Tina Memo No. 2001-008
Internal Report

Derivation of the Renormalisation Formula for the Product of Uniform Probability Distributions and Extension to Non-Integer Dimensionality.

P. A. Bromiley, T.F. Cootes and N.A. Thacker

Last updated
5 / 11 / 2001

This document forms part of the **Statistics and Segmentation** Series (2008-001) available from www.tina-vision.net.

WWW.TINA-VISION.NET

Imaging Science and Biomedical Engineering,
School of Cancer and Imaging Sciences,
University of Manchester, Stopford Building,
Oxford Road, Manchester M13 9PT, U.K.

Derivation of the Renormalisation Formula for the Product of Uniform Probability Distributions and Extension to Non-Integer Dimensionality

P. A. Bromiley, T.F. Cootes and N.A. Thacker
Imaging Science and Biomedical Engineering,
School of Cancer and Imaging Sciences,
University of Manchester, Stopford Building,
Oxford Road, Manchester M13 9PT, U.K.
paul.bromiley@man.ac.uk

Abstract

A standard formula, given below, exists for renormalising (reflattening) a quantity that is the product of several quantities having uniform probability distributions. However, extensive literature searches have failed to reveal an acceptable derivation for the equation. Therefore, we present derivations using a variety of techniques: derivations for problems of fixed dimensionality using integration in the sample space defined by the problem, derivations in log space leading to a derivation for problems of arbitrary dimensionality, and a short derivation building on the ideas in the previous sections which is suitable for use in publications. Finally, we extend the original renormalisation formula to cope with non-integer values for the dimensionality.

1 Introduction

A standard technique exists to renormalise the probability distribution of the product of several random variates drawn from a uniform probability distribution. If n quantities ω, each having a uniform probability distribution, are multiplied to produce a product P,

$$P = \prod_{i}^{n} \omega_i, \tag{1}$$

then this product can be normalised to produce a new quantity P', which has a uniform probability distribution, using

$$P' = P \sum_{j=0}^{n-1} \frac{(-\ln P)^j}{j!}. \tag{2}$$

This process is potentially nestable, providing a simple yet statistically principled method for data fusion. Extensive literature searches have failed to reveal an acceptable proof or derivation. This document provides derivations for this equation, by direct integration for the 1D, 2D, 3D and 4D cases, and then via a geometrical method for the 2D and nD cases.

Equation 2 can be expanded to give

$$P' = P(1 - \ln P + \frac{(\ln P)^2}{2} - \frac{(\ln P)^3}{6} ....), \tag{3}$$

where the highest term needed of those inside the brackets is determined by the dimensionality of the problem i.e. a 1D problem gives

$$P' = P, \tag{4}$$

a 2D problem gives

$$P' = P(1 - \ln P), \tag{5}$$

and so on. The proof for the 1D case is trivial: if P has a uniform distribution, then $P' = P$ also has a uniform distribution.

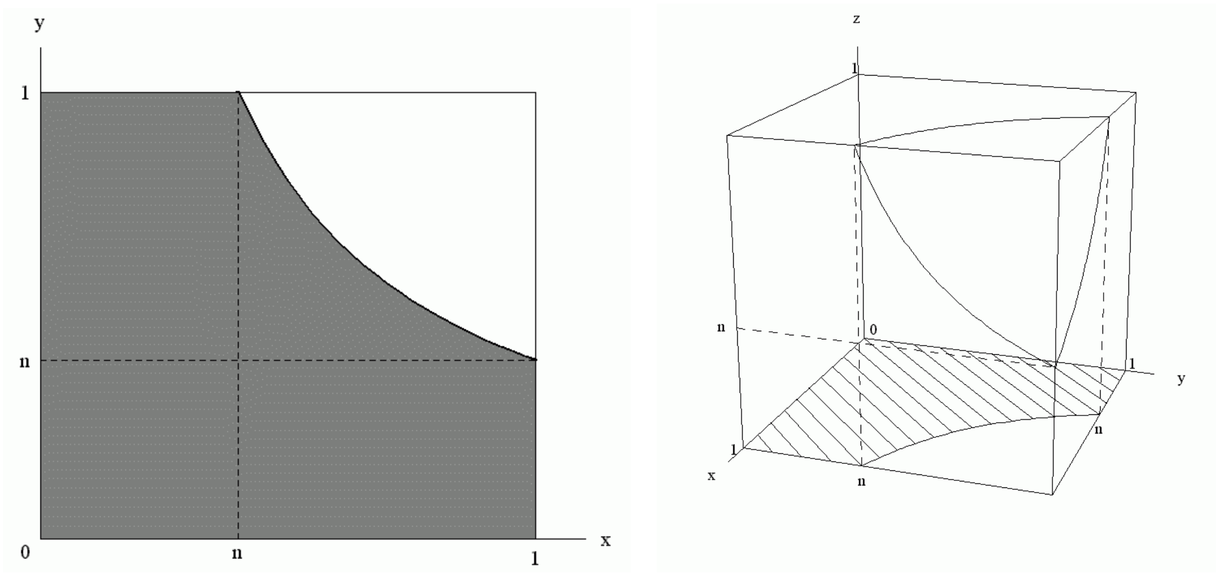

Figure 1: Sample spaces for the 2D and 3D cases.

2 Derivation by Integration in Sample Space

This section provides derivations for the renormalisation equation in problems of fixed dimensionality, for the 2, 3 and 4 dimensional cases, by integration in the sample space defined by the individual quantities ω_i.

2.1 The 2D Case

Let the individual quantities $\omega_{i=1,2}$ be called x and y. We can then plot the sample space shown in Fig. 1a, where x and y vary from 0 to 1. We can also plot a contour line of constant P, $xy = n$, where n is a constant also varying from 0 to 1. The transformation to P' such that P' has a uniform probability distribution can be achieved by replacing all P along the contour line with the integral of the area under the contour line (the shaded area). The result must have a uniform distribution, since 10% of the points lie in the lowest 10% of the space, 20% in the lowest 20% etc. This procedure scales the product $P = n$ by the number of ways of achieving $P = n$ or less. The results from the non-parametric images subtraction technique have a uniform probability distribution for the same reason. Note that this implicitly assumes no spatial correlation (c.f. the spatial correlation probe used with non-parametric images subtraction, based on renormalising the product of each pixel in the subtraction result with the four neighbouring pixels), since it assumes that the sample space is uniformly populated.

The integral can be conveniently divided into two regions, $x < n$ and $x > n$, giving

$$\int_0^n 1dx + \int_n^1 \frac{n}{x}dx = n - n\ln n, \tag{6}$$

or

$$P' = P(1 - \ln P), \tag{7}$$

which is the required 2D result (Eq. 5).

