# 統計学入門-第8章 大数の法則と中心極限定理

張 梁

2007年5月22日

## 目 次

# 8　大数の法則と中心極限定理

## 8.1　大数の法則 (law of large numbers)

### 8.1.1　真の値への集中

正しいコインを 10 回投げるようなベルヌーイ試行を行い，表を 1 とし，裏は 0 とする確率変数 $x_i$ を考える。
表が出た回数 (**頻度**) は，和

$$r = x_1 + x_2 + \cdots + x_{10} \tag{1}$$

である。表が出た回数の割合 $\hat{p} = r/10$ は**観測された成功率**である。一般に $n$ をコイン投げの回数とするとき，$r/10$ は**相対頻度**である。$r$ は確率変数で，二項分布 $Bi(10, 0.5)$ に従い，期待値と分散は，

$$E(r) = np = 5, \qquad V(r) = np(1-p) = 2.5$$

割合 $r/10$ の期待値，分散は，

$$E(r/n) = p = 0.5, \qquad V(r/n) = (1/n^2)V(r) = p(1-p)/n = 0.025 \tag{2}$$

である。ここで，期待値である $p = 0.5$ は**真の成功率**である。
$n$ が大きくなると，事実上 $\hat{p} \fallingdotseq 0.5$ となっていく。$n = 100$ では，表の観測された成功率 $\hat{p} = r/10$ が 0.4 から 0.6 までの確率は 96 %を超え，ほとんどの値が $p = 0.5$ の周囲に集中する (教科書 P157 参照)。式で表現すると，

$$P(|r/n - 0.5| \leqq 0.1) \longrightarrow 1 \quad (n \to \infty) \tag{3}$$

である。一般に，$\varepsilon$ はどのように小さい正数であっても

$$P(|r/n - 0.5| \leqq \varepsilon) \longrightarrow 1 \quad (n \to \infty) \tag{4}$$

となることが保証されるが，大数の法則の一つの形である。

**大数の法則**　元の確率分布 (母集団) から $n$ 個の標本を採る。その $n$ 個の標本は確率変数 $X_1, X_2, \cdots, X_n$ であり，お互いは独立，それぞれが平均 $\mu$，分散 $\sigma^2$ の確率分布に従うとする。

$$\begin{aligned} E(\bar{X_n}) &= E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) \\ &= \frac{1}{n}\{E(X_1) + E(X_2) + \cdots + E(X_n)\} \\ &= \frac{1}{n}\{\mu + \mu + \cdots + \mu\} = \mu \end{aligned} \tag{5}$$

$$
\begin{aligned}
V(\bar{X_n}) &= V\left(\frac{X_1+X_2+\cdots+X_n}{n}\right) \\
&= \left(\frac{1}{n}\right)^2\{V(X_1)+V(X_2)+\cdots+V(X_n)\} \\
&= \frac{1}{n^2}\{\sigma^2+\sigma^2+\cdots+\sigma^2\} = \frac{\sigma^2}{n} \qquad (6)
\end{aligned}
$$

分散の式から分かることは，$n$ が小さいとき分散が大きくなる。つまり観測された値の期待値 (＝標本平均) は真の期待値 (＝母平均) と懸け離れた値である可能性が高い。ところで，$n$ が大きいとき分散が小さくなり，標本平均は「ほぼ」母平均と同じ値である。

任意の $\varepsilon > 0$ に対して，

$$\lim_{n\to\infty} P\left(|\bar{X}-\mu| \geqq \varepsilon\right) = 0 \qquad (7)$$

$$\left(\lim_{n\to\infty} P\left(|\bar{X}-\mu| < \varepsilon\right) = 1\right) \qquad (8)$$

が成り立つ。このとき $\bar{X}$ は $\mu$ に確率収束 (convergence in probability) するという。

### 8.1.2 大数の法則の証明

**証明** チェビシェフの不等式を利用して証明。
**チェビシェフの不等式** 任意の正数 $k$ について，

$$P\left(|X-E(X)| \geqq k\sqrt{V(X)}\right) \quad \leqq \quad 1/k^2 \qquad (9)$$

この不等式を利用して，式 (5)(6) の結果を式 (9) に代入すると，

$$P\left(|\bar{X}_n-\mu| \geqq k\sqrt{\frac{\sigma^2}{n}}\right) \quad \leqq \quad 1/k^2 \qquad (10)$$

が任意の正数 $k$ について成り立つ。ここで，任意の $\varepsilon > 0$,

$$k = \frac{\varepsilon}{\sqrt{\sigma^2/n}} = \frac{\sqrt{n}\varepsilon}{\sigma} \qquad (11)$$

