1 Simulation

1.1 Generation of Random Variables

1. Pseudo-random number generators (i.i.d.): Uniform, $\mathcal{N}(0,1)$. Each value is a realization of the random variable.

2. Other random variables: X and Y are independent $\mathcal{N}(0,1)$.

$$3 + X \sim \mathcal{N}(3,1); \qquad 2X \sim \mathcal{N}(0,4);$$

$$X^2 \sim \chi^2(1); \qquad X^2 + Y^2 \sim \chi^2(2);$$

$$X/\sqrt{\chi^2(n)/n} \sim t(n); \qquad \frac{\chi^2(m)/m}{\chi^2(n)/n} \sim F(m,n);$$

$$\exp(X) \sim \text{log normal}.$$

3. Gaussian random walk: $y_t = y_{t-1} + u_t$, $y_0 = 0$, $t = 1, \ldots, T$.

   (a) Generate T values of $u \sim \mathcal{N}(0,1)$.

   (b) Taking cumulated sums of u_i: $y_1 = u_1$, $y_2 = u_1 + u_2$, ...

4. Gaussian AR(1): $y_t = 0.5\, y_{t-1} + u_t$, $y_0 = 0$, $t = 1, \ldots, T$.

   (a) Generate T values of $u \sim \mathcal{N}(0,1)$.

   (b) $y_1 = u_1$, $y_2 = 0.5\, u_1 + u_2$, $y_3 = 0.5^2\, u_1 + 0.5 u_2 + u_3$, ...

1.2 Simulations of LLN and CLT

1. Simulation of a weak law of large numbers:

   (a) Generate a sample of T values, x_i, and calculate their sample average, $\bar{x}$, which is a realization of the sample mean $\bar{X}_T$.

   (b) Replicating the first step n times yields n values of $\bar{x}$. You may then plot the empirical (sampling) distribution of $\bar{X}$.

   (c) By varying T in the first step, you get different sampling distributions of $\bar{X}_T$ and should observe the LLN effect.

2. Simulation of a central limit theorem:

   (a) Generate a sample of T values, x_i (say $\mathcal{N}(0.5, 4)$), and calculate the normalized sample average, $\sqrt{T}(\bar{x} - 0.5)/2$, which is a realization of the normalized sample mean.

   (b) Repeating the first step n times yields n values of the normalized sample average. You may then plot the resulting empirical distribution.

   (c) By varying T in the first step, you get different sampling distributions and should observe the CLT effect.

1.3 Simulations of a t Test

1. Simulation of test size (empirical type I error):

   (a) Generate a sample of T values: x_i and u_i.

   (b) Generate $y_i = 1 + 2\,x_i + u_i$ and test the null hypothesis $\beta = 2$.

   (c) Regress y_i on a constant and x_i and obtain the estimate $\hat{\beta}_T$ and $\hat{\sigma}^2 = \sum_{i=1}^{T} \hat{e}_i^2/(T-2)$.

   (d) Compute the t statistic:

   $$(\hat{\beta}_T - 2)\Big[\sum_{i=1}^{T}(x_i - \bar{x})^2\Big]/\hat{\sigma}.$$

   (e) Replicate the steps above for n times and obtain n values of the t statistic. The number of t values that are greater than 5% critical value (e.g., $T = 20$, 5% critical value of $t(18)$ is 2.1 for a two sided test) is the empirical test size.

2. Simulation of test power:

   (a) Every step is the same as above except in (b) y_i are generated as $y_i = 1 + (2+\delta)\,x_i + u_i$, where δ is a number determining the deviation from the null hypothesis $\beta = 2$.

   (b) The number of t values that are greater than 5% critical value is the empirical test power.

   (c) By fixing T but varying the value of δ, you can evaluate the test power of different alternatives that are close to and far away from the null.

(d) By fixing δ but varying the value of T, you can evaluate the test power of a given alternative under different samples.