a simulation exercise and show how CLT may work.

3 (Optional) LLN and CLT

Do Exercise 1 and 2 of chapter 17 of the Sargent's Book

4 Probability Simulator

In this problem we want to see how simulation can help us to find the area of bounded curve. To do so, suppose we have a function $y = f(x)$ defined in $[0, 1]$.

1. Use Monte-Carlo simulation method by drawing samples and calculate the area under the curve $f(x)$.

2. Use your code and find an estimation for number π.

5 Monte-Carlo in Higher Dimensions

1. Suppose x_i for $i = 1..N$ has a normal distribution $N(0, 1)$. Calculate $E\left[\sum_{i=1}^{N} x_i^2\right]$.

2. Use Monte-Carlo method and calculate $\int_{z=0}^{1}\int_{y=0}^{1}\int_{x=0}^{1}\sqrt{x^2+3y^2+5z^2}dxdydz$. How many samples are needed to receive a 10^{-3} preciseness.

6 Simulating a Household

In this problem we would like to see how simulations can help us address aggregation issues in macroeconomics where heterogeneity exists or simple classic aggregation assumptions do not hold.

Consider Household i (HH_i) with preferences over consumption and leisure

$$U_i = \frac{C^{1-\sigma}}{1-\sigma} - \gamma\frac{l^{1+\phi}}{1+\phi}$$

and maximizes it subject to

$$pC = wl + T$$

where l is the labor supplied. Take $T = 1$.

1. Solve the HH_i's problem analytically.

2. Assume that $\left(\log \frac{w}{p}\right)$ is randomly distributed such that $\left(\log \frac{w}{p}\right)\tilde{\ }N(5,1)$. Again simulate this economy and find the aggregate consumption. You can take $\sigma = 0.5, \phi = 1.3, \gamma = 1.1$. Plot the histogram of consumptions in this economy.

3. Now instead, assume that ϕ is randomly distributed such that $\phi\tilde{\ }Uniform(0.2, 3.2)$. Again simulate this economy and find the aggregate consumption. You can take $\sigma = 0.5, \gamma = 1.1, \log \frac{w}{p} = 5$. Plot the histogram of consumptions in this economy.

4. Now assume all $\frac{w}{p}, \phi$ are random with the above distributions and are independent of each other. Simulate this economy and find the aggregate consumption. You can take $\sigma = 0.5, \gamma = 1.1$. Plot the histogram of consumptions in this economy.

5. Now redo part2, with the assumption that T has an independent normal distribution $N(2,3)$.

7 Apple-Orange choice of a HH

Consider a household that maximizes preference

$$U = \alpha \log X + (1-\alpha) \log Y$$

subject to the B.C.:

$$pX + qY = I$$

where X, Y are apple and oranges consumed by a household.

1. Solve the HH problem analytically.

2. Now we want to simulate the economy. Suppose α and I have beta and log-normal distributions respectively. Simulate household i's behavior and calculate the aggregate consumption of apples and oranges in the economy.

8 Regression: Behind the Scene

Recall what you have learned about ordinary least squares and what we reviewed in class. In this problem you will estimate OLS coefficients and its standard errors in a number of ways. You will experiment with how sample size, optimizer options, and other parameters change the estimates. But first we need to fabricate a data generating process (DGP).

8.1 Data Generating Process

Write a Julia/Python code to fabricate the following process:

$$y = 2 + 3x_1 - 5x_2 + \epsilon$$

where $x_1 \sim N(20, 10)$, $x_2 \sim Binomial(10, .5)$ and $\epsilon \sim N(0, 1)$ and set sample size 200. To make sure that we all work with exactly the same data, set the random generator's seed number equal to 1395, then mean of y should be 37.04900170095231. Draw the following histograms.

8.2 Ordinary Least Squares (OLS)

Write down the minimization problem that OLS solves. Then write down the first order condition and solve for β. Hint: $\hat{\beta} = (X'^{-1}X'y$. What are the dimensions of y_i, X_i and β?

8.3 Maximum Likelihood (ML)

Here we want to derive a formula for ML estimator and show that it is similar to OLS estimator in a very special case. Assume $\epsilon_i \sim N(0, \sigma^2)$.

1. write the (log) likelihood contribution for a single individual observation i.
2. Write (log) likelihood for the whole sample.
3. Write the optimization problem which characterizes the maximum likelihood. Highlight why $\hat{\beta}_{ML} = \hat{\beta}_{OLS}$. Hint: $\max_{\beta,\sigma^2} LL(\beta, \sigma^2|y, X)$.
4. How many variables are being optimized? Is this a constraint optimization? Why?

5. Write the first order condition for σ^2 and derive a formula for standard error of $\hat{\beta}_{ML}$.

8.4 Estimation

Here we want to estimate β using the following 4 methods:

1. Julia/Python method for OLS regression: Call it $\hat{\beta}_{py}$
2. Use algebra to find $\hat{\beta}_{algebra} = (X'^{-1}X'y$
3. Use Julia/Python optimizer to solve for $\hat{\beta}_{optim}$ that minimizes the sum of squared residual for the sample: $\sum_i (y_i - X_i\beta)^2$. You may use `lambda` functions in Julia/Python to define SSR as a function of β. Then find $\hat{\beta}$.
4. Use Julia/Python optimizer to maximize the likelihood function in order to find $\hat{\beta}_{ML}$. Again you should define a `lambda` function in for log likelihood function, then maximize it.

Use sample size of $n = 10, 20, 50, 100, 1000, 1000$ to estimate β and plot $[\hat{\beta}_{py}, \hat{\beta}_{algebra}, \hat{\beta}_{optim}, \hat{\beta}_{ML}]$ as a function of n. You should draw 3 plots for intercept, β_1 and β_2. Interpret.

9 Monte-Carlo Simulation

We will later study Monte-Carlo simulations and their applications in this course. But here we just want to introduce this useful method. Statistical inference (significance levels and hypothesis testing) relies heavily on asymptotic properties of estimators. Most of the important results are actually implications of Law of Large Numbers (LLN) and Central Limit Theorem (CLT) that we studied in the first problem set. I hope that you are convinced that LLN and CLT actually hold in practice, but then may need really large numbers. For small samples analytical results are rarely available, aside from tests of linear restrictions in the linear regression model under normality. Small-sample results can nonetheless be obtained by performing a Monte Carlo study. In this exercise we simulate a simple Monte Carlo study to depict small sample properties of an OLS estimator. Our goal is to first see how Monte Carlo simulations work and then to assure that asymptotic properties of OLS are actually work, but in asymptotic sense!

9.1 Small-Sample Properties

Imagine the true data generating process is $y = 3 + 5x_1 - 2x_2 + \epsilon$. Each time we draw x and y variables, estimated β parameters are themselves random variables. Assume $x_1 \sim N(7, 3)$, $x_2 \sim Binomial(5, .4)$ and $\epsilon \sim N(0, 1)$.

What is the mean and standard deviation of these random variables? (Hint: short answer is that we don't know!) Why are we interested in these statistics?

Monte Carlo simulation is a solution to find properties of the estimators. Here is how to run MC simulation: For $R = 10,000$ times, draw a sample of size $N = 50$ observations from this DGP. Then use OLS to estimate parameters of the model, $(\beta_0, \beta_1, \beta_2)$, while we know that the true parameters are $(3, 5, -2)$. Then draw histogram of your estimated parameters and compare them with expected normal distribution.

Even for this very simple model, it is not easy to compute small sample properties of the OLS estimators. You need all regressors to be distributed normally to have nice distribution. What is the implied distribution of the estimators in this example? Compare the simulated distributions with Normal distribution.

In general, it would be impossible to compute small sample properties of more complicated estimators. These so called "Monte-Carlo simulations" help us to calculate not only mean and standard deviation of estimators, but the complete distribution of estimators.

9.2 Asymptotic versus Small Sample

For the model we studied so far, increase sample size from 40 to 10,000 and compare the distribution of estimated parameters. Interpret.

9.3 True Size of Test

Hypothesis tests lead to either rejection or non-rejection of the null hypothesis. Correct decisions are made if H_0 is rejected when H_0 is false or if H_0 is not rejected when H_0 is true.

There are also two possible incorrect decisions: (1) rejecting H_0 when H_0 is true, called a type I error, and (2) non-rejection of H_0 when H_0 is false, called a type II error. Ideally the probabilities of both errors will be low, but in practice decreasing the probability of one type of error comes at the expense of increasing the probability of the other. The classical hypothesis testing solution is to fix the probability of a type I error at a particular level, usually 0.05, while leaving the probability of a type II error unspecified. Define the size of a test or significance level

$$\alpha = \Pr[\text{type I error}] = \Pr[\text{reject } H_0 | H_0 \text{ true}]$$

The power of a test is defined to be

$$\text{Power} = 1 - \Pr[\text{type II error}] = \Pr[\text{reject } H_0 | H_a \text{ true}]$$

Now think about the 10,000 draw from the DGP and imagine that each time you want to test the null hypothesis that $\beta_1 = 5$. Each time you simulate the data, you may use t-test for this hypothesis. What is the true size and power of this test? If you know the distribution of the error terms you may find analytical solution for the power and size of the test. Or if you have large enough sample, you may use the asymptotic theories to determine the size of the test. The problem arises when neither are available. For instance, in this case you have only $N = 50$ observations, not large enough to use asymptotic properties, and you do not know the distrbution of the error term. Therefore, you need Monte Carlo Simulations to find the true size and power of the test.