とおき，式 (11) を (10) に代入すると，

$$P\left(|\bar{X}_n-\mu| \geqq \frac{\varepsilon}{\sqrt{\sigma^2/n}}\cdot\sqrt{\frac{\sigma^2}{n}}\right) \quad \leqq \quad \frac{1}{\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)^2} \qquad (12)$$

$$\Longrightarrow P\left(|\bar{X}_n-\mu| \geqq \varepsilon\right) \quad \leqq \frac{\sigma^2}{n\varepsilon^2} \qquad (13)$$

となり，$n \to \infty$ のとき右辺 $\to 0$ なので，式 (7) が成り立つことが分かる。

### 8.1.3 コンピューターシミュレーション (computer simulation)

人間ではできないこと，例えば，成功の確率 $p = 0.4$ であるようなベルヌーイ試行を 2 万回、2 万回以上繰り返し行うなどのことは，代わりにコンピューターを利用して実験を行う。

コンピューターに 0 から 1 までの一様乱数 $U$ を発生させ，(上の例)$U < 0.4$ であれば成功，$U \geqq 0.4$ であれば失敗とすれば良い。(実際の実験結果は教科書 158 ページ参照)
このような模擬実験のことを，コンピューターシミュレーションと呼ぶ。

### 8.1.4 統計学上の意義

大数の法則から分かることは，十分な大きさの標本を調べれば，母集団の様々な特性をかなり正確に知ることができる。**統計的推測**の理論を生み出す。

大数の法則は，一般的に，**大標本では，観察された標本平均を母集団の真の平均とみなしてよい**と言う意識を，数学的に厳密に証明したものである。

## 8.2 中心極限定理 (central limit theorem)

中心極限定理は，大数の法則より詳しく，母集団分布が何であっても，**和 $X_1 + \cdots + X_n$ の確率分布の形は，$n$ が大なるときには，大略正規分布と考えてよい**ということである。

### 8.2.1 和の正規性

大数の法則で述べた確率分布 ( 2 ページ参照) と同じように，$n$ 個の確率変数 $X_1, X_2, \cdots, X_n$ は，お互いに独立で平均 $\mu$，分散 $\sigma^2$ の同じ確率分布に従う確率分布とする。このとき，任意の $x$ に対して，

$$\lim_{n\to\infty} P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \leqq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \tag{14}$$

が成り立つ。すなわち，$\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}}$ が従う確率分布は，$n \to \infty$ のとき標準正規分布 $N(0,1)$ に収束する。$n$ が十分大きいとき $\bar{X}_n$ は正規分布 $N(\mu, \sigma^2/n)$ で近似できる。教科書では，$n \to \infty$ のとき，

$$\lim_{n\to\infty} P\left(a \leqq \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \leqq b\right) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-\frac{x^2}{2}} dx$$

が成り立つという。$n$ が大きければ，

$$P\left(a \leqq \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \leqq b\right) \fallingdotseq \Phi(b) - \Phi(a) \tag{15}$$

である。(注：$\Phi$ は標準正規分布の累積分布関数である)

### 8.2.2　中心極限定理の証明

**証明**　モーメント母関数を用いて証明する。

$$Y_1 = \frac{X_1 - \mu}{\sigma}, Y_2 = \frac{X_2 - \mu}{\sigma}, \cdots, Y_n = \frac{X_n - \mu}{\sigma}$$

とおき，各々$X_1, X_2, \cdots, X_n$ の標準化変数であり，

$$E(Y_1) = E(Y_2) = \cdots = E(Y_n) = 0, V(Y_1) = V(Y_2) = \cdots = V(Y_n) = 1$$

となる。$Y_1$ のモーメント母関数を $M_{Y_1}(t)$ とする，

$$\begin{aligned} M'_{Y_1}(t) \mid_{t=0} &= E(Y_1) = 0, \\ M''_{Y_1}(t) \mid_{t=0} &= E({Y_1}^2) = V(Y_1) + (E(Y_1))^2 = 1 \end{aligned} \tag{16}$$

従って，

$$\begin{aligned} M_{Y_1}(t) &= 1 + tE(Y_1) + \frac{t^2 E({Y_1}^2)}{2!} + \frac{t^3 E({Y_1}^3)}{3!} + \cdots \\ &= 1 + \frac{t^2}{2} + \cdots \end{aligned} \tag{17}$$

