Generations of Random Probability-Vectors

Generations of Random Probability-Vectors

Toshiro KUTSUWA,* Ken-ichi MORITA,** Hideo KOSAKO***
and Yoshiaki KOJIMA***

(Received June 15, 1972)

This paper describes the methods for generating random probability-vectors. The vectors are used to simulate the discrete stochastic systems, for example, markov chains, probabilistic automata and stochastic sequential machines.

The value of the vector components corresponds to the occurrence probability for each state, and the distribution of those components is the normal, exponential or uniform distribution in the range of [0, 1].

1. Introduction

Various transition matrices or probability vectors need to be used to analyze the statistical characteristics of markov chains, stochastic sequential machines[4] and others.

Therefore, the generation of probability-vectors whose components have the density function of the normal, exponential or uniform distribution is discussed in this paper.

Generally, uniform random numbers (denoted by the symbol **UN**) are made by the multiplicative congruential or the mixed congruential method on the computer. Other distributions are obtained by **UN** with the rejection or the inverse transformation method.

In the generation of the probability-vectors, random numbers viz. the vector components are generated by the above ways, and are transformed so that they may satisfy the following conditions i) and ii).

i) Each component of the vector $u_j (j=1, \cdots, n)$ corresponds to the state-occurrence probability p_j,

$$0 \leq p_j \leq 1, \quad \sum_{j=1}^{n} p_j = 1 \quad (n; \text{ the number of states}).$$

ii) The distribution of the generated vector components is the normal, exponential or uniform distribution in the range of [0, 1].

2. Probability Vectors with Normally Distributed Components

2.1. Vector generation by the rejection method

In the case of n states, we take independently n numbers out of the one dimensional normal random numbers in the range of $[A, B]$, and divide each of n numbers by those sum, where $A=\mu_x - c_p \cdot \sigma_x$, $B=\mu_x + c_p \cdot \sigma_x$.

A coefficient $c_p (>0)$ denotes the range that random number is taken. The $\mu_x (=1/n)$

* Graduate Student, Department of Electronics, College of Engineering.
** Undergraduate, Department of Electronics, College of Engineering.
*** Department of Electronics, College of Engineering.

and σ_x are the mean and the standard deviation of the one dimensional normal distribution function, respectively. If $A=\mu_x-c_p\cdot\sigma_x<0$, then A is fixed at 0.

Each of n numbers generated is assigned to each component of the n-state probability vector.

If a large number of the probability vectors would be generated by these procedures, then the distribution of the vector components is well approximated to the normal distribution in the range of $[0,1]$. The results of analysis and experiments are shown in the section **2.4.**

2.2. Vector generation by the inverse transformation method

The procedures of this method are as follows:

n numbers of the triangular distribution represented by Eq. (1) are generated by the inverse transformation.[1] Each of n numbers is divided by those sum, and assigned to each component of the vector.

If these procedures would be repeated for a great many probability-vectors, then the vector components have the pseudo-normal distribution in the range of $[0,1]$.

$$\left.\begin{aligned} g(x)&=g_1(x)+g_2(x) \\ g_1(x)&=0.181105\left(\frac{x-\mu_x}{\sigma_x^2}\right)+\frac{0.426427}{\sigma_x} \qquad (A\leq x\leq W) \\ g_2(x)&=-0.181105\left(\frac{x-\mu_x}{\sigma_x^2}\right)+\frac{0.426427}{\sigma_x} \qquad (W<x\leq B) \end{aligned}\right\} \tag{1}$$

where $g(x)$ denotes the density function of value x.

$$A=\mu_x-2.355\,\sigma_x$$
$$W=\mu_x$$
$$B=\mu_x+2.355\,\sigma_x$$

where A is zero for all negative A.

An approximate function Eq. (1) for the normal distribution $N(\mu_x, \sigma_x^2)$ is derived from the function for the $N(0,1)$.[3]

2.3. Mathematical formulation

In order to analyze the distribution of the vector components, this section gives mathematical formulations for methods discussed in the section **2.1** and **2.2.**

The dividing operation of the generating method is expressed by the transformation of probability variables X and Y,

$$U=X/(X+Y)$$

where X corresponds to a value of the n numbers, Y corresponds to the sum of n numbers except X, and U corresponds to a component of the n-state probability vector.