If you simulate the DGP enough times, then the **true size** or **actual size** of the test statistic is simply the fraction of replications for which the test statistic falls in the rejection region. Ideally, this is close to the nominal size, which is the chosen significance level of the test. For example, if testing at 5% the nominal test size is 0.05 and the true size is hopefully close to 0.05. The **power** of a test is calculated as the fraction of replications for that the null hypothesis is rejected.

9.4 Number of Replications

Numerous simulations are needed to determine actual test size, because this depends on behavior in the tails of the distribution rather than the center. If R simulations are run for a test of true size α, then the proportion of times the null hypothesis is correctly rejected is an outcome from R binomial trials with mean α. What is the variance of this binomial trial?

Find the 95% interval for the test size α? Hint: You may use CLT.

Then show that a mere 100 simulations is not enough since, for example, this interval is $(0.007, 0.093)$ when $\alpha = 0.05$. For $10,000$ simulations the 95% interval is much more precise, equalling $(0.008, 0.012)$, $(0.046, 0.054)$, $(0.094, 0.106)$, and $(0.192, 0.208)$ for α equal to, respectively, 0.01, 0.05, 0.10, and 0.20. This is why in this example we used $R = 10,000$ simulations.

9.5 Endogeneity

So far, we assumed that x_1 and x_2 are exogenous. Here we will study the problem that arises from endogeneity. Imagine that there is an exogenous variable that affect both x_2 and the y. Think about wage regression, in which the left hand side variable is wage of workers and x_2 is years of schooling. We know that many unobservable factors[1] such as ability, family properties,

[1] Some econometricians, prefer to call this term *unobserved* as some of them are intrinsically *observable* but in practice are not observed for the econometricians. For instance some family properties are generally observable but may not be accessible for the researcher.

etc, affect both. Call these variables z. Hence, imagine that the true data generating process is

$$
\begin{aligned}
x_1 \sim& N(7,3) \\
z \sim& \text{ Binomial}(20, .7) \\
x_2 \sim& \text{ Binomial}(3z, .6) \\
\epsilon =& 11z + \mu_2 \\
y =& 3 + 5x_1 - 2x_2 + \epsilon
\end{aligned}
$$

where μ_1, μ_2 and ϵ are assumed to independently drawn from standard normal distribution. However, you as the researcher, only observe (y, x_1, x_2) for each worker. Repeat paets 1, 2 and 3 for this dgp. Compare and discuss your results.

What is the problem of OLS estimator? Does this problem get solved asymptotically? (i.e. by increasing sample size, N.)

10 Simulator Class

10.1 Monte-Carlo Simulation

1. Write a Simulator class in Julia/python that can run Monte-Carlo simulation for a given function using a determined pdf.

10.2 Frequency Simulator Class

1. Now inherit a Frequency Simulator class from your Simulator class that can run a frequency simulator.

10.3 Important Sampling Simulator Class

1. Now inherit a Frequency Simulator Simulator class from your Simulator class that can run a frequency simulator.

11 Linear Regression Model

Assume the linear regression model $y = \mathbf{x}\beta + u$, where $\mathbf{x}$ and β are $K \times 1$ vectors and y and u are scalars. Suppose the error term u has zero mean conditional on regressors.

11.1 OLS as GMM

Use law of iterated expectations to show that the single conditional moment restriction $E[u|\mathbf{x}] = 0$ leads to K unconditional moment conditions $E[\mathbf{x}u] = \mathbf{0}$. Then please show that OLS is special case of GMM. (i.e. specify the $g(.)$ function and the minimization problem then show that its solution is the same as OLS estimator, $\hat{\beta}_{OLS} = \frac{\mathbf{x}'y}{\mathbf{x}'\mathbf{x}}$.)

11.2 Additional Moment Restrictions

Using additional moments can improve the efficiency of estimation but requires adaptation of regular method of moments if there are more moment conditions than parameters to estimate.

A simple example of an inefficient estimator is the sample mean!

$$\hat{\mu} = \frac{1}{N}\sum_i x_i$$

This is an inefficient estimator of the population mean unless the data are a random sample from the normal distribution or some other member of the exponential family of distributions. One way to improve efficiency is to use alternative estimators. The sample median, consistent for μ, if the distribution is symmetric, may be more efficient. Obviously the MLE could be used if the distribution is fully specified, but here we instead improve efficiency by using additional moment restrictions.

Consider estimation of β in the linear regression model. The OLS estimator is inefficient even assuming homoskedastic errors, unless errors are normally distributed. In the previous section you showed that the OLS estimator is an MM (method of moment) estimator based on $E[\mathbf{x}u] = 0$. Now make the additional moment assumption that errors are conditionally symmetric, so that $E[u^3|\mathbf{x}] = 0$ and hence $E[\mathbf{x}u^3] = \mathbf{0}$. How many parameters are to be estimated? How many moments do we have? Please explain why we can not use method of moments anymore.

12 Instrumental Variables Regression

Instrumental variables estimation is a leading example of generalized method of moments estimation. Consider the linear regression model $y = \mathbf{x}\beta + u$, with the complication that some components of $\mathbf{x}$ are correlated with the error term so that OLS is inconsistent for β.

12.1 Exclusion Restriction

Assume the existence of instruments $\mathbf{z}$ that are correlated with $\mathbf{x}$ but satisfy $E[u|\mathbf{z}] = 0$. Assume that $\mathbf{z}$ is $J \times 1$ vector. Then $E[y - \mathbf{x}'\beta|\mathbf{z}] = 0$. Recall the exclusion restriction and please explain these conditions in 3 sentences. You may graphically explain the exclusion restriction assumption.

12.2 Method of Moment

Again use law of iterated expectations to show that the single conditional moment restriction $E[u|\mathbf{z}] = 0$ leads to J unconditional moment conditions

$$E[\mathbf{z}(y - \mathbf{x}'\beta)] = \mathbf{0}$$

The method of moments estimator solves the corresponding sample moment conditions. Please write down those sample moment conditions. Under what condition, these equations have a unique solution? The solution is called IV estimator.

12.3 GMM

No unique solution exists if there are more potential instruments than regressors. Please explain why? One possibility is to use just K instruments only, but there is then an efficiency loss. The GMM estimator instead chooses β to make the sample moments as small as possible using quadratic loss. Using a weighting matrix $\mathbf{W}$ write down the quadratic loss function that GMM solves. What is the size of the weighting matrix $\mathbf{W}$?

13 Duration of Unemployment and Hazard of Job Finding

One of the most important macroeconomic indicators is the average duration of unemployment. Longer periods of joblessness are an indicator of poor labor market policies. It is crucial in labor and macroeconomics to study possible causes and effects of long periods of unemployment. In this problem we study hazard rate models and estimate a proportional hazard model using MLE and GMM methods. But first we need some new definitions.

13.1 Mean duration and Survivor Function

The cumulative distribution function of T is denoted $F(t)$ and the density function is $f(t) = dF(t)/dt$. Then the probability that the duration or spell length (unemployment) is less than t is

$$F(t) = \Pr[T \leq t] = \int_0^t f(s)ds$$

A complementary concept to the cdf is the probability that duration equals or exceeds t, called the survivor function , which is defined by

$$S(t) = \Pr[T \geq t] = 1 - F(t)$$

The survivor function is monotonically declining from one to zero since the cdf is monotonically increasing from zero. If all individuals at risk of leaving unemployment eventually do so then $S(\infty) = 0$. Otherwise, $S(\infty) > 0$ and the duration distribution is called defective. Please show that the sample mean of unemployment duration equals the area under the survival curve:

$$E[T] = \int_0^\infty S(u)du$$

13.2 Hazard Function

Another key concept is the hazard function, which is the instantaneous probability of leaving a state conditional on survival to time t. For instance, conditional on being unemployed for 3 months, what is the probability of job finding? This is defined as

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \le T \le t + \Delta t | T \ge t]}{\Delta t}$$

Please show that

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{d \ln(S(t))}{dt}$$

In regression analysis of transitions the conditional hazard rate, $\lambda(t|\mathbf{x})$, is of central interest. This contrasts with more standard regression approaches in which the conditional mean function, $E[T|\mathbf{x}]$, is of chief interest. The latter approach has the disadvantage that in practice the durations are often censored (i.e some workers are still unemployed at the time of the study, or some of them leave the labor market).

13.3 Proportional Hazards Model

In a proportional hazard model, the conditional hazard rate $\lambda(t|\mathbf{x})$ is factored into separate functions of

$$\lambda(t|\mathbf{x}) = \lambda_0(t, \alpha)\phi(\mathbf{x}, \beta)$$

where $\lambda_0(t, \alpha)$ is called the baseline hazard and is a function of t alone, and $\phi(\mathbf{x}, \beta)$ is a function of $\mathbf{x}$ alone. Usually $\phi(\mathbf{x}, \beta) = e^{\mathbf{x}'\beta}$. Thus the hazard for each individual has the same shape over time, but is shifted up or down proportionally by the covariates $\mathbf{x}_i$. Please explain the structural assumptions we are making and indicate the parameters to be estimated for the proportional hazards model. Also suggest methods for estimating the parameters.