2.2 The 3D Case

Extending to three dimensions we can plot the three dimensional sample space shown in Fig. 1b. Again we want to integrate the area of the cube below the contour surface $xyz = n$. Again this can be conveniently divided into two regions, for $xy < n$ and $xy > n$. First inspect the shaded area of the xy plane, where $xy < n$. In this part of the integral z varies from 0 to 1, and this is the integral performed in the 2D case. This leaves the region of the cube where $xy > n$, so

$$volume = n(1 - \ln n) + \int_n^1 \int_{\frac{n}{y}}^1 \frac{n}{xy}dxdy. \tag{8}$$

Concentrating on the integral,

$$\int_{n}^{1}\int_{\frac{n}{y}}^{1}\frac{n}{xy}dxdy=\int_{n}^{1}-\frac{n}{y}\ln\frac{n}{y}dy \tag{9}$$

$$=\int_{n}^{1}-\frac{n}{y}\ln n+\frac{n}{y}\ln y dy. \tag{10}$$

Using

$$\int\frac{\ln\alpha}{\alpha}=\frac{1}{2}(\ln\alpha)^{2}, \tag{11}$$

this gives

$$=n(\ln n)^{2}-\frac{n}{2}(\ln n)^{2}=\frac{n}{2}(\ln n)^{2}. \tag{12}$$

Collecting terms,

$$volume=n(1-\ln n)+\frac{n}{2}(\ln n)^{2} \tag{13}$$

or

$$P'=P(1-\ln P+\frac{(\ln P)^{2}}{2}), \tag{14}$$

which, inspecting Eq. 3, is the required 3D result.

2.3 The 4D Case

The proof can be extended to 4D in the same way. Using a, b, c and d as the dimensions, the hypervolume below the contour $abcd=n$ is the 3D result plus the integral

$$\int_{n}^{1}\int_{\frac{n}{c}}^{1}\int_{\frac{n}{bc}}^{1}\frac{n}{abc}dadbdc \tag{15}$$

$$=\int_{n}^{1}\int_{\frac{n}{c}}^{1}-\frac{n}{bc}\ln(\frac{n}{bc})dbdc \tag{16}$$

$$=\int_{n}^{1}\int_{\frac{n}{c}}^{1}-\frac{n}{bc}\ln n+\frac{n}{bc}\ln b+\frac{n}{bc}\ln c dbdc \tag{17}$$

$$=\int_{n}^{1}\frac{n\ln n}{c}\ln(\frac{n}{c})-\frac{n}{2c}(\ln(\frac{n}{c}))^{2}-\frac{n\ln c}{c}\ln(\frac{n}{c})dc \tag{18}$$

$$=\int_{n}^{1}\frac{n(\ln n)^{2}}{2c}-\frac{n\ln n\ln c}{c}+\frac{n(\ln c)^{2}}{2c}dc \tag{19}$$

$$=-\frac{n(\ln n)^{2}\ln n}{2}+\frac{n\ln n(\ln n)^{2}}{2}-\frac{n}{2}\frac{1}{3}(\ln n)^{3}=-\frac{n}{6}(\ln n)^{3}, \tag{20}$$

which is the extra term required in Eq. 2 for the 4D case.

3 Derivation by Integration in Log Space

In the above derivations the highest order term emerges last from the iterated integrals. This implies that a different technique is required to derive the equation in the general, jD case[1], and also indicates that the correct approach might be a geometrical one. Therefore, we will first derive the 2D result using a geometrical method, and then extend the same method to jD.

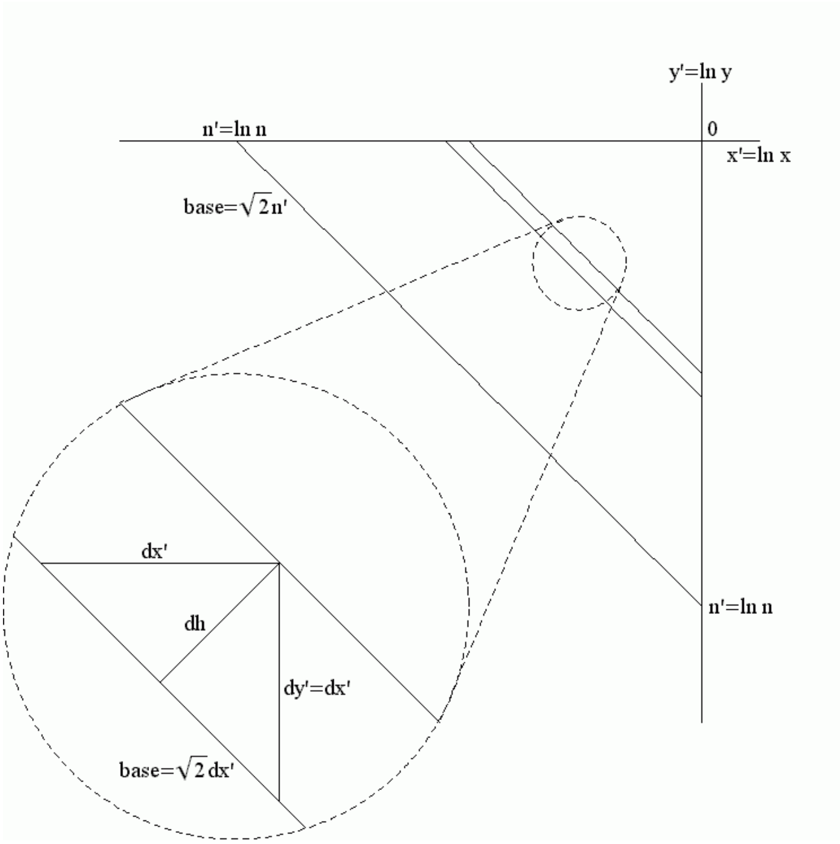

Figure 2: The 2D case in log space. The curve of $y = \frac{n}{x}$ is linear in log space, and thus encloses a right isosceles simplex. The shaded area shows the element of integration. This will in general be an equilateral simplex in (j-1)D: a line in this case. The expanded region shows the construction used to obtain dx' from dh, again consisting of a right isosceles simplex.

3.1 The 2D Case

The basis of this approach is the same as in the previous sections: to integrate the area under the curve $xy = n$, bounded by a unit cube, and then replace all probabilities lying along that curve with the result of the integration. Transforming into log space flattens this curve, since

$$y = \frac{n}{x} \tag{21}$$

$$\ln y = \ln(\frac{n}{x}) = \ln n - \ln x \tag{22}$$

Let

$$\ln x = x' \tag{23}$$

$$\ln y = y' \tag{24}$$

$$\ln n = n' \tag{25}$$

Therefore,

$$y' = n' - x' \tag{26}$$

This equation is linear: in fact, the form in jD log space is the corner form of a hypercube (i.e. an jD simplex with one corner consisting entirely of edges meeting at right angles), and so is a right isosceles triangle in the 2D case, as shown in Fig. 2. This is convenient since a standard result exists for the hypervolume of such a form (see Appendix). If the side length is a, then

$$hypervolume = \frac{(a)^j}{j!} \tag{27}$$

[1]The symbol j is being used to indicate arbitrary dimensionality, since the more usual symbol n has already been used as the product of the uniform probabilities.

In the two dimensional case this reduces to the familiar equation for the area of a triangle (see Appendix A),

$$area = \frac{1}{2}(a)^2 \tag{28}$$

i.e. the area is one half multiplied by the base length multiplied by the height. Of course, this form lies above and not below the curve, but since the integration is constrained within the unit hypercube, we can perform the integration of the area above the curve using the standard result and then subtract the result from one to give the area under the curve.

