THE AREA OF A RANDOM TRIANGLE IN A SQUARE.

JOHAN PHILIP

Abstract. We determine the distribution function for the area of a random triangle in a unit square. The result is not new, [8], [12]. The method presented here is worked out to shed more light on the problem.

1. Introduction

We shall denote the square by $K$ and the random triangle by $T$ and shall consider the random variable $X = \text{area}(T)/\text{area}(K)$. It is well known that an affine transformation will preserve the ratio $X$. This follows from the fact that the area scaling is constant for an affine transformation. The scale equals the determinant of the homogeneous part of the transformation. This means that our results hold when $K$ is a parallelogram.

Various aspects of our problem have been considered in the field of geometric probability, see e.g. [11]. J. J. Sylvester considered the problem of a random triangle $T$ in an arbitrary convex set $K$ and posed the following problem: Determine the shape of $K$ for which the expected value $\kappa = E(X)$ is maximal and minimal. A first attempt to solve the problem was published by M. W. Crofton in 1885. Wilhelm Blaschke [3] proved in 1917 that $\frac{35}{48\pi^2} \leq \kappa \leq \frac{1}{12}$, where the minimum is attained only when $K$ is an ellipse and the maximum only when $K$ is a triangle. The upper and lower bounds of $\kappa$ only differ by about 13%. It has been shown, [2] that $\kappa = \frac{11}{144}$ for $K$ a square.

A. Rényi and R. Sulanke, [9] and [10], consider the area ratio when the triangle $T$ is replaced by the convex hull of $n$ random points. They obtain asymptotic estimates of $\kappa$ for large $n$ and for various convex $K$. R. E. Miles [7] generalizes these asymptotic estimates for $K$ a circle to higher dimensions. C. Buchta and M. Reitzner, [4], has given values of $\kappa$ (generalized to three dimensions) for $n \geq 4$ points in a tetrahedron. H. A. Alikoski [2] has given expressions for $\kappa$ when $T$ is a triangle and $K$ a regular $r$-polygon.

Here, we shall deduce the distribution function for $X$. We have done this before in a simpler way than in this paper, [8]. We hope that the
method presented here shall be applicable if the square $K$ is replaced by a pentagon or hexagon.

2. Notation and formulation.

As $K$, we will take the unit square $(0 \leq x \leq 1, 0 \leq y \leq 1)$. We use a constant probability density in $K$ for generating three random points in $K$. The coordinates of the points will be denoted $(x_k, y_k)$ for $1 \leq k \leq 3$. Each $x_k$ and $y_k$ is evenly distributed in $(0, 1)$ and they are independent. Let $T$ be the convex hull of the three points. We shall determine the probability distribution of the random variable $X = \text{area}(T)$.

Our method will be to shrink the square around its midpoint until one of its sides hits a triangle point. The shrunk square is denoted $B$. The random variable $X$ that we study will be written as the product of two random variables

$$V = \text{area}(B) \text{ and } W = \frac{\text{area}(T)}{\text{area}(B)}.$$ 

We have six independent variables $x_k$ and $y_k$, $(1 \leq k \leq 3)$. One of them stops the shrinking and determines $V$. The remaining five variables determine $W$. It follows from the independence of the $x_k$ and $y_k$ that $V$ and $W$ are independent.

We shall determine the distributions of $V$ and $W$ and combine them to get the distribution of $X = VW$.

3. The distribution of $V$.

$V$ is the area of the shrunk square $B$. The size of $B$ is determined by the largest of the six variables $|x_k - \frac{1}{2}|$ and $|y_k - \frac{1}{2}|$, $(1 \leq k \leq 3)$. Each of these variables has the distribution function $K(t) = 2t$, $(0 \leq t \leq \frac{1}{2})$. The largest of the six has the distribution function $K(t_{\text{max}})^6 = (2t_{\text{max}})^6$. The area of $B$ is $v = (2t_{\text{max}})^2$. We get

$$G(v) = \text{Prob}(V < v) = (2t_{\text{max}})^6 = v^3, \quad 0 \leq v \leq 1.$$ 

