J. Appl. Math. & Informatics Vol. **28**(2010), No. 5 - 6, pp. 1573-1581
Website: http://www.kcam.biz

GENERALIZED CONVOLUTION OF UNIFORM DISTRIBUTIONS†

JONG SEONG KANG, SUNG LAI KIM, YANG HEE KIM AND YU SEON JANG*

Abstract. we investigate the n-fold convolution of the uniform distributions. First, we are concerned with the explicit distribution function of the partial sum ζ_n when the random variables are independent and has identically uniform distribution, next, we determine the n-fold convolution distribution of ζ_n when the identically distributed condition is not satisfied.

AMS Mathematics Subject Classification : 60E05, 60G50, 62E15, 94A17
key words and phrases : uniform distribution, support, convolution distribution, n-fold convolution

1. Introduction

The sum ζ_n of n mutually independent random variables $\xi_1, \xi_2, \cdots, \xi_n$ plays an important role in many fields ; information theory, insurance mathematics, economics, engineering and medical applications, etc. [2,4,5,7,9]. In information theory, entropy is the measure of the amount of information that is missing before reception and is sometimes referred to as Shannon entropy [1,4]. Shannon entropy is a broad and general concept which finds applications in information theory. It was originally devised by Claude Shannon in 1948 to study the amount of information in a transmitted message [1]. The entropy of the probability density $p(x)$ is defined by

$$\int_{-\infty}^{\infty} p(x) \log[p(x)]^{-1} dx. \tag{1}$$

Uniform distribution has maximum entropy among all distributions of continuous type with finite support [4].

Received May 10, 2010. Accepted July 5, 2010. *Corresponding author.
†This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0093815).

Mathematically, a convolution is defined as the integral over all space of one function f_1 at u times another function f_2 at $x-u$. The integration is taken over the variable u, typically from minus infinity to infinity over all the dimensions. So the convolution is a function $f_1 * f_2$ of a new variable x, as shown in the following equation

$$f_1 * f_2(x) = \int_{-\infty}^{\infty} f_1(u) f_2(x-u) du. \tag{2}$$

Let F_1 and F_2 be two distribution functions, then

$$G_2^*(x) = \int_{-\infty}^{\infty} F_1(x-u) dF_2(u). \tag{3}$$

is also a distribution function. Equation (3) defines an associative and commutative operation, and the result distribution is called the convolution of F_1 and F_2, and we write the equation (3) in the symbolic form $G_2^* = F_1 * F_2$. If one of the F_1 and F_2 is absolutely continuous, then G_2^* is also absolutely continuous and if we let ${F_1}'(x) = f_1(x)$ and ${F_2}'(x) = f_2(x)$, then $G_2^*{}'(x) = f_1 * f_2(x)$, and write also $g_2^* = f_1 * f_2$.

Let F_i and f_i denote the distribution function and the density function of the summand ξ_i, $i = 1, 2, \cdots, n$ and let G_n^* be the distribution function of the partial sum $\zeta_n = \xi_1 + \cdots + \xi_n$. Then the $G_n^*(x)$ equals the n-fold convolution $F_1 * F_2 * \cdots * F_n(x)$ and $g_n^*(x) = f_1 * f_2 \cdots * f_n(x)$.

In this paper, we investigate the n-fold convolution of the uniform distributions. First, we are concerned with the explicit distribution function of ζ_n when the random variables are independent and has identically uniform distribution, next, we determine the n-fold convolution distribution of ζ_n when the identically distributed condition is not satisfied.

2. When $\xi_1, \xi_2, \cdots, \xi_n$ are independent and identically distributed

Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent and identically distributed random variables with the common distribution function F and probability density function f. Then the distribution function of the sum ζ_n is the n-fold convolution of itself F such as

$$F^{n*}(x) = F^{(n-1)*} * F(x) \quad (n \geq 2) \tag{4}$$

where $F^{1*}(x) = F(x)$ and its probability density function is

$$f^{n*}(x) = f^{(n-1)*} * f(x) \quad (n \geq 2) \tag{5}$$

where $f^{1*}(x) = f(x)$.

