Evaluation of a Multiple Integral of Tefera via Properties of the Exponential Distribution

Yaming Yu

Department of Statistics
University of California
Irvine 92697, USA

yamingy@uci.edu

Submitted: Jul 12, 2008; Accepted: Jul 21, 2008; Published: Jul 28, 2008
Mathematics Subject Classification: 26B12, 05A19, 60E05

Abstract

An interesting integral originally obtained by Tefera ("A multiple integral evaluation inspired by the multi-WZ method," Electron. J. Combin., 1999, #N2) via the WZ method is proved using calculus and basic probability. General recursions for a class of such integrals are derived and associated combinatorial identities are mentioned.

1 Background

The integral in question reads

$$\int_{[0,\infty)^k} (e_2(\mathbf{x}))^m (e_1(\mathbf{x}))^n e^{-e_1(\mathbf{x})}\, d\mathbf{x} = \frac{m!(2m+n+k-1)!(k/2)_m}{(2m+k-1)!} \left(\frac{2(k-1)}{k}\right)^m T_k(m), \tag{1}$$

where k is a positive integer, m and n are nonnegative integers, $\mathbf{x} = (x_1, \ldots, x_k)$, $e_1(\mathbf{x}) = \sum_{i=1}^k x_i$, $e_2(\mathbf{x}) = \sum_{1\le i<j\le k} x_i x_j$, $(y)_m = \prod_{i=0}^{m-1}(y+i)$, and $T_k(m)$ is defined recursively by

$$T_k(m) - T_k(m-1) = \frac{(k(k-2))^m((k-1)/2)_m}{(k-1)^{2m}(k/2)_m} T_{k-1}(m), \quad m \ge 1,\ k \ge 2, \tag{2}$$

and

$$T_1(m) = 0, \quad m \ge 0,$$
$$T_k(0) = 1, \quad k \ge 2.$$

Note that we are using an uncommon convention $0^0 = 0$ for the case $m = n = 0,\ k = 1$.

In [1], Tefera gave a computer-aided evaluation of (1), demonstrating the power of the WZ [2] method. It was also mentioned that a non-WZ proof would be desirable, especially if it is short. This note aims to provide such a proof.

2 A short proof

This is done in two steps – the first does away with the integer n using properties of the exponential distribution, while the second builds a recursion using integration by parts. In this section we denote the left hand side of (1) by $I(n,m,k)$.

Proposition 1. *For $n \geq 1$ we have $I(n,m,k) = (2m+n+k-1)I(n-1,m,k)$.*

Proof. Let $Z_1,\ldots,Z_k$ be independent random variables each having a standard exponential distribution, i.e., the common probability density is $p(z) = e^{-z},\ z > 0$. Denoting $\mathbf{Z} = (Z_1,\ldots,Z_k)$ we have

$$\begin{aligned} I(n,m,k) &= E(e_2(\mathbf{Z}))^m (e_1(\mathbf{Z}))^n \\ &= E(e_1(\mathbf{Z}))^{2m+n} \left(\frac{e_2(\mathbf{Z})}{(e_1(\mathbf{Z}))^2}\right)^m \\ &= E(e_1(\mathbf{Z}))^{2m+n} E\left(\frac{e_2(\mathbf{Z})}{(e_1(\mathbf{Z}))^2}\right)^m \\ &= \frac{(2m+n+k-1)!}{(k-1)!} E\left(\frac{e_2(\mathbf{Z})}{(e_1(\mathbf{Z}))^2}\right)^m \end{aligned}$$

where we have used two properties of the exponential distribution: (i) $e_1(\mathbf{Z})$ is independent of $(Z_1,\ldots,Z_k)/e_1(\mathbf{Z})$ and hence independent of $e_2(\mathbf{Z})/(e_1(\mathbf{Z}))^2$, and (ii) $e_1(\mathbf{Z})$ has a gamma distribution $\mathrm{Gam}(k,1)$ whose jth moment is $(j+k-1)!/(k-1)!$. The claim readily follows. □

