Problem Set 3: Random Variables and Simulation Methods

Albert Rodriguez-Sala and Jacob Adenbaum
Programming for Economics—ECNM10106
The University of Edinburgh

2023

Exercise 1. Monte Carlo simulation: OLS asymptotic behavior. (20 points) Prove that the OLS estimator is distributed asymptotically normal using Monte Carlo simulation.

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \overset{as}{\sim} N\left(\beta, \sigma^2\left(\mathbb{E}(x_i x_i')\right)^{-1}\right)$$

To prove that, consider the model or data generating process seen in class

$$y_i = 2 + 0.5x_{2,i} + e_i$$
$$e_i \sim N(0, 25)$$

and, as a preliminary step, using the following code create a sample with size N=100, a fictitious $x_2 \sim N(0,1)$, and estimate the model by OLS.

```
# import statsmodels for OLS regression
import statsmodels.api as sm
# Let's create a sample (a data), and estimate B* by OLS
N= 100 # Sample size
sigma_e = 5
#  fictisiously create our regressor x2.
x2 =  np.random.normal(0,1,N);  #x2 follows normal (0,1)
# e is an error term (what we don't observe in the data)
e = np.random.normal(0,sigma_e,N)
# And our dependent variable is:
y1 = 2 + 0.5*x2 +e
# Let's estimate the model by OLS
X = sm.add_constant(x2)  # We need to add constant, B0
ols1 = sm.OLS(y1,X).fit()
print(ols1.summary())
```

a. Now, following the steps in lecture 4, slide 12, run a Monte Carlo simulation to show the asymptotic properties of $\hat{\beta}_{OLS}$ and **replicate figure 1 and figure 2 from slide 13**. Set the number of simulations equal to T=10000. For each repetition of the experiment (step 3) use a sample size of $N = 100$.

b. Redo exercise a) with a sample size $N = 1000$. How do the $\hat{\beta}_{OLS}$ asymptotic distributions change when the sample size increases? **Hint:** For a visual demonstration, keep the same x-axis in the plots in a) and b) using *plt.xlim()*.

Exercise 2. Solving the earnings distribution in an Aiyagari economy (40 points). In the Aiyagari model—one of the fundamental models in quantitative macroeconomics—"there is a very large number of households" whose log of labor earnings (y_t) follows the next AR(1) process

$$y_t = \rho y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{1}$$

For this exercise let's use the following parameterization of the AR(1) process:
$\rho = 0.95$, $\sigma_\varepsilon = 0.25$ and an initial value $y_0 = 0$.

a. Simulate and plot the AR(1) process given by equation (1) for T=50 periods.

b. Create a function that simulates N AR(1) processes for T periods.
   - The function should have the following inputs $(N, T, rho, y0, sigma_e)$
   - The output of the function should be a 2-D array with size (N,T)

c. Simulate and plot 5 AR(1) processes given by equation (1) for t=50 periods.

d. **The stationary earnings distribution.** Simulate for a large T (like T=1000) the AR(1) process of 10000 individuals. Use the result of the last period to plot the stationary distribution. From the stationary distribution, compute the variance of the log of earnings (y_t), and the Gini coefficient of earnings (e^{y_t}).

e. For computational reasons, Aiyagari discretizes the AR(1) process into a Markov process of 7 states. Using the Rouwenhorst method, discretize the AR(1) process of this exercise into a 3 states Markov process. What is the resulting transition matrix P? What is the resulting stationary distribution ψ^*?

Exercise 3. Simulating distributions and computing expectations (20 points).

a. Simulate a binomial distribution with $n = 4$ and $p = 0.5$. Plot the resulting distribution.

b. Compute the expected value of a function $g(x) = x^2$ where x follows a Poisson distribution with $\lambda = 2$. Use Monte Carlo integration.
   Is your result equal to $g(\mathbb{E}(x))$? where $\mathbb{E}(x) = \lambda = 2$ if $x \sim Poisson(\lambda)$. Why not?

Exercise 4. Simulating and computing expectations in an economy Consider an economy where individual's income (y) follows a log-normal distribution. That is $log(y) \sim N(\mu, \sigma^2)$ where $\mu = 7.5, \sigma = 0.8$.

a. Simulate y for $N = 100000$ and compute the average, the variance, and the Gini of y.

b. Now consider that individuals follow a consumption rule that takes the following functional form:

$$c_i = (y_i)^{0.8} + 0.5y_i + 500$$

   Compute the average, the variance, and the Gini coefficient of c.

c. Plot the distribution of y and c in the same graph.