# THE DISTRIBUTION AND THE EXPECTATION OF THE DISTANCE BETWEEN TWO RANDOM POINTS ON DIFFERENT FACES OF A UNIT CUBE IN THREE AND FOUR DIMENSIONS.

JOHAN PHILIP

Abstract. We calculate the distribution and the expectation of the distance between two random points on different faces of a unit cube. The expectation is known before, but the distribution is not. Our method enables us to carry out the calculations also for a four-dimensional cube.

## 1. Introduction

The expected distance between two random points on different sides of a square and the analouge in three dimensions was treated numerically by James D. Klein, [Borwein et. al. 2004, page 33]. Our attention to the problem was caught by [Bailey et. al. 2005]. See also [Borwein et. al. 2006] from which we cite problem 8:

Calculate the expected distance between two random points on different faces of the unit cube. Hint: This can be expressed in terms of integrals as

$$
\begin{aligned}
E_3 :=& \frac{4}{5}\int_0^1\int_0^1\int_0^1\int_0^1 \sqrt{x^2+y^2+(z-w)^2}\,dw\,dx\,dy\,dz \\
&+\frac{1}{5}\int_0^1\int_0^1\int_0^1\int_0^1 \sqrt{1+(y-u)^2+(z-w)^2}\,du\,dw\,dy\,dz.
\end{aligned}
\tag{1}
$$

Our method to find $E_3$ is to calculate the distribution functions for the two expressions under the square roots and then calculate the expectations. With some coaching, *Maple 10* is able to do the evaluations. Any higher moments can be calculated, when the distributions are known. The *Maple* worksheet `surfacedist.mw` is available at `www.math.kth.se/~johanph`.

## 2. Notation and formulation.

We start by calculating the distribution function $F(s) = \mathrm{Prob}(X^2 + Y^2 \le s)$ where $(X, Y)$ is evenly distributed in a unit square and the distribution function $G(t) = \mathrm{Prob}((Z - W)^2 \le t)$ for $(Z, W)$ evenly distributed in a unit square. Then, the corresponding densities $f$ and

*Key words and phrases.* Distance on unit cube.

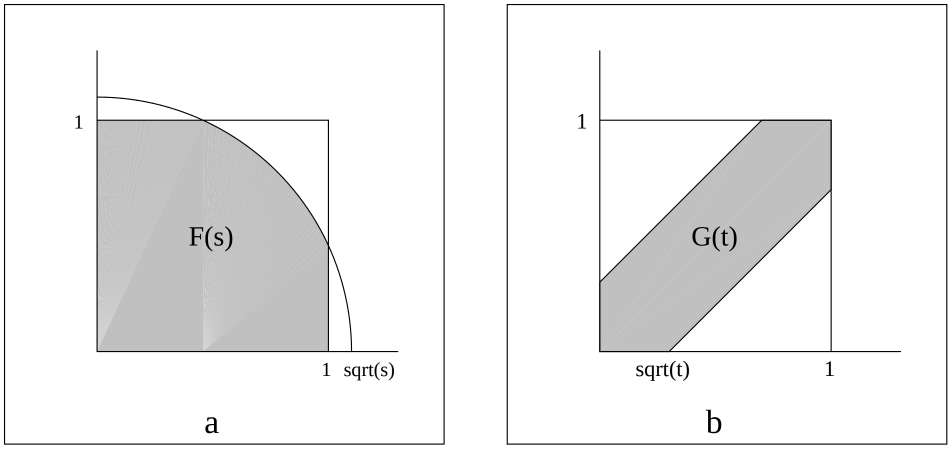


FIGURE 1. The areas $F(s)$ and $G(t)$.

$g$ are convolved to form the density $h$ of the whole expression under the first square root. Convolving $g$ by itself gives the density function $k$ for the random part under the second square root.