Let $g(x)$ and $h(y)$ be the density functions of X and Y, respectively. The density function $f(u)$ of the variable U will be derived as follows:

The joint density function $f(x,y)$ for independent variables X and Y is expressed by

$$f(x,y)=g(x)\cdot h(y).$$

Accordingly, joint density function $f(u, v)$ for the variables U and V is obtained by

$$f(u, v) = g(v) \cdot h(v/u - v) \cdot v/u^2 \tag{2}$$

with $U = X/(X+Y)$ and $V = X$.

The density function $f(u)$ equals the definite integral of the $f(u, v)$ of v with limits A and B,

$$\begin{aligned} f(u) &= \int_A^B f(u, v) dv \\ &= \int_A^B g(v) \cdot h(v/u - v) v/u^2 dv. \end{aligned} \tag{3}$$

The $f(u)$ represents the distribution of the probability vector components.

2.3.1. The case of the rejection method

In this case, the $g(x)$ is denoted by

$$g(x) = \frac{c_1}{\sqrt{2\pi}\,\sigma_x} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \qquad (A \leq x \leq B) \tag{4}$$

where $A = \mu_x - c_p \cdot \sigma_x$, $B = \mu_x + c_p \cdot \sigma_x$, and c_1 is a constant to satisfy the relation $\int_A^B g(x) = 1$.

Similarly, $h(y)$ is obtained by the following function

$$h(y) = \frac{c_2}{\sqrt{2\pi}\,\sigma_y} \exp\left\{-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right\} \qquad (0 \leq y) \tag{5}$$

where c_2 is a constant which has the same mean as c_1, $\mu_y = (n-1) \cdot \mu_x'$ and $\sigma_y^2 = (n-1) \cdot \sigma_x'^2$. The μ_x' is mean value and $\sigma_x'^2$ is the variance of $g(x)$. If $A = \mu_x - c_p \cdot \sigma_x \geq 0$, then it is clear that $\mu_x' = \mu_x$ and

$$\sigma_x'^2 = \sigma_x^2 \left\{1 - \frac{2 \cdot c_1 \cdot c_p}{\sqrt{2\pi}} \exp\left(-\frac{c_p^2}{2}\right)\right\}.$$

By exchanging x for v and y for $(v/u - v)$, respectively, and substituting (4) and (5) into (3), the density function is

$$f(u) = c \int_A^B \frac{c_1 c_2}{2\pi \sigma_x \sigma_y} \cdot \frac{v}{u^2} \exp\left\{-\frac{(v-\mu_x)^2}{2\sigma_x^2} - \frac{(v/u - v - \mu_y)^2}{2\sigma_y^2}\right\} dv. \tag{6}$$

By putting $C = cc_1 c_2$, the mean μ_u and the variance σ_u^2 of $f(u)$ can be written as follows:

$$\left.\begin{aligned} \mu_u &= \frac{C}{2\pi\sigma_x\sigma_y} \int_0^1 u \int_A^B \frac{v}{u^2} \exp\left\{-\frac{(v-\mu_x)^2}{2\sigma_x^2} - \frac{(v/u - v - \mu_y)^2}{2\sigma_y^2}\right\} dv du \\ \sigma_u^2 &= \int_0^1 u^2 f(u) du - (\mu_u)^2 \end{aligned}\right\} \tag{7}$$

2.3.2. The case of the inverse transformation method

The density function $f(u)$ for u is denoted by

$$f(u) = \frac{cc_2}{\sqrt{2\pi}\,\sigma_y} \int_A^B \frac{v}{u^2} g(v) \exp\left\{-\frac{(v/u - v - \mu_y)^2}{2\sigma_y^2}\right\} dv$$

$$=\frac{C}{\sqrt{2\pi}\,\sigma_y}\int_A^W \frac{v}{u^2}g_1(v)\exp\left\{-\frac{(v/u-v-\mu_y)^2}{2\sigma_y^2}\right\}dv$$
$$+\frac{C}{\sqrt{2\pi}\,\sigma_y}\int_W^B \frac{v}{u^2}g_2(v)\exp\left\{-\frac{(v/u-v-\mu_y)^2}{2\sigma_y^2}\right\}dv \quad (8)$$
$$(C=cc_2 \text{ is a constant}).$$