However, the transformation to log space introduces an additional complexity by changing the data density. The data density in the original space is uniform, since the quantity shown on each axis in a uniform probability distribution. The transformation into log space will result in non-uniform data density, and so we must find an expression for the data density at each point in the log space. If the transformation is given by

$$x = e^{x'} \tag{29}$$

$$y = e^{y'} \tag{30}$$

then we can consider some unit of area at coordinates x', y' with sides of length $\delta x'$ and $\delta y'$, which has an area in log space of $\delta x'\delta y'$. In normal space the area of this element is given by

$$(e^{x'+\delta x'} - e^{x'})(e^{y'+\delta y'} - e^{y'}) = e^{x'}e^{y'}(e^{\delta x'} - 1)(e^{\delta y'} - 1) \tag{31}$$

Now, the exponential function has the series expansion

$$e^{\alpha} = 1 + \alpha + \frac{\alpha^2}{2} + \frac{\alpha^3}{3} + ... \tag{32}$$

Assuming that $\delta x'^2$ and higher powers are negligible, this gives a normal space volume of

$$e^{x'}e^{y'}(1 + \delta x - 1)(1 + \delta y - 1) = e^{x'}e^{y'}(\delta x \delta y) \tag{33}$$

and so the data density of a unit area in log space is $e^{x'}e^{y'}$ times the data density in the unit area in normal space. Furthermore, the data density in log space is constant along any line for which this expression is constant, so

$$e^{x'}e^{y'} = constant = e^{(x'+y')} = xy \tag{34}$$

or

$$constant = x' + y' \tag{35}$$

Therefore the surface which forms the base of the corner form in jD i.e. the surface which bounds the area we want to integrate, is also a surface of constant data density, and so we can easily integrate the area by splitting it into elements parallel to this base i.e. elements perpendicular to the height.

The corner form of a hypercube in j dimensions has one vertex consisting entirely of edges meeting at right angles: this is the vertex that was originally the corner of the hypercube. The base, i.e. the $(j-1)$ dimensional form which contains all of the vertices except this one, is an equilateral $(j-1)$ dimensional simplex i.e. all sides of this $(j-1)$ dimensional form are equal. This is difficult to see in 2D, since the base is then one dimensional i.e. a line (or being needlessly technical a digon). It is easier to picture in 3D: the base of a right isosceles tetrahedron is an equilateral triangle. So, the jD right isosceles simplex which is the corner form of a hypercube can be decomposed into elements normal to its height which take the form of $(j-1)D$ equilateral simplexes. We can then find the volume of the corner form by integrating these elements, suitably weighted by the data density term. However, Eq. 27 applies only to the right isosceles type forms. Equilateral forms must have a smaller hypervolume (this can easily be seen by considering 2D triangles). At this point we introduce a second standard geometrical result: the hypervolume of an jD regular (equilateral-type) simplex of side length a is given by (see Appendix A)

$$hypervolume = \frac{a^j}{j!}\sqrt{\frac{(j+1)}{2^j}} \tag{36}$$

Now, if the distance from the right vertex of the original corner form to the position of the equilateral simplex element along one of the axes is x, then the length of the sides of the equilateral simplex is $\sqrt{2}x$ (see Fig. 2). This will always be true in any number of dimensions, since we can always pick any two of those dimensions and form a right isosceles triangle, the two perpendicular sides of which will have one vertex the original hypercube corner vertex, and the hypotenuse of which will be one of the sides of the equilateral simplex used as the element of integration. So, in 2D the area of the (1D) element will be

$$\frac{(\sqrt{2}x')^1}{1!}\sqrt{\frac{2}{2}} = \sqrt{2}x' \tag{37}$$

In fact, closer consideration of this equation shows that there is one more complexity. In log space we will be integrating these elements in the negative x direction, since the quantities plotted along the axes of our original space are probabilities (i.e. lie between 0 and 1) and the log of a number less than 1 is negative. However, the area we want to find remains a positive quantity: if we calculate it as negative quantity, then the answer returned for the renormalised probability will be negative and thus clearly unphysical. Therefore, the variable in Eq. 37 should be $-x$ rather than x.

There is only one more consideration before the integration can be performed. It might appear that, since we are integrating elements perpendicular to the height, then the natural direction to integrate along is the direction of the height vector. However, this would involve manipulation of the equation for the element of integration to rotate into the new coordinate system. The alternative is to integrate with respect to one of the existing axes of the log space, with suitable consideration of the relationship of an element of height, δh to an element of x', $\delta x'$. Clearly the two manipulations will be identical, and so neither has an advantage: here we use the original log space axis. The derivation of an expression linking $\delta x'$ to δh requires a third standard result from geometry: the hypervolume of any triangular-faced figure in j dimensions is given by (see Appendix A)

$$hypervolume = \frac{1}{j}bh \tag{38}$$

where b is the area/volume/hypervolume of the base and h is the height. In fact this holds for any pyramid, regardless of the form of the base, but we will only use it for triangular faced figures. In 2D it reduces to the familiar formula for the area of a triangle: one half multiplied by the base multiplied by the height. Inspecting Fig. 2, we can construct a right isosceles triangle where the perpendicular sides are of length $\delta x'$, and the height is of length δh. In this 2D case it is easy to relate δh to $\delta x'$ using Pythagoras' Theorem, giving

$$\delta h = \frac{\delta x'}{\sqrt{2}} \tag{39}$$

but in general this will be more difficult. Therefore, we can again use Eqs. 27, 36 and 39: if the area of the right isosceles form defined by $\delta x'$ and $\delta y'$ is given by Eq. 27, this can be equated to the area in terms of the base, an equilateral form whose area is given by Eq. 39, and the height, $\delta h'$, using Eq. 36. This will be explored in detail in the next section.

We now have all three terms required to construct an integral. These will be referred to as the simplex term, giving the base of the corner form, the density term, giving the data density along the simplex, and the element term, translating δh into $\delta x'$. Only one small point remains: since the data density is constant along the element of integration, it can be equated to the data density at the point $y' = 0$ on the element, and so is simplified to $e^{x'}$. The integral that must be performed for the 2D case is therefore

$$\int_{\ln n}^{0} e^{x'}\sqrt{2}(-x')\frac{1}{\sqrt{2}}dx' = \int_{\ln n}^{0} (-x')e^{x'}dx' \tag{40}$$

This integral responds to treatment by parts. Let

$$u = -x' \Rightarrow du = -dx' \tag{41}$$

and

$$dv = e^{x'}dx' \Rightarrow v = e^{x'} \tag{42}$$

giving

$$\int_{n'}^{0} -x'e^{x'}dx' = uv - \int vdu = -x'e^{x'}\Big|_{n'}^{0} + \int_{n'}^{0} e^{x'}dx' \tag{43}$$

$$= -x'e^{x'}\Big|_{n'}^{0} + e^{x'}\Big|_{n'}^{0} = n'e^{n'} - e^{n'} + 1 \tag{44}$$

Remembering that the area we want to find is one minus this area,

$$1 - n'e^{n'} + e^{n'} + 1 = e^{n'} - n'e^{n'} \tag{45}$$

Then, substituting for $n' = lnn$, and then for $P = n$ as in the previous section, we obtain for the probability of any point P' along the curve $xy = n$

$$P' = e^{\ln n} - \ln n e^{\ln n} = n - n \ln n = P(1 - \ln P) \tag{46}$$

which is the desired result for the 2D case.