13.4 GMM

Suppose that we observe a sample of 200,000 workers who find job after no later than 24 months after jobloss. For simplicity we assume that there is no censoring (i.e. all workers find job in less than 24 months and no one leave the labor market while unemployed). Hence $1 \le t \le 24$. Also assume that from an economic structural model we derived the following baseline hazard function

$$\lambda_0(t, \alpha) = \frac{1}{t^{\alpha}}$$

13.4.1 Baseline Hazard

Please interpret this function and specify role of the parameter α. What do you think about proper range for α. Use your economic intuition to answer this questions. For instance, do you expect increasing, decreasing or constant baseline hazard function?

13.4.2 Estimation: MLE not possible

Moreover, we observe the following covariates for each worker: years of schooling before jobloss, gender, GDP growth rate at the time of jobloss, unemployment rate at the time of jobloss, job vacancy rate at the industry and time of jobloss. Assume

$$\phi(\mathbf{x}_i, \beta) = \exp[\beta_0 + \beta_1 school_i + \beta_2 gender + \beta_3 ur_i + \beta_4 \gamma_i + \beta_5 vac_i] \tag{2}$$

Please explain why MLE is the ideal estimator, but it is not feasible with this assumptions. Your answer should not be more than one sentence!

13.4.3 Moments

While you can not use MLE, still you can use GMM. But first you should form the moments you want to use. For simplicity, assume that the only covariate is schooling. Furthermore, assume that there are only two levels for schooling: 0 for no college and 1 for any college degree. Hence parameters to be estimated are $\Theta = [\alpha, \beta_0, \beta_1]$. What is the smallest number of moments you need for MM estimation?

Now please find the $\lambda(t = s|school = 1)$ and $\lambda(t = s|school = 0)$ for all $1 \leq s \leq 24$ as a function of the parameters of the model. Then please identify the sample counterpart of these two terms to form 48 moments you may use in GMM. Write down the minimization problem.

13.4.4 Estimation: GMM

Use the synthesized data, stored in `hazard.dta` to estimate $\Theta = [\alpha, \beta_0, \beta_1]$ using GMM. Notice that there is not a unique way for this estimation. You may decrease number of moments by categorizing duration of unemployment. For instance, you may use $\Pr[T \leq 12|school]$ and $\Pr[T > 12|school]$ as 4 independent moments instead of the 48 moments we introduced in previous section.

13.5 Weibull Proportional Hazard Model

Now assume that the baseline duration of unemployment has a Weibull distribution:

$$F(t) = 1 - e^{-\gamma t^{\alpha}}$$

Please use the earlier definitions to find surviver and hazard functions. Form the log likelihood function and derive the first order conditions. (optional) Then use MLE to estimate $\Theta = (\gamma, \alpha, \beta_0, \beta_1, \beta_2)$. Maximum likelihood can be viewed as the GMM estimator with moment conditions equal to the first derivatives of the likelihood with respect to Θ. (these are called the score equations).

You may use Stata's `streg` command to make sure that your estimates are accurate. Now estimate the full specification in equation 2. What is the economic interpretation of your results.

14 Monte-Carlo Simulation of firms behavior (1)

In this problem we want to study how firms behave in their exportation decisions.

To do so, suppose a firm employs labor l at wage w and produces the output $y = Al^{\alpha}$ where A is chosen from a probability distribution function (PDF) that is Pareto. In other words

$$P(A \geq a) = \left(\frac{a}{\bar{A}}\right)^{\theta} \text{ for } a \geq \bar{A}$$

and otherwise zero.

The firm sells its good in a monopolistic competition environment. Specifically a firm has a monopoly power over its good with price p and the demand for it is $C = Y\left(\frac{p}{P}\right)^{-\sigma}$. where Y is the aggregate demand and P is the aggregate price. The firm needs to pay a fixed cost f_o to operate. If its profit is negative, it does not operate.

1. Setup a firm's problem..
2. Solve for the firm's price (Hint: it's constant markup over marginal cost.
3. Determine the threshed $(A_o)$the a firm enters the domestic market.
4. Now write a code to simulate the firms' behavior using the Monte-Carlo Simulation.
5. Suppose the number of firms is $N = 1000$. Set $\theta = 3.5, \bar{A} = 1, \sigma = 2.4, P = 1, Y = 100, \alpha = .5, f_o = 1.2,$. Generate N samples of firms and save them in your memory.
6. Plot the distribution of these simulated firms' sales, exports.
7. Now suppose that as an econometrician, we observe these firms' characteristics like Sale, production, export, employment. We want to calibrate the model and estimate the parameters. Use SMM to estimate the parameter α. Calibrate the other parameters as above. What are the differences? How many samples do you need so that the model is not rejected in your SMM? Discuss about the power of the esti[redacted]ation.
    (a) Use these moments: the average sales of the firms, average labor productivities.
    (b) Use these moments: the average sales of the firms, average labor productivities, standard deviation of labor productivities.
    (c) Use these moments: average employment of producers, and average employment payment to value added.
    (d) Use these moments: average employment of producers, and average employment payment to value added, standard deviation of employment of producers.
    (e) Use these moments: the average sales of the firms, average employment of domestic producers, average labor productivity, and average employment payment to value added.

(f) Use these moments: the average sales of the firms, average employment of domestic producers, average labor productivity, average employment payment to value added, standard deviation of labor productivities, standard deviation of employment of producers.

8. Now use the estimation-verification method and use one or two moments for testing. Does your model get rejected?

9. Redo 7_a and 7_b but assume that the econometrician mistakenly takes $\theta = 3$. What happens to your results? Does your model get rejected? What if there is a mistake in f_o too such that he takes it $f_o = 1.8$.

10. (Optional) Use GMM and repeat (7).

11. Now assume that the true PDF is Pareto but the econometrician mistakenly assumes it log-normal ($\log A\tilde{\ }N\left(\bar{A},\theta\right)$) and tries to estimate the model. Does the model get rejected? Discuss about the power of the estimation.

12. Now repeat the above exercise (7) and (9), but suppose you don't know f_0 , θ and they should be estimated along with α, too.

15 Consumption-Saving with Fixed Costs of Borrowing

Consider a household consumption-saving problem: period utility is CRRA with parameter σ. HH has asset a, labor income $wl = 1$, and decides to consume and save (or borrow).

HH faces a stochastic return R on his asset saving such that $R = \bar{R} + \mu$ with probability $p \geq 0.5$ and $R = \bar{R} - \mu$ with probability $1 - p$. Make an assumption about β the HH preference dicount factor such that you are confortable and have finite answers.

1. Setup the bellman equation. (Be careful)

2. Write the FOCs.

3. Numerically solve the Bellman equation.

4. What happens to the HH saving (Policy Function) if μ increases. Plot HH's saving and consumption for $\bar{R} = 0.05$ and μ varies in $(0, 0.05)$.

5. Find the stationary distribution of assets and consumption.

6. Simulate an agent behavior over time using Monte-Crlo Simulation.

   Now suppose that borowing has a fixed cost; such that HH should pay f if he want to borrow in this period.

7. Redo the problem under this new senario.How does the policy function look like?.

8. Now additionaly suppose that the HH faces a borrowing constraint such that $a \geq -B$. Now solve the above problem.

16 Logit derivation

Goal of this question is to study properties of Logit function and its derivation.

16.1 Gumbel distribution

Use the pdf and CDF function for Gumbel distribution to derive the following graph in Python. Recall that for type I extreme value (Gumbel) distribution we have:

$$g(x) = e^{-x}e^{-e^{-x}}$$
$$G(x) = e^{-e^{-x}}$$

and mean is γ, the Euler-Mascheroni constant, and variance is $\frac{\pi^2}{3}$.

16.2 Gumbel - Gumbel = Logistics ≈ Normal!

Show that the difference between two extreme value variables is distributed logistic. That is, if ϵ_{ni} and ϵ_{nj} are i.i.d. extreme value, then $\epsilon = \epsilon_{ni} - \epsilon_{nj}$ follows the logistic distribution:

$$f_\epsilon(s) = \frac{e^s}{(1+e^s)^2}$$
$$F_\epsilon(s) = \frac{e^s}{1+e^s}$$

Hint:

$$\begin{aligned}
F_\epsilon(s) &= \text{Prob}(\epsilon_{ni} - \epsilon_{nj} < s) = \text{Prob}(\epsilon_{ni} < s + \epsilon_{nj}) \\
&= \int_{\epsilon_{nj}} \text{Prob}(\epsilon_{ni} < s + \epsilon_{nj} | \epsilon_{nj}) g(\epsilon_{nj}) d\epsilon_{nj} \\
&= \int_{\epsilon_{nj}} G(s + \epsilon_{nj}) g(\epsilon_{nj}) d\epsilon_{nj} \\
&= \int_{\epsilon_{nj}} e^{-e^{-(s+\epsilon_{nj})}} e^{-\epsilon_{nj}} e^{-e^{-\epsilon_{nj}}} d\epsilon_{nj} \\
&\dots \\
&= \frac{e^s}{1+e^s}
\end{aligned}$$

Now draw pdf and CDF for Logistic distribution and compare it with Normal. Convince yourself that the difference between extreme value and independent normal errors is indistinguishable empirically. (Note: Don't forget the variance of the logistic function. What is the variance of Gumbel - Gumbel, if they are independent?)

