A Geometric Derivation of the Irwin-Hall Distribution

James E. Marengo, David L. Farnsworth, and Lucas Stefanic

School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY 14623, USA

Correspondence should be addressed to David L. Farnsworth; dlfsma@rit.edu

Received 23 May 2017; Revised 4 August 2017; Accepted 17 August 2017; Published 18 September 2017

1. Introduction

The simple continuous uniform or rectangular distribution Uniform(0, 1) with probability density function (PDF) \( f(x) = 1 \) for \( 0 < x < 1 \) and \( f(x) = 0 \) otherwise is very important. Two applications arise in numerical simulation and Bayesian analysis of proportions. If \( F \) is the cumulative distribution function (CDF) of the continuous random variable \( X \), then the random variable \( Y = F(X) \) has a Uniform(0,1) distribution. The random variable \( X \) can be simulated by first simulating \( Y \) and then letting \( X = F^{-1}(Y) \). This is called the inversion method ([1, page 295], [2, pages 194–196]). The transformation is called the probability integral transformation ([3], [4, pages 203-204]). The uniform distribution is a Bayesian noninformative prior distribution for the distribution of a random variable defined on the unit interval, such as a beta distribution for a proportion ([2, page 33], [5, pages 82–90]). For other applications and generalizations of the uniform distribution, see [6–8].

The present goal is to derive the CDF and the PDF of the sum \( T = \sum_{i=1}^{n} X_i \), where \( X_i \) are independent identically distributed Uniform(0, 1) random variables for \( i = 1, 2, \ldots, n \). The CDF and PDF are

\[
F(t) = \sum_{i=0}^{n} \left[ (-1)^i \binom{n}{i} \frac{(t-i)^{n-1}}{(n-1)!} s_a(t) \right],
\]

(1)

\[
f(t) = \sum_{i=0}^{n} \left[ (-1)^i \binom{n}{i} \frac{(t-i)^{n-1}}{(n-1)!} s_a(t) \right],
\]

(2)

respectively, where \( s_a(t) \) is the unit step function

\[
s_a(t) = \begin{cases} 
0 & t < a \\
1 & a \leq t.
\end{cases}
\]

(3)

The derivation in Section 2 is geometric and explicitly uses the inclusion-exclusion principle.

Derivations of the distribution, which more recently acquired its name Irwin-Hall, go back to Lagrange and Laplace in the latter 18th century and the early 19th century. Lagrange used generating functions based on \( a^x \) to obtain the distribution of \( T \) ([9, pages 603–612], [10, page 283]). Those generating functions are a predecessor of characteristic functions ([10, page 286]). Laplace often revisited the problem of finding the distribution of \( T \) and employed many methods ([9, pages 714–715], [10, pages 286–301]). The distribution is described in [1, pages 296–300], where it is called the Irwin-Hall distribution.

Some derivations employ characteristic functions in a variety of ways, since the characteristic function of a sum of independent random variables is the product of each summand’s characteristic function and the inverse transform...
Direct integration techniques can be used to obtain the distribution of a linear combination of Uniform(0, 1) random variables ([15, pages 358–360], [24, 25]). Similar techniques are used in [26] for uniform distributions whose domains are intervals with zero as their left endpoints. The distribution of the mean is obtained when all the constants are 1/n. In this case, the distribution is called the Bates distribution ([1, page 297], [27]), which can also be found by a simple transformation of the Irwin-Hall distribution ([15, page 359], [25, page 241]). Using moment generating functions, instead of characteristic functions, Gray and Odell [28] found the distribution of any linear combination of uniform random variables with different domains allowed. In Section 3, the present method or style of proof is extended to those cases giving the same distributions.

Because T is a sum, the Irwin-Hall distribution approximates a normal distribution with a spline, since the Irwin-Hall distribution in (2) is composed of polynomials. The support of T is the interval [0, n]; the mean, mode, and median of T are n/2; and its variance is n/12. By symmetry, all odd central moments are zero, including skewness. The kurtosis is 3 − 6/5n [1, page 300]. This is the measure of kurtosis that is 3 for a normal distribution, so Irwin-Hall distributions are platykurtic, and the kurtosis is close to 3 for large n. According to the Central Limit Theorem,

\[ Z = \frac{T - n/2}{\sqrt{n/12}} \xrightarrow{D} \text{Normal}(0, 1) \text{ as } n \to \infty \]  

([4, pages 280–283], [1, pages 213–218 and 245], [29, pages 220–222]). Figure 1 contains a normal distribution with mean n/2 and variance n/12 = 3/12 = 1/4 and its approximating Irwin-Hall distribution with n = 3. The approximation is very good even for this small value of n [30]. The uniform error bound for the normal(0, 1) CDF \( \Phi(z) \) is

\[ |F(z) - \Phi(z)| \leq \frac{\sqrt{3}}{20\sqrt{n}} \]  

([31], [32, page 51]). Approximations with spline fitting can be useful with or without complete information about the distributional shape [33, 34].

