平成 19 年 6 月 22 日

プログラム中のある乱数の評価 その 2

新潟工科大学 情報電子工学科 竹野茂治

1 はじめに

C 言語の、0 から RAMD_MAX[1] までの整数の乱数を返す rand() という関数を使って、例えばサイコロなどのように 1 から L まで ($L \ll$ RAND_MAX) の整数をランダムに作り出したい場合、

```
y1=rand()%L+1;
```
(A%B = A を B で割った余り) (1)

とするのが手っ取り早そうであるが、実はこれには問題がある。

以前 (少なくとも一部の rand() の実装では)、偶数と奇数が交互に生成される仕様になっていて、例えばサイコロの $L = 6$ の場合でも、1,3,5 のいずれかと 2,4,6 のいずれかが交互に現われてしまうことになる。これではランダムとは言えないだろう。

これはもちろん rand() の実装自体に問題があるのであるが、この場合 (1) の代わりに、

```
y2=(int)((double)rand()/(RAND_MAX+1.0)*L+1);
```
(2)

のような方法がよく用いられる。これは、x =rand() に対し、

$$p_2 = \frac{x}{\texttt{RAND\_MAX} + 1.0} \tag{3}$$

によって $0.0 \le p_2 < 1.0$ の範囲の実数の乱数 p_2 を生成し、

$$y_2 = \lfloor p_2 L + 1 \rfloor \quad (\lfloor t \rfloor \text{ は } t \text{ 以下の最大の整数}) \tag{4}$$

によって 1 から L までの乱数を得る方法である。今回は、この方法の妥当性について検討してみることにする。

[1] RAMD_MAX は、stdlib.h 等のヘッダーファイルで定義されている値で、例えば 16 進数の 7fffffff などのように定義されている。

2 比較する確率変数

前回の [1] での考察からすると、(3) は、

$$p_3 = \frac{x + 0.5}{R_M + 1.0} \tag{5}$$

(以後、R_M =RAND_MAX とする) とした方がいいのでは、と思うかもしれないし、また、

$$p_4 = \frac{x}{R_M} \tag{6}$$

によって、$0.0 \leq p_4 \leq 1.0$ の範囲の実数の乱数 p_4 を作り、

$$y_4 = \lfloor p_4(L-1) + 1 \rfloor \tag{7}$$

としたら、と考える人もいるかもしれない。

これら y_2, $y_3 (= \lfloor p_3 L + 1 \rfloor)$, y_4 を数学的に比較してみることにする。

なお、ここでは x =rand() は 0 から R_M までの一様な乱数、すなわち 0 から R_M までのどの整数になる確率も $1/(R_M + 1)$ に等しい確率変数である、とする。

今回は、1 から L までの値を取る確率変数 y_j を考えるわけであるが、望まれるのは、それらがなるべく均等に現われること、すなわち、$1 \leq k \leq L$ であるすべての整数 k に対して、

$$\mathrm{Prob}\{y_j = k\} \ (=「y_j = k \text{ である確率}」) \ \approx \frac{1}{L} \tag{8}$$

を満たすことである ('≈' はほぼ等しいことを表わす)。

3 y_2 で $L = 2$ の場合

簡単のため、まず y_2 で $L = 2$ の場合の確率を考えてみることにする。

$L = 2$ なので、$y_2 = 1$ または $y_2 = 2$ であるが、(3), (4) より、

$$\begin{aligned}
\mathrm{Prob}\{y_2 = 1\} &= \mathrm{Prob}\{\lfloor 2p_2 \rfloor = 0\} = \mathrm{Prob}\{0 \le 2p_2 < 1\} \\
&= \mathrm{Prob}\{0 \le 2x/(R_M + 1) < 1\} \\
&= \mathrm{Prob}\{0 \le x < (R_M + 1)/2\}, \qquad (9) \\
\mathrm{Prob}\{y_2 = 2\} &= \mathrm{Prob}\{\lfloor 2p_2 \rfloor = 1\} = \mathrm{Prob}\{1 \le 2p_2 < 2\} \\
&= \mathrm{Prob}\{1 \le 2x/(R_M + 1) < 2\} \\
&= \mathrm{Prob}\{(R_M + 1)/2 \le x < R_M + 1\} \qquad (10)
\end{aligned}$$