Finally, explain why binary Logit model is derived as follows:

$$P_{n1} = \frac{e^{(X_{n1}-X_{n0})\beta}}{1+e^{(X_{n1}-X_{n0})\beta}}$$

$$P_{n0} = 1 - P_{n1}$$

16.3 Derivation of Multinomial Logit Formula

Show that if unobserved component of utility is distributed i.i.d. extreme value (i.e. Gumbel) for each alternative, then the choice probabilities take the form of:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

Hint: Start from the conditional probability $P_{ni}|\epsilon_{ni}$ that is the probability of alternative i is chosen given the unobserved utility of choice i, and use independence assumption to show that

$$P_{ni}|\epsilon_{ni} = \prod_{j \neq i} e^{-e^{-(\epsilon_{ni}+V_{ni}-V_{nj})}}$$

Then use the same method used for binary Logit model in problem 1.2 to derive the multinomial Logit formula of equation **??**.

16.4 Derivatives and Elasticities

How does choice probability P_{ni} responds to a change in one of the characteristics of i's alternative? For instance, we want to predict how market share of a product responds to improvement in its performance. To answer these sort of questions we need derivatives and elasticities. For the Logit model of equation (**??**) show that:

$$\frac{\partial P_{ni}}{\partial z_{ni}} = \frac{\partial V_{ni}}{\partial z_{ni}} P_{ni}(1-P_{ni})$$

where V_{ni} depends on z_{ni}.

However, economists often measure response by elasticities rather than derivatives, since elasticities are normalized for the variables' units. An elasticity is the percentage change in one variable that is associated with a one-percent change in another variable. Show that the elasticity of P_{ni} with respect to z_{ni} , a variable entering the utility of alternative i, is

$$\begin{aligned} E_{i,z_{ni}} =& \frac{\partial P_{ni}}{\partial z_{ni}} \frac{z_{ni}}{P_{ni}} \\ =& \frac{\partial V_{ni}}{\partial z_{ni}} z_{ni}(1 - P_{ni}) \end{aligned}$$

Another useful term is cross derivative and cross elasticities which capture responds of P_{ni} to changes in characteristics of alternative $j \neq i$. Show that:

$$\begin{aligned} \frac{\partial P_{ni}}{\partial z_{nj}} =& - \frac{\partial V_{nj}}{\partial z_{nj}} P_{ni} P_{nj} \\ E_{i,z_{ni}} =& \frac{\partial P_{ni}}{\partial z_{nj}} \frac{z_{nj}}{P_{ni}} = - \frac{\partial V_{nj}}{\partial z_{nj}} P_{ni} P_{nj} \end{aligned}$$

17 Maximum Likelihood Estimation

A sample of N decision makers is obtained for the purpose of estimation.

17.1 Log Likelihood function

Since the logit probabilities take a closed form, the traditional maximum-likelihood procedures can be applied. Show that the log likelihood function is then

$$LL(\beta) = \sum_{n=1}^{N} \sum_{i} y_{ni} \ln P_{ni} \tag{3}$$

17.2 First order condition

Show that the first order condition for the problem of $\max_\beta LL(\beta)$ is given by:

$$\sum_{n=1}^{N} \sum_{i} (y_{ni} - P_{ni}) x_{ni} = 0 \tag{4}$$

What is the interpretation of this equation? Show that the maximum likelihood estimates of β are those that make the predicted average of each explanatory variable equal to the observed average in the sample.

17.3 Goodness of fit

The likelihood ratio index is defined as

$$\rho = 1 - \frac{LL(\hat{\beta})}{LL(0)}$$

In two sentences interpret this index. Is it similar to R^2 in linear regression? For instance, can you compare likelihood ratio index from two models on two datasets to each other? (Hint: the answer is NO! Explain why.)

18 Logit in action

We analyze data on supplementary health insurance coverage. Even you can do all of this exercise in Python, but it is more efficient to do it Stata. You may want to replicate this exercise in Julia/Python later. The data come from wave 5 (2002) of the Health and Retirement Study (HRS), a panel survey sponsored by the National Institute of Aging. The sample is restricted to Medicare beneficiaries. The HRS contains information on a variety of medical service uses. The elderly can obtain supplementary insurance coverage either by purchasing it themselves or by joining employer-sponsored plans. We use the data to analyze the purchase of private insurance (ins) from any source, including private markets or associations. The insurance coverage broadly measures both individually purchased and employer-sponsored private supplementary insurance, and includes Medigap plans and other policies.

Explanatory variables include health status, socioeconomic characteristics, and spouse-related information. Self-assessed health-status information is used to generate dummy variable (`hstatusg`) that measures whether health status is good, very good, or excellent . Other measures of health status are the number of limitations (up to five) on activities of daily living (`adl`) and the total number of chronic conditions (`chronic`). Socioeconomic variables used are age, gender, race, ethnicity, marital status, years of education, and retirement status (respectively, `age, female, white, hisp, married, educyear, retire`); household income (`hhincome`); and log household income if positive (`linc`). Spouse retirement status (`sretire`) is an indicator variable equal to 1 if a retired spouse is present.

18.1 Logit Estimation

Using `logit` command in Stata, estimate a logit model for Pr(`ins` = 1) when value of having insurance depends linearly on household income (`hhincome`) and socioeconomic variables (`age, female, white, hisp, married, educyear, retire`). Interpret the estimated coefficients and the reported log likelihood.

18.2 Comparing Predicted Outcome with Actual Outcome

Now we want to compare the predicted outcome from this simple logit model with the actual outcome. Use Stata's `predict` command to store predicted values for the probability of having insurance from a simple logit model that the explanatory variable is only `hhincome`. Then estimate predicted outcome from OLS regression of the same model and compare the results. Interpret.

18.3 Logit in Python (optional)

Repeat the same exercise using Julia/Python.

18.4 Mixed Logit

Now try to replicate the exercise using a mixed logit model. Consider some relevant distribution for the parameters of interest and run the mixed logit. Use the Monte-Carlo simulation to calculate the mixed-logit terms.

Then use the Maximum simulated likelyhood to estimate the model and the parameters.

Compare your results with the Logit model.

19 Consumer Surplus

For policy analysis, the researcher is often interested in measuring the change in consumer surplus that is associated with a particular policy. For example, if a new alternative is being considered, such as building a light rail system in a city, then it is important to measure the benefits of the project to see if they warrant the costs. Similarly, a change in the attributes of an alternative can have an impact on consumer surplus that is important to assess. Degradation of the water quality of rivers harms the anglers who can no longer fish as effectively at the damaged sites. Measuring this harm in monetary terms is a central element of legal action against the polluter. Under the logit assumptions, the consumer surplus associated with a set of alternatives takes a closed form that is easy to calculate. By definition, a person's consumer surplus is the utility, in dollar terms, that the person receives in the choice situation. The decision maker chooses the alternative that provides the greatest utility. Consumer surplus is therefore $CS_n = (1/\alpha_n)\max_j(U_{nj})$, where α_n is the marginal utility of income: $dU_n/dY_n = \alpha_n$.

The researcher does not observe U_nj and therefore cannot use this expression to calculate the decision maker's consumer surplus. Instead, the researcher observes V_nj and knows the distribution of the remaining portion of utility. With this information, the researcher is able to calculate the expected consumer surplus:

$$E(CS_n) = (1/\alpha_n)E[\max_j(V_{nj} + \epsilon_{nj})]$$

19.1 log-sum term

Assume utility is linear in income. (Hence α_n is constant with respect to income). Show that if each ϵ_{nj} is i.i.d extereme value then

$$E(CS_n) = (1/\alpha_n)\ln\left(\sum_{j=1}^{J} e^{V_{nj}}\right) + C$$

where C is an unknown constant that represents the fact that the absolute level of utility cannot be measured. **Hint**: read chapter 4 of Small and Rosen (1981)[2].

[2] http://www.jstor.org/stable/1911129

Note that the argument in parentheses in this expression is the denominator of the logit choice probability. Aside from the division and addition of constants, expected consumer surplus in a logit model is simply the log of the denominator of the choice probability. It is often called the log-sum term.

19.2 Marginal Utility of Income

It is important to measure marginal utility of income, α_n, for welfare analysis. In many choice models we use price of the products as one of the attributes of the product. Do you have any suggestion for measuring α_n? Explain assumptions you need to measure marginal utility of income.

20 Nested Logit Model (Optional)

Assume the vector of $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nJ})$ has the following type of GEV cumulative distribution

$$F(\epsilon_n) = exp\left[-\sum_{k=1}^{K}\left(\sum_{j\in B_k} e^{-\epsilon_{nj}/\lambda_k}\right)^{\lambda_k}\right]$$

20.1 Marginal Distributions

What is the distribution of ϵ_{nj}? Hint: recall $F_X(x) = F_{XY}(x, \infty)$.

20.2 Marginal Joint Distributions

What is the joint distribution of the vector $(\epsilon_{ni}, \epsilon_{nj})$? What is the correlation of (ϵ_{ni} and ϵ_{nj})? Hint: your answer should depend on 4whether i and j are within the same nest.