4. The distribution of $W$.

$W$ is the area of a random triangle having one vertex on the boundary of a square $(=B)$ and the other two vertices in the interior of the square. Since the area ratio $W$ is independent of the size of $B$, we will take $B$ as the original unit square $K$. Without loss of generality, we will number the three triangle vertices so that vertex one is the one sitting on the boundary and we let this boundary be the x-axis, so that vertex one is $(x, 0)$. The position of the second vertex is $(x_2, y_2)$. Let $l_0$ be the line through vertices one and two. It contains one side of the triangle. Our calculations will be divided into three cases depending on where $l_0$ intersects $B$. 
4.1. Case 1. Case 1, depicted in Figure 1, occurs when \( l_0 \) intersects the square side along the y-axis in the point \((0, y), 0 \leq y \leq 1\). The equation for \( l_0 \) is

\[
l_0 : \eta = -\frac{y}{x} \xi + y.
\]

Let \( s = \sqrt{(x - x_2)^2 + y_2^2} \) be the distance between vertices one and two. For fixed \( x \) and \( y \), the maximal value of \( s \) is \( r_1 = \sqrt{x^2 + y^2} \).

The area of the triangle \( T \) will be less than \( w \) if the distance between \( l_0 \) and the third vertex is less than \( 2w/s \). To avoid the factor 2 in numerous places below, we shall use the double area \( u = 2w \) in the calculations. The lines \( l_1 \) and \( l_2 \) have the distance \( u/s \) to \( l_0 \).

\[
l_1 : \eta = -\frac{y}{x} \xi + y - \frac{ur_1}{sx},
\]

\[
l_2 : \eta = -\frac{y}{x} \xi + y + \frac{ur_1}{sx}.
\]

This means that the conditional probability \( P(W \leq u/2 \mid x, y, s) \) equals the area between the lines \( l_1 \) and \( l_2 \) in the unit square in Figure 1. We shall use the formula \( 1 - T_1 - S_1 \) (see Figure 1) for this area and we shall average \( T_1 \) and \( S_1 \) over \( x, y, \) and \( s \) to get the contribution to \( P(W \leq u/2) \) from Case 1. In fact, when we consider all possible directions of \( l_0 \) in all our cases, it follows from a symmetry argument.
Figure 2. Area to integrate $s$ and $y$ over in Case 1 when $u = 1/6$ and $x = 1/2$.

that the areas to the left of $l_1$ will be the same as those to the right of $l_2$. This implies that it suffices to calculate the areas to the left and then double the result. Thus, we shall average $2T_1$ over $x$, $y$, and $s$ and neglect $S_1$.

From the equation of $l_1$, we get

$$2T_1 = \frac{x}{y} \left( y - \frac{u r_1}{s x} \right)^2 \text{ if } s > \frac{u r_1}{x y}, \text{ otherwise } 0.$$  

We shall determine the densities of $x$, $y$, and $s$. Obviously, $x$ is evenly distributed over $(0, 1)$. The area to the left of $l_0$ is $xy/2$, so for fixed $x$, the density is the differential $\frac{x}{2} \, dy$. For fixed $x$ and $y$ consider the small triangle with vertices in $(x, 0)$, $(0, y)$, and $(0, y + dy)$. The fraction of the small triangle below $s$ is $\left( \frac{s}{r_1} \right)^2$ and the density is the differential $\frac{2x}{r_1^2} \, ds$. Notice that the integral of the combined density $\rho_1 = x s / r_1^2$ over the whole range of $(x, y, s)$ is not 1 but $\frac{1}{4}$, which is the probability for Case 1.

Figure 2 shows the range in $(s, y)$-space to integrate over for fixed $u$ and $x$. The increasing curve is the upper bound $r_1$ for $s$ and the decreasing curve is its lower bound $s_0 = \frac{u r_1}{x y}$. The intersection of the lower and upper $s$-bounds is the lower bound $y_0 = u/x$ for $y$. We have $y_0 < 1$ when $x > u$. 


Figure 3. Case 2. The line $l_0$ through vertices one and two intersects the top side of the square, while $l_1$ intersects the left side in Case 2a and the top side in Case 2b. The figure is drawn with $z < x$.

