19. 乱数

概要: 我々は乱数をモンテカルロシミュレーションで利用する。モンテカルロシミュレーションでは、その中身を理解する必要がないため、よく用いられる。ここでは、そのモンテカルロシミュレーションで利用する乱数の紹介をする。

キーワード:乱数; 均一分布; ポアソン分布; 正規分布; 指数分布。

19.1. 序

モンテカルロシミュレーションにおいては、乱数を利用する。様々な確率密度関数が存在するため、それに応じて乱数を発生させる必要がある。ここでは、その乱数の発生のさせ方を議論する。

19.2. 均一乱数モデル

均一の乱数を発生させることを検討する。

まず、以下を設定する。

$$x_{n+1} \equiv 15 \times x_n \quad \left(\bmod 10^6 + 1\right) \tag{1}$$

$$x_0 = 1 \tag{2}$$

ここで x_{n+1} は $15x_n$ を 10^6+1 で割ったあまりとする。これは、以下のようになる。

$$\begin{aligned}
&x_0 = 1 \\
&x_1 = 15 \\
&x_2 \equiv 15 \times 15 \left(\bmod 10^6 + 1\right) \\
&x_2 = 225 \\
&x_3 \equiv 15 \times 225 \left(\bmod 10^6 + 1\right) \\
&x_3 = 3375 \\
&x_4 \equiv 15 \times 3375 \left(\bmod 10^6 + 1\right) \\
&x_4 = 50625 \\
&x_5 \equiv 15 \times 50625 \left(\bmod 10^6 + 1\right) \\
&x_5 = 759375 \\
&x_6 \equiv 15 \times 759375 \left(\bmod 10^6 + 1\right) \\
&x_6 = 390614 \\
&\cdots
\end{aligned} \tag{3}$$

したがって、以下の数字を得る。

$$1, 15, 225, 3375, 50625, 759375, 309614, \cdots \tag{4}$$

これは、0 から 10^6 までの値を 10^6 で割った余りである。これから、これらの数値を 10^6 で割ると、$[0,1]$ の乱数を得ることができる。

これをより一般化できる。以下のようなプロセスに変える。

$$x_{n+1} \equiv a x_n \quad \left(\mathrm{mod}\, 10^m + 1 \right) \tag{5}$$

$$x_0 = b \tag{6}$$

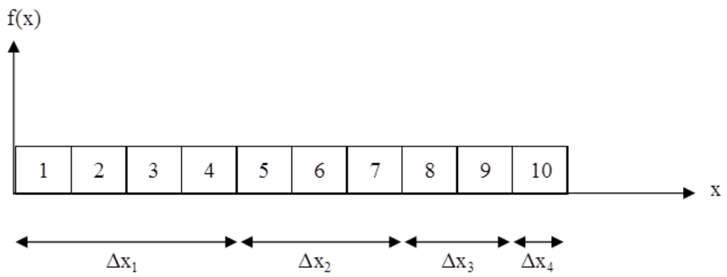
これから、以下の数値を得ることができる。

$$x_0, x_1, x_2, x_3, x_4, \cdots \tag{7}$$

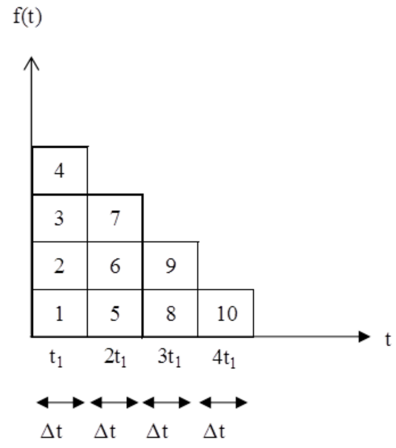
これは、n から 10^m の数値の列である。これを 10^m で割ると、$[0,1]$ の乱数を得ることができる。

19.3. 様々な確率密度分布に対する乱数

均一な乱数を使い、それを様々な確率密度分布に適用することを考える。



(a)



(b)

図 1　分布 f の場合のランダム数の定義。(a) 均一な乱数　(b)　変換された乱数

均一な乱数は範囲 $[0,1]$ の間の値を取る。 均一な乱数の場合、ある値になる確率はすべて同じとみなすことができる。

これを $1,2,3,\cdots,10$ に割り当てることを考える。

次に図 1(b)のような確率分布を考える。すなわち、1,2,3,4 は t_1 、5,6,7 は t_2 、8,9 は t_3 、10 は t_4 とする。すなわち、我々は均一な乱数を発生させると、それが $1,2,3,\cdots,10$ のうちどれに対応しているかわかり、さらにそれが t_1,t_2,t_3,t_4 のどれに対応しているかが分かる。すなわち、均一な乱数を発生させ、それから t_1,t_2,t_3,t_4 への対応をとれば、それは t の分布に対応していることになる。これが、任意の離散的な関数に対する乱数の変換になる。