## 3. THE DISTRIBUTIONS $F$ AND $G$

The probability that $X^2 + Y^2 \leq s$ equals the area in Fig. 1a. that is the intersection between the disk with radius $= \sqrt{s}$ and the unit square. We get

$$F(s) = \begin{cases} \frac{1}{4}\pi s, & 0 < s \leq 1\ ; \\ s(\frac{1}{4}\pi - \arccos\frac{1}{\sqrt{s}}) + \sqrt{s-1}, & 1 < s \leq 2; \\ 1 & 2 < s. \end{cases} \tag{2}$$

The probability that $(Z - W)^2 \leq t$ equals the area of the diagonal strip in Fig. 1b.

$$G(t) = \begin{cases} 1 - (1 - \sqrt{t})^2, & 0 < t \leq 1\ ; \\ 1 & 1 < t. \end{cases} \tag{3}$$

Differentiating, we obtain the corrsponding densities

$$f(s) = \begin{cases} \frac{1}{4}\pi, & 0 < s \leq 1\ ; \\ \frac{1}{4}\pi - \arccos\frac{1}{\sqrt{s}}, & 1 < s \leq 2, \end{cases} \tag{4}$$

and

$$g(t) = \frac{1}{\sqrt{t}} - 1, \qquad 0 < t \leq 1\ . \tag{5}$$

## 4. The first square root

The probability density for $X^2+Y^2+(Z-W)^2 \le u$ is the convolution $h$ of $f$ and $g$

$$h(u) = \int g(u-s)\, f(s)\, ds.$$

*Maple* evaluates the the two integrals involved directly. After some manual simplifications, we get

$$h(u) = \begin{cases} \frac{1}{2}\pi\sqrt{u} - \frac{1}{4}\pi u, & 0 < u \le 1\ ; \\ \frac{5}{4}\pi - \pi\sqrt{u} - \sqrt{u-1} + u\arctan\sqrt{u-1}, & 1 < u \le 2\ ; \\ -1 + \sqrt{u}(\frac{\pi}{2} - 3\arcsin(\frac{1}{u-1})) + \sqrt{u-2} & \\ +(u+5)(\frac{\pi}{4} - \arctan(\sqrt{u-2}), & 2 < u \le 3. \end{cases} \tag{6}$$

The expectation of the first square root is

$$E_1 = \int_0^3 \sqrt{u}\, h(u)\, du.$$

Told to use the substitution $u = v^2$, *Maple* evaluates this integral to

$$\begin{aligned} E_1 = -\frac{11}{180}\pi + \frac{7}{30}\sqrt{2} - \frac{1}{30}\sqrt{3} - \frac{8}{15}\ln 2 + \frac{1}{10}\ln(1+\sqrt{2}) \\ +\frac{16}{15}\ln(1+\sqrt{3}) \approx .8707768237. \end{aligned} \tag{7}$$

## 5. The second square root

The probability density for $(Y-U)^2+(Z-W)^2 \le u$ is the convolution $k$ of $g$ by itself

$$k(u) = \int g(u-t)\, g(t)\, dt.$$

*Maple* gives the following distribution

$$k(u) = \begin{cases} \pi + +u - 4\sqrt{u}, & 0 < u \le 1\ ; \\ -2 + \pi - u + 4\sqrt{u-1} - 4\arctan\sqrt{u-1}, & 1 < u \le 2\ . \end{cases} \tag{8}$$

The expectation of the second square root is

$$E_2 = \int_0^2 \sqrt{1+u}\, k(u)\, du.$$

*Maple* evaluates this integral directly to

$$\begin{aligned} E_2 = \frac{4}{15} - \frac{2}{9}\pi + \frac{1}{5}\sqrt{2} - \frac{4}{15}\sqrt{3} - \frac{2}{3}\ln 2 + \ln(1+\sqrt{2}) \\ +\frac{4}{3}\ln(1+\sqrt{3}) \approx 1.148842981. \end{aligned} \tag{9}$$