The contribution from Case 1 is the weighted average of $2T_1$:

$$h_1(u) = \int_u^1 x \, dx \int_{u/x}^1 r_1^{-2} \, dy \int_{u/r_1/xy}^{r_1} \frac{x}{y} \left( y - \frac{u r_1}{s x} \right)^2 \, s \, ds.$$

Maple is helpful in solving integrals of this kind and delivers the result

$$h_1(u) = -\frac{1}{3} u^3 + \frac{5}{4} u^2 - u + \frac{1}{12} - \frac{1}{2} u^2 \log(u) \left( 1 - \log(u) \right).$$

4.2. Cases 2. Case 2 occurs when $l_0$ intersects the top side of the square. It has two subcases a, and b, depicted in Figures 3a and 3b. Case 2a occurs when $l_1$ intersects the left side of the square and Case 2b when $l_1$ intersects the top side.

We let $z$ stand for the $x$-coordinate of the intersection between the top side and $l_0$.

In Cases 2, the maximal value of $s$ is $r_2 = \sqrt{1 + (x - z)^2}$. The area to the left of $l_0$ is $\frac{x + z}{2}$, so the $z$-density is $\frac{1}{2}$ and like in Case 1, the $s$-density is $2s/r_2^2$. The combined density is $\rho_2 = s/r_2^2$.

We have the two subcases $z < x$ and $z > x$. By symmetry, we can do the calculations for $z < x$ and then double the result.

The equations for $l_0$ and $l_1$ are

\[
\begin{align*}
l_0 : \quad \eta_0 & = -\frac{1}{x - z} (\xi - x), \\
l_1 : \quad \eta_1 & = -\frac{1}{x - z} (\xi - x) - \frac{u r_2}{s (x - z)}. \end{align*}
\]
Figure 4. Areas to integrate $s$ and $z$ over in Case 2 when $u = 1/6$ and $x = 1/2$.

We have Case 2a when $0 < \eta_1 < 1$, which occurs when

$$\frac{u r_2}{x} = s_1 < s < s_2 = \frac{u r_2}{z}.$$  

Case 2b occurs when $\eta_1 > 1$ i.e. when $s > s_2$.

The expression for $2T_{2a}$ and $2T_{2b}$ are obtained from the equation of $l_1$:

$$2T_{2a} = \frac{(x - u r_2/s)^2}{x - z},$$

$$2T_{2b} = x + z - 2u r_2/s.$$  

(5)

Figure 4 shows the range in $(s, z)$-space to integrate over for fixed $u$ and $x$. The lower and upper bounds for $s$ are $s_1 = \frac{u r_2}{z}$ and $r_2$. The curve $s_2 = \frac{u r_2}{z}$ is upper bound for Case 2a and lower bound for Case 2b. $r_2$ and $s_2$ intersect at $z = u$.

The contributions from Cases 2a and 2b are the doubled weighted averages of $2T_{2a}$ and $2T_{2b}$:
\[ h_{2a}(u) = 2 \int_u^1 dx \left( \int_0^u r_2^{-2} dz \int_{ur_2/x}^{r_2} 2T_{2a} \ s ds \right) \]
\[ + \int_u^x r_2^{-2} dz \int_{ur_2/x}^{ur_2/z} 2T_{2a} \ s ds \right), \]

\[ h_{2b}(u) = 2 \int_u^1 dx \int_u^x r_2^{-2} dz \int_{ur_2/z}^{r_2} 2T_{2b} \ s ds. \]

Evaluation of the integrals gives

\[ h_{2a}(u) = \frac{1}{6}(2u + u^2 - 3u^3) + \frac{1}{6}(9 + 2u) u^2 \log(u) \]
\[ - \frac{1}{6}(1 - 5u - 2u^2)(1 - u) \log(1 - u) \]
\[ + u^2 \log(u) \log(1 - u) + \text{dilog}(u), \]

\[ h_{2b}(u) = \frac{1}{4}(1 - 5u - 2u^2)(1 - u) - \frac{3}{2} u^2 \log(u). \]

Here, \text{dilog} is Maple’s dilog function. See Appendix A.