3.2 The jD Case

Unlike the derivation by direct integration in normal space, the derivation by integration along one axis in log space using geometrical arguments can be generalised to arbitrary numbers of dimensions easily. We must deal with three terms, identified in the previous section as the simplex term, the data density term and the element term.

3.2.1 The Data Density Term

The generalisation of the data density term is trivial: it is obvious given the rules of logarithms and the arguments in the previous section, but we will go through it in detail for the sake of completeness. So, if we have j dimensions $x'_{i=1...j}$,

$$x_i = e^{x'_i} \tag{47}$$

The volume in normal space of a hypercube element in log space at coordinates $x'_1 ... x'_j$ with side lengths $\delta x'_1 ... \delta x'_j$ is given by

$$\prod_{i=1}^{j} (e^{x'_i + \delta x'_i} - e^{x'_i}) = \prod_{i=1}^{j} e^{x'_i}(e^{\delta x'_i} - 1) \tag{48}$$

Using the series expansion of the exponential function given in Eq. 32, and assuming that any term in $(\delta x'_i)^2$ or higher is negligible, this gives

$$\prod_{i=1}^{j} e^{x'_i}(1 + \delta x'_i - 1) = \prod_{i=1}^{j} e^{x'_i} \prod_{i=1}^{j} \delta x_i \tag{49}$$

The second product on the R.H.S. is the volume of the element in normal space: the first is the required scaling for the data density. As in the previous section, since we will be integrating along one axis (which we are at liberty to choose: choose x_1), and since $e^a e^b = e^{a+b} = e^a$ if $b = 0$, we obtain

$$e^{x'_1} \tag{50}$$

as the required data density term for the jD case.

3.2.2 The Simplex Term

The form that we are integrating is the form obtained by chopping the corner off of a jD hypercube, such that the distance between the remaining vertex of the original hypercube and all of the other vertices, i.e. the distance between the remaining hypercube vertex and the points of intersection of the chopping plane with the edges of the hypercube, are all equal. Thus we might call the shape a right isosceles simplex by analogy to the right isosceles triangle obtained in the 2D case. The shape therefore has two types of side: those for which one vertex is the original hypercube vertex that was retained, and sides for which neither vertex is this vertex. Sides of the first kind have length $\ln n = n'$. Sides of the second type form the hypotenuse of a right isosceles triangle in some pair of the j dimensions, and the other two sides of this projected triangle will be sides of the first type. Therefore, sides of the second type have length $\sqrt{2}n'$.

The form of the hypervolume used as the element of integration is a $(j-1)D$ simplex consisting entirely of type 2 sides, and so we might call this an equilateral simplex in analogy to the equilateral triangle we would obtain in the 3D case. Using Eq. 36 we obtain

$$\frac{(\sqrt{2}(-x'_1)^{(j-1)}}{(j-1)!}\sqrt{\frac{(j-1)+1}{2^{(j-1)}}} = \sqrt{2}^{j-1}\frac{(-x'_1)^{j-1}}{(j-1)!}\sqrt{\frac{(j-1)+1}{2^{(j-1)}}} \tag{51}$$

as the simplex term in jD, using $-x'_1$ as the variable since we will be integrating along the negative x'_1 direction (the result will be invariant under change of this variable, so we could pick any of the dimensions to integrate along, but getting the signs right is essential for the reasons outlined in section 5).

3.2.3 The Element Term

Referring to Fig. 2, it can be seen that a jD hypercube corner form can again be constructed to find the relationship between $\delta x_1'$ and δh. Again the base of this figure is an equilateral type $(j-1)D$ simplex with side length $\sqrt{2}\delta x_1'$. As stated in Section 5, we can express $\delta x_1'$ in terms of δh using Eqs. 27, 36 and 39. Firstly, note that the hypercube corner form has two types of side: those for which one vertex is the vertex of the original hypercube, and which have length $\delta x_1'$, and those for which neither vertex is the original hypercube vertex, and which have length $\sqrt{2}\delta x_1'$. The volume of the hypercube corner form can be expressed using both Eqs. 27 and 39: equating these and solving for the height gives

$$\frac{\delta x_1'^{\,j}}{j!} = \frac{1}{j} bh \Rightarrow h = \frac{\delta x_1'^{\,j}}{(j-1)!}\frac{1}{b} \tag{52}$$

The base is an equilateral type simplex in (j-1)D, consisting entirely of sides of the second type i.e. with length $\sqrt{2}\delta x_1'$, and so its hypervolume can be expressed using Eq. 36,

$$b = \frac{(\sqrt{2}\delta x_1')^{j-1}}{(j-1)!}\sqrt{\frac{(j-1)+1}{2^{(j-1)}}} \tag{53}$$

Substituting Eq. 54 into Eq. 53 gives the element term,

$$\delta h = \frac{\delta x_1'}{\sqrt{2}^{j-1}\sqrt{\frac{(j-1)+1}{2^{j-1}}}} \tag{54}$$

The interesting point here is that, comparing the element term Eq. 54 with the simplex term Eq. 52, most of the numerical quantities will cancel out in the product of these terms, leaving

$$simplex \times element = \frac{(-x_1')^{(j-1)}}{(j-1)!}\delta x_1' \tag{55}$$

This is expected since the construction of the element term is almost the inverse of the construction of the simplex term, and simplifies the expression for the integration (although it does not simplify the integration itself, since none of the cancelled terms involve the variable of integration).

3.2.4 The Integration

Collecting the three terms, we have an integration for the hypervolume under the curve in the jD case of,

$$1 - \int_{\ln n}^{0} e^{x_1'}\sqrt{2}^{j-1}\frac{(-x_1')^{j-1}}{(j-1)!}\sqrt{\frac{(j-1)+1}{2^{(j-1)}}}\frac{dx_1'}{\sqrt{2}^{j-1}\sqrt{\frac{(j-1)+1}{2^{j-1}}}} \tag{56}$$

$$= 1 - \int_{\ln n}^{0} e^{x_1'}\frac{(-x_1')^{(j-1)}}{(j-1)!}dx_1' \tag{57}$$

This integration responds to treatment by parts. Let

$$u = \frac{(-x_1')^{(j-1)}}{(j-1)!} \Rightarrow du = -\frac{(-x_1')^{(j-2)}}{(j-2)!}dx_1' \tag{58}$$

and

$$dv = e^{x_1'}dx_1' \Rightarrow v = e^{x_1'} \tag{59}$$

so the result for the hypervolume is

$$1 - e^{x_1'}\frac{(-x_1')^{(j-1)}}{(j-1)!}\bigg|_{\ln n}^{0} + \int_{\ln n}^{0} e^{x_1'}\frac{(-x_1')^{(j-2)}}{(j-2)!}dx_1' \tag{60}$$

This expression also contains an integration which can be performed by parts in the same way, and thus the sequence is clear: the integrations will continue to lower the power to which x_1' is raised by one in each of the sequence of terms, and also lower the highest term in the factorial in the denominator by 1. The signs require some consideration: every term will be negative before substitution of the limits. However, every term except the last will also contain a power of x, and so the upper limit of 0 will have no effect. Since substituting the lower limit reverses the sign of the term, every term will become positive. The upper limit on the integration will only produce a term when applied to the last term in the sequence, in which case it will produce -1 (due to the term originally being negative), which removes the 1 in the equations above. Therefore, once the negative sign in front of the $-x'_1$ in the simplex term has been considered, all the terms in odd powers of x'_1 will be negative, and all the terms in even powers of x'_1 will be positive. Every term in the sequence will contain $e^{x'_1}$. Therefore, the final sequence after all the integrations have been performed, and substituting for $n' = \ln n$, will be

$$\sum_{i=0}^{j-1} e^{\ln n} \frac{(-\ln n)^i}{i!} = \sum_{i=0}^{j-1} n \frac{-\ln n^i}{i!} \tag{61}$$

Substituting for $n = P$ and equating the result to P', the renormalised probability, gives

$$P' = \sum_{i=0}^{j-1} P \frac{(-\ln P)^i}{i!} \tag{62}$$

QED.