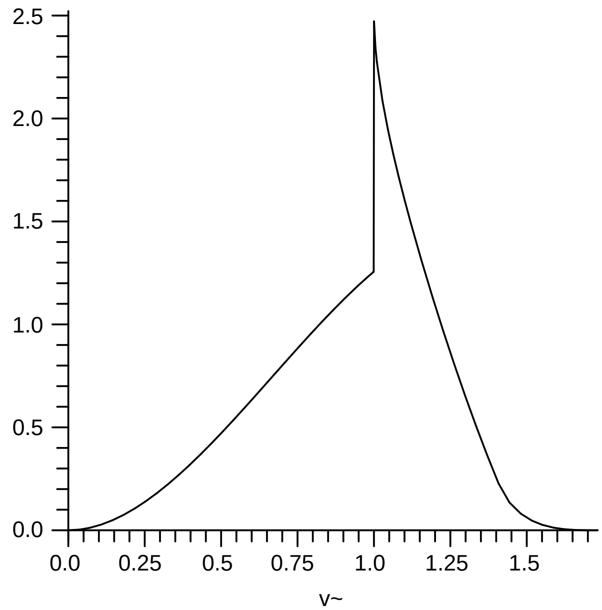


FIGURE 2. The combined density function in 3D.

## 6. The combined distribution.

Let $V$ be the distance between two random points on different faces of a unit cube. We get the combined probability density $c_3(v)$ for $V \leq v$ as

$$c_3(v) = \frac{4}{5}h(v^2) + \frac{1}{5}k(v^2 - 1). \tag{10}$$

The density $c_3(v)$ is shown in Figure 2.

Combining $E_1$ and $E_2$, we get the expection $E(V) = E_3$ in (1)

$$\begin{aligned} E(V) =& \frac{4}{5}\,E_1 + \frac{1}{5}\,E_2 \\ =& \frac{4}{75} - \frac{7}{75}\pi + \frac{17}{75}\sqrt{2} - \frac{2}{25}\sqrt{3} - \frac{14}{25}\ln 2 + \frac{7}{25}\ln{(1+\sqrt{2})} \\ &+ \frac{28}{25}\ln{(1+\sqrt{3})} \approx .9263900552. \end{aligned} \tag{11}$$

## 7. The four-dimensional problem.

The are eight three-dimensional “faces” of the unit cube in 4D and each such “face” is a unit cube in 3D. A particular “face” has one opposite “face” and six adjacent “faces”.

The four-dimensional distance between points in adjacent "faces" will consist of two distances in different direction with density $f$ and two distances between points in the same direction, each with density $g$. The combined density is $l = f * g * g = h * g$. The random part of the distance between points in opposite "faces" will consist of three distances between points in the same direction each with density $g$. The combined density is $m = g * g * g = k * g$.

The calculation of of these convolution integrals is considerably more complicated than those in three dimensions. *Maple* requires help with splitting the expressions into parts that must be handled with different methods. At many stages, manual simplifications are required for continuing the calculations. The resulting expressions contain unsolvable integrals. The *Maple* worksheets `L.mw, M.mw,` and `ELM.mw` describe the calculations and are available at `www.math.kth.se/~johanph`. We get

(12)