のようになる。ここで、次の補題を利用する。

補題 1

$\lceil t \rceil$ を t 以上の最小の整数を表わすとするとき、$-1 < A < B \le R_M + 1$ となる A, B に対して

$$\mathrm{Prob}\{A \le x < B\} = \frac{\lceil B \rceil - \lceil A \rceil}{R_M + 1} \qquad (11)$$

証明

$-1 < A < B \le R_M + 1$ ならば

$$0 \le \lceil A \rceil \le \lceil B \rceil \le R_M + 1 \qquad (12)$$

であり、一方、

$$\begin{aligned}
&\mathrm{Prob}\{A \le x < B\} \\
&\quad = \mathrm{Prob}\{x \ge A\} - \mathrm{Prob}\{x \ge B\} = \mathrm{Prob}\{x \ge \lceil A \rceil\} - \mathrm{Prob}\{x \ge \lceil B \rceil\} \\
&\quad = \mathrm{Prob}\{x = \lceil A \rceil\} + \mathrm{Prob}\{x = \lceil A \rceil + 1\} + \cdots + \mathrm{Prob}\{x = \lceil B \rceil - 1\}
\end{aligned}$$

であり、(12) より、これらの確率はすべて $1/(R_M + 1)$ となる。よって、(11) が成り立つ。■

この補題 1 を用いれば、(9), (10) より、

$$\begin{aligned}\mathrm{Prob}\{y_2=1\} &= \frac{1}{R_M+1}\left\lceil\frac{R_M+1}{2}\right\rceil,\\ \mathrm{Prob}\{y_2=2\} &= \frac{1}{R_M+1}\left(R_M+1-\left\lceil\frac{R_M+1}{2}\right\rceil\right)=1-\frac{1}{R_M+1}\left\lceil\frac{R_M+1}{2}\right\rceil\end{aligned}$$

となる。

R_M が奇数の場合、$(R_M+1)/2$ は整数なので、

$$\mathrm{Prob}\{y_2=1\}=\mathrm{Prob}\{y_2=2\}=\frac{1}{R_M+1}\,\frac{R_M+1}{2}=\frac{1}{2} \tag{13}$$

となり、R_M が偶数の場合は、

$$\left\lceil\frac{R_M+1}{2}\right\rceil=\left\lceil\frac{R_M}{2}+\frac{1}{2}\right\rceil=\frac{R_M}{2}+1$$

なので、

$$\mathrm{Prob}\{y_2=1\}=\frac{1}{R_M+1}\left(\frac{R_M}{2}+1\right),\quad \mathrm{Prob}\{y_2=2\}=\frac{1}{R_M+1}\,\frac{R_M}{2} \tag{14}$$

となることになる。

R_M が偶数の場合は、わずかに違いが出ることになるが、この場合 (R_M+1) が奇数になるので、それをなるべく等しく 2 つに分けるには $R_M/2$ と $(R_M/2)+1$ に分けるのが最善であるから、(13), (14) は、いずれの場合もある意味で (8) のように確率を均等にする最良の分け方 (のひとつ) であることになる。なお、(14) の確率が逆の場合も同じ意味で最良であり、このような最良の分配方法は一意に決まるわけではない。

とりあえず、$L=2$ の場合は、y_2 はそれなりに妥当なものであることが言えたことになる。

4 y_2 の一般の場合

この節では、3 節と同様にして、一般の L に対する y_2 を考えてみることにする。

$1 \leq k < L$ に対し、

$$\begin{aligned} \mathrm{Prob}\{y_2 = k\} &= \mathrm{Prob}\{\lfloor p_2 L \rfloor = k-1\} = \mathrm{Prob}\{k-1 \leq p_2 L < k\} \\ &= \mathrm{Prob}\{(k-1)/L \leq p_2 < k/L\} \\ &= \mathrm{Prob}\{(k-1)(R_M+1)/L \leq x < k(R_M+1)/L\} \end{aligned}$$