4 Derivation using Iterative Techniques

The above sections clearly illustrate the techniques applied in the derivation of the renormalisation equation. A shorter derivation, based on an iterative technique, is also possible. This does not illustrate the problem as clearly, but is more suitable for inclusion in publications.

4.1 The nD Case

Given n quantities each having a uniform probability distribution $p_{i=1,n}$, the product $p = \prod_{i=1}^{n} p_i$ can be renormalised to have a uniform probability distribution $F_n(p)$ using

$$F_n(p) = p \sum_{i=0}^{n-1} \frac{(-\ln p)^i}{i!} = p + p \sum_{i=1}^{n-1} \frac{(-\ln p)^i}{i!} \tag{63}$$

The quantities p_i can be plotted on the axes of an n dimensional sample space, bounded by the unit hypercube. Since they are uniform, and assuming no spatial correlation, the sample space will be uniformly populated. Therfore, the transformation to $F_n(p)$ such that this quantity has a uniform probability distribution can be achieved using the probability integral transform, replacing any point in the sample space p with the integral of the volume under the contour of constant p passing through this point, which obeys $\prod_{i=1}^{n} p_i = p = constant$. This can be expressed in terms of the volume of a hyper-region of one lower dimension by integrating over one dimension (let this be called x)

$$F_n(p) = p + \int_p^1 F_{n-1}(\frac{p}{x}) dx \tag{64}$$

This is equivalent to dividing the integration into two regions using a plane perpendicular to the x axis which intersects the axis at $x = p$. Fig. 3 shows the element of integration that would be used in the 3D case, to relate the volume of the unit cube under the contour of constant probability to the 2D case.

Now, in the simplest case of $n = 1$, clearly $F_n(p) = p$, as no renormalisation is required. The solution for higher dimensions can then be derived by iterative application of Eq. 13. This involves integration of terms in $(p/x)[-\ln(p/x)]^n$ which enter in the n=3 and higher cases. This integration can be performed using a simple substitution $x = pu$, $dx = pdu$

$$\int_p^1 (\frac{p}{x})[-\ln(\frac{p}{x})]^n dx = p \int_1^{1/p} (\frac{1}{u})[\ln u]^n du \tag{65}$$

$$= p[\frac{1}{n+1}[\ln u]^{n+1}]^{n+1} = \frac{p}{n+1}[-\ln p]^{n+1} \tag{66}$$

Iterative application of Eq. 13 therefore produces the series

$$F_n(p) = p - p \ln p + p\frac{(\ln p)^2}{2} - p\frac{(\ln p)^3}{6} + p\frac{(\ln p)^4}{24} .... \tag{67}$$

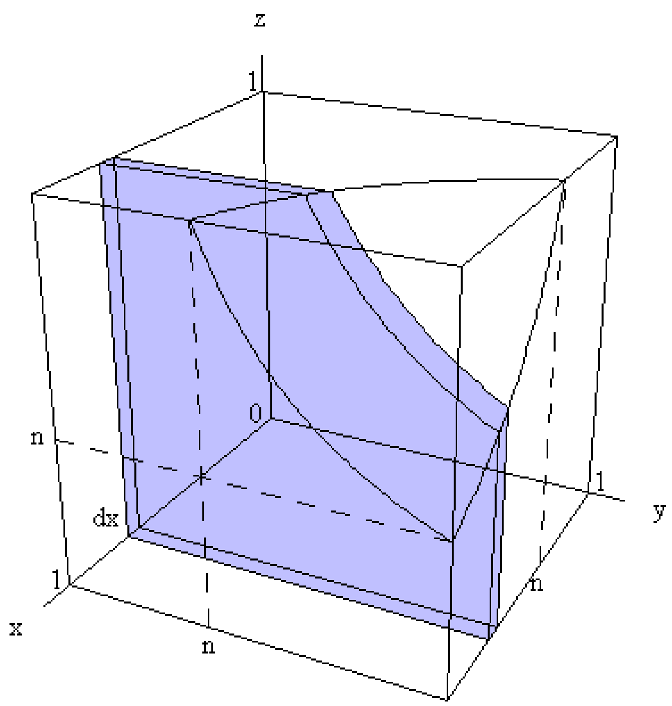

Figure 3: The sample space for the probability renormalisation in 3D, showing the element of integration (the shaded region) used to relate this to the 2D problem. The contour of constant probability is shown by the curved surface in the upper corner of the unit cube.

which can be written as

$$F_n(p) = p \sum_{i=0}^{n-1} \frac{(-\ln p)^i}{i!} \tag{68}$$

QED.

5 Extension to Non-Integer Dimensionality

The renormalisation formula has been applied in conjunction with the non-parametric image subtraction technique [2, 3, 4, 5, 6] to produce a spatial correlation analysis method. The renormalisation formula assumes no spatial correlation in p, since the derivation assumes that the sample space defined by the original quantities multiplied to produce p is uniformly populated. The non-parametric image subtraction technique produces a difference image with a uniform probability distribution, with randomly distributed values for background pixels and spatially correlated regions of low values for localised difference regions. A new image can be produced by taking the product of each pixel value in the non-parametric image subtraction result with the values of the four nearest pixels. Since this is equivalent to the product of five quantities each having a uniform probability distribution, the result can be renormalised using the expression given above. Pixels that do not form part of a localised difference region will reflatten correctly. Pixels that do form part of a localised difference region will form spatially correlated clusters of low probability pixels in the difference image, and so will form very low probability products that will not reflatten correctly. The probability distribution for the reflattened image will therefore feature a uniform distribution for background pixels, together with a spike close to zero for localised difference pixels.

There is an additional consideration that will modify the behaviour described above. The presence of spatial correlation will result in an effective reduction in the number of degrees of freedom, since the images will not be fully independent, and therefore applying the reflattening formula with $n = 5$ will result in overflattening. Examples of this effect are given in [5]. If the correct value for n to produce a uniform probability distribution for background pixels could be found, than the difference between this and $n = 5$ could provide a quantitative measure of the total amount of spatial correlation in the difference image. Furthermore, measurements of this quantity for Monte-Carlo data to which Gaussian blurring with various kernel widths had been applied would relate the quantity to a meaningful scale. However, it is highly likely that the reduction in the number of effective degrees of freedom would be fractional, and so the first requirement of this technique would be a generalisation of Eq. 2 to non-integer values of n.

As an aside, there seems to be some inconsistency in the literature over the naming convention for the various members of the family of gamma functions (the incomplete, regularised etc.). In order to avoid any confusion, the naming convention used here is laid out in full in Appendix B.