Suppose that the random variables $\xi_1, \xi_2, \cdots, \xi_n$ have the uniform distribution on $(\alpha, \beta), \beta > \alpha$, that is, the common probability density function is given by

$$f(x) = \frac{1}{\beta - \alpha} I_{(\alpha,\beta)}(x), \tag{6}$$

where $I_U(x)$ is 1 for $x \in U$ and 0 otherwise. Then the sum $\zeta_2 = \xi_1 + \xi_2$ have a symmetric triangle distribution and its density function is

$$f^{2*}(x) = \frac{x - 2\alpha}{(\beta - \alpha)^2}\, I_{(2\alpha,\alpha+\beta)}(x) + \frac{2\beta - x}{(\beta - \alpha)^2}\, I_{(\alpha+\beta,2\beta)}(x). \tag{7}$$

By Renyi [8] the density function of the sum of n mutually independent random variables with uniform distribution on $(-1, 1)$ is

$$\frac{1}{2}\sum_{j=0}^{\left[\frac{n+u}{2}\right]}(-1)^j \frac{n}{j!(n-j)!}\left(\frac{n+u}{2} - j\right)^{n-1} I_{(-n,n)}(u), \tag{8}$$

where $[u]$ is the largest integer less than or equal to u. Taking $x = (-\alpha+\beta)u/2+(\alpha+\beta)n/2$, it is possible to gain the probability density function f^{n*} of the sum ζ_n as following ;

$$f^{n*}(x) = \frac{1}{\beta - \alpha}\sum_{j=0}^{\tau}(-1)^j \frac{n}{j!(n-j)!}\left(\frac{x - n\alpha}{\beta - \alpha} - j\right)^{n-1} I_{(n\alpha,n\beta)}(x), \tag{9}$$

where τ is a function of n, x, α and β which is the largest integer less than or equal to $(x - n\alpha)/(\beta - \alpha)$. The distribution function of ζ_n is obtained by

$$F^{n*}(x) = \int_0^x f^{n*}(t)dt I_{(n\alpha,n\beta)}(x) + I_{[n\beta,\infty)}(x). \tag{10}$$

So we have the following theorem.

Theorem 1. *The distribution function of the sum of n mutually independent random variables with uniform distribution on (α, β) is*

$$F^{n*}(x) = \sum_{j=0}^{\tau}(-1)^j \frac{1}{j!(n-j)!}\left(\frac{x - n\alpha}{\beta - \alpha} - j\right)^{n} I_{(n\alpha,n\beta)}(x) + I_{[n\beta,\infty)}(x), \tag{11}$$

where τ is a function of n, x, α and β which is the largest integer less than or equal to $(x - n\alpha)/(\beta - \alpha)$.

Proof. The proof follows immediately from the equations (9) and (10). □

3. When $\xi_1, \xi_2, \cdots, \xi_n$ are independent but not identically distributed

Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uniformly distributed random variables with $\xi_i \sim U(\alpha_i, \beta_i), \beta_i > \alpha_i, i = 1, 2, \cdots, n$ and let F_i and f_i the distribution function and the density function of $\xi_i, i = 1, 2, \cdots, n$, respectively. Then the random variable $\xi_i - \alpha_i$ has the uniform distribution on $(0, \ell_i), \ell_i = \beta_i - \alpha_i, i = 1, 2, \cdots, n$ and independent property. We determine the explicit form of the distribution of the partial sum $\zeta_n = \xi_1 + \cdots + \xi_n$ throughout the distribution of $\zeta_n - (\alpha_1 + \alpha_2 + \cdots + \alpha_n)$.

Let $\mathbf{W}_n$ be the set of all n-tuples of elements of $\{0,1\}$ with the standard definitions of addition and scalar multiplication. Then $\mathbf{W}_n$ is a subspace of the n-dimensional Euclidean vector space $\mathbb{R}^n$. Let S_p be the set of all scalar products between $\mathbf{w}_n \in \mathbf{W}_n$ and $\mathbf{t}_n \in \mathbb{R}^n$, that is,

$$S_p = \{< \mathbf{w}_n, \mathbf{t}_n > \ | \ \mathbf{w}_n \in \mathbf{W}_n,\ \mathbf{t}_n \in \mathbb{R}^n\}, \tag{12}$$

