“Composition and Rejection” 法に関する注意

志　村　利　雄

（1967年11月受付）

A Note on Composition and Rejection Method

Toshio SHIMURA

We study the problem how to get the sample value of random variable depending on a given probability density function $f(x)$ $(-\infty<x<+\infty)$ from the sample values of uniformly distributed random variable in the interval [0, 1].

One of such methods, that is composition and rejection method, was found by Butcher and Messel [1] and it is very useful to many Monte Carlo calculations.

We study the method theoretically for making clear the theoretical relations and show an example of gamma distribution how to sample the random numbers depending on the density having a singularity.

The Institute of Statistical Mathematics

確率密度関数 $f(x)$ $(-\infty<x<+\infty)$ があたえられたとき，この確率密度に従う確率変数のサンプル値を，区間 [0, 1] の中で値をとる一様分布する確率変数のサンプル値をもとにしてもとめる問題を考える．[0, 1] の中の値をとる確率変数を ξ とすれば $f(x)$ の分布関数 $F(x)$ を用いて．変換

$$X=F^{-1}(\xi) \tag{1}$$

を行えば，X が $f(x)$ に従う確率変数になることはよく知られている．ここで $F^{-1}(\cdot)$ は $F(\cdot)$ の逆関数である．もし積分

$$F(x)=\int_{-\infty}^{x} f(t)\,dt \tag{2}$$

が explicit にもとめることができれば，(1) に ξ のサンプル値を代入すれば，$f(x)$ に従う X のサンプル値を得るから，理論的には [0, 1] 上の一様分布する確率変数のサンプル値から，$F(x)$ に従う確率変数のサンプル値が計算できる．しかし一般的には，(2) の不定積分が explicit にもとまらないことが多い．そのような場合に極めて有用な方法 (composition and rejection method) を Butcher and Messel [1] が提案し，乱数をもとめる問題に応用している．

ここでは，まず最初に特別な場合の確率空間と確率変数の存在について再考し，あたえられた頻度関数に従う確率変数をある確率空間上で構成し，最後に区間 [0, 1] 上で一様分布する確率変数のサンプル値から目的の確率変数のサンプル値をもとめる方法を [1] にならって述べる．特に分布関数の不定積分が explicit にもとめることができず，しかも確率密度 $f(x)$

が直線上有限のところに "特異点" をもつ場合にも，少くとも理論的にはその乱数を一様乱数からもとめることができることをガンマ分布の例を用いて説明する.

1. 直線 $R^1=(-\infty,+\infty)$ 上で定義された関数 $G(x)$ がつぎの条件をみたすとき広義 (improper) 分布関数という.

(i) $G(-\infty)=0$,

(ii) $G(\cdot)$ は単調非減少で右連続,

(iii) $G(+\infty)\leqq 1$.

$\boldsymbol{B}^1$ を R^1 上のボレル集合の全体からなる完全加法族，$R^*=R^1\cup\{\partial\}$ は R^1 に孤立点として一点を添加した空間とし，$\boldsymbol{B}^*$ は $\boldsymbol{B}^1$ と $\{\partial\}$ とから生成される完全加法族とする. R^* に順序を入れるために $R^1\subset R^*$ の元素は R' の順序をもつと考え，すべての $a\in R^1$ に対しては $a<\partial$ であると仮定する.

R^1 上の広義分布関数 $G(x)$ には，R^1 上の測度 μ が対応して $\mu(a,b]=G(b)-G(a)$ が成立つことはよく知られている. このとき，