Proposition 2. *For $k \geq 2$ and $m \geq 1$ we have*

$$I(0,m,k) = I(0,m,k-1) + \frac{m(k-1)(k+2(m-1))}{k} I(0,m-1,k). \tag{3}$$

Proof. Denote $\mathbf{x}_{-1} = (x_2,\ldots,x_k)$. Using integration by parts and exploiting the symmetry we obtain

$$\begin{aligned} I(0,m,k) &= \int\int (e_2(\mathbf{x}))^m e^{-e_1(\mathbf{x})} dx_1 d\mathbf{x}_{-1} \\ &= \int -e^{-e_1(\mathbf{x})}(e_2(\mathbf{x}))^m\Big|_{x_1=0}^{\infty} d\mathbf{x}_{-1} + \int\int m e^{-e_1(\mathbf{x})}(e_2(\mathbf{x}))^{m-1} e_1(\mathbf{x}_{-1}) dx_1 d\mathbf{x}_{-1} \\ &= \int e^{-e_1(\mathbf{x}_{-1})}(e_2(\mathbf{x}_{-1}))^m d\mathbf{x}_{-1} + \frac{m(k-1)}{k}\int e^{-e_1(\mathbf{x})}(e_2(\mathbf{x}))^{m-1} e_1(\mathbf{x}) d\mathbf{x} \\ &= I(0,m,k-1) + \frac{m(k-1)}{k} I(1,m-1,k) \end{aligned}$$

where the limits of integration are omitted to save space. The claim now follows by Proposition 1. □

To finish the proof of (1), we note that (i) by Proposition 1 it suffices to prove (1) for $n = 0$, (ii) if we denote the right hand side of (1) by $J(n,m,k)$, then based on (2), after simple algebra $J(0,m,k)$ satisfies the recursion (3) as $I(0,m,k)$ does, and (iii) the boundary values of $I(0,m,k)$ and $J(0,m,k)$ match, i.e., $I(0,m,1) = J(0,m,1) = 0$ for $m \geq 0$ and $I(0,0,k) = J(0,0,k) = 1$ for $k \geq 2$. Thus $I(n,m,k) \equiv J(n,m,k)$.

3 General recursions

This argument applies to a general class of integrals involving elementary symmetric functions. Specifically, let $e_j(\mathbf{x}) = \sum_{1\le i_1<\ldots<i_j\le k} x_{i_1}\ldots x_{i_j},\ j = 1,\ldots,k$, and consider the integral

$$I_k(n_1,\ldots,n_k) = \int_{[0,\infty)^k} e^{-e_1(\mathbf{x})} \prod_{j=1}^{k} (e_j(\mathbf{x}))^{n_j} d\mathbf{x} \tag{4}$$

for $n_j \ge 0,\ 1 \le j \le k,\ k \ge 1$. Relation (1) corresponds to $n_1 = n,\ n_2 = m$ and $n_3 = \ldots = n_k = 0$. The following recursions are obtained by trivial modifications in the proofs of Propositions 1 and 2.

Proposition 3. *For $n_1 \ge 1$ we have*

$$I_k(n_1,n_2,\ldots,n_k) = \left(k - 1 + \sum_{j=1}^{k} jn_j\right) I_k(n_1 - 1, n_2, \ldots, n_k).$$

Proposition 4. *For $k \ge 2$ we have*

$$\begin{aligned} I_k(0,n_2,\ldots,n_k) =& \delta_k I_{k-1}(0,n_2,\ldots,n_{k-1}) \\ &+ n_2\frac{k-1}{k}\left(k + 2(n_2-1) + \sum_{j=3}^{k} jn_j\right) I_k(0,n_2-1,n_3,\ldots,n_k) \\ &+ \sum_{j=3}^{k} n_j \frac{k-j+1}{k} I_k(0,\ldots,n_{j-1}+1,n_j-1,n_{j+1},\ldots,n_k) \end{aligned}$$

where $\delta_k = 1$ if $n_k = 0$ and $\delta_k = 0$ if $n_k > 0$.

Note that $I_k(n_1,\ldots,n_k)$ is given an arbitrary value if some $n_j < 0$; this does not affect the recursion in Proposition 4.