である。

$T = Y_1 + Y_2 + \cdots + Y_n$ とおき,「確率変数 $X$ と $Y$ が独立であるとき，$M_{X+Y}(t) = M_X(t) \bullet M_Y(t)$」の性質から，

$$M_T(t) = M_{Y_1}(t) \cdot M_{Y_2}(t) \cdot \cdots \cdot M_{Y_n}(t) = \{1 + \frac{t^2}{2} + \cdots\}^n \tag{18}$$

$T/\sqrt{n}$ のモーメント母関数は，(17) 式の $t$ を $t/\sqrt{n}$ に置き換え，

$$M_T\left(\frac{t}{\sqrt{n}}\right) = \{1 + \frac{\left(\frac{t}{\sqrt{n}}\right)^2}{2} + \cdots\}^n = \{1 + \frac{\left(\frac{t^2}{2}\right)}{2} + \cdots\}^n \tag{19}$$

であることがわかる。$n \to \infty$ のとき，

$$M_T\left(\frac{t}{\sqrt{n}}\right) \longrightarrow exp\left(\frac{t^2}{2}\right) \tag{20}$$

$exp(t^2/2)$ は標準正規分布 $N(0,1)$ のモーメント母関数である。

$$\begin{aligned} T/\sqrt{n} &= \frac{Y_1 + Y_2 + \cdots + Y_n}{\sqrt{n}} \\ &= \frac{\frac{X_1 - \mu}{\sigma} + \frac{X_2 - \mu}{\sigma} + \cdots + \frac{X_n - \mu}{\sigma}}{\sqrt{n}} \\ &= \frac{(X_1 + X_2 + \cdots + X_n) - n\mu}{\sqrt{n\sigma^2}} = \frac{n\bar{X}_n - n\mu}{\sqrt{n\sigma^2}} = \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \end{aligned} \tag{21}$$

以上により，中心極限定理が証明された。

理論的には $n \to \infty$ のときに定理が成り立つが，実用的には $n \geqq 25$ 程度あれば標準正規分布と考えても問題がない。

### 8.2.3 コンピューターシミュレーション

二項分布を例として「R」でシミュレーションしてみる (プログラムは「R/S-PLUS による統計解析入門」より)。

**二項分布のシミュレーション** $x_1, x_2 \cdots, x_n$ はそれぞれ独立にベルヌーイ分布 $Bi(1,p)$ に従う確率変数である。

$$\mu = E(x_i) = p, \quad \sigma^2 = V(x_i) = p(1-p)$$

である。たたみこみの公式から得られた性質

$$Bi(n+m,p) = Bi(n,p) * Bi(m,p)$$

を利用し,

$$Bi(n,p) = Bi(1,p) * Bi(1,p) * \cdots\cdots (n\text{ 個})$$

であり，$Bi(n,p)$ に従う確率変数 $r$ は，$r = x_1 + x_2 + \cdots + x_n$ である。中心極限定理により，標準化変数

$$z = \frac{r - n\mu}{\sqrt{n\sigma^2}} = \frac{r - np}{\sqrt{np(1-p)}} \tag{22}$$

は $n$ が大きいとき，標準正規分布 $N(0,1)$ に近づく $\Longrightarrow$ 二項分布 $Bi(n,p)$ は $n$ が大きいとき，正規分布に近づくことをコンピュータ・シミュレーション (「R」) により実験を行ってみる。

シミュレーションでは, 区間 (0,1) の一様乱数 $U$ を用い, 成功の確率を p=0.1(すなわち $U < 0.1$ のとき成功 $U \geqq 0.1$ のとき失敗とする) として，$n = 1, 2, \cdots, 15 (= nmax)$ に対してそれぞれ，二項乱数 $r$ を $500 (= nsim)$ 個ずつ発生させた。それらをそれぞれ $z = (r - np)/\sqrt{np(1-p)}$ によって標準化し，分布をグラフにしたものは 8 ページのように示してある。.