4.3. Case 3. This case, depicted in Figures 5 a-d, occurs when \( l_0 \) intersects the right side of the square in the point \((1, y), \) \( 0 \leq y \leq 1. \) When dealing with Case 3, we shall use the variable \( x_1 = 1 - x. \) The density \( \rho_3 \) is obtained by replacing \( x \) by \( x_1 \) in \( \rho_1. \) The equations for \( l_0 \) and \( l_1 \) are

\[ l_0 : \eta = \frac{y}{x_1} (\xi - 1) + y, \]
\[ l_1 : \eta = \frac{y}{x_1} (\xi - 1) + y + \frac{ur_3}{sx_1}. \]

Here, \( s \) ranges from 0 to \( r_3 = \sqrt{x_1^2 + y^2}. \) Depending on the values of \( x_1, y, \) and \( u/s, \) the area to the left of \( l_1 \) takes four different shapes as demonstrated in Figure 5. The four areas are:

\[ 2T_{3a} = \frac{1}{x_1 y} \left( x_1 + y - x_1 y - \frac{r_3 u}{s} \right)^2, \]
\[ 2T_{3b} = 2 - 2x_1 + \frac{x_1}{y} - \frac{2r_3 u}{sy}, \]
\[ 2T_{3c} = 2 - 2y + \frac{y}{x_1} - \frac{2r_3 u}{sx_1}, \]
\[ 2T_{3d} = 2 - \frac{1}{x_1 y} \left( x_1 y + \frac{r_3 u}{s} \right)^2. \]

\( T_{3a} \) occurs for the largest values of \( u/s, \) i.e. for the smallest \( s. \) \( T_{3a} = 0 \)

when \( s < s_3 = \frac{r_3 u}{x_1 + y - x_1 y}. \) When \( s \) increases, \( l_1 \) moves to the right and if
y > \( x_1 \) it passes the origin so that \( T_{3b} \) replaces \( T_{3a} \). This happens when \( s \) passes \( s_4 = \frac{r_3 u}{y(1-x_1)} \). If, on the other hand, \( y < x_1 \), \( l_1 \) passes the point \((1,1)\) and \( T_{3c} \) replaces \( T_{3a} \) when \( s \) passes \( s_5 = \frac{r_3 u}{x_1(1-y)} \). \( T_{3d} \) occurs when \( s > \max(s_4, s_5) \). The upper bound for \( s \) is \( r_3 \), so the various subcases occur only where \( s_4 \) and \( s_5 \) are smaller than \( r_3 \).

Figure 7 displaying the boundaries in \((y,s)\)-space for \( u = \frac{1}{8} \) and \( x_1 = \frac{1}{2} \) gives an idea of the situation. The corresponding boundaries for \( u = \frac{3}{8} \) and \( x_1 = \frac{1}{2} \) are given in Figure 8.

The intersections of the curves in the figures are at:

\[ y_3 = \frac{u - x_1}{1 - x_1} \text{ between } r_3 \text{ and } s_3 \]

\[ y_4 = \frac{u}{1 - x_1} \text{ between } r_3 \text{ and } s_4 \]

\[ y_5 = 1 - \frac{u}{x_1} \text{ between } r_3 \text{ and } s_5. \]
We always have $y_3 \leq y_4$, and we have $y_3 > 0$ when $u > x_1$, and $y_4 < 1$ when $x_1 < 1 - u$, and $y_4 < y_5$ when $u < x_1 (1 - x_1)$. This means that we have different subcases depending on the values of $u$ and $x_1$. We have drawn the boundaries $x_1 = u$, $x_1 = 1 - u$, and $u = x_1 (1 - x_1)$ in Figure 6 and indicated where the configurations in Figures 7 and 8 occur.

We shall show that, even though the areas to integrate $s$ and $y$ over are very different in Figures 7 and 8, the resulting integrals are the same, meaning that we don’t have to carry out the integrations in Figure 7. Later, we shall show that we don’t have to calculate the integrals in Figure 8 either because the integrals in Figure 9 give the same result.

When going from Figure 7 to 8, i.e. when increasing $u$ past $x_1 (1 - x_1)$, $s_3$, $s_4$, and $s_5$ rise and $y_5$ becomes smaller then $y_4$. The intersection of $s_4$ and $s_5$ passes $r_3$ so that the area in Figure 7 marked $T_{3d}$ disappers and the area marked 0 in Figure 8 is created. Denote the area where $T_{3d}$ is valid in Figure 7 by $A$ and consider the integral of $T_{3b}$. This integral goes in Figure 8 from $s_4$ to $r_3$ in the s-direction and from $y_4$ to 1 in the y-direction. In Figure 7, it goes over the same area minus $A$. A corresponding argument holds for $T_{3c}$. The missing integral of $T_{3a}$ over the area marked 0 in Figure 8 is
Figure 7. Areas in \((y,s)\)-space, where the different \(T_{3i}\) occur when \(u = \frac{1}{8}\) and \(x_1 = \frac{1}{2}\).