Together with the boundary condition $I_k(0,\ldots,0) = 1,\ k \ge 1$, Propositions 3 and 4 determine $I_k(n_1,\ldots,n_k)$ for all $k \ge 1$ and $n_j \ge 0,\ 1 \le j \le k$. It is doubtful whether these recursions are solvable in a simpler form. At any rate, we may calculate $I_k(0,n_2,\ldots,n_k),\ k \ge 2$, by building up a table of $I_l(0,m_2,\ldots,m_l)$ for values of l and m_i's that satisfy $l \le k,\ \sum_{j=2}^{l} m_j \le \sum_{j=2}^{k} n_j$, and $m_k \le n_k$ if $l = k$; this range can be further restricted if the largest j for which $n_j \ne 0$ is less than k. We omit the details but include some values of $I_3(0,n_2,n_3)$ calculated this way in Table 1.

It is reassuring to see that Table 1 contains only integer entries. This is not obvious from Proposition 4 but is so from (4), after expanding the product $\prod_{j=1}^{k}(e_j(\mathbf{x}))^{n_j}$ inside the integral. Alternatively, $I_k(n_1,\ldots,n_k)$ is a sum of products of various moments of the standard exponential distribution, and these moments are all integers.

Table 1: Values of $I_3(0, n_2, n_3)$ for $n_2 + n_3 \leq 4$.

$n_2 \backslash n_3$	0	1	2	3	4
0	1	1	8	216	13824
1	3	12	216	10368	
2	24	252	8640		
3	372	8208			
4	9504				

4 Associated combinatorial identities

It would be interesting to know whether there exists a direct combinatorial interpretation of $I_k(n_1, \ldots, n_k)$ as defined by (4). In this direction we mention two associated binomial sum identities.

Let $Z_1, Z_2, \ldots,$ be independent standard exponential random variables. For $n, m \geq 0$ we have

$$\begin{aligned} I_2(n,m) &= E(Z_1 + Z_2)^n (Z_1 Z_2)^m \\ &= \sum_{k=0}^{n} E\binom{n}{k} Z_1^{k+m} Z_2^{n-k+m} \\ &= \sum_{k=0}^{n} \binom{n}{k} (k+m)!(n-k+m)!. \end{aligned}$$

On the other hand, (1) gives

$$I_2(n,m) = \frac{(2m+n+1)!}{(2m+1)!}(m!)^2.$$

Thus we obtain a familiar looking identity

$$\binom{2m+n+1}{n} = \sum_{k=0}^{n} \binom{k+m}{m} \binom{n-k+m}{m}, \quad m, n \geq 0. \tag{5}$$

Another instance of (1) is

$$I_3(0, m, 0) = \frac{(2m+1)!}{3^m} \sum_{k=0}^{m} \frac{3^k (k!)^2}{(2k+1)!}, \quad m \geq 0.$$

We also have

$$\begin{aligned} I_3(0,m,0) &= E(Z_1Z_2+Z_1Z_3+Z_2Z_3)^m \\ &= \sum_{0\le i,\,0\le j,\,i+j\le m} E\frac{m!}{i!j!(m-i-j)!}(Z_1Z_2)^i(Z_1Z_3)^j(Z_2Z_3)^{m-i-j} \\ &= \sum_{0\le i,\,0\le j,\,i+j\le m} \frac{m!(i+j)!(m-j)!(m-i)!}{i!j!(m-i-j)!}, \end{aligned}$$

and after rewriting we get a less familiar but interesting identity

$$\frac{(2m+1)!}{3^m(m!)^2}\sum_{k=0}^{m}\frac{3^k(k!)^2}{(2k+1)!} = \sum_{0\le i,\,0\le j,\,i+j\le m}\binom{m-j}{i}\binom{m-i}{j}\binom{m}{i+j}^{-1}, \quad m\ge 0. \qquad (6)$$

Of course, (5) and (6) can be derived via alternative methods, for example the WZ method; the purpose of presenting them is mainly to draw attention to the potential of $I_k(n_1,\dots,n_k)$ as combinatorial entities.

References

[1] A. Tefera, A multiple integral evaluation inspired by the multi-WZ method, *Electron. J. Combin.* **6** (1999), #N2.

[2] H.S. Wilf and D. Zeilberger, An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities, *Invent. Math.* **108** (1992), 575–633.