$$G^*(x)=\begin{cases} G(x) & x\in R^1 \\ 1 & x=\partial \end{cases}$$

とおけば，$G^*(x)$ は

(i) $G^*(-\infty)=0$

(ii) $G^*(\cdot)$ は単調非減少でかつ R^1 の中で右連続,

(iii) $G^*(\partial)=1$

をみたすことは明らかである. また $(R^*,\boldsymbol{B}^*)$ 上の測度 μ^* が存在して,

$$\mu^*(a,b]=G^*(b)-G^*(a)\,.$$

である. なぜなら $E\in\boldsymbol{B}^1$ のとき

$$\mu^*(E)=\mu(E)$$
$$\mu^*(E\cup\{\partial\})=1-\mu(R^1-E)$$

によって μ^* を定義すれば μ^* は有限加法的である. すなわち，$E_1,E_2\in\boldsymbol{B}^1$ が互に素 ($E_1\cap E_2=\phi$) ならば

$$\begin{aligned}\mu^*(E_1\cup E_2)&=\mu^*(E_1)+\mu^*(E_2)\\ \mu^*(E_1\cup\{\partial\}\cup E_2)&=\mu^*((E_1\cup E_2)\cup\{\partial\})\\ &=1-\mu(R^1-(E_1\cup E_2))\\ &=1-\mu(R^1-E_1)+\mu(E_2)\\ &=\mu^*(E_1\cup\{\partial\})+\mu^*(E_2)\end{aligned}$$

から有限加法的であることがわかる. 全く同じような方法で μ^* が完全加法的であることもわかる. つぎに，$a,b\in R^1$ のとき

$$\mu^*(a,b]=\mu(a,b]=G(b)-G(a)=G^*(b)-G^*(a)\,.$$

また,

$$\mu^*(a,\partial]=1-\mu(-\infty,a]=1-G(a)=G^*(\partial)-G^*(a)\,.$$

となる.

この μ^* を F^* に対応する確率測度とよぶ.

したがって R^1 上の広義分布関数はいつも R^* 上の "分布関数" に延長できる. (R^1 上の分布関数は $\{\partial\}$ に測度0をあたえることによって R^* 上に延長できる).

前節の意味で $F^*(x)$ が $R^*=R^1\cup\{\partial\}$ 上の分布関数であれば，確率空間 $(\Omega,\mathfrak{B},\boldsymbol{P})$ および $(\Omega,\mathfrak{B})$ 上の可測関数 X^* が存在して，$F^*(x)$ は X^* の分布関数になり，もし $F^*(\infty)=1$

ならば $\boldsymbol{P}\{X^*=\partial\}=0$ である．なぜなら，$\Omega=R^*$，$\mathfrak{B}=\boldsymbol{B}^*$ とし $\omega\in\Omega$ に対して，$X^*(\omega)=\omega$ とおき，$\boldsymbol{P}$ は F^* に対応する確率測度とすれば，

$$\boldsymbol{P}\{X^*\leqq x\}=F^*(x)\ .$$

したがって $(\Omega,\mathfrak{B},\boldsymbol{P})$ と X^* がもとまりそのうえ，$F^*(\infty)=1$ ならば，

$$\boldsymbol{P}\{X^*=\partial\}=F^*(\partial)-F^*(\infty)=1-1=0$$

である．

ここで Kolmogorov の拡張定理に相当する Ionescu Tulcea の定理（[2] p. 166）を用いれば，$F_k^*(x)$ $(k=1,2,\cdots)$ が可算個の R^* 上の分布関数であるとき，確率空間 $(\Omega^*,\mathfrak{B}^*,\boldsymbol{P}^*)$ とその上の確率変数列 $(X_1^*,X_2^*,\cdots)$ が存在して，

$$\boldsymbol{P}^*\{X_1^*\leqq x_1,\cdots,X_n^*\leqq x_n\}=F_1^*(x_1)\cdots F_n^*(x_n)$$

であり，もし $F_k^*(x)$ が R^1 上の（proper な）分布関数からの延長であれば，

$$\boldsymbol{P}^*\{X_k^*\in R^1\}=1$$

であることがわかる．このとき，確率変数 $X_1^*,X_2^*,\cdots$ は互に独立である．

2. R^1 上の“頻度関数”が

$$f(x)=\sum_{i=1}^{\infty}\alpha_i f_i(x)\quad(-\infty<x<\infty) \tag{3}$$

という形をしている場合に，この頻度関数に従う確率変数と，α_i，$f_i(x)$，$g_i(x)$ との関係をしらべてみる．ここで，$\alpha_i>0$，$\sum_{i=1}^{\infty}\alpha_i<\infty$；$f_i(x)$ は確率密度関数；$g_i(x)$ は $0\leqq g_i(x)\leqq 1$ で，R^1 上で可積分であると仮定する．