表 1　変換された乱数

<table>
<tr><th>x</th><th>t</th></tr>
<tr><td>1</td><td rowspan="4">t_1</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td rowspan="3">$2t_1$</td></tr>
<tr><td>6</td></tr>
<tr><td>7</td></tr>
<tr><td>8</td><td rowspan="2">$3t_1$</td></tr>
<tr><td>9</td></tr>
<tr><td>10</td><td>$4t_1$</td></tr>
</table>

19.4. 乱数発生の逆関数法

均一の乱数の定義域は$[0,1]$ であり、それは均一である。その分布を $g(x)$ とする。$g(x)$は均一であるから、それをaと置くと、

$$\int_0^1 g(x)dx = \int_0^1 adx = a = 1 \tag{8}$$

よって、

$$g(x) = 1 \tag{9}$$

である。

あ確率密度を考えると、以下の関係になる。

$$g(x)\Delta x = f(t)\Delta t \tag{10}$$

$g(x)=1$,であるから、これは以下になる。

$$\Delta x = f(t)\Delta t \tag{11}$$

これを積分して

$$\int_0^x dx = \int_0^t f(t)dt \tag{12}$$

したがって、

$$x = F(t) \tag{13}$$

となる。ここで、

$$F(t) \equiv \int_0^t f(t)\,dt \tag{14}$$

である。これから、確率密度 $f(t)$ に対応する t は以下になる。

$$t = invF(x) \tag{15}$$

ここで、x は $[0,1]$ の均一な乱数であり、それを

$$x = Rand[1] \tag{16}$$

とすると、

$$t = invF(x) = invF(Rand[1]) \tag{17}$$

となる。これが、任意の確率密度 $f(t)$ を反映した乱数となる。

19.5. 指数分布と関連する乱数

前節で任意の確率密度分布に対する乱数の発生法を示した。ある、特定の関数の場合は、より簡単になる場合がある。その中の関数として指数分布がある。ここでは指数分布に関連する乱数の発生法を示す。

指数分布は以下で与えられる。

$$f(t) = \mu \exp(-\mu t) \tag{18}$$

この累積積分は以下である。

$$F(t) = \int_0^t \mu \exp(-\mu t)dt = 1 - \exp(-\mu t) \tag{19}$$

したがって、以下の関係になる。

$$x = 1 - \exp(-\mu t) \tag{20}$$

よって、指数分布と関連する変数 t は以下のように発生させることができる。

$$t = -\frac{1}{\mu}\ln[1-x] = -\frac{1}{\mu}\ln[1 - Rand[1]] \tag{21}$$

$Rand[1]$は $[0,1]$ の均一分布を持つ乱数である。したがって、$1 - Rand[1]$ も同様に $[0,1]$ の均一分布を持つ乱数である。したがって、これは以下のようにもなる。

$$t = -\frac{1}{\mu}\ln[Rand[1]] \tag{22}$$

$Rand[1]$ は通常

$$0 \le Rand[1] < 1 \tag{23}$$

である。つまり、$Rand[1] = 0$ は起こるが、$Rand[1] = 1$ は起こらない。$Rand[1]$ が 0 の場合、Eq. (22)はより簡便な表式であるが、発散してしまう。したがって、 Eq. (21) がこの理由で用いられる。

19.6. 正規分布と関連する乱数

正規分布は以下で与えられる。

$$f(t)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)^2\right] \tag{24}$$

この積分は、以下である。

$$\begin{aligned}F(t)&=\int_{-\infty}^{t}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\left(\frac{t'-\mu}{\sqrt{2}\sigma}\right)^2\right]dt'\\&=\frac{1}{2}\left[1+Erf\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)\right]\end{aligned} \tag{25}$$