となり、$(k-1)(R_M+1)/L \geq 0$, $k(R_M+1)/L \leq R_M+1$ なので、補題 1 により

$$\mathrm{Prob}\{y_2 = k\} = \frac{1}{R_M+1}\left\{\left\lceil \frac{R_M+1}{L}k \right\rceil - \left\lceil \frac{R_M+1}{L}(k-1) \right\rceil\right\} \tag{15}$$

となる。今、(15) の分子 (中カッコ内) を α_k とすれば、

$$\alpha_k = \left\lceil \frac{R_M+1}{L}k \right\rceil - \left\lceil \frac{R_M+1}{L}(k-1) \right\rceil \tag{16}$$

となるが、これが 2 つの整数のいずれかとなることを示す。

補題 2

整数 $x = \lceil A+B \rceil - \lceil A \rceil$ は、

$$\lceil B \rceil - 1 \leq x \leq \lceil B \rceil$$

を満たし、よって、$x = \lceil B \rceil - 1$ か $x = \lceil B \rceil$ かのいずれかである。

証明

今、任意の実数 t に対して

$$\langle t \rangle = \lceil t \rceil - t$$

と書くことにすると、$t \leq \lceil t \rceil < t+1$ より $0 \leq \langle t \rangle < 1$ であり、よって、

$$\begin{aligned} x &= \lceil A+B \rceil - \lceil A \rceil = \lceil \lceil A \rceil - \langle A \rangle + B \rceil - \lceil A \rceil \\ &= \lceil A \rceil + \lceil -\langle A \rangle + B \rceil - \lceil A \rceil = \lceil B - \langle A \rangle \rceil \end{aligned}$$

となる。よって、$-1 < -\langle A \rangle \leq 0$ より、

$$\lceil B \rceil - 1 \leq \lceil B - \langle A \rangle \rceil = x \leq \lceil B \rceil$$

が言える。■

この補題 2 と (16) より、

$$\left\lceil \frac{R_M + 1}{L} \right\rceil - 1 \leq \alpha_k \leq \left\lceil \frac{R_M + 1}{L} \right\rceil \tag{17}$$

であること、および、α_k が $\lceil (R_M + 1)/L \rceil - 1$ か $\lceil (R_M + 1)/L \rceil$ かのいずれかであることがわかる。

それは、α_1,α_2,...,α_L に、高々 1 だけしか違いがないことを意味し、また、もちろん、

$$\sum_{k=1}^{L} \alpha_k = R_M + 1$$

であるから、これらは、$(R_M + 1)$ を L 個へできるだけ均等に分割する方法として、(16) の α_k が最適 (なものの一つ) であることを意味する。

これにより、y_2 の妥当性の保証が得られたことになる。

5 y_3 の場合

次は、y_3 の場合を考えてみる。$1 \leq k \leq L$ に対し (5) より、

$$\begin{aligned}
\mathrm{Prob}\{y_3 = k\} &= \mathrm{Prob}\{\lfloor p_3 L \rfloor = k - 1\} = \mathrm{Prob}\{k - 1 \leq p_3 L < k\} \\
&= \mathrm{Prob}\{(k-1)/L \leq p_3 < k/L\} \\
&= \mathrm{Prob}\{(k-1)(R_M+1)/L \leq x + 0.5 < k(R_M+1)/L\} \\
&= \mathrm{Prob}\{(k-1)(R_M+1)/L - 0.5 \leq x < k(R_M+1)/L - 0.5\}
\end{aligned}$$

となるが、$(k-1)(R_M+1)/L - 0.5 \geq -0.5$, $k(R_M+1)/L - 0.5 < R_M + 1$ なので、やはり補題 1 により、

$$\mathrm{Prob}\{y_3 = k\} = \frac{1}{R_M + 1} \left\{ \left\lceil \frac{R_M + 1}{L} k - 0.5 \right\rceil - \left\lceil \frac{R_M + 1}{L} (k-1) - 0.5 \right\rceil \right\}$$