For the function $g(x)$ is represented by Eq. (1) in the section **2.2**, and $h(y)$ is determined by using μ_x' and $\sigma_x'^2$ of $g(x)$,

$$h(y)=\frac{c_2}{\sqrt{2\pi}\,\sigma_y}\exp\left\{-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right\}.$$

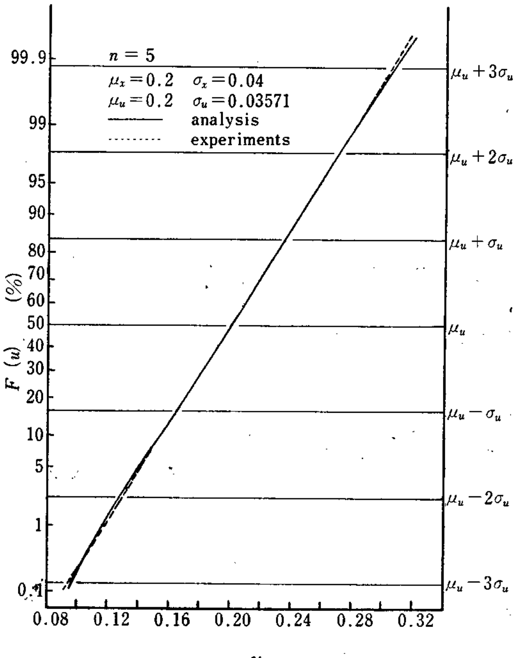
2.4. Results of analysis and experiments

As the representative examples, for $n=5$, $\mu_x=0.2$ and $\sigma_x=0.04$, the results of numerical analysis and experiments (5000 trials) are shown on the normal probability papers of Fig. 1, (a) to (c).

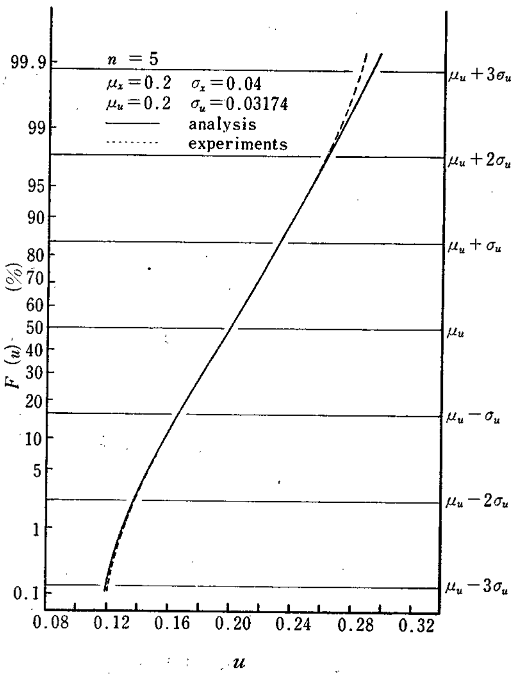
Fig. 1, (a) and (b) are the results of the rejection method $c_p=3.0$ and $c_p=2.0$, respectively. Fig. 1, (c) is of the inverse transformation.

Following facts are obvious from Fig. 1. In the case of the rejection $c_p=3.0$, the density function of the vector components is extremely close to the normal distribution $N(\mu_u, \sigma_u^2)$. A degree of the conformity between $N(\mu_u, \sigma_u^2)$ and the $f(u)$ of the inverse transformation is much the same as that between $N(\mu_u, \sigma_u^2)$ and the $f(u)$ of the rejection $c_p=2.0$.