20.3 Logit

Recall that for the logit model unobserved terms where assumed to be independent from each other. In other words, we need to have $corr(\epsilon_{ni}, \epsilon_{nj}) = 0$ for any i and j. What is the sufficient condition for a nested logit model to reduce to a simple logit model? Hint: use your answer to the previous question!

20.4 Decomposition into Two Logits

It is straightforward to use equation (**??**) to show that choice probabilities under nested logit assumptions are

$$P_{ni} = \frac{e^{V_{ni}/\lambda_k}(\sum_{j\in B_k} e^{V_{nj}/\lambda_k})^{\lambda_k - 1}}{\sum_{l=1}^{K}(\sum_{j\in B_l} e^{V_{nj}/\lambda_l})^{\lambda_l}} \tag{5}$$

Verify that $0 < P_{ni} < 1$ and $\sum_i P_{ni} = 1$.

Often it is useful to decompose the observed component of utility into two parts: $V_{ni} = W_{nk} + Y_{ni} + \epsilon_{ni}$: a part W that is constant for all alternatives within a nest and another part Y which varies over alternatives within a nest. Verify that $P_{ni} = P_{ni|B_k} P_{nB_k}$ where

$$\begin{aligned}
P_{nB_k} &= \frac{e^{W_{nk}+\lambda_k I_{nk}}}{\sum_{l=1}^{K} e^{W_{nl}+\lambda_l I_{nl}}} \\
P_{ni|B_k} &= \frac{e^{Y_{ni}/\lambda_k}}{\sum_{l=1}^{K} e^{Y_{ni}/\lambda_l}} \\
I_{nk} &= \ln \sum_{j\in B_k} e^{Y_{nj}/\lambda_k}
\end{aligned}$$

Hint: you may use the fact that $e^x b^c = e^{x + c\ln b}$.

20.5 Log Likelihood Function

For a sample $Y = (y_1, \ldots, y_N)$ of size N that independently drawn from a population which is distributed (or at least assumed to be distributed) nested logit, form the log likelihood function. Hint: you can review the last problem set!

21 The GEV family (Optional)

McFadden (1978) developed GEV models and showed under some conditions the choice probabilities are consistent with utility maximization. In this exercise we want to verify whether these conditions hold for Logit and Nested Logit models. Recall that G must satisfy:

1. $G > 0$ for all positive values of Y_j

2. G is homogenous of degree 1

3. $G \to \infty$ as $Y_j \to \infty$ for any j

4. $G_i \geq 0$, $G_{ij} \leq 0$, $G_{ijk} \geq 0$ and so on

21.1 Logit

First show that $G = \sum_{j=1} Y_j$ all these four conditions. Then show that $P_i = \frac{Y_i G_i}{G}$ is actually the Logit model that we derived in last lectures.

21.2 Nested Logit

The J alternatives are partitioned into K nests labeled $B_1, \ldots, B_K$. Let

$$G = \sum_{l=1}^{K} \left(\sum_{j \in B_l} Y_j^{1/\lambda_l} \right)^{\lambda_l}$$

Find G_i and G_{ij}, G_{ijk}. What is the sufficient condition that these derivatives satisfy the above conditions? Finally show that

$$P_i = \frac{Y_i G_i}{G} = \frac{e^{V_i/\lambda_k} (\sum_{j \in B_k} e^{V_j/\lambda_k})^{\lambda_k - 1}}{\sum_{l=1}^{K} (\sum_{j \in B_l} e^{V_j/\lambda_l})^{\lambda_l}}$$

22 A Search Model for Young Workers

As young workers enter labor market, they receive a job offer that they have to accept or reject. If they accept the job they will earn the offered wage and become employed. If they reject the offer they will stay unemployed but enjoy their leisure time. Jobs are different in two dimensions: hourly wage rate, and hours needed for the job. Employed workers remain employed forever and also unemployed workers remain unemployed forever. Unemployed workers receive an observable monetary subsidy of z as unemployment insurance. This life is really boring, but we need these assumptions to make your life much easier![3]

[3] Don't worry, we will study much more intersting life style once you become masters of *Dynamic Programing*!

Then utility of workers can be modeled with

$$U_{ni} = \alpha I_{ni} + \beta L_{ni} + \epsilon_{ni}$$

where $i = 1$ represent employment and $i = 0$ represent unemployment. Also income of employed workers are $I_{n1} = w_n h_h$ and unemployed workers' income is only the UI: $I_{n0} = z$. Their leisure are $L_{n1} = (1 - h_n)$ and $L_{n0} = 1$ for employed and unemployed ones respectively.

Of course, as researchers, we don't observe U_{ni}. What is observable is y_{ni} which is a binary variable. Imagine that we also observe all job offers (w_n, h_n) for everyone, either employed or unemployed. (This is not usually the case in practical studies. In the next problem set we will study the Roy model and selection problem that arises in empirical research.) Also notice that we assume that all workers have exactly the same taste toward income and leisure (α and β are similar for all workers). We will later relax this assumption. Young individuals accept the wage offer if and only if $U_{n1} > U_{n0}$.

To make the problem set easier, assume that $\beta = 0$.

22.1 Data Generating Process

Let's create such economy. Set random generator's seed number equal 1395 so we all get similar results. Assume ϵ_{ni} are iid and normally distributed: $\epsilon_{ni} \sim N(0, \sigma^2)$. Then use Python/Julia to simulate a sample of 1,000 workers, once the true parameters of the model are

$$\Theta \equiv (\alpha, \sigma) = (1, 3)$$

and wage offers are distributed exponentially with parameter $\lambda = .2$: $f(w) = .2e^{-.2w}$. Also $z = 12$ is observable by the researcher. Notice that here I assumed a very specific covariance matrix for error terms. We will relax this assumption later. Draw scatter plot of workers decisions versus their wage offers.

22.2 Identification

Which of the parameters of the model, α and σ, are identified? What about α/σ? Explain.

22.3 OLS, Logit and Probit

Use Python/Julia's `statsmodels` or a similar commande to fit OLS, Logit and Probit model to the decision of workers. Compare the estimated coefficient for wage rate in these three models, $\hat{\alpha}_{OLS}$, $\hat{\alpha}_{Logit}$ and $\hat{\alpha}_{Probit}$, with each other and with the true parameter of interest. Why are they different? Does it mean that logit and probit are different? Interpret.

Hint: Compute

$$\frac{3\hat{\alpha}_{Logit}}{\sqrt{\frac{\pi^2}{6}}} \text{ and } \hat{\alpha}_{Probit} \times \sqrt{2 \times 3^2}$$

and compare them with the true parameter α. Recall from problem set 2 that the Gumble distribution has variance of $\frac{\pi^2}{6}$. Finally, notice that even if the true model is not logit, but logit provides a very good approximation for the true parameters. Why?

Compare predicted probabilities with the "true" decisions. What is the problem with OLS? Do Logit and Probit have different predictions?

23 Maximum Simulated Likelihood Estimator

So far we have simulated a structural model for behavior of young workers. The next step is to estimate the parameters of the structural model. In this very simple (and boring!) model there are multiple software packages and very efficient methods to estimate the parameters of interest. Hence, it may seem useless to write our own program to do the same job. However, as it will become clear later during the course, it is very valuable pedagogically. In this problem, we write down our own Probit estimator and will compare our results with those of professional software such as Python/Julia and Stata. But first we need to understand the necessary structure needed to be imposed on the data such that parameters are identified.

23.1 Identification

Imagine that we are given data on y_n and w_n. We assume that the true model is the one that we explained in the previous problem. In other words, we impose structure on the data to better understand the relation between observed variables. Benefit of this exercise is that we already know the true parameters. So we can start from a legitimate initial guess and we expect to find estimates close enough to the true parameters. Our goal is to estimate $\Theta = (\alpha, \Omega)$ where

$$\Omega = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ . & \sigma_{11} \end{pmatrix}$$

and σ_{ij} is the covariance between the error term for alternative i and j. To take account of the fact that the level of utility is irrelevant, we take utility differences. Show that the only identified term is $\sigma^* = \sigma_{00} + \sigma_{11} - 2\sigma_{01}$. Suggest more structure to be imposed such that we can estimate the parameters. (Hint: There is not a unique answer! You just have to be able to justify and defend your identification assumption.)

23.2 Accept-Reject Simulator

One possible identification assumption is to set $\sigma_{01} = 0$ and $\sigma_{00} = \sigma_{11}$ which may not be the best assumption depending on the context, but if we make this assumption then $\sigma_{00} = \sigma_{11} = \sigma^*/2$ will be identified. For the rest of this problem let's impose this structure on the covariance matrix of error terms. Create a two-dimensional grid for (α, σ) which includes the true parameters: $(1, 9)$. For each of the points on this grid do the following:

1. Use the accept and reject simulator, discussed in class, to calculate $\check{P}_{ni}$ using $R = 100$ simulations.

2. Then use the observed data, (y_n, w_n), to calculate the simulated log likelihood. Recall that the the log-likelihood function is $\mathcal{LL} = \sum_n \sum_j d_{nj} \log P_{nj}$, where $d_{nj} = 1$ if n chose j and 0 otherwise. When the probabilities cannot be calculated exactly, as in the case of Probit, the simulated log-likelihood function is used instead, with the true probabilities replaced with the simulated probabilities: $\mathcal{SLL} = \sum_n \sum_j d_{nj} \log \check{P}_{nj}$. [4]

Then find the parameters $(\hat{\alpha}, \hat{\sigma})$ that maximizes the above simulated likelihood function on this gird. If maximum takes at a point close enough to (1,9), then congratulations! The value

[4] In our binomial Logit model where $y_n = 0, 1$ represent choice of workers whether to reject or accept the job offer, simulated log likelihood function is much easier and faster to compute:

$$\mathcal{SLL} = \sum_n y_n \log \check{P}_{n1} + (1 - y_n) \log \check{P}_{n0} = \sum_{n \in Accept} \log \check{P}_{n1} + \sum_{n \in Reject} \log(1 - \check{P}_{n1})$$

of the parameters that maximizes $\mathcal{SLL}$ is called the maximum simulated likelihood estimator (MSLE). You have successfully created your first MSLE!