いま，

$$p_i=\frac{\alpha_i}{\sum_{j=1}^{\infty}\alpha_j},\qquad F_0(x)=\sum_{i\leqq x}p_i$$

とすれば $F_0(x)$ は R^1 上の分布関数である．つぎに

$$\left.\begin{aligned}F_i(x)&=\int_{-\infty}^{x}f_i(t)dt\\ G_i(x)&=\int_{-\infty}^{x}f_i(t)g_i(t)dt\end{aligned}\right\}\quad(i=1,2,\cdots)$$

とおけば $F_i(x)$ は R^1 上の分布関数，$G_i(x)$ は R^1 上の広義分布関数になる．いま $F_i(x)$，$F_i(x)$，$G_i(x)$ $(i=1,2,\cdots)$ を R^* 上に延長したものを $F_0^*(x)$，$F_i^*(x)$，$G_i^*(x)$ とおけば ∂ に質量をもつものは $G_i^*(x)$ に対応する測度のみである．Ionescu Tulcea の定理を適用して得られる確率関数を $(\Omega^*,\mathfrak{B}^*,\boldsymbol{P}^*)$ 確率変数を N，X_i，Y_i（N は F_0 に，X_i は F_i に，Y_i は G_i にそれぞれ対応している）とおく．作り方から，N，X_i，Y_i $(i=1,2,\cdots)$ は独立である．Y_N は Y_i と N との複合確率変数とすれば，$a\in R^1$ のとき

$$\begin{aligned}\boldsymbol{P}^*\{Y_N\leqq a\}&=\sum_{i=1}^{\infty}\boldsymbol{P}^*\{Y_N\leqq a|N=i\}\boldsymbol{P}^*\{N=i\}\\ &=\sum_{i=1}^{\infty}\boldsymbol{P}^*\{Y_i\leqq a\}p_i\\ &=\sum_{i=1}^{\infty}p_i\int_{-\infty}^{a}f_i(x)g_i(x)dx\\ &=\int_{-\infty}^{a}\sum_{i=1}^{\infty}p_i f_i(x)g_i(x)dx\end{aligned} \tag{4}$$

$$\begin{aligned}\boldsymbol{P}^*\{Y_N=\partial\} &= \sum_{i=1}^{\infty}\boldsymbol{P}^*\{Y_i=\partial\}\boldsymbol{P}^*\{N=i\}\\ &= \sum_{i=1}^{\infty}p_i\left[1-\int_{-\infty}^{\infty}f_i(x)g_i(x)dx\right]\\ &= \sum_{i=1}^{\infty}p_i\int_{-\infty}^{\infty}(1-g_i(x)f_i(x))dx\\ &= \int_{-\infty}^{\infty}\sum_{i=1}^{\infty}p_if_i(x)(1-g_i(x))dx\end{aligned}$$

となるから，

$$\boldsymbol{P}^*\{Y_N\leqq a\,|\,Y_N\in R^1\}=\frac{\displaystyle\int_{-\infty}^{a}\sum_{i=1}^{\infty}p_if_i(x)g_i(x)dx}{\displaystyle\int_{-\infty}^{\infty}\sum_{i=1}^{\infty}\alpha_if_i(x)g_i(x)dx}$$

したがって Y_N を R^1 に限定したその確率変数を $Y_N|R^1$ とかくときの分布関数は (3) 式を (R^1 上の分布関数になるように) 正則化したものになるから，$Y_N|R^1$ は頻度関数 (3) に従うことがわかる.