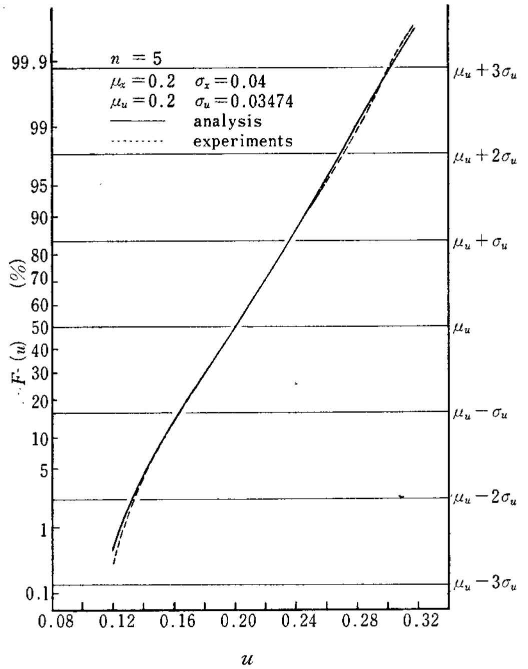
The relation between σ_x and σ_u for $n=5$ is shown in Fig. 2, where σ_u is the standard deviation of the generated vector components. If $A\geq 0$, then σ_u is nearly proportional to



(a) Rejection method $c_p=3.0$



(b) Rejection method $c_p=2.0$

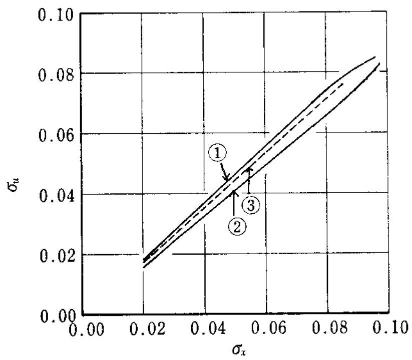


(c) Inverse transformation method

Fig. 1. Value of vector components u vs. the cumulative distribution $F(u)$ on the normal probability paper.

Table 1. Value $\xi=\delta_u/\delta_x$ for n and c_p.

		$n=3$	$n=5$	$n=8$	$n=10$
$\xi=\delta_u/\delta_x$	Rejection $c_p=3.0$	0.8291	0.9068	0.9332	0.9395
	Rejection $c_p=2.0$	0.7328	0.8024	0.8292	0.8390
	Inverse transformation	0.8167	0.8991	0.9150	0.9233



① Rejection method $c_p=3.0$
② Rejection method $c_p=2.0$
③ Inverse transformation method

Fig. 2. Relation between σ_x and σ_u for $n=5$.

σ_x. Table 1 shows the value $\xi=\sigma_u/\sigma_x$ in the proportional range for n and c_p.

2.5. Generating efficiency of random vector components

The generating efficiency (**GE**) of the rejection method is illustrated in Fig. 3. In Fig. 3, a point in the quadrilateral $ABQR$ is selected by 2 **UN**, and if the point is not in the hatched part, then those 2 **UN** are rejected. Only when the point is in the hatched part, a random number, namely, a vector component is obtained from 2 **UN**. On the other hand, in the inverse transformation method, a vector component has one to one correspondence with each of **UN**.

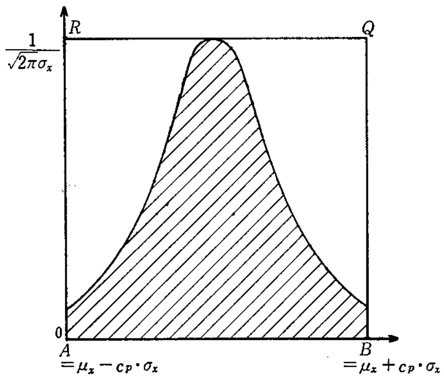


Fig. 3. Illustration of generating efficiency.

Let the **GE** of the inverse transformation be one, then the **GE** of the rejection is expressed by

$$\mathbf{GE}=\frac{\text{the area of hatched part in Fig. 3}}{\text{the area bounded by } A,\, B,\, Q \text{ and } R}\times\frac{1}{2}$$

$$=\frac{\dfrac{1}{\sqrt{2\pi}\,\sigma_x}\displaystyle\int_A^B \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}dx}{2\cdot c_p\cdot\sigma_x\cdot\dfrac{1}{\sqrt{2\pi}\,\sigma_x}}\times\frac{1}{2}. \tag{9}$$

Now, if $c_p=2.0$,

$$\mathbf{GE}=\frac{0.9544}{4\times 1/\sqrt{2\pi}}\times\frac{1}{2}=0.2990$$

and if $c_p=3.0$, then $\mathbf{GE}=0.2083$.