5.1 Non-Integer n Extension to the Renormalisation Formula

The aim here is to generalise the equation

$$p' = p\sum_{j=0}^{n-1}\frac{(-\ln p)^j}{j!} \tag{69}$$

to non-integer values of n. The presence of the factorial in the denominator immediately demands the use of the complete gamma function to generalise it to non-integer arguments. This in turn suggests the possibility of rewriting the whole expression in terms of gamma functions. The lower incomplete gamma function $\gamma(a, z)$ has a series expansion

$$\gamma(a, z) = (a-1)!(1 - e^{-z}\sum_{i=0}^{a-1}\frac{z^i}{i!}) \tag{70}$$

Also

$$\Gamma(a) = (a-1)! \tag{71}$$

so

$$\frac{\gamma(a, z)}{\Gamma(a)} = 1 - e^{-z}\sum_{i=0}^{a-1}\frac{z^i}{i!} \tag{72}$$

Rearranging and using Eq. 6 gives

$$e^{-z}\sum_{i=0}^{a-1}\frac{z^i}{i!} = 1 - \frac{\gamma(a, z)}{\Gamma(a)} = \frac{\Gamma(a) - \gamma(a, z)}{\Gamma(a)} = \frac{\Gamma(a, z)}{\Gamma(a)} \tag{73}$$

Notice that the R.H.S. is itself a gamma function: the regularised upper incomplete gamma function $Q(a, z)$. Now, putting $a = n$ and $z = -\ln(p)$ gives

$$p' = p\sum_{j=0}^{n-1}\frac{(-\ln p)^j}{j!} = Q(n, -\ln(p)) \tag{74}$$

The original probability renormalisation equation can therefore be expressed in terms of the regularised upper incomplete gamma function, generalising it to non-integer values of n.

Figure 4 shows plots of the two formulations of the probability renormalisation equation against n for a range of values of p. The points were plotted using the original formulation (Eq. 2) and show the values at integer n for $n = 1$ to 5. The lines were plotted using the regularised upper incomplete gamma function $Q(n, -\ln(p))$. The two formulations of the equation are equal at integer values, and $Q(n, -\ln(p))$ is smooth and continuous between integer values of n.

5.2 Reason for the Link to Gamma Functions

In the course of providing a non-integer extension to the renormalisation formula, the proof given above raises a further question: why does the renormalisation formula take the form of a gamma function? In order to answer this question, a further derivation for the renormalisation function, this time using purely statistical arguments, must be explored.

The original quantity p, which the renormalisation equation seeks to transform into a uniform distribution, is the product of a number of quantities drawn from uniform distributions

$$p = \prod_{i}^{n}\omega_i = \omega_1\omega_2\omega_3...\omega_n \tag{75}$$

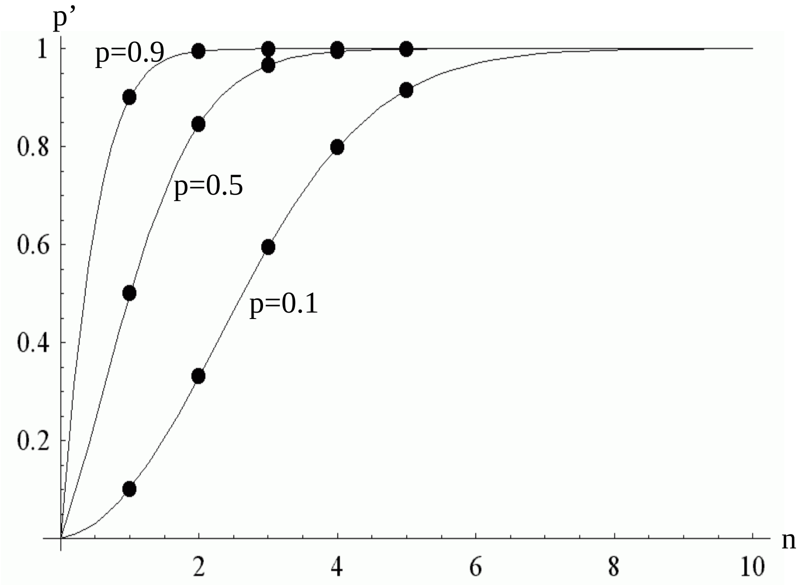

Figure 4: Plots of the probability reflattening equation against n, for p=0.1, 0.5 and 0.9. The points show the values at integer n, plotted using the original equation (expressed as a sum) and the lines were plotted using $Q(n, -\ln(p))$.

Transforming into log space, this can be expressed as a sum

$$\log p = \sum_{i}^{n} \log \omega_i = \log \omega_1 + \log \omega_2 + \log \omega_3 + ... + \log \omega_n \tag{76}$$

The first stage in the proof is to find an expression for the distribution of $\log p$, which in turn requires a distribution for $\log \omega_i$. If ω_i has a uniform distribution in x,

$$P(x) = 1 \tag{77}$$

then in the histogram of ω_i, there will be equal numbers of data points in bins of equal width in x. Considering a bin of width δx at position x,

$$width = (x + \delta x) - x \tag{78}$$

Transforming into log space, such that $x' = \log x$ and $x = e^{x'}$

$$width = e^{x' + \delta x'} - e^{x'} = e^{x'}(e^{\delta x'} - 1) \tag{79}$$

The exponential function has the series expansion

$$e^{\alpha} = 1 + \alpha + \frac{\alpha^2}{2!} + \frac{\alpha^3}{3!} ... \tag{80}$$

Ignoring terms in $\delta x'^2$ or higher gives

$$width = e^{x'}(1 + \delta x' - 1) = e^{x'} \delta x' \tag{81}$$

So in log space, bins of equal width $\delta x'$ in x' contain $e^{x'}$ data points, and so

$$P(x') = e^{x'} \quad for \quad x' = -\infty \ to \ 0 \tag{82}$$

This is an exponential distribution i.e. it has the form

$$P(x) = \lambda e^{-\lambda x'} \tag{83}$$

with $\lambda = 1$ and $x' = -\log x$.

In order to find the distribution of a quantity being the sum of n quantities drawn from exponential distributions, we must perform the n-fold convolution of the exponential distribution, i.e. repeat

$$if \ Z = X + Y \quad P_Z(z) = \int_{-\infty}^{+\infty} P_X(z - y) P_Y(y) dy \tag{84}$$

$n-1$ times. Using the general form of the exponential distribution

$$P_2(x') = \int_0^{x'} \lambda e^{-\lambda(x'-y')} \lambda e^{-\lambda y'} dy' = \int_0^{x'} \lambda^2 e^{-\lambda y'} dy' = \lambda^2 x' e^{-\lambda x'} \tag{85}$$

$$P_3(x') = \int_0^{x'} \lambda e^{-\lambda(x'-y')} \lambda^2 y' e^{-\lambda y'} dy' = \int_0^{x'} \lambda^3 y' e^{-\lambda y'} dy' = \lambda^3 \frac{x'^2}{2} e^{-\lambda x'} \tag{86}$$

and

$$P_4(x') = \int_0^{x'} \lambda e^{-\lambda(x'-y')} \lambda^3 \frac{y'^2}{2} e^{-\lambda y'} dy' = \int_0^{x'} \lambda^4 \frac{y'^2}{2} e^{-\lambda y'} dy' = \lambda^4 \frac{x'^3}{6} e^{-\lambda x'} \tag{87}$$