◎使用した関数は下記である：

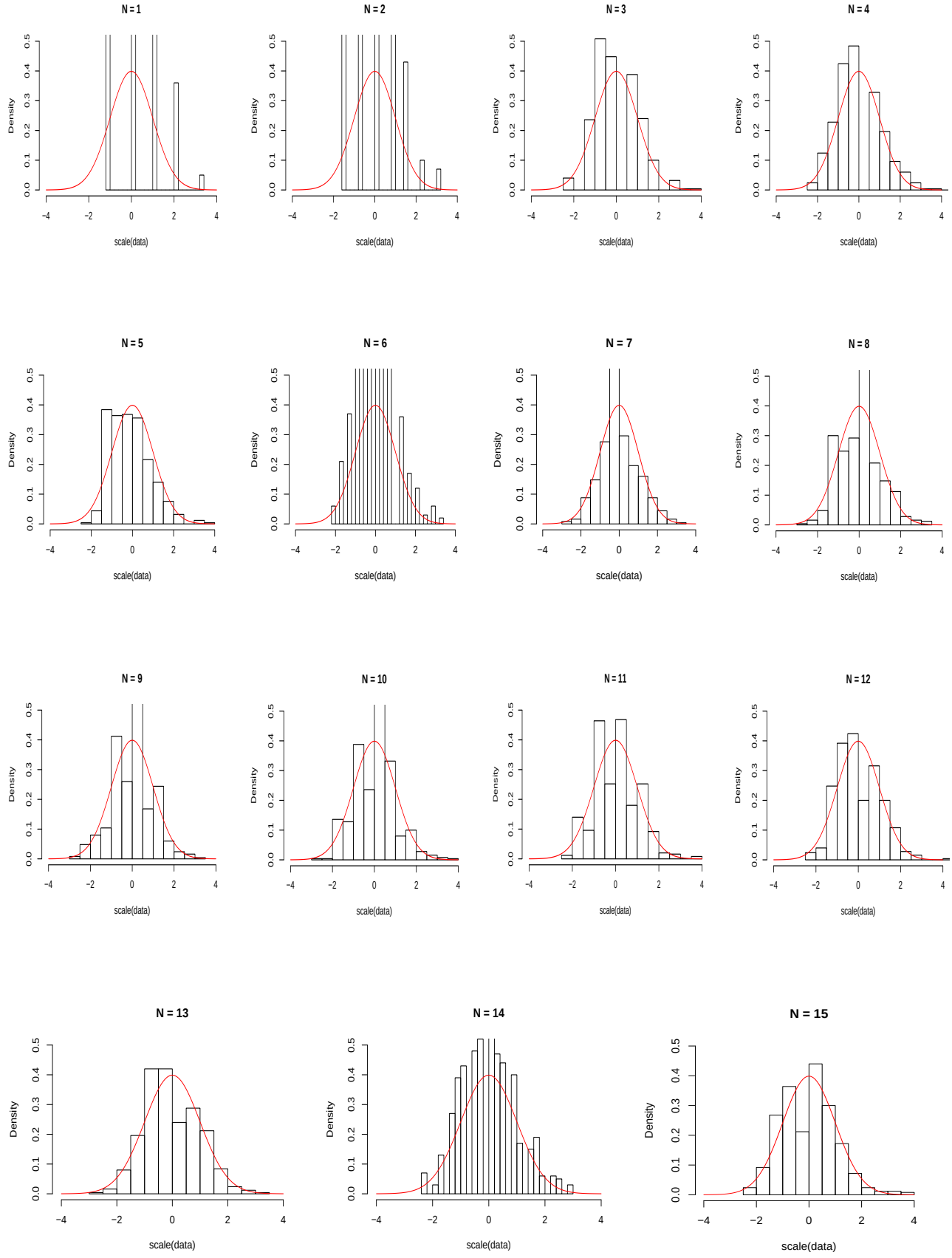
```
cltplot <- function(result, n){

   hist(scale(result), nclass = 20, xlim = c(-4, 4),
  ylim = c(0, 0.5), freq = F, main = paste("N =" , n))
  curve(dnorm, -4, 4, col = 2, add = T)

}


cltrand <- function(n, nsim, rdist){
```

```
  if(n == 1)
    result <- apply(matrix(sapply(rep(n, nsim), rdist), n, nsim), 2, mean)
  else
    result <- apply(sapply(rep(n, nsim), rdist), 2, mean)
  cltplot(result, n)

}



clt <- function(nmax, rdist){

   for(i in 1:nmax){
    if (options()$device == "X11")
      X11()
    if (options()$device == "windows")
      win.graph()
    cltrand(i, nsim, rdist)
   }

}
```

# scale(result):データの標準化 (平均 0、分散 1 に)
# dnorm:標準正規分布の密度関数 (赤)
# result:n 個一様乱数を発生させ，平均値を求めるシミュレーションを nsim 回繰り返した結果
# nmax:データ数ｎの最大値
# rdist:乱数を発生させる関数を指定する
# X11():UNIX,Linux のデフォルトドライバ
# win.graph():Windows 版のドライバ

.