さて，$a\in R^1$ のとき

$$E^*\{g_N(X_N)\,;\,X_N\leqq a\}$$

を考えよう．ここで，E^* は $\boldsymbol{P}^*$ についての平均値，X_N は X と N との複合確率変数である.

$$\begin{aligned}E^*\{g_N(X_N)\,;\,X_N\leqq a\} &= \sum_{i=1}^{\infty}E^*\{g_N(X_N)\,;\,X_N\leqq a,\ N=i\}\\ &= \sum_{i=1}^{\infty}E^*\{g_i(X_i)\,;\,X_i\leqq a,\ N=i\}\\ &= \sum_{i=1}^{\infty}E^*\{g_i(X_i)\,;\,X_i\leqq a\}\boldsymbol{P}^*\{N=i\}\\ &= \sum_{i=1}^{\infty}p_i\int_{-\infty}^{a}g_i(x)f_i(x)dx\,.\end{aligned}$$

したがって，

(5) $$E^*\{g_N(X_N)\,;\,N_N\leqq a\}=\boldsymbol{P}^*\{Y_N\leqq a\}$$

頻度関数 (3) に従う確率変数 Y_N の存在がわかり，またその分布関数と，f_i, g_i との関係が (5) 式であたえられることもわかった.

3. (5) 式を使って具体的に (3) 式に従う確率変数のサンプル値をもとめることができる (Butcher and Messel [1]).

(I) N のサンプル値

[0,1] 上の一様乱数をとり，

$$x_0=\xi$$
$$x_i=x_{i-1}-p_i \qquad (i=1,2,\cdots)$$

とし，$x_{i_0-1}>0$, $x_{i_0}\leqq 0$ であるような番号 i_0 が N の1つのサンプル値になる.

(II) X_{i_0} のサンプル値

(I) で得た N のサンプル値 i_0 に対して $F_{i_0}(x)$ に従う確率変数 X_{i_0} のサンプル値 x_{i_0} をとり，$g_{i_0}(x_{i_0})$ を計算する.

(III) 再び [0,1] 上の一枚乱数 η をサンプルし，

$$\eta<g_{i_0}(x_{i_0})$$

ならば x_{i_0} を Y_N の1つのサンプル値とし，そうでなければ捨てる (reject する).

このような段階の操作で得られた x_{i_0} が Y_N のサンプル値になることは (4), (5) 式から明らかであろう.

［注　意］

1)　この方法で頻度関数 (3) のサンプル値として取り出したものは R^1 の中の値だけであるから,

$$
(6)\qquad P^*\{Y_N\in R^1\}=\int_{-\infty}^{\infty}\sum_{i=1}^{\infty}p_i f_i(x)g_i(x)dx=\frac{\int_{-\infty}^{\infty}\sum_{i=1}^{\infty}\alpha_i f_i(x)g_i(x)dx}{\sum_{i=1}^{\infty}\alpha_i}
$$

が大きいほど, サンプルされる能率はよいことになる. したがって, 実際にモンテカルロ法の計算等を行う場合には, 頻度関数 $f(x)$ を分解する際には, (6) 式の値が大きくなるようにすればよい. もし $f(x)$ が分布関数ならば (6) 式は $1\Big/\sum_{i=1}^{\infty}a_i$ になるから, この値が大きくなるように分解すればよい.

2)　(II) の段階で X_{i_0} のサンプル値は直接 $F_{i_0}(x)$ から求めなければならない. そのために $f(x)$ の分解に際しては $f_i(x)$ はなるべく容易にサンプルできる形のものにしなければならない.

