LECTURE 21

§24. SIMULATION

As our starting point in simulation, we suppose that we can simulate from the uniform $(0,1)$ distribution and we use the term "random numbers" to mean independent random variables from this distribution. To simulate the observation of continuous random variables, we usually start with uniform random numbers, and relate these to the distribution function of interest. We could use two- or three-digit random integers, perheps selected from the Table of "random numbers", but most software programs have a continuous uniform random number generator.

A general method for simulating a random variable having a continuous distribution — called the *inverse transformation method* — is based on the following Lemma.

Lemma 8. *Let $\eta(\omega)$ be a random variable with continuous distribution function $F(x)$. Then the random variable*

$$\zeta(\omega) = F(\eta(\omega))$$

has uniform distribution over the interval $(0,1)$.

Proof: For any $x \le 0$ the distribution function F_ζ is equal to 0 and for $x > 1$ F_ζ is equal to 1. Let us consider $x \in (0,1)$.

$$F_\zeta(x) = P\{\omega\colon\ \zeta(\omega) < x\} = P\{\omega\colon\ F(\eta(\omega)) < x\} = P\{\omega\colon\ \eta(\omega) < F^{-1}(x)\} = F(F^{-1}(x)) = x$$

Above $F^{-1}(x)$ is defined to equal that value y for which $F(y) = x$.

Corollary 14. Let U be the uniform $(0,1)$ random variable. For any continuous distribution function F if we define the random variable $\eta(\omega)$ by

$$\eta(\omega) = F^{-1}(U)$$

then the random variable $\eta(\omega)$ has distribution function $F(x)$.

The proof immediately follows from Lemma 8.

Example 42.1. (Simulating an Exponential Random Variable). If $F(x) = 1 - e^{-\lambda x}$, then $F^{-1}(u)$ is that value of x such that $1 - e^{-\lambda x} = u$ or

$$x = -\frac{1}{\lambda}\log(1-u).$$

Hence, if U is a uniform $(0,1)$ variable, then

$$F^{-1}(U) = -\frac{1}{\lambda}\log(1-U)$$

is exponentially distributed with mean λ. Since $1-U$ is also uniformly distributed on $(0,1)$, it follows that $-\frac{1}{\lambda}\log U$ is exponential with mean λ. Since $c\eta$ is exponential with mean c when η is exponential with mean 1, it follows that $-c\log U$ is exponential with mean c.

Suppose we wish to simulate an observation he exponential distribution

$$F(x) = 1 - e^{-0.3x}, 0 < x < +\infty.$$

The computer would firs produce the value u from the uniform distribution. The we solve

$$u = F(x) = 1 - e^{-0.3x}$$

so

$$x = -\frac{1}{0.3}\ln(1-u)$$

is the corresponding value of an exponential random variable. For example, if $u = 0.45$, then

$$x = -\frac{1}{0.3}\ln(1-0.45) = 1.993.$$

Example 42.2. (Simulating a Weibull Random Variable). A similar procedure applies to the simulation of observations from a Weibull distribution. Starting with the value of a uniform variable u, we now solve the equation

$$u = F(x) = 1 - e^{-\alpha x^{\beta}}$$

so

$$x = \left[-\frac{1}{\alpha}\ln(1-u)\right]^{1/\beta},$$

which is the corresponding value of a Weibull random variable.

Example 42.3. Simulate five observations of a random variable having Weibull distribution with $\alpha = 0.05$ and $\beta = 2$.