となる。よってこの分子を β_k とすれば、補題 2 によりこの β_k も (17) 同様

$$\left\lceil \frac{R_M+1}{L} \right\rceil - 1 \le \beta_k \le \left\lceil \frac{R_M+1}{L} \right\rceil$$

を満たすことがわかり、よって y_3 も最適な均等配分のひとつであることがわかる。

なお、$\alpha_k = \lceil (R_M+1)/L \rceil$ となる k の個数と $\beta_k = \lceil (R_M+1)/L \rceil$ となる k の個数は同じになるが、そのような k の集合自体が一致するわけではないので (例えば $L=3$ で考えよ)、y_2 と y_3 が一致するわけではない。

同様にすれば、一般に、$0 \le \mu < 1$ である実数定数 μ に対し、

$$p_5 = \frac{x+\mu}{R_M+1}, \quad y_5 = \lfloor p_5 L + 1 \rfloor \tag{18}$$

も、y_3 と同じ議論により、y_2 と同じく最良の均等配分になることがわかる。

6 y_4 の場合

y_4 は、(6), (7) によって 1 から L までの整数が作られるので、この場合は、$1 \le k \le L$ に対し、

$$\begin{aligned} \mathrm{Prob}\{y_4 = k\} &= \mathrm{Prob}\{\lfloor p_4(L-1) \rfloor = k-1\} = \mathrm{Prob}\{k-1 \le p_4(L-1) < k\} \\ &= \mathrm{Prob}\{(k-1)/(L-1) \le p_4 < k/(L-1)\} \\ &= \mathrm{Prob}\{(k-1)R_M/(L-1) \le x < kR_M/(L-1)\} \end{aligned} \tag{19}$$

となる。ここで、$(k-1)R_M/(L-1) \ge 0$ であるが、$kR_M/(L-1)$ は、$k=L$ のとき、

$$\frac{R_M}{L-1}L = R_M + \frac{R_M}{L-1} > R_M + 1 \quad (R_M + 1 \gg L \text{ より})$$

となってしまうので、$k=L$ に対しては補題 1 を適用できず、別に考える必要がある。ただし、$k < L$ ならば、

$$\frac{R_M}{L-1}k \le \frac{R_M}{L-1}(L-1) = R_M$$

なので、補題 1 は適用できる。この場合は、補題 1 と (19) より、

$$\mathrm{Prob}\{y_4 = k\} = \frac{1}{R_M + 1}\left\{\left\lceil \frac{R_M}{L-1}k \right\rceil - \left\lceil \frac{R_M}{L-1}(k-1) \right\rceil\right\}$$

となる。$k = L$ のときは、(19) より、

$$\begin{aligned}\mathrm{Prob}\{y_4 = k\} &= \mathrm{Prob}\{(L-1)R_M/(L-1) \le x < R_M/(L-1) + R_M\} \\ &= \mathrm{Prob}\{x = R_M\} = \frac{1}{R_M + 1}\end{aligned}$$

となるので、y_4 に関しては、

$$\gamma_k = \begin{cases} \left\lceil \frac{R_M}{L-1}k \right\rceil - \left\lceil \frac{R_M}{L-1}(k-1) \right\rceil & (1 \le k < L) \\ 1 & (k = L) \end{cases}$$

に対して、

$$\mathrm{Prob}\{y_4 = k\} = \frac{\gamma_k}{R_M + 1}$$

となることになる。

$1 \le k < L$ の場合は、補題 2 より

$$\left\lceil \frac{R_M}{L-1} \right\rceil - 1 \le \gamma_k \le \left\lceil \frac{R_M}{L-1} \right\rceil \tag{20}$$

であることがわかる。

これは、γ_1,γ_2,...,γ_{L-1} が R_M の $(L-1)$ 個への最適な均等配分であることを意味し、よっていずれも $R_M/(L-1)$ $(\gg 1)$ に近く、γ_L のみ 1 となっていることになる。

よって、y_4 は、$y_4 = L$ となる確率以外は均等だが、$y_4 = L$ となる確率だけかなり小さい、ということになる。

元々、$p_4(L-1)+1$ を考えてみれば、これは確かに

$$1 \le p_4(L-1) + 1 \le L$$

となり、それぞれの等号が成立する場合もあるのであるが、この $p_4(L-1)+1$ の整数部分 $(=y_4)$ が L になるのは、$p_4=1$、すなわち x が丁度 R_M になるときのみであり、それ以外の整数部分とは明らかに割合が違うことが想像できると思う。