$$l(u) = \begin{cases} l_1 = \frac{1}{4}\pi^2 u - \frac{2}{3}\pi u^{3/2} + \frac{1}{8}\pi u^2, & 0 < u \le 1 \ , \\ l_2 = -\frac{1}{8}\pi(-6 + 2u(10-\pi) + u^2) & \\ +\frac{1}{6}(-2 + 5u + (8+4u)\pi)\sqrt{u-1} & \\ +\frac{4}{3}\pi u^{3/2} - \frac{1}{2}u(4\pi + u)\operatorname{arcsec}(\sqrt{u}), & 1 < u \le 2 \ , \\ l_3 = -\frac{2}{3} + \frac{13}{4}\pi + u - \frac{1}{2}\pi^2 u - \frac{1}{8}\pi u^2 & \\ -\frac{4}{3}\pi(2+u)\sqrt{u-1} - \frac{5}{3}(1+u)\sqrt{u-2} & \\ +u\pi \operatorname{arcsec}(\sqrt{u}) + (\frac{4}{3}\pi - 4\operatorname{arcsec}(u-1))u^{3/2} & \\ +(-3 + 10u + u^2)\operatorname{arcsec}(\sqrt{u-1}) & \\ +\frac{3}{2}u\int_2^u \frac{\operatorname{arcsec}(t-1)}{\sqrt{u-t}\sqrt{t}}\,dt, & 2 < u \le 3 \ , \\ l_4 = -\frac{4}{3} + \frac{5}{2}\pi + \frac{1}{2}(-2 + 5\pi - \pi^2)u + \frac{1}{8}\pi u^2 & \\ +4(\operatorname{arcsec}(u-2) - \frac{1}{3}\pi)(2+u)\sqrt{u-1} & \\ +(2 + \frac{5}{6}u)\sqrt{u-3} & \\ +\pi u(\operatorname{arcsec}(\sqrt{u}) + \operatorname{arcsec}(\sqrt{u/3})) & \\ -(10 + 10u + \frac{1}{2}u^2)\operatorname{arcsec}(\sqrt{u-2}) & \\ +\frac{3}{2}u\int_{u-1}^3 \frac{\operatorname{arcsec}(t-1)}{\sqrt{u-t}\sqrt{t}}\,dt, & 3 < u \le 4 \ . \end{cases}$$

The density $m$ can very well be calculated as $k * g$. We will use that it equals the density for the distance between two random points in a unit cube. We copy this density from [4]. Remembering that $u$ is the square of the distance, we have

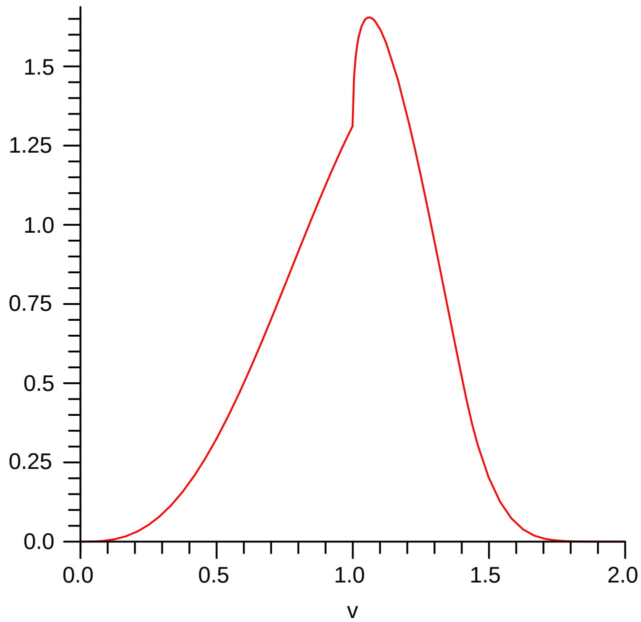


FIGURE 3. The combined density function in 4D.

$$m(u) = \begin{cases} m_1 = 2\pi\sqrt{u} - 3\pi u + 4u^{3/2} - \frac{1}{2}u^2, & 0 < u \le 1 \ , \\ m_2 = -\frac{1}{2} + 3\pi - 4\pi\sqrt{u} + u^2 & \\ +3u(1 + 4\,\mathrm{arcsec}(\sqrt{u})) - 4(1+2u)\sqrt{u-1}, & 1 < u \le 2 \ , \\ m_3 = \frac{1}{2}(-5 + 6\pi - u)(1+u) - 4\pi\sqrt{u} & \\ +4(1+u)\sqrt{u-2} + 12\sqrt{u}\ \mathrm{arcsec}(u-1) & \\ -12(1+u)\,\mathrm{arcsec}(\sqrt{u-1}), & 2 < u \le 3 \ . \end{cases} \tag{13}$$

The combined probability density for the distance $v$ between two random points on different "faces" of a cube in 4D is

$$c_4(v) = \frac{6}{7}\ l(v^2) + \frac{1}{7}\ m(v^2 - 1). \tag{14}$$

The density $c_4(v)$ is shown in Figure 3.