\[
\int_{y_5}^{x_1} \int_{r_3}^{s_5} T_{3a} \rho_3 ds + \int_{x_1}^{y_4} \int_{r_3}^{s_4} T_{3a} \rho_3 ds = \\
= \int_{x_1}^{y_4} \int_{s_3}^{r_3} T_{3a} \rho_3 ds + \int_{y_4}^{x_1} \int_{s_4}^{r_3} T_{3a} \rho_3 ds.
\]

The latter two integrals are integration of \(T_{3a}\) over \(A\). This means that the difference between Figure 8 and 7 is integration over \(A\) of \(\Delta = T_{3a} + T_{3d} - T_{3b} - T_{3c}\). Insertion of the \(T_{3i}\) from equation (8) shows that \(\Delta = 0\), implying that integration in Figure 7 gives the same result as in Figure 8.

Now, consider the area in Figure 6 marked Fig10 and also Figure 10 as well as Figure 8. These Figures show that \(T_{3b}\) exists when \(y_4 < 1\), i.e. when \(x_1 < 1 - u\). The contribution from \(T_{3b}\) to the distribution function is
The Figures 8, and 11 show that $T_{3c}$ exists when $y_5 > 0$, i.e. when $x_1 > u$. The contribution from $T_{3c}$ to the distribution function is

\begin{equation}
(11)
  h_{3c}(u) = \int_{u}^{1} dx_1 \int_{y_5}^{y_{s_5}} dy \int_{s_5}^{r_3} T_{3c} \rho_3 ds = \frac{1}{2} (1 - u + u \log(u))^2.
\end{equation}

The contribution from $T_{3a}$ is more complicated since it exists for all $x_1$. $T_{3a}$ is present in Figures 8, 9, 10, and 11. We shall show that the contribution is the same for $u < \frac{1}{2}$ and $u > \frac{1}{2}$. For $u < \frac{1}{2}$, we have $u < 1 - u$ and shall integrate over the areas in Figures 10, 8, and 11. Omitting the integrand $2T_{3a} \rho_3$ and the differentials and just writing the integration boundaries, we have for $u < \frac{1}{2}$:
Figure 9. Areas in (y,s)-space, where $T_{3a}$ is valid when $u = \frac{5}{8}$ and $x_1 = \frac{1}{2}$. Note that $1 - u < x_1 < u$.

(12) $I_1 = \int_0^u \left( \int_{y_3}^{y_4} \int_{s_3}^{r_3} + \int_{y_4}^1 \int_{s_3}^{s_4} \right) + \int_1^{1-u} \left( \int_0^{y_5} \int_{s_3}^{s_5} + \int_{y_5}^{r_3} \right) + \int_0^1 \left( \int_{s_3}^{r_3} + \int_{s_3}^{s_3} \right)$

For $u > \frac{1}{2}$, we have $1 - u < u$ and shall integrate over the areas in Figures 10, 9, and 11 and get:

(13) $I_2 = \int_0^{1-u} \left( \int_{y_3}^{y_4} \int_{s_3}^{r_3} + \int_{y_4}^1 \int_{s_3}^{s_4} \right) + \int_{1-u}^u \int_{y_3}^{y_5} \int_{s_3}^{s_5} + \int_u^1 \left( \int_{s_3}^{r_3} + \int_{s_3}^{s_3} \right)$

We shall show that the integrals $I_1$ and $I_2$ are the same.