Since round-off errors for random variables that are rounded to the nearest integer are distributed Uniform(−1/2, 1/2), the sum of round-off errors is a linearly transformed Irwin-Hall distribution [12]. For large n, the sum of round-off errors is easily described with a normal distribution [29, page 222]. For small n, the Irwin-Hall distribution is also appropriate and not too complicated.

Lee et al. [35] use the Irwin-Hall distribution to examine the efficacy of goodness-of-fit tests. Heinrich et al. [36] adapt the Irwin-Hall distribution in consideration of the accumulated accuracy of round-off errors. Inequalities for

linear combinations of independent random variables whose domains have an upper bound are given in [37].

### 2. Derivation of the Irwin-Hall Distribution

**Theorem 1.** Let \( X_i \) for \( i = 1, 2, \ldots, n \) be independent random variables, each having the continuous uniform distribution on the unit interval, and let \( T = \sum_{i=1}^{n} X_i \). Then, the CDF and PDF of \( T \) are given by (1) and (2), respectively.

**Proof.** For \( m \in \{0, 1, 2, \ldots, n-1\} \) and \( t \in [m, m+1) \), let

\[ A_n(t) = \{ (x_1, x_2, \ldots, x_n) : x_i \geq 0 \text{ for } i \in \{1, 2, \ldots, n\} \} \]

and

\[ B_j(t) = \{ (x_1, x_2, \ldots, x_n) \in A_n(t) : x_j > 1 \} \]

for \( j \in \{1, 2, \ldots, n\} \). The hypervolume of the \( n \)-dimensional solid \( A_n(t) \) has value

\[ \text{Vol}(A_n(t)) = \frac{t^n}{n!} \]  

[38], since the solid is a standard orthogonal simplex from the corner of an \( n \)-cube. Similarly, if \( k \in \{1, 2, \ldots, m\} \), then the hypervolume of \( \bigcap_{j=1}^{k} B_j(t) \) is

\[ \text{Vol}\left( \bigcap_{j=1}^{k} B_j(t) \right) = \frac{(t-k)^n}{n!} \]  

For \( k \in \{m+1, m+2, \ldots, n\} \),

\[ \bigcup_{j=1}^{k} B_j(t) = \emptyset, \]

\[ \text{Vol}\left( \bigcup_{j=1}^{k} B_j(t) \right) = 0, \]

since the sum of nonnegative coordinates exceeds the number of coordinates which are greater than 1.
By the inclusion-exclusion principle,
\[ F(t) = P(T \leq t) = Vol(A_n(t) \cap C_n) = Vol(A_n(t)) \]
\[ - Vol \left( \frac{n}{n!} \sum_{k=1}^{m} (-1)^{k-1} \right. \]
\[ \cdot \sum_{1 \leq j_1 < j_2 < \cdots < j_n \leq m} \text{Vol}(B_{j_1}(t) \cap B_{j_2}(t) \cap \cdots \cap B_{j_n}(t)) \]
\[ = \frac{t^n}{n!} \sum_{k=1}^{m} (-1)^{k-1} \]
\[ \cdot \left( \binom{n}{k} \text{Vol}(B_{j_1}(t) \cap B_{j_2}(t) \cap \cdots \cap B_{j_k}(t)) \right) \]
\[ - \sum_{k=0}^{m} (-1)^{k-1} \binom{n}{k} \left( t-k \right)^n \frac{k!}{n!} = \sum_{k=0}^{m} (-1)^{k} \binom{n}{k} \left( t-k \right)^n \frac{k!}{n!} \]
\[ \cdot \left( t-k \right)^n \frac{k!}{n!}. \]  

In (1), \( F(n) \) is the Stirling number of the second kind with both parameters equal to \( n \) and has numerical value 1 [39, pages 38-39]. If \( t \geq n \), then \( C_n \subset A_n(t) \), so \( F(t) = 1 \) in this case. Since \( F \) is a polynomial, \( \sum_{k=0}^{\infty} (-1)^k \left( \binom{n}{k} \left( t-k \right)^n / n! \right) = 1 \) for all real-valued \( t \). Introducing the unit step function gives (1), and differentiation with respect to \( t \) gives (2).