where$< \cdot, \cdot >$ is the scalar product on $\mathbb{R}^n$ and let for any given $\alpha_i, \beta_i, i = 1, 2, \cdots, n$

$$S = \{w_1\ell_1 + w_2\ell_2 + \cdots + w_n\ell_n \mid w_i \in \{0,1\}, i = 1, 2, \cdots, n\} \subset S_p, \tag{13}$$

where $\ell_i = \beta_i - \alpha_i,\ i = 1, 2, \cdots, n$. Then S has the 2^n values $0, \ell_1, \ell_2, \cdots, \ell_1 + \cdots + \ell_n$. We obtain the rearrangement set $S = \{s_{n1}, s_{n2}, \cdots, s_{n2^n}\}$ with $s_{n1}(= 0) \leq s_{n2} \leq \cdots \leq s_{n2^n}(= \ell_1 + \ell_2 + \cdots + \ell_n$, say $L_n)$

For any $x \in \mathbb{R}^1$ we define

$$\tau_n^+ = \max\{j \ \mid \ x - s_{nj} > 0 \text{ and } s_{nj} \in S \text{ for } j = 1, 2, \cdots, 2^n\}. \tag{14}$$

Then the τ_n^+ is the number of index j with $x - s_{nj} > 0, j = 1, 2, \cdots, 2^n$. Now we know that for any $\alpha_i, \beta_i,\ (\beta_i - \alpha_i > 0,\ i = 1, 2, \cdots, n)$ and index $j,\ j = 1, 2, \cdots, 2^n$ there exist a vector $\mathbf{w}_n = (w_1, w_2, \cdots, w_n) \in \mathbf{W}_n$ such that

$$s_{nj} = w_1\ell_1 + w_2\ell_2 + \cdots + w_n\ell_n \quad (j = 1, 2, \cdots, 2^n). \tag{15}$$

In this time, for any given index j we let $\mathbf{w}_{nj}$ be the vector $\mathbf{w}_n$ which is satisfied the following condition

$$s_{nj} =< \mathbf{w}_n, (\ell_1, \cdots, \ell_n) > \tag{16}$$

where $< \cdot, \cdot >$ is the scalar product on $\mathbb{R}^n$.

For example, If $n = 2$ and $\max\{\ell_1, \ell_2\} < x < \ell_1 + \ell_2 (= L_2)$, then

$$\begin{aligned} &\tau_2^+ = 3 \\ &\mathbf{w}_{21} = (0,0),\ \mathbf{w}_{22} = (0, \ell_2),\ \mathbf{w}_{23} = (\ell_1, 0) \ \ (\text{if } \ell_2 < \ell_1) \\ &s_{21} = 0,\ s_{22} = \min\{\ell_1, \ell_2\},\ s_{23} = \max\{\ell_1, \ell_2\} \end{aligned} \tag{17}$$

and the distribution function $G_2(x)$ of $\zeta_2 - A_2$ is

$$P(\zeta_2 - A_2 \leq x) = \int_{-\infty}^{\infty} F_1(x_1 - x_2) dF_2(x_2), \tag{18}$$

where F_i is the distribution function of $\xi_i - \alpha_i, i = 1, 2$ and $A_2 = \alpha_1 + \alpha_2$. Thus we have the following lemma.

Lemma 1. *Let the random variables ξ_1, ξ_2 be independent and uniformly distributed on $(0, \ell_i), \ell_i > 0, i = 1, 2$. Then the distribution function G_2^* of the sum $\zeta_2 = \xi_1 + \xi_2$ is*

$$G_2^*(x) = \frac{1}{2\ell_1\ell_2} \sum_{j=1}^{\tau_2^+} (-1)^{\|\ \mathbf{w}_{2j}\ \|} (x - s_j)^2 I_{(0, L_2)}(x) + I_{[L_2, \infty)}(x), \tag{19}$$

where τ_2^+ is defined by equation (14) and $||\cdot||$ is the norm of the vector.