4.　例

1)　正規乱数 (Hammersely and Handscomb [3])

これはすでによく知られている例である.

$$
f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\qquad(-\infty<x<\infty)
$$

$$
f(x)=\alpha_1 f_1(x)+\alpha_2 f_2(x)g_2(x)\,.
$$

$\lambda,\mu>0$ として

$$
\alpha_1=\mu\sqrt{\frac{2}{\pi}},\qquad \alpha_2=\frac{1}{\lambda}\sqrt{\frac{2}{\pi}}e^{\frac{1}{2}\lambda^2-\lambda\mu}
$$

$$
f_1(x)=\frac{1}{2\mu}\chi_{(-\mu,\mu]}(x)
$$

$$
f_2(x)=\frac{\lambda}{2}e^{-\lambda(|x|-\mu)}\chi_{(\mu,\infty)\cup(-\infty,-\mu]}(x)
$$

$$
g_1(x)=e^{-\frac{x^2}{2}},\qquad g_2(x)=e^{-\frac{1}{2}(|x|-\lambda)^2}
$$

$$
\chi_A(x)=\begin{cases}1 & x\in A\\ 0 & x\notin A\end{cases}
$$

“サンプリング能率” を最大にするためには $\alpha_1+\alpha_2$ を最少にすればよい, そのとき $\lambda=\sqrt{2}$, $\mu=\frac{2}{\sqrt{2}}$ である. “能率” は $\frac{\sqrt{2}}{3}\doteqdot 0.44$.

2)　ガンマ分布

$$
f(x)=\frac{x^{s-1}e^{-x}}{\Gamma(s)}\qquad(0<s<1)\ (x\geqq 0)
$$

$0<\tau_1\leqq\tau_2<\infty$ として, つぎのように分解する;

$$
f(x)=\alpha_{-2}f_{-2}(x)g_{-2}(x)+\alpha_{-1}f_{-1}(x)g_{-1}(x)+\sum_{n=0}^{\infty}\alpha_n f_n(x)g_n(x)\,.
$$

$$
\alpha_{-2}=\frac{\tau_2^{s-1}e^{-\tau_2}}{\Gamma(s)},\qquad \alpha_{-1}=\frac{\tau_1^{s-1}e^{-\tau_1}(\tau_2-\tau_1)}{\Gamma(s)}
$$

$$\alpha_n=\frac{\tau_1{}^s}{\Gamma(s)\,(2^s)^{n+1}} \qquad (n\geq 0)$$

$$f_{-2}(x)=e^{-(x-\tau_2)}\chi_{(\tau_2,\infty)}(x)$$

$$f_{-1}(x)=\frac{\chi_{(\tau_1,\tau_2)}(x)}{\tau_2-\tau_1}$$

$$f_n(x)=\frac{2^{n+1}}{\tau_1}\chi_{\left(\frac{\tau_1}{2^{n+1}},\ \frac{\tau_1}{2^n}\right]}(x) \qquad (n\geq 0).$$

$$g_{-2}(x)=\frac{x^{s-1}}{\tau_2{}^{s-1}}$$

$$g_{-1}(x)=\frac{x^{s-1}e^{-x}}{\tau_1{}^{s-1}e^{-\tau_1}}$$

$$g_n(x)=\frac{e^{-x}x^{s-1}(2^{s-1})^{n+1}}{\tau_1{}^{s-1}} \qquad (n\geq 0)$$

この場合のサンプリング能率は

$$\frac{1}{\tau(s)}\tau_1{}^s\left\{\frac{1}{2^s-1}+e^{-\tau_1}\left(\frac{\tau_2}{\tau_1}-1\right)+\left(\frac{\tau_2}{\tau_1}\right)^s\frac{e^{-\tau_2}}{\tau_2}\right\}$$

で$\tau_1=\tau_2=1$, $s=\frac{1}{2}$ のときおよそ 0.6 ぐらいである.

統計数理研究所

参 考 文 献

[1] Butcher, J. C. and Messel, H.: Electron number distribution in electron-photon showers in air and aluminium absorbers, Nuclear physics 20 (1960).

[2] Neveu, J.: Mathematical foundations of the calculus of probability, Holden-Day Inc. (1965).

[3] Hammersely, J. M. and Handscomb, D. C.: Monte Carlo methods, Methuen's monographs. (1965).