3. Discussion and a Generalization

Figures 2 and 3 reveal the structure of the CDF
\[ F(t) = \frac{1}{2} t^2 s_0(t) - \frac{1}{2} (t - 1)^3 s_1(t) + \frac{1}{2} (t - 2)^3 s_2(t) \]
for \( n = 2 \). Figure 2 demonstrates how the hyperplane (line), which is the line of a constant sum of the values of the random variables and is perpendicular to the \( n \)-cube's (square’s) main diagonal, accrues volume (area) below it. Figure 3 illustrates the regions that are included and excluded for various positions of the hyperplane (line) and how vertices are meet in sets. For \( n = 2 \), the binomial coefficients, which provide the counts of the vertices, are 1 for \((0,0)\), 2 for \((1,0)\) and \((0,1)\), and 1 for \((1,1)\), as seen in Figures 2 and 3. In (11), the first term is the area of the large triangle in Figures 3(a), 3(b), and 3(c); the second term is the sum of the areas of the two hatched triangles in Figure 3(b), where exactly one of \( x_1, x_2 \) is greater than 1, and in Figure 3(c); and the third term is the area of the crosshatched triangle in Figure 3(c), where both \( x_1 \) and \( x_2 \) are greater than 1.

Figure 4 shows the same geometric interpretation for \( n = 3 \). In its CDF
\[ F(t) = \frac{1}{6} t^3 s_0(t) - \frac{1}{2} (t - 1)^3 s_1(t) + \frac{1}{2} (t - 2)^3 s_2(t) + \frac{1}{2} (t - 3)^3 s_3(t), \]
the first term is the volume using (7) of the large orthogonal simplex in Figures 4(a), 4(b), and 4(c) with edges of length \( t \). The second term is the sum of the volumes using (8) of the three orthogonal simplexes, where exactly one of \( \{x_1, x_2, x_3\} \) is greater than 1. In Figure 4(b), the vertices \( P_1, P_2, P_3, \) and \( P_4 \) of the simplex with \( x_1 > 1 \) are labeled. Their coordinates are \( P_1 : (t, 0, 0), P_2 : (1, 0, t - 1), P_3 : (1, t - 1, 0), \) and \( P_4 : (1, 0, 0) \). The lengths of the edges \( P_1 P_2, P_2 P_3, \) and \( P_3 P_4 \) are \( t - 1 \). The third term of (12) is the sum of the three volumes using (8), where exactly two of \( \{x_1, x_2, x_3\} \) are greater than 1. In Figure 4(c), the vertices are labeled \( P_3, P_5, P_6, \) and \( P_7 \) in the region where both \( x_1 \) and \( x_2 \) are greater than 1. Their
coordinates are $P_3 : (1, t-1, 0), P_5 : (t-1, 1, 0), P_6 : (1, 1, t-2),$ and $P_7 : (1, 1, 0).$ The lengths of the edges $P_3 P_5, P_3 P_7,$ and $P_6 P_7$ are $t-2.$ The fourth term is the region that is shared by all the other regions, analogous to the crosshatched region in Figure 3(c).

In the same way, for any $n$, the terms are the $n$-volumes of orthogonal $n$-simplexes, whose multiplicity is counted by binomial coefficients determined by the number of vertices of the $n$-cube in sets as the “moving” $n-1$-dimensional hyperplane “passes” them as $t$ increases. The hyperplane is perpendicular to the diagonal line $x_1 = x_2 = x_3 = \ldots = x_n$. The volumes of the simplexes are computed using (7) and (8).

The Website [40] has a free simulator for $T$, where selecting $n$ yields the PDF (2). Other calculators are at [41, 42].

The method of proof in Section 2 can be extended to linear combinations of uniform random variables on different intervals. Suppose that $X_1, X_2, \ldots, X_n$ are independent, that $X_k$ is uniformly distributed on the interval $[a_k, b_k]$, and that $c_1, c_2, \ldots, c_n$ are real constants. Also,

$$P\left(\sum_{k=1}^{n} c_k X_k \leq t\right) = P\left(\sum_{k=1}^{n} d_k Y_k \leq t'\right),$$

(13)

where

$$Y_k = \frac{X_k - a_k}{b_k - a_k},$$

$$d_k = c_k (b_k - a_k),$$

(14)

$$t' = \sum_{k=1}^{n} a_k c_k.$$

Then, $Y_1, Y_2, \ldots, Y_n$ are independent uniform random variables on $[0, 1]$, and $P(\sum_{k=1}^{n} c_k X_k \leq t)$ can be interpreted as
the hypervolume of the solid that consists of all points that lie inside the unit hypercube $[0, 1]^n$ and on one side of the hyperplane $\sum_{k=1}^n d_k Y_k = t'$. Now, proceed by inclusion-exclusion as in Section 2. In general, the formula for $P(\sum_{k=1}^n q_i X_k \leq t)$ is complicated because of the lack of symmetry that is caused by the presence of $d_1, d_2, \ldots, d_n$. This increases the number of cases and removes the congruence of the solids of each size whose hypervolumes need to be added or subtracted at each stage of the inclusion-exclusion process. Nevertheless, the correct distribution is obtained in this manner. A special case in which these problems disappear is $d_1 = d_2 = \cdots = d_n = d$, so that

$$P\left( \sum_{k=1}^n q_i X_k \leq t \right) = \begin{cases} F\left( \frac{t'}{d} \right) & \text{for } d > 0 \\ 1 - F\left( \frac{t'}{d} \right) & \text{for } d < 0, \end{cases} \quad (15)$$

where $F$ is given in (1).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2, John Wiley & Sons, New York, NY, USA, 2nd edition, 1995.