Hence, in a few words, 5 components of the vector will be generated by the inverse transformation during the time required for generating only one component by the rejection $c_p=3.0$.

Therefore, one should choose a suitable method from section **2.1** and **2.2** according to the results of **2.4** and **2.5** for one's objective simulations.

3. Probability Vectors with Exponentially Distributed Components

The density function of one dimensional exponential random numbers is $f(x)=ae^{-ax}$ $(o\leq x<\infty)$, the mean μ_x and the variance σ_x^2 are $1/a$ and $1/a^2$, respectively.

Since the variability of the probability vector components is $[0, 1]$, then its density function $f(u)$ should be defined by following equation.

$$f(u)=cae^{-au} \qquad (0\leq u\leq 1) \tag{10}$$

where c and a are determined as follows:

$$\int_0^1 cae^{-au}du=1$$

$$\int_0^1 caue^{-au}du=1/n.$$

If $n=5$, then $a=4.8010$ and $c=1.00829$.

Exponential random number x $(0\leq x)$ is obtained by the inverse transformation with the cumulative distribution function $F(x)=[1-e^{-ax}]$,

$$x=-\frac{1}{a}\log(1-\gamma)$$

where $1-\gamma$ is **UN** in the range of $[0,1]$.

If the same method would be used for the vector generation and each value of random numbers were divided by the sum of those n numbers, then the vector component u_j becomes irrelevant to the coefficient a,

$$u_j=\frac{-\frac{1}{a}\log\eta_j}{-\frac{1}{a}\log\eta_1-\cdots-\frac{1}{a}\log\eta_n}=\frac{\log\eta_j}{\log\eta_1+\cdots+\log\eta_n}\qquad(j=1,\cdots,n)$$

where $\eta=1-\gamma$.

The density function of the above is shown by

$$f(u)=(n-1)\cdot(1-u)^{n-2} \tag{11}$$

The Eq. (11) is derived as follows:

Since each of n numbers is a variable of the exponential random numbers, then the density function is

$$g(x)=ae^{-ax}\qquad(A\leq x<B) \tag{12}$$

where $A=0$, $B=\infty$.

Now, the moment generating function (**MGF**) for the $g(x)$ is

$$M_X(t)=a/(a-t).$$

Hence, the **MGF** for the sum of $(n-1)$ numbers in the n states can be written by

$$M_Y(t)=[M_X(t)]^{n-1}=\left(\frac{a}{a-t}\right)^{n-1}.$$

This equation is equal to the **MGF** for the gamma distribution function having parameters a and $(n-1)$.[2)]

Then,

$$h(y)=\frac{a^{n-1}}{\Gamma(n-1)}e^{-ay}\cdot y^{n-2}. \tag{13}$$

By substituting (12) and (13) into (3), we have a density function $f(u)$,

$$f(u) = \int_0^\infty ae^{-av} \frac{a^{n-1}}{\Gamma(n-1)} e^{-a(v/u-v)} (v/u-v)^{n-2} \frac{v}{u^2} dv$$
$$= (n-1)(1-u)^{n-2}.$$

The curve of $f(u)$ for $n=5$ is shown in Fig. 4.

Either the following way i) or ii) should be used to make the $f(u)$ approach to cae^{-au}.

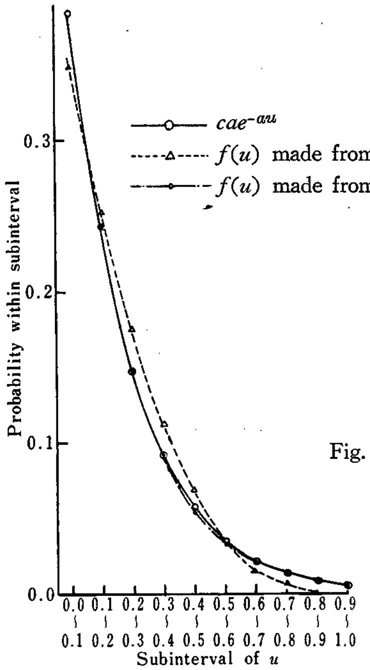


Fig. 4. Distribution $f(u)$ made from $g(x)$.