Clearly the series will become

$$P_n(x') = \lambda^n \frac{x'^{(n-1)}}{(n-1)!} e^{-\lambda x'} \tag{88}$$

A more formal proof that this is the form of the series can be performed by induction, by assuming Eq. 88 and convolving with an exponential distribution to show that $P_{n+1}(x')$ also obeys the series form

$$P_{n+1}(x') = \int_0^{x'} \lambda e^{-\lambda(x'-y')} \lambda^n \frac{y'^{(n-1)}}{(n-1)!} e^{-\lambda y'} dy' = \lambda^{n+1} \frac{x'^n}{n!} e^{-\lambda x'} \tag{89}$$

In order to transform a the distribution of a quantity having a distribution given by Eq. 88 into a uniform distribution, we can simply apply the probability integral transform: that is, replace the value of any given data point x' with the integral of the distribution below that point

$$\int_0^{-x'} \lambda^n \frac{x'^{(n-1)}}{(n-1)!} e^{-\lambda x'} dx' \tag{90}$$

Note that the limits correspond to the range over which the original exponential distributions are defined in the specific case at hand. This integral responds to treatment by parts, differentiating the term in x and integrating the term in e^x

$$\frac{\lambda^n}{(n-1)!} \int_0^{-x'} x'^{(n-1)} e^{-\lambda x'} dx' = \frac{\lambda^n}{(n-1)!} \left[\left. \frac{x'^{n-1} e^{-\lambda x'}}{\lambda} \right|_0^{-x'} - \int_0^{-x'} -\frac{1}{\lambda} e^{-\lambda x'} (n-1) x'^{(n-2)} dx' \right] \tag{91}$$

$$= \frac{\lambda^{n-1}}{(n-1)!} x'^{n-1} e^{-\lambda x'} + \int_0^{-x'} \frac{\lambda^{n-1}}{(n-2)!} x'^{(n-2)} e^{-\lambda x'} dx' \tag{92}$$

$$= \frac{(\lambda x')^{n-1}}{(n-1)!} e^{-\lambda x'} + \frac{\lambda^{n-1}}{(n-2)!} \int_0^{-x'} x'^{(n-2)} e^{-\lambda x'} dx' \tag{93}$$

which also responds to parts in the same way, and so the sequence becomes

$$D_n(x') = e^{-\lambda x'} \sum_{i=0}^{n-1} \frac{(\lambda x')^i}{i!} \tag{94}$$

Substituting in the values $\lambda = 1$ and $x' = -\log p$, and equating the solution to the reflattened variable p' gives

$$p' = p \sum_{i=0}^{n-1} \frac{(-\log p)^i}{i!} \tag{95}$$

QED.

In a Poisson process the distribution of waiting times to the next Poisson event is an exponential distribution, with λ in the above equations corresponding to the rate parameter. Therefore, the distribution of waiting times to the n^{th} next event is given by the sum of n quantities drawn from exponential distributions i.e. the n-fold convolution of the exponential distribution, and this distribution is known as the gamma distribution, since it can be expressed in terms of gamma functions. The product of n drawings from uniform distributions can also be expressed as the sum of n drawings from exponential distributions by transforming into log space, and the probability renormalisation equation is the integral of the distribution of this sum. So, the distribution defined by Eq. 88 is the gamma distribution, and its cumulative density function Eq. 94 gives the probability renormalisation equation Eq. 95. The coincidence occurs because both situations involve summing quantities drawn from exponential distributions.

6 Summary

This document has provided derivations for the standard probability renormalisation technique in 1D (trivial), 2D, 3D and 4D by direct integration, and in 2D and nD using a more complex geometrical method, and in nD using an iterative method.

As an aside, it should in theory be possible to derive the jD result by an equivalent approach to the geometrical method, but performed in the original space. The expression for the element of integration will involve the arc length of the curve $x_1 x_2 ... = n$, which in 2D can be derived using the standard equation

$$L_a^b = \int_a^b \sqrt{1 - [f'(x)]^2} dx = \int_n^1 \sqrt{1 - \frac{n^2}{x^4}} dx \tag{96}$$

However, this integration is not trivial and so this approach has not been explored further.

This report has also shown that the probability renormalisation equation can be expressed in terms of the regularised upper incomplete gamma function,

$$p' = p \sum_{j=0}^{n-1} \frac{(-\ln p)^j}{j!} = Q(n, -\ln(p)) \tag{97}$$

extending it to non-integer values of n. The function is equal to the previous expression at integer n, and smooth and continuous in the ranges between these values. Thus it is possible to correct for the reduction in the effective number of degrees of freedom when the equation is applied to a quantity p formed from quantities ω featuring spatial correlation, and thus also a method for estimating the extent of this spatial correlation. It has also shown the reason for the link between uniform probabilities and gamma functions, providing another proof for the renormalisation formula.

A Proofs for Standard Results

Three standard geometrical results have been used in the jD derivation above:

- The hypervolume of any triangular-faced figure in jD can be expressed as

  $$hypervolume = \frac{1}{j} bh \tag{98}$$

  where b is the hypervolume of the base and h is the height.

- As a special case of (1), if the corner of a jD hypercube is chopped off such that the distances between the vertex of the hypercube that is retained and the new vertices formed at the intersections of the chopping hyperplane with the edges of the hypercube are all equal (let this distance be called a), the hypervolume of the resulting figure (which might be called a right isosceles form, see above), can be expressed as

  $$hypervolume = \frac{a^j}{j!} \tag{99}$$

- The hypervolume of an equilateral (or regular) simplex in jD, with side length a, can be expressed as

  $$hypervolume = \frac{a^j}{j!} \sqrt{\frac{j+1}{2^j}} \tag{100}$$

Since these are standard results, detailed derivations or proofs will not be presented here. However, some brief comments can be made.

A.1 Standard Result 1

Here is a brief proof by induction for this formula for the case of triangular-faced simplices. A jD simplex can be decomposed into (j-1)D simplex elements along its height e.g. a tetrahedron can be decomposed into a stack of triangular elements. Now, each dimension of the element will vary linearly with the height, so the hypervolume of

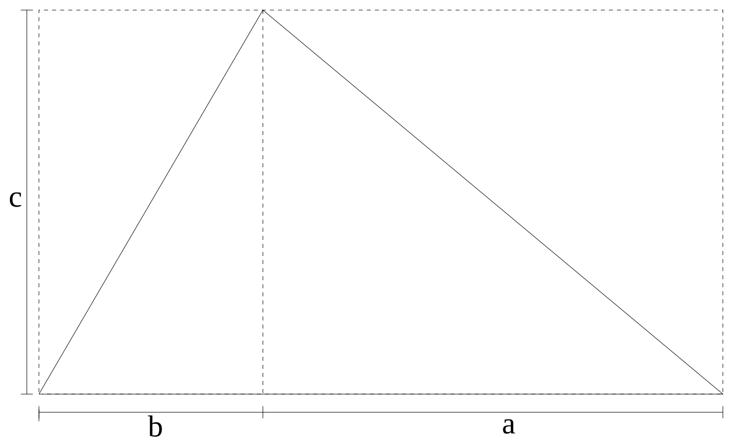

Figure 5: Construction used to prove that the area of a general triangle can be expressed as one half of the base times the height, by considering the two right-angled triangles into which it is divided by its height, each of which has half of the area of its exscribed rectangle.