First, notice that in both $I_1$ and $I_2$, the $x$-integration of $\int_{y_4}^1 \int_{s_3}^{s_4}$ goes from 0 to $1 - u$ and that of $\int_{y_5}^1 \int_{s_3}^{s_3}$ goes from $u$ to 1.
The three remaining integrals in $I_1$ and $I_2$ are all of type $\int_{s_3}^{r_3}$ in the $s$-direction. In the $x_1$- and $y$- directions, they are

$$I_1^* = \int_0^u \int_{y_3}^{y_4} + \int_u^{1-u} \int_{y_5}^{y_4} + \int_{1-u}^{1} \int_{y_5}^{1},$$

and

$$I_2^* = \int_0^{1-u} \int_{y_3}^{y_4} + \int_{1-u}^{1} \int_{y_3}^{1} + \int_{1}^{1-u} \int_{y_3}^{1} + \int_{1}^{1} \int_{y_3}^{1}.$$

$I_2^*$ can be split up into:

$$I_2^* = \int_0^u \int_{y_3}^{y_4} + \int_u^{1-u} \int_{y_5}^{y_4} - \int_u^{1-u} \int_{y_3}^{1} + \int_{1-u}^{1} \int_{y_5}^{1} + \int_{1}^{1-u} \int_{y_5}^{1}.$$

Comparing terms, it is easily seen that $I_1^*$ and $I_2^*$ are the same. This implies that the contribution $h_3$ to the distribution of $u$ from Case 3, has the same formula for $u < \frac{1}{2}$ and $u > \frac{1}{2}$.

We shall use the expression for $u > \frac{1}{2}$ to calculate $h_3 = h_{3a} + h_{3b} + h_{3c}$, where $h_{3b}$ and $h_{3c}$ are given in (10) and (11) and $h_{3a}$ is the contribution from $T_{3a}$ which we take from $I_2$ in (13). With the notation

$$k = k(u, x_1, y, s) = 2T_{3a} \rho_3,$$

we have.
4.4. Combination of cases. By combining the calculated \( h_{nx} \), we get the distribution function for twice the area of a random triangle in a unit square, when one of the triangle vertices sits on the boundary of the square:

\[
h_{3a}(u) = \int_0^{1-u} dx_1 \left( \int_{y_3}^{y_4} dy \int_{s_3}^{r_3} k ds + \int_{y_4}^{1} dy \int_{s_3}^{s_4} k ds \right) \\
+ \int_{1-u}^{u} dx_1 \int_{y_3}^{1} dy \int_{s_3}^{r_3} k ds \\
+ \int_{u}^{1} dx_1 \left( \int_{y_5}^{y_4} dy \int_{s_3}^{s_5} k ds + \int_{y_5}^{1} dy \int_{s_3}^{r_3} k ds \right) \\
= \frac{1}{36} (1-u)(5-97u-22u^2) \\
+ \frac{1}{3} (1-u)(-1 + 5u + 2u^2) \log (1-u) \\
+ \left( \frac{1}{6} - 2u + \frac{2}{3} u^3 \right) \log (u) - u^2 \log (u)^2 \\
+ 2u^2 (\log (u) \log (1-u) + \text{dilog}(u)),
\]

Figure 11. Areas in \((y,s)\)-space, where \( T_{3a} \) and \( T_{3c} \) are valid when \( u = \frac{3}{5} \) and \( x_1 = \frac{6}{7} \). Note that \( u < x_1 < 1 \).
\[ H(u) = 1 - h_1 - h_{2a} - h_{2b} - h_{3a} - h_{3b} - h_{3c} \]
\[ = \frac{u}{3} (14 - 11 u - 4 u^2 \log (u)) \]
\[ + \frac{2}{3} (1 - u)(1 - 5 u - 2 u^2) \log (1 - u) \]
\[ - 4 u^2 (\log (u) \log (1 - u) + \text{dilog}(u)), \quad 0 \leq u \leq 1. \]

5. Combination of the V- and W-distributions.

Let \( F(x) \) be the distribution function for the triangle area \( X \). We have \( X \leq x \) when \( VW = UV/2 \leq x \). Putting \( x = y/2 \), this happens when \( UV \leq y \) and we get

\[ F(y/2) = \int_0^1 G(y/u) dH(u) = \]
\[ = [G(y/u)H(u)]_0^1 - \int_y^1 H(u) \frac{d}{du} G(y/u) du = \]
\[ = G(y) - \int_y^1 H(u) \frac{d}{du} G(y/u) du, \quad 0 \leq y \leq 1. \]

The partial integration in (16) is used to avoid integrating to the lower bound \( u = 0 \). To write the result, we need the \( \nu \) function

\[ \nu(x) = - \int_0^x \frac{\log |1 - t|}{t} dt. \]

This function is the real part of the dilogarithm function \( \text{Li}_2(x) \) discussed by Euler in 1768 and named by Hill, [5]. \( \nu(x) \) is well defined on the whole real axis. Some properties of \( \nu(x) \) are given in Appendix A.