## 8. The average distance in 4D.

In analogy with the 3D case, we have

$$E_4 = \frac{6}{7}\ E_l + \frac{1}{7}\ E_m, \tag{15}$$

where

$$E_l = \int_0^4 \sqrt{u}\; l(u)\, du \quad \text{and} \quad E_m = \int_0^3 \sqrt{u+1}\; m(u)\, du. \tag{16}$$

The evaluation of $E_m$ is straightforward but the intgration in $E_l$ is hard to carry out. We rewrite $E_l$ as

$$E_l = \int \sqrt{u} \int h(t)\, g(u-t)\, dt\, du, \tag{17}$$

and integrate over $u$ first. Thus, we form

$$p(t) = \int \sqrt{u}\; g(u-t)\, du = \int_0^1 \sqrt{s+t}\, g(s)\, ds. \tag{18}$$

This integration is easy to carry out and gives

$$p(t) = \sqrt{1+t} + \frac{2}{3}(t^{3/2} - (1+t)^{3/2}) - \frac{1}{2} t \log t + t \log(1+\sqrt{1+t}). \tag{19}$$

$p(t)$ is well defined and increasing on $(0, \infty)$, starting from $p(0) = \frac{1}{3}$, which is the average distance between two random points on the unit interval.

Then, we calculate $E_l$ as

$$E_l = \int_0^3 p(t)\, h(t)\, dt. \tag{20}$$

Both $E_l$ and $E_m$ are long espressions. See `ELM.mw`. We give there combination from (15):

(21)

$$\begin{aligned}
E_4 =& \frac{5}{3} - \frac{16}{245}\pi \\
&+ \frac{1}{5880}\Big[(876 + 4158\pi - 13212 \arcsin(\frac{1}{3}))\sqrt{2} + (576 - 12936\pi)\sqrt{3} \\
&- (3744 + 924\pi)\log 2 + (20178 + 273\pi)\log 3 \\
&+ (540 - 1932\pi)\log(1+\sqrt{2}) + (7488 + 1848\pi)\log(1+\sqrt{3})\Big] \\
&+ \frac{13}{70}\int_2^3 \frac{\operatorname{arcsec}(\sqrt{t-1})}{t\sqrt{1+t}}\, dt + \frac{66}{35}\int_2^3 \sqrt{1+t}\,\sqrt{t}\operatorname{arcsec}(t-1)\, dt \\
&\approx .9998558403.
\end{aligned}$$

## References


[1] D.H. Bailey and J.M. Borwein, *Experimental Mathematics: Examples, Methods and Implications,* Notices of the AMS, Vol. 53, Number 5, May 2005

[2] D.H. Bailey, J.M. Borwein, V Kapoor, and E.W. Weisstein *Ten Problems in Experimental Mathematics,* http://crd.lbl.gov/˜dhbailey/dhbpapers/tenproblems.pdf, March 8, 2006.

[3] J.M. Borwein, D.H. Bailey, and R. Girgensohn *Experimentation in Mathematics: Computational Paths to Discovery,* A. K. Peters, Natick, MA, 2004

[4] J. Philip, *The Probability distribution of the Distance between two Random Points in a Box*. TRITA MAT 07 MA 10, ISSN 1401-2278, ISRN KTH/R -07/01-SE, Dec. 2007, http://www.math.kth.se/~johanph/habc.pdf

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

*E-mail address*: johanph@kth.se