7 さらなる改善

4, 5 節の議論により、y_2, y_3, そして y_5 は同程度に確率を均等に配分していることがわかったが、$R_M+1\equiv 0 \pmod{L}$ の場合は完全に均等で、そうでなければ少しずれがあり、そのずれは

$$\frac{1}{R_M+1}\left\lceil\frac{R_M+1}{L}\right\rceil-\frac{1}{R_M+1}\left(\left\lceil\frac{R_M+1}{L}\right\rceil-1\right)=\frac{1}{R_M+1} \tag{21}$$

である。通常この値は非常に小さいので、十分均等であると言える。

一方で、式 (18) で定義される y_5 には $0\le\mu<1$ というパラメータが含まれ、この値を変えると、

$$\frac{1}{R_M+1}\left\lceil\frac{R_M+1}{L}\right\rceil,\quad \frac{1}{R_M+1}\left(\left\lceil\frac{R_M+1}{L}\right\rceil-1\right)$$

の現われ方が変わる。よって、μ を定数とせず、y_5 を使用する前に適当に取ってから y_5 を使用すると、適当にならされることでさらなる均等配分になることが期待される。

つまり、x, z を、ともに rand() から独立に計算される乱数値、すなわち 0 から R_M までの値を取る、独立で一様な確率変数とし、

$$\mu=\frac{z}{R_M+1},\quad p_6=\frac{x+\mu}{R_M+1},\quad y_6=\lfloor p_6L+1\rfloor \tag{22}$$

と定められる y_6 を考えてみる。なお、この y_6 の計算には rand() を 2 回使用することになる。

y_6 は x, z の 2 変数関数 $y_6=y_6(x,z)$ と考えることができ、よって、$1\le k\le L$ に対し、

$$\mathrm{Prob}\{y_6=k\}$$

$$
\begin{aligned}
&= \sum_{j=0}^{R_M} \mathrm{Prob}\{z = j \text{ かつ } y_6(x,j) = k\} \\
&= \sum_{j=0}^{R_M} \mathrm{Prob}\{z = j\}\mathrm{Prob}\{y_6(x,j) = k\} \quad (x \text{ と } z \text{ は独立}) \\
&= \sum_{j=0}^{R_M} \frac{1}{R_M+1}\mathrm{Prob}\{y_6(x,j) = k\}
\end{aligned}
$$

となるが、6 節と同様の計算により、

$$
\begin{aligned}
&\mathrm{Prob}\{y_6(x,j) = k\} \\
&= \frac{1}{R_M+1}\left\{\left\lceil\frac{R_M+1}{L}k - \frac{j}{R_M+1}\right\rceil - \left\lceil\frac{R_M}{L-1}(k-1) - \frac{j}{R_M+1}\right\rceil\right\}
\end{aligned}
$$

となるので、

$$
\begin{aligned}
&\mathrm{Prob}\{y_6 = k\} && (23)\\
&= \frac{1}{(R_M+1)^2}\sum_{j=0}^{R_M}\left\{\left\lceil\frac{R_M+1}{L}k - \frac{j}{R_M+1}\right\rceil - \left\lceil\frac{R_M+1}{L}(k-1) - \frac{j}{R_M+1}\right\rceil\right\} && (24)
\end{aligned}
$$

となる。

補題 3

整数 A、および $R_M + 1 > L$ に対して、

$$
\sum_{j=0}^{R_M}\left\lceil\frac{A}{L} - \frac{j}{R_M+1}\right\rceil = \left\lceil\frac{R_M+1}{L}A\right\rceil \tag{25}
$$