[2] K. R. Koch, *Introduction to Bayesian Statistics*, Springer, Berlin, Germany, 2nd edition, 2010.

[3] C. P. Quesenberry, "Probability integral transformations," in *Encyclopedia of Statistical Sciences*, S. Kotz, N. L. Johnson, and C. B. Read, Eds., vol. 7, pp. 225–231, Wiley, New York, NY, USA, 1986.

[4] V. K. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, New York, NY, USA, 1976.

[5] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer Series in Statistics, Springer, New York, NY, USA, 2nd edition, 1985.
[6] P. C. Silva, J. O. Cerdeira, M. J. Martins, and T. Monteiro-Henriques, "Data depth for the uniform distribution," *Environmental and Ecological Statistics*, vol. 21, no. 1, pp. 27–39, 2014.

[7] K. Jayakumar and K. K. Sankaran, "On a generalization of uniform distribution and its properties," *Statistica*, vol. 76, no. 1, pp. 83–91, 2016.

[8] C. P. Dettmann and M. K. Roychowdhury, "Quantization for uniform distributions on equilateral triangles," *Real Analysis Exchange*, vol. 42, no. 1, pp. 149–166, 2017.

[9] E. S. Pearson, *The History of Statistics in the 17th and 18th Centuries*, Against the Changing Background of Intellectual, Scientific and Religious Thought, Lectures by Karl Pearson given at University College London during the Academic Sessions, Macmillan, New York, NY, USA, 1978.

[10] O. B. Sheynin, "Finite random sums (a historical essay)," *Archive for History of Exact Sciences*, vol. 9, no. 4-5, pp. 275–305, 1972/1973.

[11] H. Cramer, *Mathematical Methods of Statistics*, Princeton, Princeton, NJ, USA, 1946.

[12] S. K. Mitra and S. N. Banerjee, "On the probability distribution of round-off errors propagated in tabular differences," *The Australian Computer Society: The Australian Computer Journal*, vol. 3, no. 2, pp. 60–68, 1971.

[13] J. O. Irwin, "On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's Type II," *Biometrika*, vol. 19, no. 3/4, pp. 225–239, 1927.

[14] A. N. Lowan and J. Laderman, "On the distribution of errors in Nth tabular differences," *The Annals of Mathematical Statistics*, vol. 10, no. 4, pp. 360–364, 1939.

[15] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, vol. 1, Oxford, New York, NY, USA, 5th edition, 1987.

[16] V. M. Kruglov, "On one identity for distribution of sums of uniformly distributed random variables," *Theory of Probability and its Applications*, vol. 58, no. 2, pp. 329–331, 2014.

[17] H. Potuschak and W. G. Müller, "More on the distribution of the sum of uniform random variables," *Statistical Papers*, vol. 50, no. 1, pp. 177–183, 2009.

[18] E. G. Olds, "A note on the convolution of uniform distributions," *Annals of Mathematical Statistics*, vol. 23, no. 2, pp. 282–285, 1952.

[19] S. K. Mitra, "On the probability distribution of the sum of uniformly distributed random variables," *SIAM Journal on Applied Mathematics*, vol. 20, no. 2, pp. 195–198, 1971.

[20] D. M. Bradley and R. C. Gupta, "On the distribution of the sum of n non-identically distributed uniform random variables," *Annals of the Institute of Statistical Mathematics*, vol. 54, no. 3, pp. 689–700, 2002.

[21] S. M. Sadooghi-Alavandi, A. R. Nematzadeh, and R. Habibi, "On the distribution of the sum of independent uniform random variables," *Statistical Papers*, vol. 50, no. 1, pp. 171–175, 2009.

[22] H. Murakami, "A saddlepoint approximation to the distribution of the sum of independent non-identically uniform random variables," *Statistica Neerlandica. Journal of the Netherlands Society for Statistics and Operations Research*, vol. 68, no. 4, pp. 267–275, 2014.

[23] G. S. Lo, H. Sangare, and C. Ndiaye, "A review on asymptotic normality of sums of associated random variables," *Afrika Statistika*, vol. 11, no. 1, pp. 855–867, 2016.

[24] D. L. Barrow and P. W. Smith, "Classroom Notes: spline notation applied to a volume problem," *American Mathematical Monthly*, vol. 86, no. 1, pp. 50–51, 1979.
Submit your manuscripts at https://www.hindawi.com