したがって、以下を得る。

$$t=\mu+\sqrt{2}\sigma Erf^{-1}\left[2Rand[1]-1\right] \tag{26}$$

正規分布の乱数は、以下のようにも単純に得ることができる。

$Rand[1]$ の平均値は以下である。

$$\int_0^1 x\bullet 1dx=\frac{1}{2} \tag{27}$$

その分散は以下である。

$$\begin{aligned}&\int_0^1\left(x-\frac{1}{2}\right)^2\bullet 1dx=\int_0^1\left(x^2-x+\frac{1}{4}\right)dx\\&=\frac{1}{3}-\frac{1}{2}+\frac{1}{4}\\&=\frac{4-6+3}{12}\\&=\frac{1}{12}\end{aligned} \tag{28}$$

n 変数を利用すると、その分布は平均 $n/2$ 分散 $n/12$ の正規分布に近づいていく。したがって、

変数

$$z=\frac{\left(\sum_{i=1}^{n}Rnad_i[1]\right)-\frac{n}{2}}{\sqrt{\frac{n}{12}}} \tag{29}$$

は、標準正規分布に従うと考えることができる。

変換された

$$t=\sigma\frac{\left(\sum_{i=1}^{n}Rnad_i\left[1\right]\right)-\frac{n}{2}}{\sqrt{\frac{n}{12}}}+\mu \tag{30}$$

は平均μ 標準偏差 σの正規分布に従う。$n=12$と置くと、 以下を得る。

$$t=\sigma\left[\left(\sum_{i=1}^{12}Rnad_i\left[1\right]\right)-6\right]+\mu \tag{31}$$

これは、正にも負にもなることができる。しかしながら、ある場合には正の値のみが許さる。例えば、サービス時間や訪問時間間隔などがその例である。その場合には負の値が出た場合はデータを捨て、正の値の場合のみ採用する。この場合は、実際の平均値μ'は分布の平均値とは異なる。μ'は以下のように評価される。

$$\mu'=\frac{\int_0^\infty t\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)^2\right]dt}{\int_0^\infty \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)^2\right]dt} \tag{32}$$

ここで、変数

$$s=\frac{t-\mu}{\sqrt{2}\sigma} \tag{33}$$

を導入すると、上の積分は以下のようになる。

$$\begin{aligned}\mu'-\mu&=\frac{\int_0^\infty \left(t-\mu\right)\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)^2\right]dt}{\int_0^\infty \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\left(\frac{t-\mu}{\sqrt{2}\sigma}\right)^2\right]dt}\\&=\frac{\int_{-\frac{\mu}{\sqrt{2}\sigma}}^\infty s\exp\left[-s^2\right]dt}{\int_{-\frac{\mu}{\sqrt{2}\sigma}}^\infty \exp\left[-s^2\right]dt}\\&=\frac{\exp\left[-\left(\frac{\mu}{\sqrt{2}\sigma}\right)^2\right]}{\sqrt{\pi}\left[1+erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right]}\end{aligned} \tag{34}$$

つまり、以下となる。

$$\mu' = \mu + \frac{\exp\left[-\left(\frac{\mu}{\sqrt{2}\sigma}\right)^2\right]}{\sqrt{\pi}\left[1+erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right]} \tag{35}$$

これは、期待通り μ より大きくなる。

19.7. 1 から n までの整数の乱数

1 から n までの整数の乱数は以下で発生させることができる。

$$Int\left[Rand[1]\times 10n+1\right] \tag{36}$$

19.8. まとめ

この章のまとめを行う。

以下の演算を行う。

$$x_{n+1} \equiv ax_n \quad \left(\bmod 10^m+1\right)$$

$$x_0 = b$$

そして、以下の数列を得る。

$$x_0, x_1, x_2, x_3, x_4, \cdots$$

これは、0 から 10^m までの乱数となる。これを10^mで割ると、これは$[0,1]$の乱数になる。

我々は、任意の関数 $f(t)$ の分布を持つ乱数を以下のように発生させることができる。

まず、関数を積分する。

$$F(t) \equiv \int_0^t f(t)dt$$

$F(t)$ の逆関数を $invF$ とすると、この分布に関連する乱数は均一の乱数を用いて以下のように発生させることができる。

$$t = invF(x) = invF(Rand[1])$$

指数分布は以下で与えられるが、

$$f(t) = \mu\exp(-\mu t)$$

対応する乱数は以下となる。

$$t = -\frac{1}{\mu}\ln[1-x] = -\frac{1}{\mu}\ln[1-Rand[1]]$$

正規分布の乱数は以下のように発生させることができる。

$$t = \sigma\left[\left(\sum_{i=1}^{12} Rnad_i\left[1\right]\right) - 6\right] + \mu$$

1 から n の乱数は以下で与えられる。

$$Int\left[Rand\left[1\right] \times 10n + 1\right] \quad (37)$$