証明

まず、$0 \le A < L$ に対して (25) を示す。このとき、$0 \le A/L < 1$ なので、$0 \le j \le R_M$ に対し、

$$
-1 < \frac{A}{L} - \frac{j}{R_M+1} < 1
$$

となる。よって、

$$\left\lceil \frac{A}{L} - \frac{j}{R_M+1} \right\rceil$$

は、この場合 0 か 1 である。よって、これが 1 となる j の個数を数えればよい。これが 1 となるのは、

$$\frac{A}{L} - \frac{j}{R_M+1} > 0$$

のとき、すなわち、

$$0 \le j < \frac{R_M+1}{L} A$$

となる j の個数となるので、それは丁度 $\lceil (R_M+1)A/L \rceil$ 個となる。よって、$0 \le A < L$ のときは (25) が成り立つ。

一般の整数 A に対しては、A を L で割った商を Q, 余りを R とすれば、$A = QL + R$, $0 \le R < L$ で、

$$\left\lceil \frac{A}{L} - \frac{j}{R_M+1} \right\rceil = \left\lceil Q + \frac{R}{L} - \frac{j}{R_M+1} \right\rceil = Q + \left\lceil \frac{R}{L} - \frac{j}{R_M+1} \right\rceil$$

となり、R に対しては、(25) の A を R に変えたものが成り立つので、

$$\begin{aligned}
&\sum_{j=0}^{R_M} \left\lceil \frac{A}{L} - \frac{j}{R_M+1} \right\rceil = \sum_{j=0}^{R_M} \left(Q + \left\lceil \frac{R}{L} - \frac{j}{R_M+1} \right\rceil \right) \\
&= \quad Q(R_M+1) + \sum_{j=0}^{R_M} \left\lceil \frac{R}{L} - \frac{j}{R_M+1} \right\rceil = Q(R_M+1) + \left\lceil \frac{R_M+1}{L} R \right\rceil \\
&= \quad \left\lceil Q(R_M+1) + \frac{R_M+1}{L} R \right\rceil = \left\lceil (R_M+1) \left(Q + \frac{R}{L} \right) \right\rceil \\
&= \quad \left\lceil \frac{R_M+1}{L} A \right\rceil
\end{aligned}$$

となり、一般の A に対しても (25) が成り立つ。∎

この補題 3 により、(24) は、

$$\begin{aligned}
&\mathrm{Prob}\{y_6 = k\} \\
&= \frac{1}{(R_M+1)^2}\left\{\left\lceil\frac{R_M+1}{L}(R_M+1)k\right\rceil - \left\lceil\frac{R_M+1}{L}(R_M+1)(k-1)\right\rceil\right\} \\
&= \frac{\delta_k}{(R_M+1)^2}, \\
\delta_k &= \left\lceil\frac{(R_M+1)^2}{L}k\right\rceil - \left\lceil\frac{(R_M+1)^2}{L}(k-1)\right\rceil
\end{aligned}$$

となる。この δ_k は、補題 2 により、

$$\delta_k = \left\lceil\frac{(R_M+1)^2}{L}\right\rceil - 1,\ \left\lceil\frac{(R_M+1)^2}{L}\right\rceil$$

のいずれかであることもわかる。よって、$(R_M+1)^2 \equiv 0 \pmod{L}$ であれば、y_6 は均等で、そうでない場合のずれは、

$$\frac{1}{(R_M+1)^2}\left\lceil\frac{(R_M+1)^2}{L}\right\rceil - \frac{1}{(R_M+1)^2}\left\{\left\lceil\frac{(R_M+1)^2}{L}\right\rceil - 1\right\} = \frac{1}{(R_M+1)^2}$$

となり、(21) と比較すれば、こちらの方がはるかに小さいことがわかる。

ついでに、y_2, y_6 それぞれの分散も計算してみよう。いずれも平均は $1/L$ であるから、y_2 の分散 V_2 は、

$$V_2 = \frac{1}{L}\sum_{k=1}^{L}\left(\frac{\alpha_k}{R_M+1} - \frac{1}{L}\right)^2 \tag{26}$$

と定義されるが、今、

$$\left\lceil\frac{R_M+1}{L}\right\rceil = q, \quad r = qL - (R_M+1) = L\left\langle\frac{R_M+1}{L}\right\rangle$$