Solution. A computer generates the five values 0.57, 0.74, 0.26, 0.77, 0.12. We calculate

$$x_1 = -20\ \ln(1 - 0.57)]^{1/2} = 4.108,$$

$$x_2 = -20\ \ln(1 - 0.74)]^{1/2} = 5.191.$$

The last three uniform numbers yield

$$x_3 = 2.454, \qquad x_4 = 5.422, \qquad x_5 = 1.599.$$

Example 42.4 (Simulating Independent, Standard Normal Random Variables). For the important case of the normal distribution, it is not possible to give a close form expression for the distribution function $F(x)$, and so we cannot solve the equation $F(y) = x$. For this reason, special methods have been developed. One such method relies on the fact that if U and V are independent random variables with uniform densities on $(0, 1)$, then, random variables

$$\eta_1 = \sqrt{-2\log(U)}\ \cos(2\pi\, V)$$

and

$$\eta_2 = \sqrt{-2\log(U)}\ \sin(2\pi\, V)$$

are independent, and have normal distribution function with parameters $a = 0$ and $\sigma = 1$.

Example 42.5. (Simulating two values from a normal distribution). Simulate two observations of a random variable having the normal distribution with mean50 and standard deviation 5.

Solution. A computer generates the two values 0.253 and 0.531 from a uniform distribution. We first calculate the standardnormal values

$$z_1 = \sqrt{-2\log(0.531)}\ \cos(2\pi\, 0.253) = -0.021$$

$$z_2 = \sqrt{-2\log(0.531)}\ \sin(2\pi\, 0.253) = 1.125$$

and then the normal values

$$x_1 = 50 + 5\,z_1 = 50 + 5(-0.021) = 49.895$$

$$x_2 = 50 + 5\,z_2 = 50 + 5(1.125) = 55.625.$$

§33. MARKOV CHAINS

There is an increasing interest in the study of systems which vary in time in a random manner. Mathematical models of such systems are known as stochastic processes. A stochastic process can be defined quite generally as any collection of random variables $\eta(t,\omega)$, $t \in T$, defined on a common probability space, where T is a subset of $(-\infty, +\infty)$ and is thought of as the time parameter set. The process is called a *continuous parameter process* if T is an interval having positive length and a *discrete parameter process* if T is a subset of the integers. If the random variables $\eta(t,\omega)$ all take on values from the fixed set $\mathcal{G}$, then $\mathcal{G}$ is called the *state space* of the process.

Many stochastic processes of theoretical and applied interest possesses the property that, given the present state of the process, the past history does not affect conditional probabilities of events defined in terms of the future. In 1907, A. A. Markov began the study of an important new type of chance process. Such processes are called Markov processes. In the next sections we study Markov chains, which are discrete parameter Markov processes whose state space is finite or countable infinite.

Consider a system that can be in any one of a finite or countably infinite number of states. Let $\mathcal{G}$ denote this set of states. We can assume that $\mathcal{G}$ is a subset of the integers. Let the system be observed at the discrete moments of time $n = 0, 1, 2, \ldots$, and let $\eta_n(\omega) = \eta(n,\omega)$ denote the state of the system at time n. If $\eta_n(\omega) = i$, then the process is said to be in state i at time n.

Since we are interested in non–deterministic systems, we think of $\eta_n(\omega)$, $n \geq 0$, as random variables defined on a common probability space. Little can be said about such random variables unless some additional structure is imposed upon them.

The simplest possible structure is that of independent random variables. This would be a good model for such systems as repeated experiments in which future states of the system are independent of past and present states. In most systems that arise in practice, however, past and present states of the system influence the future states even if they do not uniquely determine them.

Many systems have the property that given the present state, the past states have no influence on the future. This property is called the *Markov property*, and systems having this property are called *Markov chains*. The Markov property is defined precisely by the requirement that

$$\begin{aligned} &P\left\{\eta_{n+1}(\omega)=j\,/\eta_0(\omega)=i_0,\eta_1(\omega)=i_1,...,\eta_{n-1}(\omega)=i_{n-1},\eta_n(\omega)=i\,\right\}= \\ &=P\left\{\eta_{n+1}(\omega)=j\,/\eta_n(\omega)=i\,\right\} \end{aligned} \tag{33.1}$$

for every choice of the nonnegative integer n and the numbers i_0, i_1, ..., i_{n-1}, i, j, each in $\mathcal{G}$. The conditional probabilities $P\left(\eta_{n+1}(\omega)=j\,/\eta_n(\omega)=i\,\right)$ are called the *one–step transition probabilities* of the chain. We study Markov chains having *stationary* transition probabilities, that is, those such that

$$P\left(\eta_{n+1}(\omega)=j\,/\eta_n(\omega)=i\,\right)=P_{i,j}$$

is independent of n. From now on, when we say that $\eta_n(\omega)$, $n\geq 0$ forms a Markov chain, we mean that these random variables satisfy the Markov property and have stationary transition probabilities.