Proof. The convolution G_2^* of F_1 and F_2 is represented by

$$G_2^*(x) = \int\int_{x_1+x_2\le x} dF_1(x_1)dF_2(x_2). \tag{20}$$

By integrating (20) we obtain

$$G_2^*(x) = \begin{cases} 0, \quad x<0 \\ \dfrac{1}{2\ell_1\ell_2}x^2, \quad 0\le x<m_2 \\ \dfrac{1}{2\ell_1\ell_2}\left\{x^2-(x-m_2)^2\right\}, \quad m_2\le x<M_2 \\ 1-\dfrac{1}{2\ell_1\ell_2}(L_2-x)^2, \quad M_2\le x<L_2 \\ 1, \quad x\ge L_2 \end{cases} \tag{21}$$

where $m_2=\min\{\ell_1,\ell_2\}$, $M_2=\max\{\ell_1,\ell_2\}$. Since

$$1-\frac{1}{2\ell_1\ell_2}(L_2-x)^2 = \frac{1}{2\ell_1\ell_2}\left\{x^2-(x-m_2)^2-(x-M_2)^2\right\}, \tag{22}$$

we know that for $0<x<L_2$

$$G_2^*(x) = \frac{1}{2\ell_1\ell_2}\sum_{j=1}^{\tau_2^+}(-1)^{||\mathbf{w}_{2j}\ ||}(x-s_j)^2\, I_{(0,L_2)}(x) \tag{23}$$

which completes the proof. □

Lemma 2. *Under the conditions of Lemma 1, the distribution function G_2^* of the sum $\zeta_2=\xi_1+\xi_2$ is*

$$G_2^*(x) = \frac{1}{2\ell_1\ell_2}\sum_{j=1}^{\tau_2^+}(-1)^{||\ \mathbf{w}_{2j}\ ||}(x-s_j)^2\, I_{(0,\infty)}(x). \tag{24}$$

Proof. If $x\ge L_2$, then the number $\tau_2^+=4$. By equation (22) we have

$$\begin{aligned} G_2^*(x) &= \frac{1}{2\ell_1\ell_2}\sum_{j=1}^{4}(-1)^{||\mathbf{w}_{2j}\ ||}(x-s_j)^2 \\ &= \int_0^{\ell_2}\int_0^{\ell_1} f_1(x_1)f_2(x_2)dx_1dx_2 = 1. \end{aligned} \tag{25}$$

Hence from Lemma 1 we obtain the result in lemma. □

Theorem 2. *Let the random variables $\xi_1, \xi_2, \cdots, \xi_n$ be independent and uniformly distributed on $(0, \ell_i), \ell_i > 0, i = 1, 2, \cdots, n$. Then the distribution function G_n^* of the sum $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$ is*

$$G_n^*(x) = \frac{1}{n! \prod_{i=1}^n \ell_i} \sum_{j=1}^{\tau_n^+} (-1)^{\| \mathbf{w}_{nj} \|} (x - s_j)^n I_{(0,\infty)}(x), \tag{26}$$

where τ_n^+ is defined by equation (14) and s_j ($j = 1, 2, \cdots, \tau_n^+$) is the combination form with respect to ℓ_i $(i = 1, 2, \cdots, n)$.

Proof. It is trivial for $x \le 0$. If $x \ge L_n (= \ell_1 + \cdots + \ell_n,\ n \ge 2)$, then the distribution function $G_n^*(x)$ is the integral of n-fold convolution of the density functions $f_1, f_2, \cdots$ and f_n over $\mathbb{R}^1$;

$$G_{k+1}^*(x) = \int_{\mathbb{R}^1} f_1 * f_2 * \cdots * f_n(u) du. \tag{27}$$

Since $f_1 * f_2 * \cdots * f_n$ is a density function, the distribution function $G_n^*(x)$ is equal to 1.

For $0 < x < L_n$ the proof will be performed by mathematical induction with respect to n. For $n = 2$, we obtain the result in lemma 2. Suppose that the equation (26) is true for any positive integer k $(k \le n)$ and g_k^* denote the density function of ζ_k. Then the distribution function G_{k+1}^* of ζ_{k+1} is obtained by convolution of G_k^* and F_k such as

$$G_{k+1}^*(x) = \int_{\mathbb{R}^1} G_k^*(x - x_{k+1}) dF_{k+1}(x_{k+1}) \quad (0 < x < L_{k+1}) \tag{28}$$