とすると、明らかに r は $0 \le r < L$ となる整数で、

$$R_M + 1 = qL - r, \quad q = \left\lceil\frac{R_M+1}{L}\right\rceil \tag{27}$$

となるので、この場合、α_k は、$(q-1)$ が r 個、q が $(L-r)$ 個あり、

$$(q-1)r + q(L-r) = qL - r = R_M + 1$$

となっていることがわかる。よって、V_2 は、

$$\begin{aligned} V_2 &= \frac{1}{L}\left\{\left(\frac{q-1}{R_M+1} - \frac{1}{L}\right)^2 r + \left(\frac{q}{R_M+1} - \frac{1}{L}\right)^2 (L-r)\right\} \\ &= \frac{r\{(q-1)L - (R_M+1)\}^2 + (L-r)\{qL - (R_M+1)\}^2}{L^3(R_M+1)^2} \\ &= \frac{r\{(q-1)L - (qL-r)\}^2 + (L-r)\{qL - (qL-r)\}^2}{L^3(R_M+1)^2} \\ &= \frac{r(L-r)^2 + (L-r)r^2}{L^3(R_M+1)^2} = \frac{r(L-r)}{L^2(R_M+1)^2} \end{aligned}$$

となる。

一方、y_6 の分散 V_6 は、

$$V_6 = \frac{1}{L}\sum_{k=1}^{L}\left(\frac{\delta_k}{(R_M+1)^2} - \frac{1}{L}\right)^2$$

であり、この場合は、(27) の代わりに

$$(R_M+1)^2 = \bar{q}L - \bar{r}, \quad \bar{q} = \left\lceil \frac{(R_M+1)^2}{L} \right\rceil, \quad 0 \le \bar{r} < L \tag{28}$$

と取れば、δ_k は、$(\bar{q}-1)$ が $\bar{r}$ 個、$\bar{q}$ が $(L-\bar{r})$ 個あり、よって、

$$\begin{aligned} V_6 &= \frac{1}{L}\left\{\left(\frac{\bar{q}-1}{(R_M+1)^2} - \frac{1}{L}\right)^2 \bar{r} + \left(\frac{\bar{q}}{(R_M+1)^2} - \frac{1}{L}\right)^2 (L-\bar{r})\right\} \\ &= \frac{\bar{r}\{(\bar{q}-1)L - (R_M+1)^2\}^2 + (L-\bar{r})\{\bar{q}L - (R_M+1)^2\}^2}{L^3(R_M+1)^4} \\ &= \frac{\bar{r}\{(\bar{q}-1)L - (\bar{q}L-\bar{r})\}^2 + (L-\bar{r})\{\bar{q}L - (\bar{q}L-\bar{r})\}^2}{L^3(R_M+1)^4} \\ &= \frac{\bar{r}(L-\bar{r})^2 + (L-\bar{r})\bar{r}^2}{L^3(R_M+1)^4} = \frac{\bar{r}(L-\bar{r})}{L^2(R_M+1)^4} \end{aligned}$$

となる。

$$0 \leq \frac{r(L-r)}{L^2},\ \frac{\bar{r}(L-\bar{r})}{L^2} \leq \frac{1}{4}$$

であり、もちろん $r, \bar{r}$ によって多少変わるが、一般的には V_2 よりも V_6 の方がかなり小さくなる。

8 最後に

結局今回は、(2) で (も (5) でも) それなりに妥当であることがわかったが、さらに改善したものとして、y_6 のように rand() を 2 回使う方法もあり、これは y_2 よりもはるかに均等になる。

しかし、通常はこのような方法で精度を向上しようとしても、疑似乱数 rand() の乱数としての質 (ランダムさ) が問題になるだろうし、むしろそちらの方の影響の方が強いのではないかとも思う。よって、通常は単に y_2 を使っておけば十分だろうと思う。

また、0.0 以上 1.0 未満の実数の乱数を返す drand48() のような実装を使って、

```
y=(int)(drand48()*L+1);
```

のようにすればもっと楽だし、多分それなりに品質もよいのだろうと思う。

参考文献

[1] 竹野茂治「プログラム中のある乱数の評価」(2007 年 5 月)