Equation (33.1) may be interpreted as stating that, for a Markov chain, the conditional distribution of any future state $\eta_{n+1}(\omega)$ given the past states $\eta_0(\omega)$, $\eta_1(\omega)$, ..., $\eta_{n-1}(\omega)$ and the present state $\eta_n(\omega)$, is independent of the past states and depends only on the present state.

The value $P_{i,j}$ represents the probability that a system in state i will enter state j at the next transition. Since probabilities are nonnegative and since the process must make a transition into some state, we have that

$$P_{i,j}\geq 0 \quad \text{for any} \quad i,j\in\mathcal{G} \tag{33.2}$$

and

$$\sum_j P_{i,j}=1 \quad \text{for any} \quad i\in\mathcal{G}. \tag{33.3}$$

Let ¶ denote the matrix of one–step transition probabilities $P_{i,j}$, so that

$$¶ = \|P_{i,j}\|_{i,j \in \mathcal{G}}.$$

The function $\pi_0(i)$, $i \in \mathcal{G}$, defined by $\pi_0(i) = P\{\eta_0(\omega) = i\}$, $\quad i \in \mathcal{G}$ is called the initial distribution of the Markov chain. It is such that

$$\pi_0(i) \geq 0 \quad \text{for any} \quad i \in \mathcal{G} \tag{33.4}$$

and

$$\sum_i \pi_0(i) = 1. \tag{33.5}$$

The joint distribution of $\eta_0(\omega)$, ..., $\eta_n(\omega)$ can easily be expressed in terms of the transition function and the initial distribution. For example,

$$P\{\eta_0(\omega) = i_0, \eta_1(\omega) = i_1\} = P\{\eta_0(\omega) = i_0\} \cdot P\{\eta_1(\omega) = i_1 / \eta_0(\omega) = i_0\} = \pi_0(i_0) \cdot P_{i_0,i_1}.$$

By induction it is easily seen that

$$P\{\eta_0(\omega) = i_0, \eta_1(\omega) = i_1, ..., \eta_n(\omega) = i_n\} = \pi_0(i_0) \cdot P_{i_0,i_1} \cdot P_{i_1,i_2} \cdot ... \cdot P_{i_{n-1},i_n}. \tag{33.6}$$

It is usually more convenient, however, to reverse the order of our definitions. We say that $P_{i,j}$, $i \in \mathcal{G}$ and $j \in \mathcal{G}$, is a ***transition function*** if it satisfies (33.2) and (33.3), and we say that $\pi_0(i)$, $i \in \mathcal{G}$, is an ***initial distribution*** if it satisfies (33.4) and (33.5). It can be shown that given any transition function $P_{i,j}$ and any initial distribution $\pi_0(i)$, there is a probability space and random variables $\eta_n(\omega)$, $n \geq 0$, defined on that space satisfying (33.6). It is not difficult to show that, these random variables form Markov chain having transition function $P_{i,j}$ and initial distribution $\pi_0(i)$.

Example 33.1. Consider the Markov chain consisting of the three states $1, 2, 3$ and having transition probability matrix

$$¶ = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{pmatrix}$$

It is possible to go from state 1 to state 3 since $P_{1,2} = 1/2$, $P_{2,3} = 1/4$. That is, one way of getting from state 1 to state 3 is to go from state 1 to state 2 (with probability $1/2$) and then go from state 2 to state 3 (with probability $1/4$).