and

$$g_{k+1}^*(x) = \frac{d}{dx} G_{k+1}^*(x) \quad (0 < x < L_{k+1}). \tag{29}$$

Since

$$G_{k+1}^*(x) = \frac{1}{\ell_{k+1}} \int_0^{\ell_{k+1}} G_k^*(x - x_{k+1}) dx_{k+1}, \tag{30}$$

the density function of ζ_{k+1} is

$$\begin{aligned} g_{k+1}^*(x) &= \frac{1}{\ell_{k+1}} \left[G_k^*(x) - G_k^*(x - \ell_{k+1}) \right] \\ &= \frac{1}{k! \prod_{i=1}^{k+1} \ell_i} \left[\sum_{j=1}^{\tau_k^+} (-1)^{\| \mathbf{w}_{kj} \|} (x - s_{kj})^k - (x - \ell_{k+1} - s_{kj})^k \right] \\ &= \frac{1}{k! \prod_{i=1}^{k+1} \ell_i} \sum_{j=1}^{\tau_{k+1}^+} (-1)^{\| \mathbf{w}_{(k+1)j} \|} \left(x - s_{(k+1)j} \right)^k \quad (0 < x < L_{k+1}). \end{aligned} \tag{31}$$

and 0 otherwise. Thus we have the distribution function

$$G^*_{k+1}(x) = \frac{1}{(k+1)!\prod_{i=1}^{k+1}\ell_i}\sum_{j=1}^{\tau^+_{k+1}}(-1)^{\|\ \mathbf{w}_{(k+1)j}\ \|}\left(x - s_{(k+1)j}\right)^{k+1} I_{(0,L_{k+1})} \qquad (32)$$

which completes the proof. □

Corollary 1. *Under the conditions of Theorem 2, the density function* g^*_n *of the sum* ζ_n *is*

$$g^*_n(x) = \frac{1}{(n-1)!\prod_{i=1}^{n}\ell_i}\sum_{j=1}^{\tau^+_n}(-1)^{\|\ \mathbf{w}_{nj}\ \|}\left(x - s_j\right)^{n-1} I_{(0,L_n)}(x), \qquad (33)$$

where $L_n = \sum_{i=1}^n \ell_i$.

Proof. Since for $0 < x < L_n$

$$G^*_n(x) = \frac{1}{n!\prod_{i=1}^{n}\ell_i}\sum_{j=1}^{\tau^+_n}(-1)^{\|\ \mathbf{w}_{nj}\ \|}\left(x - s_j\right)^{n} \qquad (34)$$

and $G^*_n{}'(x) = g_n(x)$, the density function is

$$g^*_n(x) = \frac{1}{(n-1)!\prod_{i=1}^{n}\ell_i}\sum_{j=1}^{\tau^+_n}(-1)^{\|\ \mathbf{w}_{nj}\ \|}\left(x - s_j\right)^{n-1}. \qquad (35)$$

And also since $G^*_n(x) = 1$ for $x \geq L_n$, we obtain the result in corollary. □

Theorem 3. *Let the random variables* $\xi_1, \xi_2, \cdots, \xi_n$ *be independent and uniformly distributed on* (α_i, β_i) *with* $\beta_i - \alpha_i > 0, i = 1, 2, \cdots, n$. *Then the density function* h^*_n *of the sum* $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$ *is*

$$h^*_n(x) = \frac{1}{(n-1)!\prod_{i=1}^{n}(\beta_i - \alpha_i)}\sum_{j=1}^{\tau^+_n}(-1)^{\|\ \mathbf{w}_{nj}\ \|}\left(x + A_n - s_j\right)^{n-1} I_{(A_n,B_n)}(x), \qquad (36)$$

where $A_n = \sum_{i=1}^n \alpha_i$, $B_n = \sum_{i=1}^n \beta_i$ *and* s_j ($j = 1, 2, \cdots, \tau^+_n$) *is the combination form with respect to* $\beta_i - \alpha_i$ $(i = 1, 2, \cdots, n)$.