23.2.1 Start Small

This process seems very easy, but I bet none of you find a smooth likelihood function which has a maximum at a point close to the desired parameters, at least in your first try! This is typically what happens in estimation of any structural model, for almost everyone. Don't get disappointed! It is very important to know how to fix this issue in real research. Here is the solution: You should start from the simplest model. For instance, in this model, let's start from one dimensional log likelihood. Imagine we know the true $\sigma^2 = 9$. Then repeat the above estimation to get log likelihood function similar to the left panel of the following figure. The right panel depicts simulated log likelihood once you take $\alpha = 1$ and try to estimate σ^2.

In order to get this result use the following:

1. Set random generator's seed number 1395.
2. Chose number of sample $N = 50$,
3. Set $R = 100000$. Smaller numbers won't work. Why? Even this R won't work for larger sample. Why?

Try to get the result for a sample with 1500 observations. What are the limitations of Accept-Reject simulator?

23.3 Smoothed Accept-Reject Simulator

Now modify your program to create the logit-smoothed accept-reject simulator. Now you should get likelihood function in much broader range. Why? Now increase the size of the sample and see if you can estimate α. Explain.

23.4 GHK Simulator (Optional)

In class we studies the GHK simulator for the case of 3 alternatives. Derive the GHK simulator for the case of 2 alternatives. Then estimate the parameters of the model. Using `timeit` command, compare the speed of your codes.

24 GHK Simulator: Choice Probabilities (Optional)

In the lecture notes, for any Probit model with three alternatives, we claimed that after using Choleski decomposition, choice probabilities could be simplified to

$$P_{n1} = \Phi(\frac{-V_{n21}}{c_a a}) \times \int_{\eta_1=-\infty}^{-\tilde{V}_{n21}/c_{aa}} \Phi\left(\frac{-(\tilde{V}_{n31} + c_{ab}\eta_1)}{c_{bb}}\right) \bar{\phi}(\eta_1) d\eta_1$$

where $\Phi(.)$ is the standard normal cdf and $\bar{\phi}(.)$ is the truncated normal density. Recall that the model was:

$$\begin{aligned} U_{nj} - U_{n1} =& (V_{nj} - V_{n1}) + (\epsilon_{nj} - \epsilon_{n1}) \\ \tilde{U}_{nj1} =& \tilde{V}_{nj1} + \tilde{\epsilon}_{nj1} \end{aligned}$$

where $\tilde{\epsilon}_{n1} = (\tilde{\epsilon}_{n21}, \tilde{\epsilon}_{n31}) \sim N(\mathbf{0}, \tilde{\Omega}_1)$ and the vector $\tilde{\Omega}_1$ is derived from Ω. And

$$L_1 = \begin{bmatrix} c_{aa} & 0 \\ c_{ab} & c_{bb} \end{bmatrix}$$

be the Choleski factor of $\tilde{\Omega}_1$: $L_1 L_1' = \tilde{\Omega}_1$. Then the original error differences, which are correlated, can be rewritten as linear functions of uncorrelated standard normal deviates:

$$\begin{aligned} \tilde{\epsilon}_{n21} =& c_{aa}\eta_1 \\ \tilde{\epsilon}_{n31} =& c_{ab}\eta_1 + c_{bb}\eta_2 \end{aligned}$$

where η_1 and η_2 are iid and $N(0,1)$. the utilities are :

$$\begin{aligned} \tilde{U}_{n21} =& \tilde{V}_{n21} + c_{aa}\eta_1 \\ \tilde{U}_{n31} =& \tilde{V}_{n31} + c_{ab}\eta_1 + c_{bb}\eta_2 \end{aligned}$$

First prove this claim, then explain why is this a progress in estimation of the Probit model. (Hint: Compare this integral with integral being simulated in AR and Smoothed AR model)

25 Random Utility Model: Mixed Logit

Now assume that the proper model for young workers is as follows:

$$U_{ni} = \alpha_n I_{ni} + \beta_n L_{ni} + \epsilon_{ni}$$

Notice that the only difference is that now the parameters of the model, namely α is different across workers. We assume that $\log(\alpha_n) \sim N(\mu, \sigma)$. Explain how to estimate μ and σ then write a Python/Julia code for estimating them.

26 Monte-Carlo Simulation of firm's behavior in exportation

In this problem qwe wantto study how firms behave in their exportation decisions.

To do so, suppose a firm employs labor l at wage w and produces the output $y = Al^{\alpha}$ where A is chosen from a probability distribution function (PDfF) that is Pareto. In other words

$$P(A \geq a) = \left(\frac{a}{\bar{A}}\right)^{\theta} \text{ for } a \geq \bar{A}$$

and otherwise zero.

The firm sells its good in a monopolistic competition environment. Specifically a firm has a monopoly power over its good with price p and the demand for it is $C = Y\left(\frac{p}{P}\right)^{-\sigma}$. where Y is the aggregate demand and P is the aggregate price. If the firm exports, there is a similar demand for it from the overseas trade partners. Thr firm needs to pay a fixed cost f_o to operate and f_x to export. If its profits is negative, it doesn not export or it may not produce if it only sells in the domestic market. Exporting requires τ percent tariffs costs.

1. Setup a firm's problem..

2. Solve for the the firm's price (Hint: it's constant mrkup over marginal cost.

3. Determine the threshod (A_o) the a firm enters the dommestic market.

4. Determine the threshod (A_x) the a firm enters the export market.

5. Now write a code in Python to simulate the firms' behavior using the Monte-Carlo Simulation.

6. Suppose the number of firms is $N = 1000$. Set $\theta = 3.5, \bar{A} = 1, \sigma = 2.4, P = 1, Y = 100, \alpha = .5, f_o = 1.2, f_x = 4$. Generate N samples of firms and save them in your memory.

7. Plot the distribution of these simulated firms' sales, exports.

8. Now suppose that as an econometrician, we observe these firms' characteristics like Sale, production, export, employment. We want to calibrate the model and estimate the parmeters. Use SMM to estimate the parameters θ, σ, f_o, f_x. Calibrate the other parameters as above. Suppose the moments are the average sales of the firms, average export if export, the fraction of firms that export, avrage employment of domestic producers.

9. Now repeat the previous problem, but now you add two other moments and re-estimate the model.

10. Again, repeat the same estimation exercise assuming that the distribution is log-normal $\log A \tilde{} N\left(\bar{A}, \gamma\right)$.(Estimate γ instead of θ).

27 Bootstrap (From Cameron)

Consider the model $y = \alpha + \beta x + \varepsilon$, where α, β, and x are scalars and $\varepsilon \sim N[0, \sigma^2]$. Generate a sample of size $N = 25$ with $\alpha = 3, \beta = 2$, and $\sigma^2 = 2$ and suppose that $x \sim N[5, 3]$. We wish to test $H_0 : \beta = 2$ against $H_a : \beta \neq 2$ at level 0.05 using the t-statistic $t = (\beta - 2)/se[\beta]$. Do as much of the following as your software permits. Use $B = 999$ bootstrap replications.

1. Estimate the model by OLS, giving slope estimate β.

2. Use a paired bootstrap to compute the standard error and compare this to the original sample estimate. Use the bootstrap standard error to test H_0.

3. Use a paired bootstrap with asymptotic refinement to test H_0.

4. Use a residual bootstrap to compute the standard error and compare this to the original sample estimate. Use the bootstrap standard error to test H_0.

5. Use a residual bootstrap with asymptotic refinement to test H_0.

28 Numerical Optimization

1. Write a code that can maximize any given function on a given interval using the

   (a) Bisection method

   (b) Gradient method

   (c) Modified gradient method with different substitutes for the Hessian.

   (d) Modified gradient method with the Variance-Covariance Matrix instead of the Hessian.

2. Now consider a simple MLE for estimation of a Logit model. Use your code to do the estimation. To do so, generate a sample data of size $N = 100$ with 3 choices and 4 control variables (as you wish, simulate it yourself). After simulating the data, use your code to maximize the Log-likelihood to estimate the coefficients of the control variables.) How is your code compared to the internal functions of Python/Julia?

3. Now repeat the previous part using Probit instead of Logit. Also use Monte-Carlo simulation to calculate the probit function. How does your answers change? Can you find the estimators?