We will not carry out the integration (16) in detail, but will just give the result

\[ (18) \quad F(x) = \frac{4x}{3} (10 - 17x) - \frac{16x^3}{3} (17 - 3 \log(2x)) \log(2x) \]
\[ + \frac{2}{3} (1 - 16x - 68x^2)(1 - 2x) \log(1 - 2x) + 16x^2(3 + 2x) \left( \nu(2x) - \frac{\pi^2}{6} \right) \]
\[ 0 \leq x \leq 1/2. \]

Figure 12 shows the density function is \( f(x) = dF/dx \):
Figure 12. Density function for the area of a random triangle in a square.

(19) \[ f(x) = 12 \left[ 1 - 2x - 4x^2(5 - \log(2x)) \log(2x) \\
- (1 + 10x)(1 - 2x) \log(1 - 2x) + 8x(1 + x) \left( \nu(2x) - \frac{\pi^2}{6} \right) \right], \quad 0 \leq x \leq 1/2. \]

The first moments and the standard deviation of the triangle area are

(20) \[ \alpha_1 = \int_0^{1/2} x \, dF(x) = \frac{11}{144} \approx .076389, \]

(21) \[ \alpha_2 = \int_0^{1/2} x^2 \, dF(x) = \frac{1}{96}, \]

(22) \[ \sigma = \sqrt{\alpha_2 - \alpha_1^2} = \frac{\sqrt{95}}{144} \approx .067686. \]

6. Concluding comment.

We have not shown any integral calculations in detail. In principle, they are elementary, which doesn’t mean that they don’t require a substantial effort. As indicated, the calculations have been done in Maple. The calculations would not have been possible without some tool for handling the large number of terms that come out of the integrations. This doesn’t mean that Maple performs the integrations automatically. Often, we had to split up the integrands in parts and treat each part in a special way. We had to do some partial integrations manually.
We will supply any interested reader with a Maple file describing the calculations.

**APPENDIX A**

The dilogarithm function $\text{Li}_2(x)$ is defined in [6] for complex $x$ as

\[
\text{Li}_2(x) = -\int_0^x \frac{\log(1-t)}{t} \, dt.
\]

When $x$ is real and greater than unity, the logarithm is complex. A branch cut from 1 to $\infty$ can give it a definite value. In this paper, we are only interested in real $x$ and the real part of $\text{Li}_2$

\[
\nu(x) = \text{Re}(\text{Li}_2(x)) = -\int_0^x \frac{\log |1-t|}{t} \, dt.
\]

We have the series expansion

\[
\nu(x) = \text{Re}(\text{Li}_2(x)) = \sum_{k=1}^{\infty} \frac{x^k}{k^2}, \quad |x| \leq 1.
\]

Although the series is only convergent for $|x| \leq 1$, the integrals in (23) and (24) are not restricted to these limits and the $\nu$ function is defined and is real on the whole real axis. We use this function for $0 \leq x \leq 1$.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{nu_function}
\caption{The function $\nu(x)$.}
\end{figure}

The definition of the dilogarithm function has varied a little from author to author. Maple has the function polylog$(2, x)$ which is defined by the series expansion (25) for $|x| \leq 1$ otherwise by analytic
continuation. Maple also has a function \( \text{dilog}(x) = \text{Li}_2(1 - x) \) defined on the whole real axis. Maple’s dilog function is the same as the dilog function given in [1], page 1004.

\( \nu(x) \) is increasing from \( \nu(0) = 0 \) via \( \nu(1) = \pi^2/6 \) to \( \nu(2) = \pi^2/4 \).

The integrals involving \( \nu(x) \) needed for calculating the moments of various distributions take rational values like

\[
\begin{align*}
\int_0^1 x \, d\nu(x) &= 1, \\
\int_0^1 x^2 \, d\nu(x) &= \frac{3}{4}, \\
\int_0^1 x^3 \, d\nu(x) &= \frac{11}{18}.
\end{align*}
\]

References


Department of Mathematics, Royal Institute of Technology, S-10044 Stockholm Sweden

E-mail address: johanph@kth.se