Proof. Let $\nu_i = \xi_i - \alpha_i,\ i = 1, 2, \cdots, n$. Then ν_i has the uniform distribution with the support $(0, \ell_i),\ \ell_i = \beta_i - \alpha_i,\ i = 1, 2, \cdots, n$ and mutually independent property between their. By corollary 1 the sum $\sum_{i=1}^n \nu_n = \zeta_n - A_n$ has the density function $g^*_n(u)$ with support $(0, L_n)$, where $L_n = \sum_{i=1}^n \ell_i$. Letting $x = u - A_n$, the density function h^*_n has the support (A_n, B_n), where $A_n = \sum_{i=1}^n \alpha_i$ and $B_n = \sum_{i=1}^n \beta_i$. By transformation method we obtain easily the result in theorem. □

Corollary 2. *Under the conditions of Theorem 3, the distribution function H_n^* of the sum ζ_n is*

$$H_n^*(x) = \frac{1}{n!\prod_{i=1}^{n}(\beta_j - \alpha_j)} \sum_{j=1}^{\tau_n^+} (-1)^{\|\ \mathbf{w}_{nj}\ \|} (x + A_n - s_j)^n \, I_{(A_n,\infty)}(x). \quad (37)$$

Proof. For any given x

$$H_n^*(x) = \int_{-\infty}^{x} h_n(u)du. \quad (38)$$

Since $h_n(x) = 0$ for $x \le A_n$, the distribution function is

$$H_n^*(x) = \int_{A_n}^{x} h_n(u)du I_{(A_n,B_n)}(x) + I_{[B_n,\infty)}(x). \quad (39)$$

But if $x \ge B_n$, then $\tau_n^+ = 2^n$. Since

$$\frac{1}{n!\prod_{i=1}^{n}(\beta_i - \alpha_i)} \sum_{j=1}^{2^n} (-1)^{\|\mathbf{w}_{nj}\ \|} (x + A_n - s_j)^n = 1, \quad (40)$$

so we have

$$H_n^*(x) = \frac{1}{n!\prod_{i=1}^{n}(\beta_i - \alpha_i)} \sum_{j=1}^{\tau_n^+} (-1)^{\|\mathbf{w}_{nj}\ \|} (x + A_n - s_j)^n \, I_{(A_n,B_n)}(x) + I_{[B_n,\infty)}. \quad (41)$$

Thus the proof is complete. □

References

1. R. Balian, *Entropy, a Protean Concept*, Poincaré Seminar 2 (2003), 119-145.
2. Y. Chow and S. Y. Huang, *A Characterization of the Uniform Distribution via Moments of n-fold Convolution Modulo One*, Sankhya : Indian Journal of Statistics, **61** (A) (1999), 148-151.
3. W. Feller, *An introduction to probability theory and its applications,* Vol. I and II, John Wiley & Sons, 1957 and 1966.
4. F. Killmann and E. v. Collani, *A Note on the Convolution of the Uniform and Related Distributions and Their Use in Quality Control*, Economic Quality Control, **16** (1) (2001), 17-41.
5. C. H. v. Lanzenauer and W. N. Lundberg, *The n-fold convolution of a mixed density and mass function*, International Actuarial Association, **8** (1) (1974), 91-103.
6. E. Lukacs, *Stochastic Convergence*, Academic Press, 1975.
7. N. D. Pril, *Recursions for Convolutions of Arithmetic Distributions*, Astin Bulletin, **15** (2) (1985), 135-139.
8. A. Rényi, *Probability Theory*, North-Holland, Amsterdam, 1970.
9. B. Sundt, *Some recursions for moments of n-fold convolutions*, Insurance: Mathematics and Economics, **33** (3) (2003), 479-486.

Jong Seong Kang is a professor at Chungnam National University.

College of Pharmacy, Chungnam National University, Daejeon, 305-764, Korea.

e-mail: kangjss@cnu.ac.kr

Sung Lai Kim is a professor at Chungnam National University.

Department of Mathematics, Chungnam National University, Daejeon, 305-764, Korea.
e-mail: slkim@cnu.ac.kr

Yang Hee Kim is a Ph.D in Mathematics.

Department of Mathematics Education, Chungnam National University, Daejeon, 305-764, Korea.
e-mail: mathkyh@cnu.ac.kr

Yu Seon Jang is a research professor at Chungnam National University.

Drug Development Research Institute, Chungnam National University, Daejeon, 305-764, Korea.
e-mail: ysjang@cnu.ac.kr