the element will vary as $(height)^{(j-1)}$. Let the height and base of the jD simplex be h_j and b_j, and the height and base of its $(j-1)D$ elements be $h_{(j-1)}$ and $b_{(j-1)}$. Then

$$b_j = \frac{1}{(j-1)} h_{(j-1)} b_{(j-1)} \tag{101}$$

So the hypervolume of the element is given by

$$volume(x) = \frac{1}{(j-1)} h_{(j-1)} b_{(j-1)} \frac{x^{(j-1)}}{h_j^{(j-1)}} dx \tag{102}$$

where x is the dimension along which the height of the figure lies. Integrating between 0 and the height, h_j

$$\int_0^{h_j} \frac{h_{j-1} b_{j-1}}{(j-1) h_j^{(j-1)}} x^{(j-1)} dx = \frac{h_{j-1} b_{j-1}}{(j-1) h_j^{(j-1)}} \frac{x^j}{j} \bigg|_0^{h_j} \tag{103}$$

$$= \frac{1}{(j-1)} h_{(j-1)} b_{(j-1)} \frac{h_j^j}{h_j^{(j-1)}} \frac{1}{j} = \frac{1}{j} b_j h_j \tag{104}$$

So, if the formula is true for (j-1)D it is true for jD. It then only remains to prove it for 2D. Consider Fig. 5. Any triangle is divided into two right-angled triangles by its height. Each of these has half of the area of its exscribed rectangle. Thus the total area of the original triangle is given by

$$area = \frac{ac}{2} + \frac{bc}{2} = \frac{c(a+b)}{2} = \frac{1}{2} base \times height \tag{105}$$

A.2 Standard Result 2

The second standard result is an extension of the first, and can be derived from it. Consider Fig. 6. Let the sides of a jD right isosceles simplex which terminate at the right-angled vertex be of length a. Let one of these sides be the height. The form of the base is then also a right isosceles simplex with right-angled vertex sides of length a, but this time it is in $(j-1)D$. In turn, the base figure will have height a and a base which is a right isosceles simplex in $(j-2)D$ with right-angled vertex sides of length a. The sequence continues until we reach the 2D case, and so the hypervolume of the original figure can be expressed as

$$hypervolume = \frac{a}{j} \frac{a}{(j-1)} \frac{a}{(j-2)} ... \frac{a}{j-(j-1)} = \frac{a^j}{j!} \tag{106}$$

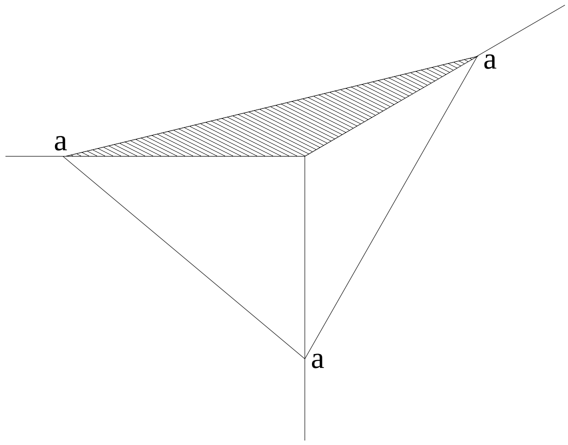

Figure 6: Construction used to prove standard result 2 for right isosceles simplexes: see main text for explanation. Notice that, if one of the right-angled vertex sides is defined as the height, then the base is itself a right isosceles simplex of right-angled vertex side length a, but with its dimensionality lowered by 1.

A.3 Standard Result 3

This formula is a special case of a much more general result, the most powerful of all those used here. For a general convex polytope in nD

$$hypervolume = \frac{1}{j!}\sqrt{\frac{1}{2^j}D_{n\times n}}u^j \tag{107}$$

Here D is the determinant of an $n \times n$ matrix containing the distances between the vertices of the polytope, and u^n denotes a unit cube in jD that countably tesselates the jD space. This is known as the Cayley-Menger determinant formula, and is in fact itself a special case of the more general Heron's formula. No attempt will be made to prove or derive this equation here. However, it can easily be shown to give the well-known equations for the area of an equilateral triangle

$$area = \frac{\sqrt{3}a^2}{4} \tag{108}$$

and for the volume of a regular tetrahedron

$$volume = \frac{a^3}{6\sqrt{2}} \tag{109}$$

where a is the side length in both cases. These formulae can easily be proven in a number of ways (geometrical construction, integration or using standard result 1) if desired.

B Gamma Functions

There seems to be some inconsistency in the literature over the naming (though not the definitions) of the gamma function and related functions. Therefore, the definitions used here are as follows. The complete gamma function $\Gamma(a)$ extends the definition of the factorial to non-integer values and is given by

$$\Gamma(a) = (a-1)! = \int_0^\infty t^{a-1}e^{-t}dt \tag{110}$$

This can be generalised to the upper incomplete gamma function $\Gamma(a,z)$ and the lower incomplete gamma function $\gamma(a,z)$,

$$\Gamma(a,z) = \int_z^\infty t^{a-1}e^{-t}dt \tag{111}$$

$$\gamma(a,z) = \int_0^z t^{a-1}e^{-t}dt \tag{112}$$

Clearly

$$\Gamma(a) = \Gamma(a,z) + \gamma(a,z) \tag{113}$$

Two more functions, the regularised incomplete gamma functions $P(a, z)$ and $Q(a, z)$, can also be defined,

$$P(a, z) = 1 - Q(a, z) = \frac{\gamma(a, z)}{\Gamma(a)} \tag{114}$$

$$Q(a, z) = 1 - P(a, z) = \frac{\Gamma(a, z)}{\Gamma(a)} \tag{115}$$

These will be referred to here as the regularised lower incomplete gamma function $P(a, z)$ and the regularised upper incomplete gamma function $Q(a, z)$.

References

[1] ALEPH Collaboration. "A Precise Measurement of $\Gamma_{Z \to b\bar{b}} / \Gamma_{Z \to hadrons}$." *Physics Letters* **B313**, pp. 535-548, 1993.

[2] P.A. Bromiley, N.A. Thacker, and P. Courtney, Non-parametric Image Subtraction using Grey Level Scattergrams. Proc. BMVC 2000, pp. 795-804. BMVA, 2000.

[3] P.A. Bromiley, N.A. Thacker and P. Courtney, Non-parametric Image Subtraction using Grey Level Scattergrams. Image and Computer Vision, BMVC 2000 Special Edition, In press.

[4] P.A. Bromiley, N.A. Thacker, and P. Courtney, Non-parametric Image Subtraction for MRI. Proc. MIUA 2001, pp. 105-108, BMVA, 2001.

[5] P.A. Bromiley, N.A. Thacker, M.J.L. Scott, M. Pokrić, A.J. Lacey and T.F. Cootes, Bayesian and Non-Bayesian Probabilistic Models for Image Analysis, IVC Special Edition: The use of Probabilistic Models in Computer Vision, in press, 2001.

[6] P.A. Bromiley, M. Pokric, N.A. Thacker and A. Jackson, Detection of MS Lesions in MRI scans using Non-Parametric Image Subtraction, Proc. ISMRM 2002, submitted.