**その他の分布，一様分布，$\chi^2$ 分布，$t$ 分布や $F$ 分布などは ほかの関数で乱数を発生させ，シミュレーションをする。**


N = 1
N = 2
N = 3
N = 4
N = 5
N = 6
N = 7
N = 8
N = 9
N = 10
N = 11
N = 12
N = 13
N = 14
N = 15
Density
scale(data)

**逆関数法による乱数の作り方**　まず，区間 (0,1) の一様乱数 $U$ から，$X = F^{-1}(U)$ は

$$P(X \leqq x) = P(F^{-1}(U) \leqq x) = P(U \leqq F(x)) = F(x) \tag{23}$$

となることより $X$ は分布関数 $F$ に従う。したがって，分布関数の逆関数の引数に一様乱数を入れば必要な分布の乱数が得られる。

**指数分布のシミュレーション**　指数分布に従う乱数は，逆変換法を用いて区間 (0,1) の一様乱数 $U$ から $-logU/\lambda$ で発生させる。母数 $\lambda$ の指数分布 $E_x(\lambda)$ に対しては，$F(x) = 1 - e^{-\lambda x}$ であり，逆関数 $F^{-1}(y)$ を求めると，

$$F^{-1}(y) = -log(1-y)/\lambda$$

となる。したがって，

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

が $E_x(\lambda)$ に従う。$1-U$ も [0,1] 上の一様分布に従うから，$-logU/\lambda$ でよい。

教科書の実験では，母数を $\lambda = 1/2$ とした。期待値 $E(x) = 1/\lambda = 2$, 分散は $V(X) = 1/\lambda^2 = 4$ である。$n = 1, 2, 10, 30$ に設定し，それぞれの $n$ について標本平均 $\bar{x}_n$ を 250 個ずつ求めた。それぞれ得られた度数分布を観察し，分布の形が正規分布に近づいていくことが結果である。

## 8.3　中心極限定理の応用

### 8.3.1　二項分布の正規分布による近似

二項分布 $Bi(n,p)$ における成功の回数 $S$ は，それぞれが二項分布 $Bi(1,p)$ に従う確率変数 $X_1, X_2, \cdots, X_n$ の和である。$n$ が大きいときに $Bi(n,p)$ は正規分布に近づくので，それで近似することができる。

$n$ 回試行のうち，$n$ は大きいときの成功回数が $k$ 回以上 $k'$ 回以下である確率は，標準正規分布の累積分布関数 $\Phi$ により，

$$\begin{aligned} P\left(k \leqq S \leqq k'\right) &= P\left(\frac{k-np}{\sqrt{np(1-p)}} \leqq z \leqq \frac{k'-np}{\sqrt{np(1-p)}}\right) \\ &\fallingdotseq \Phi\left(\frac{k'-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-np}{\sqrt{np(1-p)}}\right) \end{aligned} \tag{24}$$

で求めることができる。

二項分布で $n$ の値がどのくらい大きければ，正規分布による近似を用いてよいか，通常言われている必要条件は，$np > 5$ かつ $n(1-p) > 5$ である。したがって，$p$ が 1/2 のときには $n$ は 10 以上であればよいが，$p$ が 0 や 1 に近いときは，$n$ が相当大きくなければならない。

### 8.3.2 正規乱数の発生

まず，(0,1) 上の一様乱数を $n$ 個発生させ，それらを $r_1, r_2, \cdots, r_n$ とする。期待値，分散はそれぞれ 1/2,1/12 である。中心極限定理から，

$$z = \frac{\sum_{i=1}^{n} r_i - \frac{n}{2}}{\sqrt{n/12}} \tag{25}$$

は $n$ が大きいときほぼ標準正規分布に従う。予定の正規乱数 $x$ の期待値を $\mu$，分散を $\sigma^2$ とすると，$x$ の標準化の逆変換 $x = \sigma z + \mu$ を代入して

$$x = \sigma\sqrt{\frac{12}{n}}\left(\sum_{i=1}^{n} r_i - \frac{n}{2}\right) + \mu \tag{26}$$

となる。ほぼ正規分布 $N(\mu, \sigma^2)$ に従う正規乱数である。

# 参考文献


[1] 垂水　共之・飯塚　誠也「R/S-PLUS による統計解析入門」共立出版 (2006)

[2] 石井　博昭・塩出　省吾・新森　修一「確率統計の数理」(1995) 裳華房

[3] 浅野　晃「情報統計学」　http://kuva.mis.hiroshima-u.ac.jp/ asano/Kougi/06a/IS/IS13pr.pdf

[4] 水本　久夫「微分積分学の基礎」(2003) 培風館