For $g(x)$,

i) decrease the probability that the assigned value falls in the neighbourhood of the average,

or

ii) bring relatively the distribution of $f(u)$ to the lower part by increasing the probability that the value of $g(x)$ is higher than the average value.

In the case of $n=5$, let

$$g(x) = \begin{cases} 8\,e^{-8x} & (0 \leq x \leq 0.374467) \\ 0.4 & (0.8 \leq x \leq 0.9) \\ 0.01/3 & (2.5 \leq x \leq 5.5) \\ 0 & (\text{otherwise}) \end{cases} \tag{14}$$

then the χ^2-value is less than 16.919 that is $\chi_\alpha{}^2$ of the significance level $\alpha=0.05$ and 9 degrees of freedom. The above $\chi^2=7.7263$ is calculated for the number of u in same ten subintervals of [0, 1] for the vector generations of 5000 times.

Hence, the density function of the vector components generated by this way is considerably near to cae^{-au} $(0 \leq u \leq 1)$.

Similarly, for $n=8$,

$$g(x)=\begin{cases} 8\,e^{-8x} & (0\leq x\leq 0.4) \\ -0.01852x+0.04074 & (0.4<x\leq 2.2) \\ 0 & (\text{otherwise}) \end{cases}$$

then $\chi^2=6.6362<16.919$.

For $n=10$,

$$g(x)=\begin{cases} 10\,e^{-10x} & (0\leq x\leq 0.351) \\ -0.09375x+0.11250 & (0.4\leq x\leq 1.2) \\ 0 & (\text{otherwise}) \end{cases}$$

then $\chi^2=5.8067<16.919$.

4. Probability Vectors with Uniformly Distributed Components

The mean value of the probability vector components is $1/n$ for the n states. So, it will be difficult that the distribution of the vector components is uniformalized in the range of $[0,1]$. Then, uniformalize the distribution within limited interval.

Consider that a basic value B_p is given previously for each state j $(j=1,\cdots,n)$.

A generating procedure of vectors whose components have the uniform distribution within the interval of $(B_p, 1-(n-1)B_p]$ is shown by the flow-chart in Fig. 5.

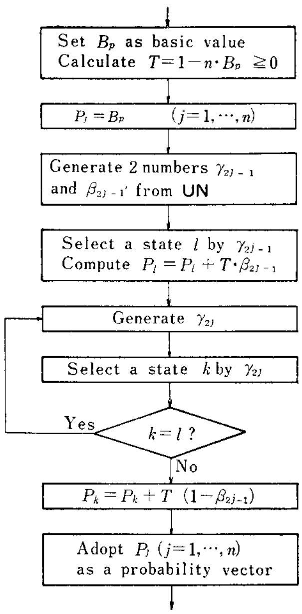


Fig. 5. Generating procedure.

Let M be the number of the probability vectors, then $2\cdot M/n$ denotes the frequency that the assigned value falls in the range of $(B_p, 1-(n-1)B_p]$. The frequency that the assigned value is B_p is denoted by $M\cdot(n-2)/n$.

5. Conclusion

Various transition matrices or probability vectors are made by these generating methods. By using those generated matrices or vectors, we will be able to simulate the higher order transition probability of markov chains, the optimal sequence of stochastic sequential machines, the learning of probabilistic automata and others.

Furthermore, we hope, in games against nature, a valuation for various criterions of behaviour will be derived from the simulations of games using thus generated matrices or vectors.

References

1) T. Tsuda, Monte Carlo Method and Simulation, Baifukan (1969).
2) Y. Morita, Introduction to Statistical Mathematics, Nihonhyoronsha (1968).
3) M. Boll, Tables Numeriques Universelles, (1957).
(Translation; S. Iyanaga and K. Yano, Hakusuisha (1965)).
4) T. Kutsuwa, T. Ikai, E. Kinoshita, H. Kosako and Y. Kojima, Trans. of IECE of Japan, **55**-D, 1 (1972).