4. Use bootstrap to estimate the standard error of the estimation, for the case of probit.

29 Value Function: Guess and Verify

1. Consider the following Bellman equation:

$$V(k) = \max_{k'} \left\{ \ln \left(Ak^a - k' \right) + \beta V\left(k'\right) \right\}$$

We guess that the solution is of the following form:

$$\begin{aligned} V(k) &= d + c\log k \\ k' &= eAk^{\alpha} \end{aligned}$$

Verify this guess and solve for c, d, e. Explain the intuitions.

2. Consider the following Bellman equation:

$$V(k) = \max_{k'} \left\{ \left(Ak^{a} - k'\right) + \beta V\left(k'\right) \right\}$$

We guess that the solution is of the following form:

$$\begin{aligned} V(k) &= cAk^{\alpha} + d \\ k' &= k^{*} \end{aligned}$$

Verify this guess and solve for c, d, k^{*}. Explain the intuitions.

30 Value Function Calculation

The Bellman equation of the planner's problem in the Neoclassical Growth Model is:

$$V(K) = \max_{K' \in \Gamma(K)} \left\{ U\left(AK^{\alpha} + p_k(1-\delta)K - p_k K'\right) + \beta V\left(K'\right) \right\}$$

where

$$\begin{aligned} \Gamma(K) &= \left[0, \frac{AK^{\alpha}}{p_k} + (1-\delta)K\right] \\ U(c) &= \frac{c^{1-\sigma}}{1-\sigma} \end{aligned}$$

Use the algorithm in the attached document and calculate the value function and the policy function $g(K) = K'(K)$.

Write your code such that it implements the Value Function Iteration and it can solve any bellman problem of the form

$$V(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta V(y)\}$$

31 Policy Function Iteration

The Bellman equation of the planner's problem in the Neoclassical Growth Model is:

$$V(K) = \max_{K' \in \Gamma(K)} \left\{U\left(AK^{\alpha} + p_k (1 - \delta) K - p_k K'\right) + \beta V\left(K'\right)\right\}$$

where

$$\begin{aligned} \Gamma(K) &= \left[0, \frac{AK^{\alpha}}{p_k} + (1 - \delta) K\right] \\ U(c) &= \frac{c^{1-\sigma}}{1 - \sigma} \end{aligned}$$

Use the **Policy Function Iteration Method** and calculate the policy function $g(K) = K'(K)$.

Write your code such that it can solve any bellman problem of the form

$$V(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta V(y)\}$$

32 Bellman Equation with Shocks

The Bellman equation of the planner's problem in the Stochastic Neoclassical Growth Model is:

$$V(y) = \max_{c} \{U(c) + \beta E[V(\xi (y - c)^{\alpha})]\}$$

where

$$U(c) = \frac{c^{1-\sigma}}{1-\sigma}$$
$$\log \xi \tilde{} N(0, \omega)$$

1. Solve the Bellman Equation using a Value Function Iteration Method

2. Solve the Bellman Equation using a Policy Function Iteration Method

3. Write a solver class such that it can solve any bellman problem of the form

$$V(y) = \max_{u \in \Gamma(y)} \{r(y, u) + \beta V(h(y, u, \xi))\}$$

where ξ follows any given distribution.

33 Transitional Dynamics

1. Use your Value Function Iteration class written the previous problem sets and solve the following Bellman equation numerically:

$$\begin{aligned} V(y) &= \max_{c} \left\{ \left(\frac{c^{1-\sigma}}{1-\sigma} \right) + \beta V(k') \right\} \\ c + k' &= Ak^{\alpha} + (1-\delta) k \end{aligned}$$

2. Write a function to calculate the half-life time (the number of periods that it takes to reach to half of the distance to the steady state. Plot the Half-Life values for $k_0 \in (0, k_{ss})$

3. Write a function to calculate the number of periods that it takes to reach to 1% distance of the steady state. Plot the results for $k_0 \in (0, k_{ss})$.

34 Log-Linearization Method

1. Use Log-Linearization Method and solve the following Bellman Equation:

$$V(y) = \max_c \left\{ \frac{c^{1-\sigma}}{1-\sigma} + \beta E\left[V\left(\xi (y-c)^{\alpha}\right)\right] \right\}$$

   where $\log \xi$ is $N(0, \omega)$.

2. Compare your results with the numerical solution form the Value Function Iteration Class.

3. Use Log-Linearization Method and solve the following Bellman Equation:

$$V(k, s) = \max_c \left\{ \frac{\left(sk^{\alpha} + (1-\delta)k - k'\right)^{1-\sigma}}{1-\sigma} + \beta E\left[V\left(k', s'\right)|s\right] \right\}$$

   such that $\log s' = \rho \log s + \varepsilon$ where ε is $N(0, \omega)$.

35 Discrete Dynamic Programming

1. Write a class in Python and implement the consumption saving problem we had in class using discrete Dynamic programming method.

2. Compare your result with the results from the Value Function Iteration Class you had in the previous problem sets, the one that use continuous maximization methods. With the same number of grids (say 100), which code is faster and how much?

3. Now write a class the uses the Discrete version with a few grids and then use its results as the initial guess for the value function iteration method. How much your value function iteration method becomes faster?

36 Numerical Methods tests

1. In the previous problem set you have solved the following Bellman Equation analytically:

$$V(y) = \max_c \left\{ \ln(c) + \beta E\left[V\left(\xi (y-c)^{\alpha}\right)\right] \right\}$$

where $\log \xi$ is $N(0, \sigma)$

Now use your Value Function Iteration class written the previous problem sets and solve the above Bellman equation numerically. Plot the iteration results for $N = 1, 10, 100, 1000, 10000$ and the analytical solution in one graph. With 10000 iteration what is the distance between the actual solution and the numerical one? How does your response depend on the variance of the shock?

2. Now use your Policy Function Iteration Class to do the same tasks. Which method converges faster?

3. Now suppose ξ has the Pareto distribution such that $F(\xi) = 1 - \left(\frac{\xi}{a}\right)^{-b}$ for $\xi \geq a$ and $F(\xi) = 0$ for $\xi \leq a$. Now use your code to solve this problem. How different it is from the previous problem. Plot the policy function for three different values of the variance of the shock (vary the distribution parameters). How does the policy function and the steady state depend on this variance?

37 Consumption Saving

1. Use your Value Function Iteration Class in the previous problem sets and solve the simple consumption saving problem with $U(c) = \ln C$ and a uniform shock with mean m and variance σ.

2. Plot the policy and value functions for $m = 1, 2, 3$. How do results depend on the variance m?

3. Plot the policy and value functions for $\sigma = .1, .2, .3$. How do results depend on the variance σ? Explain in terms of the precautionary saving. Do we observe it here?

4. Change the utility function to $U(c) = \frac{c^{1-\gamma}}{1-\gamma}$. Show how much the precautionary saving depend on γ. Explain Intuitively.

5. Calculate the stationary distribution function. (Optional)

38 LQ DP

Write a code to solve a Riccatti Matrix Equation. Then implement the Lyapanov Equation method to solve for an LQ DP.

39 Borrowing Constraint

Consider the standard borrowing constraint problem of a household who faces i.i.d income shocks and has $x = a + e$ cash at hand:

$$\begin{aligned} V(x) &= \max_{a'} u(c) + \beta E\left[V\left(a' + e'\right)\right] \\ s.t.\ \underline{a} &\leq a' = (1+r)(x-c) \\ c &\geq 0 \end{aligned}$$

Suppose the income shocks take two values e_l, e_h with probabilities $\pi, 1-\pi$.

1. Write a Python code to solve for this problem numerically by solving the FOC conditions numerically as explained in the lecture notes.
2. Assume a CRRA preferences and then numerically show how do the HH's optimal decisions change when $\underline{a}, \beta, r, \pi, E[e]$ vary and explain the economic intuiton of each.

40 Convex Adjustment Cost (1)

1. Consider a firm with convex adjustment cost. Firm's value maximization problem (Constant marginal cost, dynamic states) is

$$\begin{aligned} V(y,p) &= \max_{y'} \left\{ py - 0.5d\left(y'-y\right)^2 + \beta E\left[V\left(y',p'\right)|p\right]\right\} \\ p' &= Ap + B\xi \end{aligned}$$

   (a) Solve for the firm's Optimal Policy and the firm's value. (You can use the LQ method or any other mehtod of your interest).

(b) Solve for the steady state values as well.

(c) Solve for the Recurisve competitive Equilibrium assuming that the demand curve is linear in terms of p.

2. (Optional) Now redo the previous problem if the firm's problem is

$$\begin{aligned} V(k,p) &= \max_{y'} \left\{ pk'^{\alpha} - \left(I + 0.5dI^2\right) + \beta E\left[V\left(k',p'\right)|p\right]\right\} \\ p' &= Ap + B\xi \\ I &= k' - k \end{aligned}$$

41 (Optional) Convex Adjustment Cost (2): Numerical (From Robert Shimer)

There is a competitive industry consisting of many small firms, each endowed with a constant returns production technology using only capital. If the aggregate capital stock in the industry is K, aggregate output is K and the price of the good is $p(K,u) = ue^{-K}$, where u is a demand shifter.

Each firm maximizes the expected present value of profits, discounted at rate R^{-1}. Profits for a firm that starts a period with capital k and invests x is $p(K,u)k' - x$, where $k' = kh(\frac{x}{k})$. The firm starts the next period with capital k'. Assume $h(x) = \frac{(1+\delta)x}{\delta(1+x)}$ for some $\delta > 0$.

1. Assume that the demand curve is time-invariant, i.e. u is constant. Look for an industry steady state, where the industry capital stock is constant at K. Express K as a function of u, δ, and R.

2. Assume u follows a two-state Markov process, switching states with probability $\pi \leq \frac{1}{2}$ each period. To be concrete, let $u \in \{1,2\}, \pi = 0.2, R = 1.05$, and $\delta = 0.1$. Compute the expected present value of consumer's surplus net of investment costs as a function of the current industry capital stock K and the current value of u. On a computer, solve the planner's problem discussed in class. Show graphically how investment depends on K and u.

3. Suppose the economy happens to have had $u = 1$ in all past periods and so the capital stock has reached its stochastic steady state. Plot the behavior of the investment following a change in period 0 to $u = 2$, in the event that the shock is never reversed. You should use the parameter values from part 2, including the expectation that the industry switches states with probability $\pi = 0.2$.

4. Using Monte Carlo, compute the variance and autocorrelation of industry investment.

You should prepare your code so that it is straightforward to produce results with other values of u, π, R, and δ Hand in your code with your problem set.

42 Non-Convex Adjustment Cost [From Robert Shimer]

Consider the Non-Convex Adjustment cost problem in class where a single firm with diminishing returns to scale faces a fixed adjustment cost. At the start of a period, the firm has productivity A and capital k. If it keeps its capital constant, its current period output is $A^{1-\alpha}k^{\alpha}$ units of the consumption good. If it changes its capital to any $k' \neq k$, its current period output is $A^{1-\alpha}k'^{\alpha} + k - k' - \mu k$, where μ represents a fixed cost of adjustment, with a total adjustment cost proportional to the beginning of period capital stock. The gross growth rate of productivity, $\gamma \equiv A'/A$, is independently and identically distributed over time. Assume that γ can take on three possible values, e^{Δ} , 1, and $e^{-\Delta}$, with probabilities $\pi, 1 - 2\pi$ and π.

1. Let k^* denote the frictionless optimal capital stock for a given level of productivity. Let $z = k/k^*$ and $z' = k'/k^*$ denote the beginning- and end-of-period capital stock relative to the frictionless optimum. Find a Bellman equation for the ratio of the value of the firm to its frictionless capital stock in terms of the single state variable z. Let $v(z)$ denote the Bellman value.

2. Conjecture that the solution to this problem involves three numbers, $\underline{z} < z^* < \bar{z}$. If $z \in [\underline{z}, \bar{z}]$, the firm does not adjust its capital stock. If $z < \underline{z}$ or $z > \bar{z}$, the firm adjusts its capital stock to z^*. Take any $z \in [\underline{z}, \bar{z}]$ and find a second order difference equation for the

value function. Solve it, letting c_1 and c_2 denote the constants of integration. Compare your answer with the value function in the region of inaction that we found in class.

3. Prove that, if it is optimal for the firm to adjust its capital stock when it starts the period in state z, $v(z) = (1-\mu)z + \max_{z'} v(z') - z'$. Using this result, find expressions for $v'(\bar{z})$,$v'(\underline{z})$, and $v'(z^*)$. Show that there are enough equations and unknowns to solve for the two constants of integration c_1 and c_2, the two thresholds $\bar{z}$ and $\underline{z}$, and the target z^*.

4. What is marginal q when the firm invests? When does it disinvest? When it is inactive? Plot marginal q and average Q as a function of z when $R = 1.001$, $\Delta = 0.01$, $\pi = 0.5$, $\alpha = 0.9$, and $\mu = 0.01$. How does investment depend on marginal q and average Q? Plot the thresholds and targets for values of $\mu \in [0.001, 0.1]$.

5. For the remainder of the question, assume $(\log \bar{z} - \log z^*)/\Delta$ and $(\log z^* - \log \underline{z})/\Delta$ are both integers. Let $\omega(z)$ be the fraction of time spent in state z. Find a second order difference equation for $\omega(z)$ when $\underline{z}e^{\Delta} \le z \le z^* e^{-\Delta}$ and another second order difference equation when $z^* e^{\Delta} \le z \le \bar{z}e^{-\Delta}$. Solve these second order difference equations, giving four constants of integration.

6. Write down a first order difference equation relating $\omega(\underline{z})$ and $\omega(\underline{z}e^{\Delta})$ and another equation relating $\omega(\bar{z})$ and $\omega(\bar{z}e^{-\Delta})$. Use these to eliminate two of the constants of integration. What determines the other two constants of integration?

7. Assume $\Delta = 0.01$, $\pi = 0.5$, $z^* = 1$, $\bar{z} = e^{0.5}$, and $\underline{z} = e^{-0.4}$. Plot $\omega(z)$.

Prepare your code so that it is straightforward to produce results with other parameters. Hand in your code with your problem set.

43 (Optional) Menu Cost Model

Consider the model of sticky prices developed in class. There is a continuum of households of measure 1. Household $i \in [0, 1]$ is the monopoly producer of good i. The household starts each period with money $m_{i,t-1}$. It then gets a money transfer $M_t - M_{t-1}$. Next, all households set the price of their good, $p_t(i)$ for household i. Then production and consumption takes place, with each household producing the demanded amount of its good by hiring workers at a nominal wage W_t, each of whom produces one unit of the household's type of good. The household maximizes expected lifetime utility,

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(u\left(c_{i,t}\right) + l_{\,i,t} - 1\right)$$

where $c_{i,t}$ is the consumption aggregator for household i, reflecting the desire to consume the goods produced by all the other households,

$$c_{i,t} = \left(\int_0^1 c_{i,t}\left(j\right)^{\frac{\theta-1}{\theta}} dj\right)^{\frac{\theta}{\theta-1}}$$

$l_{i,t} = 1 - n_{i,t} - h\left(c_{i,t}, m_{\,i,t}/P_t\right)$ is the household's leisure, which depends on its labor supply $n_{i,t}$ and on the amount of time spent shopping h; and P_t is the price level,

$$P_t = \left(\int_0^1 p_t\left(j\right)^{1-\theta} dj\right)^{\frac{1}{1-\theta}}$$

The household has a nominal intertemporal budget constraint

$$\int_0^1 p_t\left(j\right) c_{i,t}\left(j\right) dj + m_{i,t} = W_t n_{i,t} + \left(p_t\left(i\right) - W_t\right) \int_0^1 c_{j,t}\left(i\right) dj + m_{i,t-1} + M_t - M_{t-1}$$

The left hand side is the household's nominal spending plus nominal money balance. The right hand side is its nominal labor income, plus its profits from selling good i, plus its initial money, plus the money injection.

The only role of the government is to provide the money injection. Assume that the money growth rate is zero with probability π (so $M_{t+1} = M_t$) and Δ otherwise (so $M_{t+1} = \Delta M_t$) for some $\Delta > \beta$. Money growth is independently and identically distributed.

Let $u(c) = 0.5 \log c$, $h(c, m) = 0.1c/m$, $\theta = 10$, $\beta = 0.99$, $\Delta = 1.01, \mu = 0.05$, and $\pi = 0.5$.

1. Set up the problem, write the FOCs and simplify the equations as in the class (show details of the steps, for example, show how do we derive the equation for $V_0\left(z_i\right)$ and so on)

2. Write down the steps of the algorithm on ow this problem can be solved numerically.

3. Write down a code to solve the problem numerically. You can use the calibration above for the parameters.

4. Find the equilibrium of the type described in class: at any point in time, a fraction $1/I$ of households charges each of the relative prices z_0, z_0/Δ,..., $z_0/\Delta^{I-1}$for some positive integer I and positive number z_0. A household finds it optimal to keep its nominal price fixed when its relative price is at any of these levels and to change its nominal price when its relative price reaches z_0/Δ^I. The real money supply and real wage are constant.

5. Now For $0.0001 < \mu < 0.1$, find the set of equilibria and plot the equilibrium value(s) of $I, z_0, z_0/\Delta^I, \Delta^I, c, m/P, W/P, l$ and the average frequency of price changes(average duration of price stickiness) against μ. Write your explanation for the behavior of these variables with respect to μ.

6. Redo part 5 for $\mu = 0.05$ and $1.001 < \Delta < 1.1$ and plot those outcome variables with respect to Δ.Write your explanation for the behavior of these variables with respect to Δ.

7. Redo part 5 for $\mu = 0.05$, $\Delta = 1.01$ and $0.1 < \pi < 0.9$ and plot those outcome variables with respect to π.Write your explanation for the behavior of these variables with respect to π.

8. Redo part 5 for $\mu = 0.05$, $\Delta = 1.01$ and $\pi = 0.5$ and $2 < \theta < 12$ and plot those outcome variables with respect to θ.Write your explanation for the behavior of these variables with respect to θ.

9. Does this type of equilibrium always exist? If not, is there some other stationary equilibrium of the economy?

10. Now consider the case where $\mu = 0$ (no stickiness) and prove mathematically that if $\gamma_1 < \gamma_2$ then $m_1 > m_2$ and $c_1 < c_2$. Explain the intuition. (what can you say when we have $\mu > 0$)